OPTGAME 2.0:
An Algorithm for Equilibrium Solutions of *N*-Person Discrete-Time (Non-)Linear Dynamic Games*

Doris A. Behrens[1] and Reinhard Neck[1]

[1] Department of Economics, University of Klagenfurt, Universitätsstraße 65-67, A-9020 Klagenfurt

Abstract: We present an algorithm to solve N-person discrete-time LQ games exactly, and discrete-time non-linear quadratic games approximately by means of an appropriate linearization procedure, where $N>2$. Among the different solutions of these dynamic games are open-loop and feedback Nash and Stackelberg equilibrium and Pareto-optimal solutions, where the derivation of each of these solutions is explained in detail in the present paper.

Keywords: Dynamic game theory, non-cooperative equilibrium solution, cooperative equilibrium solution, nonlinear dynamic systems, optimal economic policies.

1. Introduction

Dynamic game theory serves as mathematical instrument to illustrate questions of strategic interdependences between policy makers, in particular in the analysis of possible advantages of international policy coordination. Different solution concepts are distinguished in the literature on dynamic games according to the information patterns and the strategy spaces of the players, which correspond to the degree of commitment assumed for the players (see, e.g., Basar and Olsder (1995), Dockner and Neck (1988), Leitmann (1974), Mehlmann (1991), Petit (1990)).

Therefore, we distinguish between non-cooperative and cooperative solution concepts of dynamics games. In the former case we exclude binding agreements between the players, while these agreements are assumed to be established and to hold in the latter case. Among non-cooperative solutions we distinguish between Nash equilibrium solutions, where no player can improve her (his) performance by one-sided deviations from the equilibrium strategy, and Stackelberg equilibrium solutions, where the players have asymmetric roles. For the Nash and for the Stackelberg equilibrium solution of a dynamic game additional assumptions concerning the information pattern of the players can be made. These assumptions specify the informational basis of each player's decision. Here, we consider open-loop information patterns, where the player's strategies depend only on the initial state of the dynamic system, and feedback information patterns, where the strategies depend on the current state of the system (but not on the initial conditions).

* This report was supported by the Austrian Science Foundation under the contract no. P12745-OEK.

According to the interpretation of the distinction between open-loop and feedback Nash equilibrium solutions suggested by Reinganum and Stokey (1985), with the open-loop information pattern, each player can be imagined to commit herself (himself) at the initial period a priori to all future actions she (he) will take ("path strategies"). In the feedback information pattern, players can be imagined to observe the current values of the state variable and to react upon them by choosing their actions according to a "decision rule", i.e., a strategy specified at the initial period which makes the values of the control variables dependent on the current values of the state variables at each (future) point in time.

We present an extension of the algorithm OPTGAME 1.0 (Hager et al. (2000)) for the calculation of the approximate solutions of discrete-time non-linear quadratic games. I.e., the objective function is assumed to be quadratic in the deviations of states and control variables from their respective desired target-values, and will be optimized for a pre-specified period of time subject to a nonlinear autonomous system. The extension is solely done with respect to the number of players. Where OPTGAME 1.0 restricts the number of players to two, OPTGAME 2.0 allows the calculation of the feedback and the open-loop Nash and Stackelberg equilibrium solutions, and the cooperative Pareto-optimal solutions for an arbitrary number of players.

We present the discrete-time (non)-linear-quadratic dynamic game form in Section 2, give a detailed description of the elements of the algorithm OPTGAME 2.0 in Section 3, and conclude the present paper with a short summary of the insights generated in the previous Sections in Section 4.

2. The Game-Theoretic Problem

OPTGAME 2.0 approximates solutions for game-theoretic problems with a quadratic objective function and a non-linear dynamic model in discrete time. Thus, for the calculation of the non-cooperative solutions we consider the following intertemporal quadratic loss functions of the players $i, (i = 1,\dots, N)$,

$$J_i(T) = \sum_{t=1}^{T} L_{it} = \sum_{t=1}^{T} \tfrac{1}{2}\left(X_t - \tilde{X}_{it}\right)' Q_{it} \left(X_t - \tilde{X}_{it}\right), \qquad i = 1,\dots, N . \quad (1)$$

T denotes the terminal period of the finite planing horizon. Hence, the control variables, $u_{i(T+1)}$ $(i = 1,\dots, N)$, are not considered in the optimization. X_t denotes an r-dimensional stacked "state" vector consisting out of n_s-dimensional state variable (summarizing the information available about the dynamical system), the n_1-dimensional control variable accessible for player 1, the n_2-dimensional control variable accessible for player 2, the n_3-dimensional control variable accessible for player 3, etc.,

$$X_t = \begin{pmatrix} x_t \\ u_{1(t+1)} \\ \vdots \\ u_{N(t+1)} \end{pmatrix}_{r\times 1}, \quad r = n_s + n_1 + \ldots + n_N . \tag{2}$$

X_{it} $(i = 1, \ldots, N)$ denotes the r-dimensional stacked "desired targets" vector which contains the "ideal levels" of the state and control variables for either of the players,

$$\tilde{X}_{it} = \begin{pmatrix} \tilde{x}_{it} \\ \tilde{u}_{i1(t+1)} \\ \vdots \\ \tilde{u}_{iN(t+1)} \end{pmatrix}_{r\times 1}, \quad \begin{matrix} r = n_s + n_1 + \ldots + n_N \\ i = 1, \ldots, N \end{matrix} . \tag{3}$$

Furthermore, the penalty matrices are defined as

$$Q_{it} = \begin{pmatrix} Q_{it}^{x} & 0 & \cdots & 0 \\ 0 & Q_{i(t+1)}^{u_1} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & Q_{i(t+1)}^{u_N} \end{pmatrix}_{r\times r}, \quad \begin{matrix} r = n_s + n_1 + \ldots + n_N \\ i = 1, \ldots, N \end{matrix} . \tag{4}$$

The matrices Q_{it}^{x}, $Q_{i(t+1)}^{u_1}$, ..., and $Q_{i(t+1)}^{u_N}$ describe the penalties for deviations of the state variable, x_t, in time period t, and the control variables, $u_{1(t+1)}, \ldots, u_{N(t+1)}$, in time period $t+1$ from their desired target values. Note that the matrices containing the penalties for failing the desired control values are assumed to have full rank.

For the calculation of the cooperative (Pareto-optimal) solution we modify the objective functional to be

$$J(T) = \mu_1 L_1(T) + \ldots + \mu_N L_N(T) = \sum_{t=1}^{T} \{\mu_1 L_{1t} + \ldots + \mu_N L_{Nt}\}, \qquad \sum_{i=1}^{N} \mu_i = 1 . \tag{5}$$

The dynamic system – which may be an economic model – is assumed to be given by a system of nonlinear difference equations,

$$x_t = f(x_{t-1}, x_t, u_{1t}, u_{2t}, \ldots, u_{Nt}, y_t), \qquad t = 1, \ldots, T , \tag{6}$$

where we have the initial condition, $x(0) = x_0$; Furthermore, y_t $(t = 1, \ldots, T)$, denotes a q-dimensional vector of non-controlled, exogenous variables. f is a vector valued function, where $f^{\alpha}(\ )\,(j = 1, \ldots, n_s)$ denotes the α^{th} component of f.

The assumption of a first-order system of difference equations in (6) is not really restrictive, as higher-order difference equations can be reduced to systems of first-order difference equations by suitably redefining variables as new state variables and augmenting the state vector. Also the assumption of a quadratic objective function, although of a special form, can be interpreted as a second-order Taylor-series approximation to a more general objective function. Thus, the class of problems to be solved by our algorithm OPTGAME 2.0 is rather broad.

3. Elements of the Algorithm OPTGAME 2.0

The algorithm OPTGAME 2.0 starts from computing a tentative path of the state vector from the nonlinear system equations – using the Gauss-Seidl algorithm - with a given tentative path for the control variables. Then the algorithm linearizes the system equations at the reference values obtained before, replacing the nonlinear autonomous system by a linear non-autonomous one. Then, the algorithm calculates numerically the Nash, Stackelberg, and Pareto-optimal solutions of nonlinear, quadratic deterministic games (with a finite planing horizon) under open-loop and feedback information structure following prior work from Hager et al. (2000), Hathaway (1992), and Chow (1975).

Note that the term "OPTGAME" denotes both, the computer algorithm and its implementation, where the implementation part consists of a set of procedures which are implemented in the programming language GAUSS. GAUSS is a high level matrix programming language specializing in commands, functions, and procedures for data analysis and statistical applications. This interplay is of special interest for the application of OPTGAME 2.0 in the field of optimal short-run and long-run fiscal policies towards the European Monetary Union. Furthermore, GAUSS includes a variety of routines which perform standard matrix operations, e.g. routines to calculate determinants, matrix inverses, decompositions, eigenvalues and eigenvectors, and condition numbers.

Input of the algorithm for $t = 1,\ldots,T$ and $i, j = 1,\ldots,N$:

- length of the planing horizon	T
- system function	f
- initial values of the state variables	$x(0) = x_0$
- path of exogenous variables not subject to control	y_t
- weighting matrices of objective function	Q_{it}
- target path for state variables	$\tilde{x}_{it}$
- target path for control variables	$\tilde{u}_{ijt}$

Output of algorithm for $t = 1,\ldots,T$ and $i = 1,\ldots,N$:

- optimal path of the state variables	x_t^*
- 5 types of optimal paths for the control variables	u_{it}^*
- quadratic loss functions evaluated along the optimal paths	$J_i^*(T)$

3.1 Linearization of the System Equations

In most game-theoretical models the system dynamics of the form (6) are a convenient starting point. For the calculation of the Nash, Stackelberg, and Pareto-optimal equilibrium strategies, however, the state-space representation will be more adequate. This representation does not include x_t at the right-hand side of the implicit function (6). According to Neck and Matulka (1992) and following Chow (1975) it is easy to show how to eliminate x_t in the course of a linearization of system (6).

- With the given exogenous non-controlled vector, y_t, assumed tentative control paths, $\hat{u}_{it}$ $(i = 1,\ldots, N)$, (which are determined either by historical values or by foregoing optimization procedures), and the lagged tentative state vector, $\hat{x}_{t-1}$, (starting with $x(0) = x_0$) the autonomous non-linear system (6) can be solved for all t using the numerical Gauss-Seidl approximation (a well-known non-linear equation-solving method), and, thus, one derives a reference path for the state vector over the entire time horizon. That is, that for known values of $\hat{x}_{t-1}, \hat{u}_{1t},\ldots, \hat{u}_{Nt}$ and y_t we can compute a value $\hat{x}_t$ such that

$$\bar{x}_t - f(\bar{x}_{t-1}, \bar{x}_t, \bar{u}_{1t}, \bar{u}_{2t}, \ldots, \bar{u}_{Nt}, y_t) = 0, \qquad t = 1,\ldots T, \tag{7}$$

  by straightforward application of the Gauss-Seidl technique.

- Then, we linearize the vector-valued system function, $f(\ )$, numerically around the reference values, $\hat{x}_{t-1}, \hat{x}_t, \hat{u}_{1t}, \ldots, \hat{u}_{Nt}$, and the given reference path y_t according to the first Taylor approximation and gain the following approximate non-autonomous system equations:

$$\begin{aligned} x_t &= A_t x_{t-1} + \sum_{j=1}^{N} B_{jt} u_{jt} + s_t, \qquad t = 1,\ldots, T, \\ x_0 &= x(0), \end{aligned} \tag{8}$$

  where we have defined the $n_s \times n_s$-matrix A_t, the $n_s \times n_i$-matrices B_{it} (for $i = 1,\ldots, N$), and the n_s-dimensional vector s_t, as

$$\left.\begin{aligned} A_t &:= \left(I_{n_s} - F_{x_t}\right)^{-1} F_{x_{t-1}}, & & (9) \\ B_{it} &:= \left(I_{n_s} - F_{x_t}\right)^{-1} F_{u_{kt}}, \qquad i = 1,\ldots, N, & & (10) \\ s_t &:= \bar{x}_t - A_t \bar{x}_{t-1} - \sum_{j=1}^{N} B_{jt} \bar{u}_{jt}, & & (11) \end{aligned}\right\} \qquad t = 1,\ldots, T,$$

  where I denotes the $n_s \times n_s$ identity matrix. Here, and in what follows, we require that the first and second derivatives of the system function with respect to $x_{t-1}, x_t, u_{1t}, \ldots, u_{Nt}, s_t$ exist and are continuos, and we use the following abbreviations: $F_{x_{t-1}}$ denotes an $n_s \times n_s$ -matrix where its elements are defined by

$$\left(F_{x_{t-1}}\right)_{ij} = \frac{\partial f^i(\,)}{\partial x_{t-1}^j}, \quad \begin{matrix} i = 1,\ldots,n_s \\ j = 1,\ldots,n_s \end{matrix}, \quad t = 1,\ldots,T, \tag{12}$$

F_{x_t} denotes an $n_s \times n_s$ -matrix where its elements are defined by

$$\left(F_{x_t}\right)_{ij} = \frac{\partial f^i(\,)}{\partial x_t^j}, \quad \begin{matrix} i = 1,\ldots,n_s \\ j = 1,\ldots,n_s \end{matrix}, \quad t = 1,\ldots,T, \tag{13}$$

$F_{u_{kt}}$ denotes an $n_s \times n_k$ -matrix ($k = 1,\ldots,N$) where its elements are defined by

$$\left(F_{u_{kt}}\right)_{ij} = \frac{\partial f^i(\,)}{\partial u_{kt}^j}, \quad \begin{matrix} i = 1,\ldots,n_s \\ j = 1,\ldots,n_k \end{matrix}, \quad t = 1,\ldots,T, \quad k = 1,\ldots,N. \tag{14}$$

Hence, it has to be assumed that the term "$I - F_{x_t}$" is non-singular. The matrices and vectors defied above are time-dependent functions of the reference paths along which they have been evaluated. If these paths change, the matrices will change too.

3.2 Computation of the Equilibrium Solutions

In a deterministic optimal control setting, the solutions obtained, e.g., by the application of the minimum principle or the dynamic programming method, equal. Though deterministic dynamic game models can be solved by using essentially the same techniques as for solving deterministic optimal control models, the choice of the solution technique determines the qualitative results of the game, i.e., the information pattern. E.g., the application of the minimum principle generates the open-loop solutions, while the application of the dynamic programming technique determines the feedback solutions. Hence, using the appropriate optimization technique corresponding to the desired information structure of the game yields so-called Riccati equations which can be solved by backward integration. The terminal conditions for the Riccati matrices can be defined easily as already shown by Kendrick (1981) in control theory.

Furthermore, for the solutions derived in Sections 3.2.1 - 3.2.4 we need to define the so-called feedback matrices, G_{it} and g_{it} (for $i = 1,\ldots,N$ and $t = 1,\ldots,T$), which are described separately for each solution concept below. By further substitution of these feedback matrices into linear relations in the preceding state variable, x_{t-1}, we are able to derive the values of the optimal control variables expressed in feedback form, u_{it},

$$u_{it}^{*} = G_{it} x_{t-1}^{*} + g_{it}, \qquad t = 1,\ldots,T, \quad i = 1,\ldots,N, \tag{15}$$

and the optimal state values, x_t^{*},

$$x_t^{*} = K_t\, x_{t-1}^{*} + k_t, \qquad t = 1,\ldots,T, \quad i = 1,\ldots,N, \tag{16}$$

for each player by forward iteration, i.e. for $t = 1,\ldots,T$, using the initial values of the states, $x_0 = x(0)$, where

$$K_t = A_t + \sum_{j=1}^{N} B_{jt} G_{jt}, \qquad t = 1,\ldots,T, \tag{17}$$

$$k_t = s_t + \sum_{j=1}^{N} B_{jt} g_{jt}, \qquad t = 1,\ldots,T. \tag{18}$$

3.2.1 The Feedback Nash Equilibrium Solution

The feedback Nash equilibrium solution is generated using the dynamic programming (Jacobi-Hamilton-Bellman) technique, which is discussed in Kydland (1975) and Oudiz-Sachs (1985). The dynamic programming solution in the two-dimensional case, i.e. $N = 2$, is derived in de Zeeuw and van der Ploeg (1991), Hatheway (1992) or Hager et al. (2000).

PROPOSITION 1

The solution of the feedback Nash game with N players is given by the following procedure: Starting with the quadratic tracking form of the objective function (1) for the terminal period $t = T$ and proceeding by minimization of the cost-to-go function step by step towards the initial node, the Riccati equations, H_{it} and h_{it}, can be solved recursively backwards in time, i.e. $t = T, T-1,\ldots,2$,

$$H_{i(t-1)} = K_t' H_{it} K_t + Q_{i(t-1)}^{x} + \sum_{j=1}^{N} G_{jt}' Q_{it}^{u_j} G_{jt}, \tag{19}$$

$$H_{iT} = Q_{iT}^{x},$$

$$h_{i(t-1)} = K_t' \left(h_{it} - H_{it} k_t\right) + Q_{i(t-1)}^{x} \tilde{x}_{i(t-1)} + \sum_{j=1}^{N} G_{jt}' Q_{it}^{u_j} \left(\tilde{u}_{jit} - g_{jt}\right) \tag{20}$$

$$h_{iT} = Q_{iT}^{x} \tilde{x}_{iT},$$

using definitions (17) and (18), yielding Riccati matrices for either of the players $i\ (i = 1,\ldots,N)$, where the feedback matrices are determined by the equations

$$G_{it} = -\left(Q_{it}^{u_i}\right)^{-1} B_{it}{}' H_{it} E_t A_t \,, \tag{21}$$

$$g_{it} = -\left(Q_{it}^{u_i}\right)^{-1} B_{it}{}' \left(H_{it} E_t F_t - h_{it}\right) + \tilde{u}_{iit} \,, \tag{22}$$

where

$$E_t := \left(I + \sum_{j=1}^{N} B_{jt} \left(Q_{jt}^{u_j}\right)^{-1} B_{jt}{}' H_{jt} \right)^{-1} \text{, and} \tag{23}$$

$$F_t := s_t + \sum_{j=1}^{N} B_{jt} \left(\tilde{u}_{jjt} + \left(Q_{jt}^{u_j}\right)^{-1} B_{jt}{}' h_{jt} \right). \tag{24}$$

For proof see Appendix.

The solution of the equations (19) and (20) yields the Riccati matrices which define the feedback matrices determined by the equations (21) and (22). Then equations (15) describe a linear relation between the optimal control actions of the players in dependence from the previous state variables. With the Riccati matrices and the feedback matrices the system equations can be solved by forward iteration, which determines the optimal equilibrium solution paths for state and control variables.

3.2.2 The Open-Loop Nash Equilibrium Solution

Conceptually, the open-loop Nash solution can be viewed as follows: From the initial state each of the policy-makers chooses optimal values for her (his) controls for the entire duration of the planing horizon under the assumption that the other sovereign policy authorities act in a similar fashion. Policy decisions are performed simultaneously by all players. Thus, in the beginning of the game each of the players makes a binding commitment to stick to a chosen policy. As long as these commitments hold (and by construction they do), the solution is optimal in the sense that non of the players can improve her (his) welfare by one-sided deviations from the equilibrium path. The open-loop Nash solution is generated using the minimum principle as illustrated for the two-dimensional case, i.e. $N = 2$, in Basar and Olsder (1995), de Zeeuw and van der Ploeg (1991), Hatheway (1992) or Hager et al. (1999).

PROPOSITION 2

The solution of the open-loop Nash game with N players is determined in the following way: Definition of the (current value) Hamiltonian function with the state equations and the objective function for each player and appropriate differentiation yields the adjoint equations and the necessary conditions for the control variables.

Under the assumption of a linear relation between the co-states and the states, Riccati equations are derived which can be solved by iterating backwards in time,

$$H_{i(t-1)} = A_t{}' H_{it} E_t A_t + Q^x_{i(t-1)}, \qquad H_{iT} = Q^x_{iT}, \tag{25}$$

$$h_{i(t-1)} = A_t{}' \left(h_{it} + H_{it} E_t F_t\right) - Q^x_{i(t-1)} \tilde{x}_{i(t-1)}, \quad h_{iT} = -Q^x_{iT} \tilde{x}_{iT}, \tag{26}$$

yielding the Riccati matrices for either of the players $i\,(i = 1,\dots,N)$, where the feedback matrices are determined by the equation (21) and equation (22), with (23) and

$$F_t := s_t + \sum_{j=1}^{N} B_{jt} \left(\tilde{u}_{jjt} - \left(Q^{u_j}_{jt}\right)^{-1} B_{jt}{}' h_{jt} \right). \tag{24*}$$

For proof see Appendix.

3.2.3 The Pareto-Optimal Solutions

We will derive the Pareto solution in a similar way as the feedback Nash equilibrium solution in Section 3.2.1 using the dynamic programming technique with the only difference that a convex combination of the players' objective functions (equation (5) with the quadratic tracking form of the objective function (1)) is to be optimized. The appropriate application of the minimum principle, however, would yield the optimal cooperative solution as well - as this corresponds to a classical optimal control problem.

PROPOSITION 3

The solution of the cooperative dynamic game with N players is given by the following procedure: Starting with the objective function (5) for the terminal period $t = T$ and proceeding by minimization of the cost-to-go function step by step towards the initial node, the Riccati equations, H_t and h_t, can be solved recursively backwards in time, i.e. $t = T, T-1, \dots, 2$,

$$H_{t-1} = K_t{}' H_t K_t + P^x_{t-1} + \sum_{j=1}^{2} G_{jt}{}' P_{jt} G_{jt}, \tag{27}$$

$$H_T = P^x_T,$$

$$h_{t-1} = K_t{}' \left(h_t - H_t k_t\right) + \tilde{P}^x_{t-1} + \sum_{j=1}^{2} G_{jt}{}' \left(\tilde{P}_{jt} - P_{jt} g_{jt}\right) \tag{28}$$

$$h_T = \tilde{P}^x_T,$$

with the definitions (17) and (18), where

$$P_t := \sum_{j=1}^{N} \mu_j Q_{jt}, \qquad P_{it} := \sum_{j=1}^{N} \mu_j Q_{jt}^{u_i}, \qquad P_t^x := \sum_{j=1}^{N} \mu_j Q_{jt}^x, \tag{29}$$

$$\widetilde{P}_t := \sum_{j=1}^{N} \mu_j Q_{jt} \widetilde{X}_{jt}, \quad \widetilde{P}_{it} := \sum_{j=1}^{N} \mu_j Q_{jt}^{u_i} \widetilde{u}_{jit}, \quad \widetilde{P}_t^x := \sum_{j=1}^{N} \mu_j Q_{jt}^x \widetilde{x}_{jt} . \tag{30}$$

For $i = 1,\dots,N$ the feedback matrices are determined by the equations

$$G_{it} = -(P_{it})^{-1} B_{it}' H_t \overline{E}_t A_t , \tag{31}$$

$$g_{it} = -(P_{it})^{-1} B_{it}' \left(H_t \overline{E}_t \overline{F}_t - h_t\right) + (P_{it})^{-1} \widetilde{P}_{it} , \tag{32}$$

where

$$\overline{E}_t := \left(I + \sum_{j=1}^{N} B_{jt} \left(P_{jt}\right)^{-1} B_{jt}' H_t \right)^{-1}, \text{ and} \tag{33}$$

$$\overline{F}_t := s_t + \sum_{j=1}^{N} B_{jt} \left(P_{jt}\right)^{-1} \left(B_{jt}' h_t + \widetilde{P}_{jt}\right) \tag{34}$$

For proof see Appendix.

3.2.4 The Feedback Stackelberg Equilibrium Solution

Another kind of non-cooperative game is the Stackelberg game. The feedback Stackelberg equilibrium solution is derived similar to the feedback Nash equilibrium solution with one drastical difference, however: While the actions of the players are performed simultaneously for feedback Nash, the players act in a hierarchical way in a Stackelberg game. This leads to the additional consideration of the reactions of the followers (players $i = 2,\dots,N$) to the announced strategy of the leader (player 1) which results in an asymmetric game. Of course, this constellation of a single leader and *N-1* followers is only one possible scenario, but it turned out to be sufficient for the purpose of economic modeling in the field of optimal fiscal and monetary policies towards EMU.

This kind of hierarchical game may be relevant if one of the players owns a dominant position, i.e. the feedback Stackelberg equilibrium solution may be imagined to be valid if the leader (*player 1*) announces her (his) decision rule, $u_{1t} = \phi_1(x_{t-1})$, whereas the followers (*players* $i = 2,\dots,N$) base their actions on the current state and on the decision of the leader according to the reaction functions, $u_{it} = \phi_i(x_{t-1}, u_{1t}), (i = 2,\dots,N)$. (Note that the followers play feedback Nash among each other.) The leader in turn considers the reaction coefficients,

$\partial u_{it} / \partial u_{1t}$, $(i = 2, \ldots, N)$, as rational reactions of the followers in the optimization process (see, e.g., de Zeeuw (1984), or de Zeeuw and van der Ploeg (1991)).

PROPOSITION 4

The equilibrium solution of the non-cooperative Stackelberg game with a single leader and *N-1* followers is given by the following procedure: Again the principle of dynamic programming is used to derive feedback matrices and Riccati equations which are calculated backwards in time and determined by equations (19) and (20) – in total analogy to the feedback Nash solution. The calculation of the feedback matrices, however, results in

$$G_{1t} = -M_t \left(\overline{B}_t' H_{1t} - \sum_{j=2}^{N} \left(\overline{B}_t' H_{1t} B_{jt} + R_{jt}' Q_{1t}^{u_j} \right) \left(Q_{jt}^{u_j} \right)^{-1} B_{jt}' H_{jt} E_t \right) A_t , \qquad (35)$$

$$\begin{aligned} g_{1t} = & -M_t \left(\sum_{j=2}^{N} \left(\overline{B}_t' H_{1t} B_{jt} + R_{jt}' Q_{1t}^{u_j} \right) \left(\tilde{u}_{jjt} - \left(Q_{jt}^{u_i} \right)^{-1} B_{jt}' \left(H_{jt} E_t F_t - h_{jt} \right) \right) \right) \\ & - M_t \left(\overline{B}_t' \left(H_{1t} s_t - h_{1t} \right) - Q_{1t}^{u_1} \tilde{u}_{11t} - \sum_{j=2}^{N} R_{jt}' Q_{1t}^{u_j} \tilde{u}_{1jt} \right) \end{aligned} , \qquad (36)$$

and for $i = 2, \ldots, N$,

$$G_{it} = -\left(Q_{it}^{u_i} \right)^{-1} B_{it}' H_{it} E_t \left(A_t + B_{1t} G_{1t} \right), \qquad (37)$$

$$g_{it} = \tilde{u}_{iit} - \left(Q_{it}^{u_i} \right)^{-1} B_{it}' \left(H_{it} E_t \left(F_t + B_{1t} g_{1t} \right) - h_{it} \right), \qquad (38)$$

where for $i = 2, \ldots, N$,

$$\frac{\partial u_{it}}{\partial u_{1t}} := R_{it} = -\left(Q_{it}^{u_i} \right)^{-1} B_{it}' H_{it} E_T B_{1t} , \qquad (39)$$

$$\overline{B}_t := B_{1t} + \sum_{j=2}^{N} B_{jt} R_{jt} , \qquad (40)$$

$$M_t := \left(\overline{B}_t' H_{1t} \overline{B}_t + Q_{1t}^{u_1} + \sum_{j=2}^{N} \left(R_{jt}' Q_{1t}^{u_j} R_{jt} \right) \right)^{-1} , \qquad (41)$$

$$E_t := \left(I + \sum_{j=2}^{N} B_{jt} \left(Q_{jt}^{u_j} \right)^{-1} B_{jt}' H_{jt} \right)^{-1}, \text{ and} \tag{42}$$

$$F_t := s_t + \sum_{j=2}^{N} B_{jt} \left(\tilde{u}_{jjt} + \left(Q_{jt}^{u_j} \right)^{-1} B_{jt}' h_{jt} \right). \tag{43}$$

For proof see Appendix.

3.3 The Open-Loop Stackelberg Equilibrium Solution

Contrary to the feedback Stackelberg strategy, the open-loop Stackelberg game assumes that the *leader* (*player 1*) announces her (his) strategy for the entire planing horizon. This means that all players base their actions only on the initial state, x_0. Hence, the open-loop Stackelberg equilibrium solution may be derived in analogy to open-loop Nash – with the additional assumption of the asymmetric information structure of the game, however. The leader makes binding commitments about future policy actions, where the followers' rational reaction functions are taken into consideration. Note that, because the leader forces her (his) strategy to the followers, the leader is always better off (while the followers are worse off, respectively), compared to open-loop Nash equilibrium solution.

PROPOSITION 5

The solution of the open-loop Stackelberg game with *player 1* as leader and *players* $i = 2, \ldots, N$ as followers is given by the following procedure: Starting with the terminal conditions

$$H_T = \begin{pmatrix} Q_{1T}^x & Q_{2T}^x & \cdots & Q_{NT}^x \\ Q_{2T}^x & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ Q_{NT}^x & 0 & \cdots & 0 \end{pmatrix}, \quad \text{and} \quad h_T = -\begin{pmatrix} Q_{1T}^x \tilde{x}_{1T} \\ Q_{2T}^x \tilde{x}_{2T} \\ \vdots \\ Q_{NT}^x \tilde{x}_{NT} \end{pmatrix} \tag{44}$$

the Riccati matrices,

$$H_{t-1} = \overline{A}_t' \left(I - H_t \overline{\overline{E}}_t \right)^{-1} H_t \overline{A}_t + \overline{Q}_{t-1}^x, \tag{45}$$

$$h_{t-1} = \overline{A}_t' \left(I - H_t \overline{\overline{E}}_t \right)^{-1} \left[H_t \overline{\overline{F}}_t + h_t \right] - q_{t-1}^x, \tag{46}$$

can be solved by backward integration, where

$$\bar{A}_t = \begin{pmatrix} A_t & 0 & \cdots & 0 \\ 0 & A_t & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & A_t \end{pmatrix}, \text{ and } \bar{\bar{F}}_t = \begin{pmatrix} s_t + \sum_{j=1}^{N} B_{jt}\tilde{u}_{jjt} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad (47)$$

$$\bar{\bar{E}}_t = -\begin{pmatrix} B_{1t}\left(Q_{1t}^{u_1}\right)^{-1} B_{1t}' & B_{2t}\left(Q_{2t}^{u_2}\right)^{-1} B_{2t}' & \cdots & B_{Nt}\left(Q_{Nt}^{u_N}\right)^{-1} B_{Nt}' \\ B_{2t}\left(Q_{2t}^{u_2}\right)^{-1} B_{2t}' & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ B_{Nt}\left(Q_{Nt}^{u_N}\right)^{-1} B_{Nt}' & 0 & \cdots & 0 \end{pmatrix}, \qquad (48)$$

$$\bar{Q}_{t-1}^{x} = \begin{pmatrix} Q_{1t}^{x} & Q_{2t}^{x} & \cdots & Q_{Nt}^{x} \\ Q_{2t}^{x} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ Q_{Nt}^{x} & 0 & \cdots & 0 \end{pmatrix}, \text{ and } q_t^{x} = -\begin{pmatrix} Q_{1t}^{x}\tilde{x}_{1t} \\ Q_{2t}^{x}\tilde{x}_{2t} \\ \vdots \\ Q_{Nt}^{x}\tilde{x}_{Nt} \end{pmatrix}. \qquad (49)$$

The generalized state vector is given by

$$\hat{x}_t = \left(I - \bar{\bar{E}}_t H_t\right)^{-1}\left(\bar{A}_t\, \hat{x}_{t-1} + \bar{\bar{E}}_t\, h_t + \bar{\bar{F}}_t\right), \qquad (50)$$

and control is determined by

$$\begin{pmatrix} u_{1t} \\ u_{2t} \\ \vdots \\ u_{Nt} \end{pmatrix} = \begin{pmatrix} \tilde{u}_{11t} \\ \tilde{u}_{22t} \\ \vdots \\ \tilde{u}_{NNt} \end{pmatrix} - \begin{pmatrix} \left(Q_{1t}^{u_1}\right)^{-1} B_{1t}' & 0 & \cdots & 0 \\ 0 & \left(Q_{2t}^{u_2}\right)^{-1} B_{2t}' & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \left(Q_{Nt}^{u_N}\right)^{-1} B_{Nt}' \end{pmatrix} \left(H_t \hat{x}_t + h_t\right). \qquad (51)$$

For proof see Appendix.

4. Conclusions

OPTGAME 2.0, an algorithm to solve N-person discrete-time LQ games exactly and discrete-time non-linear quadratic games approximately by means of an appropriate linearization procedure, is presented in this paper as an extension of Hager et al (2000). The application of this technique of solving intertemporal optimization problems with an arbitrary number of decision-makers yields a real improvement with respect to decision-making support in the field of economics.

References

1. Basar, T. and Olsder, G.J. (1995): Dynamic Noncooperative Game Theory. 2nd edition. Academic Press, London.
2. Chow, G.C. (1975): Analysis and Control of Dynamic Economic Systems. John Wiley & Sons Inc., New York.
3. Dockner, E.J. and Neck, R. (1988): Time-Consistency, Subgame-Perfectness, Solution Concepts and Information Patterns in Dynamic Models of Stabilization Policies. Discussion Paper, University of Saskatchewan, Saskatoon.
4. Hager, M., Behrens, D., and Neck, R. (2000): OPTGAME 1.0: An Algorithm to Derive Equilibrium Solutions from Two-Person Discrete-Time (Non-)Linear Dynamic Games.
5. Hatheway, L.H. (1992): Modeling International Economic Interdependences: An Application of Feedback Nash Dynamic Games. Ph.D.-Thesis, University of Texas.
6. Kendrick, D. (1981): Stochastic Control for Economic Models. McGraw-Hill Inc., New York.
7. Leitmann, G. (1974): Cooperative and Non-Cooperative Many Players Differential Games. Springer, Vienna.
8. Mehlmann, A. (1991): Applied Differential Games. Plenum Press, New York.
9. Neck, R. and Matulka, J. (1991): OPTCON: An Algorithm for the Optimal Control of Nonlinear Stochastic Models. Annals of Operations Research **37**, 375 – 401.
10. Petit, M.L. (1990): Control Theory and Dynamic Games in Economic Policy Analysis. Cambridge University Press, Cambridge.
11. Reinganum, J.F. and Stokey, N.S. (1985): Oligopoly Extraction of a Common Property Natural Resource: The Importance of the Period of Commitment in Dynamic Games. International Economic Review **26**, 161 – 173.
12. de Zeeuw, A.J. (1984): Difference Games and Linked Econometric Policy Models. Ph.D. Thesis. Tillburg University.
13. de Zeeuw, A.J. and van der Ploeg, F. (1991): Difference Games and Policy Optimization: A Conceptional Framework. Oxford Economic Papers **43**, 612-636.

Appendix

Proof of **PROPOSITION 1**

To define the Riccati matrices for the terminal period we start with the quadratic tracking forms of the objective functions, (1),

$$L_{it} = \tfrac{1}{2} X_t{}' Q_{it} X_t - X_t{}' Q_{it} \widetilde{X}_{it} + \tfrac{1}{2} \widetilde{X}_{it}{}' Q_{it} \widetilde{X}_{it}, \qquad \begin{array}{l} t = 0,\ldots,T, \\ i = 1,\ldots,N, \end{array} \qquad (A1)$$

in period $t = T$, and determine the terminal conditions for the Riccati matrices,

$$\left.\begin{array}{l} H_{iT} = Q_{iT}^{x} \\ h_{iT} = Q_{iT}^{x} \widetilde{x}_{iT} \\ c_{iT} = \tfrac{1}{2} \widetilde{x}_{iT}{}' Q_{iT}^{x} \widetilde{x}_{iT} \end{array}\right\} \qquad i = 1,\ldots,N \, . \qquad (A2)$$

Then, the objective functions of the players $i \, (i = 1,\ldots,N)$, which are the cost-to-go functions for the terminal period, may be written in the following form,

$$L_{iT} = \tfrac{1}{2} x_T{}' H_{iT} x_T - x_T{}' h_{iT} + c_{iT}, \qquad i = 1,\ldots,N, \qquad (A3)$$

where the constants, $c_{iT} \; (i = 1,\ldots,N)$, are without relevance for further calculation. The state vector in the terminal period, x_T, can be derived due to the system equations (8) by the use of the optimized state vector of the preceding period in time, x_{T-1}, and the control variables, $u_{i(T-1)}, (i = 1,\ldots,N)$, which were optimized the period before. Hence, the minimization starts with period $T-1$ using Bellman's principle of optimality. The optimal cost-to-go functions, $J_{i(T-1)}$, for time period $t = T-1$ are given by the minimization with respect to $u_{iT} \; (i = 1,\ldots,N)$:

$$J_{i(T-1)}^{*} = \min_{u_{iT}} \left\{ J_{i(T-1)} \right\}, \qquad i = 1,\ldots,N \, . \qquad (A4)$$

To perform the minimization we have to replace the state in period *T*, x_T, by the right hand side of the system equation (8), and we yield for $i = 1,\ldots,N$,

$$\begin{aligned} J_{i(T-1)} = {} & \tfrac{1}{2} \left(A_T x_{T-1} + \sum_{j=1}^{N} B_{jT} u_{jT} + s_T \right)' H_{iT} \left(A_T x_{T-1} + \sum_{j=1}^{N} B_{jT} u_{jT} + s_T \right) - \\ & - \left(A_T x_{T-1} + \sum_{j=1}^{N} B_{jT} u_{jT} + s_T \right)' h_{iT} + c_{iT} + \\ & + \tfrac{1}{2} X_{T-1}{}' Q_{i(T-1)} X_{T-1} - X_{T-1}{}' Q_{i(T-1)} \widetilde{X}_{i(T-1)} + \tfrac{1}{2} \widetilde{X}_{i(T-1)}{}' Q_{i(T-1)} \widetilde{X}_{i(T-1)}. \end{aligned} \qquad (A5)$$

Then, we can carry out the minimization of the cost to go function $(A5)$ with respect to the controls which delivers the first-order conditions for the control variables for $i = 1,\dots,N$. These conditions can be transformed and be rewritten to be

$$u_{iT} = -\left(Q_{iT}^{u_i}\right)^{-1} B_{iT}{}'\left(H_{iT}x_T - h\ \right) - B_{iT}{}' h_{iT} + Q_{iT}^{u_i} = 0, \qquad i = 1,\dots,N . \qquad (A6)$$

The Nash equilibrium for the control variables, u_{iT} $(i = 1,\dots,N)$, is determined by the intersection of the N hypothetical reaction functions $(A6)$. With the optimal control variables, u_{iT} $(i = 1,\dots,N)$, depending on the state variable, x_T , however, we may write x_T as a linear function of the pre-period state variable x_{T-1} using the system equations (8). This also illustrates the evolution of the system using the corresponding Riccati matrices. Thus, we can derive the Nash equilibrium in a simpler way, namely by substitution of the optimal state vector x_T by a function of x_{T-1} using equation $(A6)$ and solving for x_T gives the optimized state variable for the terminal period T as a function of x_{T-1}:

$$x_T = E_T\left(A_T x_{T-1} + F_T\right) \qquad (A7)$$

where

$$E_T = \left(I + \sum_{j=1}^{N} B_{jT}\left(Q_{jT}^{u_j}\right)^{-1} B_{jT}{}' H_{jT} \right)^{-1}, \text{ and} \qquad (A8)$$

$$F_T = s_T + \sum_{j=1}^{N} B_{jT}\left(\tilde{u}_{jjT} + \left(Q_{jT}^{u_j}\right)^{-1} B_{jT}{}' h_{jT} \right). \qquad (A9)$$

The replacement of the optimal state vector, x_T, in $(A6)$ according to $(A7)$ with $(A8)$ and $(A9)$ yields the Nash equilibrium values for the controls.

$$u_{iT}^{*} = G_{iT}x_{T-1} + g_{iT} \qquad i = 1,\dots,N , \qquad (A10)$$

where

$$G_{iT} = -\left(Q_{iT}^{u_i}\right)^{-1} B_{iT}{}' H_{iT} E_T A_T \qquad (A11)$$

$$g_{iT} = -\left(Q_{iT}^{u_i}\right)^{-1} B_{iT}{}'\left(H_{iT}E_T F_T - h_{iT}\right) + \tilde{u}_{ii(T)}, \qquad (A12)$$

which can be interpreted as optimal linear feedback rule for each of the players $i\,(i = 1,\dots,N)$. The linear form results from the quadratic structure of the cost function, the fact that mixed penalties are not allowed according to the definition of the corresponding matrices, (4), and the linearity of the system equations (8). With $(A10)$ and the feedback matrices, $(A11)$ and $(A12)$, we can determine the optimal state vector, x_T, as a linear function of x_{T-1}. Due to the definitions (17) and (18), the optimal state vector may be expressed by

$$x_T^{*} = K_T x_{T-1} + k_T . \qquad (A13)$$

Note that from now on we omit the asterisks to simplify notation. For the derivation of the Riccati matrices of time period *T-1* we have to re-substitute the optimized state vector $(A13)$ and the optimized control vectors $(A10)$ into the cost-to-go function $(A5)$, where

$$X_{T-1}=\begin{pmatrix} x_{T-1} \\ u_{1T} \\ \vdots \\ u_{NT} \end{pmatrix}=\begin{pmatrix} x_{T-1} \\ G_{1T}x_{T-1}+g_{1T} \\ \vdots \\ G_{NT}x_{T-1}+g_{NT} \end{pmatrix}=\begin{pmatrix} I \\ G_{1T} \\ \vdots \\ G_{NT} \end{pmatrix}x_{T-1}+\begin{pmatrix} 0 \\ g_{1T} \\ \vdots \\ g_{NT} \end{pmatrix}. \qquad (A14)$$

Then, we collect all terms containing x_{T-1} and identify the Riccati matrices for $T-1$ by comparison with the equation

$$J_{i(T-1)}=\tfrac{1}{2}x_{T-1}{}'H_{i(T-1)}x_{T-1}-x_{T-1}{}'h_{i(T-1)}+c_{i(T-1)}, \qquad i=1,\dots,N \quad (A15)$$

to be defined as follows,

$$H_{i(T-1)}=K_T{}'H_{iT}K_T+\begin{pmatrix} I \\ G_{1T} \\ \vdots \\ G_{NT} \end{pmatrix}'Q_{i(T-1)}\begin{pmatrix} I \\ G_{1T} \\ \vdots \\ G_{NT} \end{pmatrix}, \qquad (A16)$$

$$\begin{aligned} h_{i(T-1)}=&K_T{}'\left(h_{iT}-H_{iT}k_T\right)- \\ &-\begin{pmatrix} I \\ G_{1T} \\ \vdots \\ G_{NT} \end{pmatrix}'Q_{i(T-1)}\begin{pmatrix} 0 \\ g_{1T} \\ \vdots \\ g_{NT} \end{pmatrix}+\begin{pmatrix} I \\ G_{1T} \\ \vdots \\ G_{NT} \end{pmatrix}'Q_{i(T-1)}\tilde{X}_{i(T-1)}. \end{aligned} \qquad (A17)$$

System Equations, $(A16)$ and $(A17)$, can be solved recursively in time to obtain the Riccati matrices for period *T-1* as a function of the Riccati matrices of the terminal period *T*. This procedure can be extended straight forward to period $t=T-2$ and generalized to any other period $t=\tau$ by induction. Note that the existence and uniqueness of the solutions of the Riccati equations for all periods $t\in\{1,\dots,T\}$ of the linear-quadratic game can readily be verified according to Basar and Olsder (1995). □

Proof of **PROPOSITION 2**

The solution of the open-loop Nash game with *N* players is given by the application of Pontryagin's minimum principle (see, e.g., Basar and Olsder (1995)) to perform the minimization of each player's loss function, (1), subject to the system equations, (8), the initial conditions, and the assumption that the other players act identically. Note that for all further calculations $t \in \{1,\dots,T+1\}$. For $i\ (i=1,\dots,N)$ we define the (current value) Hamiltonian function using the state equations and the objective function for each player,

$$\mathscr{H}_{it} = \tfrac{1}{2}\left(X_{t-1} - \widetilde{X}_{i(t-1)}\right)' Q_{i(t-1)}\left(X_{t-1} - \widetilde{X}_{i(t-1)}\right) + \\ + \lambda_{it}'\left(A_t x_{t-1} + \sum_{j=1}^{N} B_{jt} u_{jt} + s_t\right), \qquad (A18)$$

where λ_{it} is a n_s-dimensional row vector, the vector of the so-called costate variables for player $i\ (i=1,\dots,N)$. Appropriate differentiation yields the adjoint equations and the necessary conditions for the control variables. The costate equations may be derived by performing the derivative of the stacked state vector, (2), with respect to the adjoint variables of each player $i\ (i=1,\dots,N)$,

$$\lambda_{i(t-1)} = A_t'\lambda_{it} + Q^x_{i(t-1)}\left(x_{t-1} - \widetilde{x}_{i(t-1)}\right), \qquad i=1,\dots,N. \qquad (A19)$$

Then, we derive the Hamiltonian minimizing condition. Since we know that $\lambda_{i(T+1)} = 0 \quad (i=1,\dots,N)$, we can express the first order conditions for the controls as hypothetical reaction function for the players $i\ (i=1,\dots,N)$:

$$u_{it} = \widetilde{u}_{iit} - \left(Q^{u_i}_{it}\right)^{-1} B_{it}'\lambda_{it}, \qquad i=1,\dots,N. \qquad (A20)$$

We assume a linear relationship between the adjoint and the optimal state variables, i.e. $\lambda_{it} = H_{it}x_t + h_{it}\ (i=1,\dots,N)$ with the Riccati matrices H_{it} and h_{it}, respectively. The costate variables, $\lambda_{it}\ (i=1,\dots,N)$, can be substituted into the hypothetical reaction functions, $(A20)$, yielding

$$u_{it} = \widetilde{u}_{iit} - \left(Q^{u_i}_{it}\right)^{-1} B_{it}'\left(H_{it}x_t + h_{it}\right), \qquad i=1,\dots,N. \qquad (A21)$$

Then, the optimal state equation follows from the substitution of $(A21)$ into (8), where we rearrange the resulting equation with respect to x_t, which is already determined by equation $(A6)$. To solve $(A6)$, we generate the Riccati matrices for each time period $t \in \{1,\dots,T\}$. This can be done in the following way: Substitute $\lambda_{it} = H_{it}x_t + h_{it}\ (i=1,\dots,N)$ with $(A6)$ into the equations for the costates, $(A19)$, and rearrange them with respect to x_{t+1},

$$\lambda_{i(t-1)} = H_{i(t-1)}x_{t-1} + h_{i(t-1)}, \qquad i=1,\dots,N, \qquad (A22)$$

where

$$H_{i(t-1)} = A_t' H_{it} E_t A_t + Q^x_{i(t-1)}, \qquad i=1,\dots,N, \qquad (A23a)$$

$$h_{i(t-1)} = A_t'\left(H_{it}E_tF_t + h_{it}\right) - Q^x_{i(t-1)}\widetilde{x}_{i(t-1)}, \qquad i=1,\dots,N\text{, and} \qquad (A24a)$$

$$F_T = s_T + \sum_{j=1}^{N} B_{jT} \left(\widetilde{u}_{jjT} - \left(Q_{jT}^{u_j} \right)^{-1} B_{jT}{}' h_{jT} \right). \qquad (A25)$$

Now, the Riccati matrices may be solved backwards in time. Using $\wedge_{i(T+1)} = \upsilon$ equations $(A23a)$ and $(A24a)$ can be determined for the terminal period as,

$$H_{iT} = Q_{iT}^{x}, \qquad i = 1, \ldots, N, \qquad (A23b)$$

$$h_{iT} = -Q_{iT}^{x} \widetilde{x}_{iT}, \qquad i = 1, \ldots, N. \qquad (A24b)$$

After we calculated the Riccati matrices by iterating backwards in time, the optimal state variable can be determined by forward iteration. For the sake of comparison we want to write down the open-loop solution in feedback form. This form is already determined by equation $(A10)$, where the feedback matrices are determined by the equations $(A11)$ and $(A12)$. □

Proof of **PROPOSITION 3**

To determine the Riccati matrices for the terminal period $t = T$ we can expand the loss function for the terminal period according to (5) with

$$L_{iT} = \tfrac{1}{2} x_T{}' Q_{iT}^{x} x_T - x_T{}' Q_{iT}^{x} \widetilde{x}_{iT} + \tfrac{1}{2} \widetilde{x}_{iT}{}' Q_{iT}^{x} \widetilde{x}_{iT}, \qquad i = 1, \ldots, N, \qquad (A26)$$

in the following way,

$$J_T = \sum_{j=1}^{N} \mu_j L_{jT}, \qquad \sum_{j=1}^{N} \mu_j = 1. \qquad (A27)$$

In analogy to feedback Nash ,$(A27)$ can easily be re-written – since for the terminal period, $t = T$, there are no penalties for the control variables. Then, we may derive the terminal condition for the Riccati matrices,

$$\begin{aligned} H_T &= \sum_{j=1}^{N} \mu_j Q_{jT}^{x} = P_T^{x}, \\ h_T &= \sum_{j=1}^{N} \mu_j Q_{jT}^{x} \widetilde{x}_{jT} = \widetilde{P}_T^{x}, \qquad (A28) \\ c_T &= \tfrac{1}{2} \sum_{j=1}^{N} \mu_j \widetilde{x}_{jT}{}' Q_{jT}^{x} \widetilde{x}_{jT}, \end{aligned}$$

where we use the definitions (29) and (30). Hence, the objective function, which is the cost-to-go function for the terminal period, may then be written in the following form,

$$J_T = \tfrac{1}{2} x_T{}' H_T x_T - x_T{}' h_T + c_T , \qquad (A29)$$

where the constant c_T is without any relevance for further calculation. Similar to feedback Nash, the optimization starts at $t = T - 1$ with the minimization of the cost-to-go function (where we use again the definitions (29) and (30) to simplify notation),

$$\begin{aligned} J_{T-1} = J_T + \tfrac{1}{2} X_{T-1}{}' P_{T-1} X_{T-1} - X_{T-1}{}' \widetilde{P}_{T-1} + \\ + \tfrac{1}{2} \left(\sum_{j=1}^{N} \mu_j \widetilde{X}_{j(T-1)}{}' Q_{j(T-1)} \widetilde{X}_{j(T-1)} \right), \end{aligned} \qquad (A30)$$

with respect to u_{iT} $(i = 1,\ldots,N)$, which delivers the first-order conditions for the control variables for $i = 1,\ldots,N$, in analogy to feedback Nash. These conditions can be solved by performing the derivative of the stacked state vector, (2), with respect to the control variables of each player i $(i = 1,\ldots,N)$:

$$B_{iT}{}' \left(H_T x_T - h_T \right) + \left(\sum_{j=1}^{N} \mu_j Q_{jT}^{u_i} \right) u_{iT} - \left(\sum_{j=1}^{N} \mu_j Q_{jT}^{u_i} \widetilde{u}_{jiT} \right) = 0 . \qquad (A31)$$

$(A31)$ can be solved for any control vector, u_{iT} $(i = 1,\ldots,N)$, and can be written as (where we use again the definitions (29) and (30) to simplify notation),

$$u_{iT} = -\left(P_{iT}\right)^{-1} \left(B_{iT}{}' \left(H_T x_T - h_T \right) - \widetilde{P}_{iT} \right) \qquad i = 1,\ldots,N . \qquad (A32)$$

Hence, we receive by substitution that

$$x_T = \overline{E}_T \left(A_T x_{T-1} + \overline{F}_T \right) \qquad (A33)$$

where

$$\overline{E}_T = \left(I + \sum_{j=1}^{N} B_{jT} \left(P_{jT} \right)^{-1} B_{jT}{}' H_T \right)^{-1} , \text{ and} \qquad (A34)$$

$$\overline{F}_T = s_T + \sum_{j=1}^{N} B_{jT} \left(P_{jT} \right)^{-1} \left(B_{jT}{}' h_T + \widetilde{P}_{jT} \right). \qquad (A35)$$

Replacement of the optimal state vector x_T according to $(A33)$ with $(A34)$ and $(A35)$ in $(A32)$ yields the Nash equilibrium values for the controls. Since mixed penalties are not allowed according to the definition of the corresponding matrices, (4), the control variables of the opponents, u_{jT} $(j = 1,\ldots,i-1,i+1,\ldots,N)$, do not occur in the corresponding equation for u_{iT}, and we receive a linear relation between the control variables of player i, u_{iT} $(i = 1, \ldots, N)$, and the state variables x_{T-1}. This may be described by equation $(A10)$,

$$u_{iT}^{*} = G_{iT} x_{T-1} + g_{iT}, \qquad i = 1, \ldots, N,$$

where

$$G_{iT} = -(P_{iT})^{-1} B_{iT}{}' H_T \overline{E}_T A_T, \tag{A36}$$

$$g_{iT} = -(P_{iT})^{-1}\left(B_{iT}{}'\left(H_T \overline{E}_T \overline{F}_T - h_T\right) - \widetilde{P}_{iT}\right). \tag{A37}$$

With $(A10)$ and the feedback matrices, $(A36)$ and $(A37)$, we can determine the optimal state vector x_T as a linear function of x_{T-1}, and due to the definitions (17) and (18) it may be expressed by equation $(A21)$,

$$x_T^{*} = K_T x_{T-1} + k_T .$$

We omit the asterisks from now on to simplify notation. As for feedback Nash, also for the derivation of the Riccati matrices of time period *T*-1 we have to re-substitute the optimized state vector $(A21)$ and the optimized control vectors $(A10)$ into the cost-to-go function $(A30)$ and collect all terms containing x_{T-1}. By comparison with the equation

$$J_{T-1} = \tfrac{1}{2} x_{T-1}{}' H_{T-1} x_{T-1} - x_{T-1}{}' h_{T-1} + c_{T-1}, \tag{A38}$$

we can identify the Riccati matrices for *T-1* as

$$H_{T-1} = K_T{}' H_T K_T + \begin{pmatrix} I \\ G_{1T} \\ \vdots \\ G_{NT} \end{pmatrix}' P_{T-1} \begin{pmatrix} I \\ G_{1T} \\ \vdots \\ G_{NT} \end{pmatrix}, \tag{A39}$$

$$h_{T-1} = -K_T{}'\left(H_T k_T - h_T\right) - \begin{pmatrix} I \\ G_{1T} \\ \vdots \\ G_{NT} \end{pmatrix}' P_{T-1} \begin{pmatrix} 0 \\ g_{1T} \\ \vdots \\ g_{NT} \end{pmatrix} + \begin{pmatrix} I \\ G_{1T} \\ \vdots \\ G_{NT} \end{pmatrix}' \widetilde{P}_{T-1}. \tag{A40}$$

Equation system $(A39)$ and $(A40)$ can be solved recursively in time to obtain the Riccati matrices for period *T*-1 as a function of the Riccati matrices of the terminal period *T*. This procedure can be extended straight forward to period $t = T - 2$ and be generalized to any other period *t* by induction. Hence, the Riccati equations are given by

$$H_{t-1} = K_t{}' H_t K_t + P_{t-1}^x + \sum_{j=1}^{N} G_{jt}{}' P_{jt} G_{jt}, \qquad H_T = P_T^x, \quad (A41)$$

$$h_{t-1} = K_t{}' (h_t - H_t k_t) + \widetilde{P}_{t-1}^x + \sum_{j=1}^{N} G_{jt}{}' (\widetilde{P}_{jt} - P_{jt} g_{jt}), \quad h_T = \widetilde{P}_T^x, \quad (A42)$$

and can be solved by backward iteration. □

Proof of **PROPOSITION 4**

Note, that two different optimization problems occur in the derivation of the Stackelberg game. The first one determines the rational reaction of the followers (*players* $i = 2,\dots,N$), $u_{it} = \phi_i(x_{t-1}, u_{1t})$ $(i = 2,\dots,N)$, to the announced decision rule, $u_{1t} = \phi_1(x_{t-1})$, of the leader (*player 1*). **Note that the followers play feedback Nash among each other.** The second optimization problem determines the optimal action of the leader, given the rational reactions of the followers in form of the reaction coefficients, $\partial u_{it} / \partial u_{1t}$, $(i = 2,\dots,N)$.

The first optimization problem

To define the Riccati matrices for the terminal period T we follow the calculation of feedback Nash according to the dynamic programming principle and start with the quadratic tracking form of the objective function (1) for the *followers*,

$$L_{it} = \tfrac{1}{2} X_t{}' Q_{it} X_t - X_t{}' Q_{it} \widetilde{X}_{it} + \tfrac{1}{2} \widetilde{X}_{it}{}' Q_{it} \widetilde{X}_{it}, \qquad \begin{array}{l} t = 1,\dots,T, \\ i = 2,\dots,N, \end{array} \qquad (A43)$$

in the terminal period $t = T$. This yields, in analogy to Proposition 1, the first-order condition for the followers with the terminal conditions for the Riccati matrices,

$$\left. \begin{array}{l} H_{iT} = Q_{iT}^x \\ h_{iT} = Q_{iT}^x \widetilde{x}_{iT} \\ c_{iT} = \tfrac{1}{2} \widetilde{x}_{iT}{}' Q_{iT}^x \widetilde{x}_{iT} \end{array} \right\} \qquad i = 2,\dots,N \ . \qquad (A44)$$

The Nash equilibrium solution for the control variables, u_{iT} $(i = 2,\dots,N)$, as a function of the control of the leader, u_{1T} , is determined by the intersection of the N-1 hypothetical reaction functions $(A45)$. With the optimal control variables, u_{iT} $(i = 2,\dots,N)$, depending on the control of the leader, u_{1T} , and on the state variable, x_T , in equation $(A45)$, we may write the state variable x_T as a linear function of the pre-period state variable x_{T-1} using the system equations (8). Thus, we can derive the Nash equilibrium strategy of the followers by substitution of the optimal state vector $x_T(u_{1T})$ by a function of $x_{T-1}(u_{1T})$ with the equation

$$u_{iT} = \widetilde{u}_{iiT} - \left(Q_{iT}^{u_i}\right)^{-1} B_{iT}{}' \left(H_{iT} x_T - h_{iT}\right), \qquad i = 2,\dots,N, \qquad (A45)$$

where we receive after rearranging and solving for x_T the optimized state variable for the terminal period T as a function of the optimal control of the leader, u_{1T}:

$$x_T = E_T\left(A_T x_{T-1} + B_{1T} u_{1T} + F_T\right), \qquad (A46)$$

where

$$E_T = \left(I + \sum_{j=2}^{N} B_{jT}\left(Q_{jT}^{u_j}\right)^{-1} B_{jT}{}' H_{jT}\right)^{-1}, \text{ and} \qquad (A47)$$

$$F_T = s_T + \sum_{j=2}^{N} B_{jT}\left(\widetilde{u}_{jjT} + \left(Q_{jT}^{u_j}\right)^{-1} B_{jT}{}' h_{jT}\right). \qquad (A48)$$

Replacement of the optimal state vector x_T according to $(A46)$ in $(A45)$ yields the Nash equilibrium values for the controls u_{iT} $(i = 2,\dots,N)$:

$$u_{iT} = \widetilde{u}_{iiT} - \left(Q_{iT}^{u_i}\right)^{-1} B_{iT}{}' \left(H_{iT} E_T\left(A_T x_{T-1} + B_{1T} u_{1T} + F_T\right) - h_{iT}\right). \qquad (A49)$$

Hence, we receive a relation between the control variables of the followers, the announced strategy of the leader, and the state variables x_{T-1}. This may be described by the following equation,

$$\overset{*}{u}_{iT} = W_{iT} x_{T-1} + V_{iT} + R_{iT} u_{1T}, \qquad i = 2,\dots,N, \qquad (A50)$$

where

$$W_{iT} = -\left(Q_{iT}^{u_i}\right)^{-1} B_{iT}{}' H_{iT} E_T A_T, \qquad i = 2,\dots,N, \qquad (A51)$$

$$V_{iT} = \widetilde{u}_{iiT} - \left(Q_{iT}^{u_i}\right)^{-1} B_{iT}{}' \left(H_{iT} E_T F_T - h_{iT}\right), \qquad i = 2,\dots,N, \qquad (A52)$$

$$R_{iT} = \frac{\partial u_{iT}}{\partial u_{1T}} = -\left(Q_{iT}^{u_i}\right)^{-1} B_{iT}{}' H_{iT} E_T B_{1T}, \qquad i = 2,\dots,N. \qquad (A52)$$

Since the reaction coefficients for the terminal period can be solved with the reaction functions of the followers as a function of the announced strategy of the leader, and the state variables x_{T-1}, the reaction coefficients become the derivative of $(A49)$ with respect to u_{1T}, and are defined by equation set $(A52)$.

The second optimization problem

The objective function of the leader, which is the cost-to-go function for the terminal period, may then be written in the following form,

$$L_{1T} = \tfrac{1}{2} x_T{}' H_{1T} x_T - x_T{}' h_{1T} + c_{1T}. \qquad (A53)$$

The terminal conditions for the Riccati matrices are given by

$$\begin{aligned} H_{1T} &= Q_{1T}^{x}, \\ h_{1T} &= Q_{1T}^{x}\tilde{x}_{1T}, \\ c_{1T} &= \tfrac{1}{2}\tilde{x}_{1T}{}' Q_{1T}^{x}\tilde{x}_{1T}. \end{aligned} \qquad (A54)$$

In total analogy to Proposition 1, the minimization starts with the period $T-1$ using Bellman's principle of optimality. The crucial difference between the first-order condition for the optimal control variable, u_{1T}, and the corresponding equation for feedback Nash is, however, the occurrence of the terms containing the reaction coefficients, $\partial u_{iT}/\partial u_{1T}$ $(i = 2,\ldots,N)$, in the feedback Stackelberg game. Solving the first order condition for the first control variable with respect to u_{1T} weighted by the penalty for failing the desired control value makes the functional relationship more obvious:

$$\overline{B}_T{}'\left(H_{1T}x_T - h_{1T}\right) + Q_{1T}^{u_1}\left(u_{1T} - \tilde{u}_{11T}\right) + \sum_{j=2}^{N} R_{jT}{}' Q_{1T}^{u_j}\left(u_{jT} - \tilde{u}_{1jT}\right) = 0\,, \qquad (A55)$$

where

$$\overline{B}_T = B_{1T} + \sum_{j=2}^{N} B_{jT}R_{jT}\,. \qquad (A56)$$

We proceed to solve equation $(A55)$ with respect to u_{1T} by using the reaction functions of the followers $(A49)$ with $(A50)$–$(A52)$, for replacement when calculating the state equations,

$$x_T = Y x_{T-1} + \overline{B}_T u_{1T} + Z\,, \qquad (A57)$$

where

$$Y = A_T + \sum_{j=2}^{N} B_{jT}W_{jT} = \left(I - \sum_{j=2}^{N} B_{jT}\left(Q_{jT}^{u_j}\right)^{-1} B_{jT}{}' H_{jT}E_T \right) A_T\,, \qquad (A58)$$

$$\begin{aligned} Z &= s_T + \sum_{j=2}^{N} B_{jT}V_{jT} = \\ &= s_T + \sum_{j=2}^{N} B_{jT}\left(\tilde{u}_{iiT} - \left(Q_{iT}^{u_i}\right)^{-1} B_{iT}{}'\left(H_{iT}E_T F_T - h_{iT}\right)\right), \end{aligned} \qquad (A59)$$

and E_T is defined by equation $(A47)$ and F_T by equation $(A48)$. For W_{iT} defined by equation $(A50)$ and V_{iT} by equation $(A51)$, the optimal control variable for player 1 may be simplified written as

$$u_{1T} = G_{1T} x_{T-1} + g_{1T} \,, \qquad (A60)$$

where

$$G_{1T} = -M_T \left(\overline{B}_T{}' H_{1T} - \sum_{j=2}^{N} \left(\overline{B}_T{}' H_{1T} B_{jT} + R_{jT}{}' Q_{1T}^{u_j} \right) \left(Q_{jT}^{u_j} \right)^{-1} B_{jT}{}' H_{jT} E_T \right) A_T A_T \,, \qquad (A61)$$

$$\begin{aligned} g_{1T} = & -M_T \left(\sum_{j=2}^{N} \left(\overline{B}_T{}' H_{1T} B_{jT} + R_{jT}{}' Q_{1T}^{u_j} \right) \left(\tilde{u}_{jjT} - \left(Q_{jT}^{u_i} \right)^{-1} B_{jT}{}' \left(H_{jT} E_T F_T - h_{jT} \right) \right) \right) \\ & - M_T \left(\overline{B}_T{}' \left(H_{1T} s_T - h_{1T} \right) - Q_{1T}^{u_1} \tilde{u}_{11T} - \sum_{j=2}^{N} R_{jT}{}' Q_{1T}^{u_j} \tilde{u}_{1jT} \right), \end{aligned} \,, (A62)$$

$$M_T = \left(\overline{B}_T{}' H_{1T} \overline{B}_T + Q_{1T}^{u_1} + \sum_{j=2}^{N} \left(R_{jT}{}' Q_{1T}^{u_j} R_{jT} \right) \right)^{-1} \,, \qquad (A63)$$

which can be interpreted as optimal linear feedback rule for player 1. Immediately from $(A49)$ follows that

$$G_{iT} = -\left(Q_{iT}^{u_i} \right)^{-1} B_{iT}{}' H_{iT} E_T \left(A_T + B_{1T} G_1 \right), \qquad i = 2, \ldots, N \,, (A64)$$

$$g_{iT} = \tilde{u}_{iiT} - \left(Q_{iT}^{u_i} \right)^{-1} B_{iT}{}' \left(H_{iT} E_T \left(F_T + B_{1T} g_{1T} \right) - h_{iT} \right), \; i = 2, \ldots, N \,. (A65)$$

All the rest happens in total analogy to the calculation of the feedback Nash equilibrium solution. Hence, see $(A13)-(A17)$. □

Proof of **PROPOSITION 5**

The solution of the open-loop Stackelberg game for the *followers (players* $i = 2, \ldots, N$*)* is given by the application of Pontryagin's minimum principle (see, e.g., Basar and Olsder (1995)), because the followers play open-loop Nash among each other. The leader in turn faces a non-classical optimal control problem because the decision-maker has to take the dynamics of the optimal decisions of the followers (which are expressed by the costate equations) into account. Thus, from the point of view of the leader the application of the minimum principle has to be altered in the following way: In contrast to open-loop Nash, the leader has to consider N state vectors, where the additional state vectors are equivalent to the costate vectors of the followers. Note that for all further calculations $t \in \{1, \ldots, T+1\}$.

The first optimization problem

We define the (current value) Hamiltonian function for the followers – with the state equations, the objective function, and the given action of the leader – according to Basar and Olsder (1995), for $i = 2,\dots,N$:

$$\mathcal{H}_{it} = \tfrac{1}{2}\left(X_t - \widetilde{X}_{it}\right)' Q_{it}\left(X_t - \widetilde{X}_{it}\right) + \\ + \lambda_{i(t+1)}'\left(A_{t+1}x_t + \sum_{j=1}^{N} B_{j(t+1)}u_{j(t+1)} + s_{t+1}\right), \qquad (A66)$$

where $\lambda_{it}\ (i = 2,\dots,N)$ are n_s-dimensional row vectors, i.e. the vectors of the costate variables for the players $i = 2,\dots,N$. Analogous to open-loop Nash, appropriate differentiation yields the adjoint equations and necessary conditions for the control variables. First, the adjoint equations are derived as

$$\lambda_{it} = A_{t+1}'\lambda_{i(t+1)} + Q_{it}^{x}\left(x_t - \widetilde{x}_{it}\right), \qquad \lambda_{i(T+1)} = 0,\ i = 2,\dots,N. \qquad (A67)$$

Since we know that $\lambda_{i(T+1)} = 0\ (i = 2,\dots,N)$, performing the Hamiltonian minimization we derive the first order condition for the actions of the follower as hypothetical reaction function,

$$u_{it} = \widetilde{u}_{iit} - \left(Q_{it}^{u_i}\right)^{-1} B_{it}'\lambda_{it}, \qquad i = 2,\dots,N. \qquad (A68)$$

The second optimization problem

According to de Zeeuw (1984) the rational behavior of the followers is expressed in terms of the costates, determined by equations $(A67)$. Therefore, the leader has to deal with a system of state equations including the adjoint systems of the followers. Hence, for the first player we define the (current value) Hamiltonian function according to Basar and Olsder (1995, p.371),

$$\mathcal{H}_{1t} = \tfrac{1}{2}\left(X_t - \widetilde{X}_{1t}\right)' Q_{1t}\left(X_t - \widetilde{X}_{1t}\right) + \\ + \lambda_{11(t+1)}'\left(A_{t+1}x_t + B_{1(t+1)}u_{1(t+1)} + \sum_{j=2}^{N} \widetilde{E}_{j(t+1)}\lambda_{j(t+1)} + \widetilde{F}_{t+1}\right) + \\ + \sum_{j=2}^{N} \lambda_{1jt}'\left(A_{t+1}'\lambda_{j(t+1)} + Q_{jt}^{x}\left(x_t - \widetilde{x}_{jt}\right)\right) \qquad (A69)$$

where λ_{11t} is the costate variable with respect to the state vector, x_t, and $\lambda_{1it}\ (i = 2,\dots,N)$ represent the costate variables with respect to the "additional state vectors", $\lambda_{it}\ (i = 2,\dots,N)$, and where

$$\widetilde{E}_{it} = -B_{it}\left(Q_{it}^{u_i}\right)^{-1} B_{it}', \qquad i = 1,\ldots,N \text{, and} \qquad (A70)$$

$$\widetilde{F}_t = s_t + \sum_{j=2}^{N} B_{jt}\widetilde{u}_{jjt} . \qquad (A71)$$

Appropriate differentiation of $(A69)$ (see Basar and Olsder (1995, p.371)) yields the control,

$$u_{1t} = \widetilde{u}_{11t} - \left(Q_{1t}^{u_1}\right)^{-1} B_{1t}'\lambda_{11t}, \qquad (A72)$$

and the adjoint equations,

$$\lambda_{11(t-1)} = A_t'\lambda_{11t} + Q_{1(t-1)}^{x}\left(x_{t-1} - \widetilde{x}_{1(t-1)}\right) + \sum_{j=2}^{N} Q_{j(t-1)}^{x}\lambda_{1j(t-1)}, \qquad (A73)$$

$$\lambda_{11(T+1)} = 0,$$

and

$$\lambda_{1it} = A_t\lambda_{1i(t-1)} + \widetilde{E}_{it}\lambda_{11t}, \qquad i = 2,\ldots,N . \qquad (A74)$$

Now we may rewrite the state equations using the equations $(A68)$ and $(A72)$ as well as the definitions $(A70)$ and $(A71)$:

$$x_t = A_t x_{t-1} + B_{1t}\left(\widetilde{u}_{11t} - \left(Q_{1t}^{u_1}\right)^{-1} B_{1t}'\lambda_{11t}\right) + \sum_{j=2}^{N} \widetilde{E}_{jt}\lambda_{jt} + \widetilde{F}_t . \qquad (A75)$$

Finally, with the equations $(A67)$, $(A73)$, $(A74)$, and $(A75)$ we obtain a system of $2N$ coupled difference equations. N of these equations may be solved by forward iteration, the remaining ones by backward iteration. To solve the system we have to de-couple the $2N$ equations and specify the initial and terminal conditions, respectively. This yields the system $(A76)$:

- to be solved by forward iteration

$$\underbrace{\begin{pmatrix} x_t \\ \lambda_{12t} \\ \vdots \\ \lambda_{1Nt} \end{pmatrix}}_{\hat{x}_t} = \underbrace{\begin{pmatrix} A_t & 0 & \cdots & 0 \\ 0 & A_t & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & A_t \end{pmatrix}}_{\overline{A}_t} \underbrace{\begin{pmatrix} x_{t-1} \\ \lambda_{12(t-1)} \\ \vdots \\ \lambda_{1N(t-1)} \end{pmatrix}}_{\hat{x}_{t-1}} + \underbrace{\begin{pmatrix} \widetilde{E}_{1t} & \widetilde{E}_{2t} & \cdots & \widetilde{E}_{Nt} \\ \widetilde{E}_{2t} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{E}_{Nt} & 0 & \cdots & 0 \end{pmatrix}}_{\widetilde{\widetilde{E}}_t} \underbrace{\begin{pmatrix} \lambda_{11t} \\ \lambda_{2t} \\ \vdots \\ \lambda_{Nt} \end{pmatrix}}_{\hat{\lambda}_t} + \underbrace{\begin{pmatrix} \widetilde{F}_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{\overline{\overline{F}}_t}$$

- to be solved by backward iteration

$$\underbrace{\begin{pmatrix} \lambda_{11(t-1)} \\ \lambda_{2(t-1)} \\ \vdots \\ \lambda_{N(t-1)} \end{pmatrix}}_{\hat{\lambda}_{t-1}} = \underbrace{\begin{pmatrix} A_t' & 0 & \cdots & 0 \\ 0 & A_t' & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & A_t' \end{pmatrix}}_{\overline{A}_t'} \underbrace{\begin{pmatrix} \lambda_{11t} \\ \lambda_{2t} \\ \vdots \\ \lambda_{Nt} \end{pmatrix}}_{\hat{\lambda}_t}$$

$$+ \underbrace{\begin{pmatrix} Q^x_{1(t-1)} & Q^x_{2(t-1)} & \cdots & Q^x_{N(t-1)} \\ Q^x_{2(t-1)} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ Q^x_{N(t-1)} & 0 & \cdots & 0 \end{pmatrix}}_{\overline{Q}^x_{t-1}} \underbrace{\begin{pmatrix} x_{t-1} \\ \lambda_{12(t-1)} \\ \vdots \\ \lambda_{1N(t-1)} \end{pmatrix}}_{\hat{x}_{t-1}} + \underbrace{\begin{pmatrix} Q^x_{1(t-1)}\tilde{x}_{1(t-1)} \\ Q^x_{2(t-1)}\tilde{x}_{2(t-1)} \\ \vdots \\ Q^x_{N(t-1)}\tilde{x}_{N(t-1)} \end{pmatrix}}_{q^x_{t-1}}$$

where

$$\overline{\overline{\tilde{F}}}_t = s_t + \sum_{j=1}^{N} B_{jt}\tilde{u}_{jjt} \ . \qquad (A77)$$

Note that $\hat{x}_t$ and λ_t can be seen as supplementary state and costate vector of the leader, respectively. Now, we have to solve a *N*-point boundary problem, where the initial condition for the state, $\hat{x}_t$, and the terminal condition for the costate, λ_t, are required. The initial condition for the vector x_t is already given by $x_0 = x(0)$. The missing initial conditions for λ_{1it} $(i = 2,\dots,N)$ can be obtained by the definition $\lambda_{1i}(0) := 0$ $(i = 2,\dots,N)$ according to de Zeeuw and van der Ploeg (1991) or Dockner and Neck (1988). Hence,

$$\hat{x}_0 = \begin{pmatrix} x \\ \lambda_{12} \\ \vdots \\ \lambda_{1N} \end{pmatrix}_0 = \begin{pmatrix} x_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \qquad (A78)$$

The terminal condition for the supplementary costate can be defined using conditions $(A67)$ and $(A73)$,

$$\hat{\lambda}_T = \overline{Q}^x_T \ \hat{x}_T - q^x_T \ . \qquad (A79)$$

Now we proceed in analogy to open-loop Nash. We assume again that the supplementary state and costate vectors are related linearly, $\lambda_t = H_t \hat{x}_t + h_t$, which allows to derive Riccati equations to yield Riccati matrices for all periods recursively backwards in time.

$$\hat{\lambda}_t = H_t\left(\overline{A}_t \ \hat{x}_{t-1} + \overline{\overline{E}}_t \ \hat{\lambda}_t + \overline{\overline{F}}_t\right) + h_t \qquad (A80)$$

may be rearranged and solved with respect to λ_t. Substitution into the costate equation yields

$$\hat{\lambda}_{t-1} = \overline{A}_t' \left(I - H_t \tilde{\overline{\overline{E}}}_t \right)^{-1} \left(H_t \overline{A}_t \hat{x}_{t-1} + H_t \overline{\overline{\overline{F}}}_t + h_t \right) + \overline{Q}_{t-1}^x \hat{x}_{t-1} - q_{t-1}^x . \quad (A81)$$

Consequently, the lagged Riccati matrices may be identified by

$$H_{t-1} = \overline{A}_t' \left(I - H_t \tilde{\overline{\overline{E}}}_t \right)^{-1} H_t \overline{A}_t + \overline{Q}_{t-1}^x , \quad (A82a)$$

$$h_{t-1} = \overline{A}_t' \left(I - H_t \tilde{\overline{\overline{E}}}_t \right)^{-1} \left[H_t \overline{\overline{\overline{F}}}_t + h_t \right] - q_{t-1}^x . \quad (A83a)$$

The terminal conditions for the Riccati equations are determined as

$$H_T = \begin{pmatrix} Q_{1T}^x & Q_{2T}^x & \cdots & Q_{NT}^x \\ Q_{2T}^x & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ Q_{NT}^x & 0 & \cdots & 0 \end{pmatrix} \text{, and} \quad (A82b)$$

$$h_T = -\begin{pmatrix} Q_{1T}^x \tilde{x}_{1T} \\ Q_{2T}^x \tilde{x}_{2T} \\ \vdots \\ Q_{NT}^x \tilde{x}_{NT} \end{pmatrix} . \quad (A83b)$$

Substitution into the equation for the calculation of the supplementary state vector yields

$$\hat{x}_t = \left(I - \tilde{\overline{\overline{E}}}_t H_t \right)^{-1} \left(\overline{A}_t \hat{x}_{t-1} + \tilde{\overline{\overline{E}}}_t h_t + \overline{\overline{\overline{F}}}_t \right) . \quad (A84)$$

Finally, we summarize the derivation of the open-loop Stackelberg solution: Starting with the terminal conditions, $(A82b)$ and $(A83b)$, the Riccati matrices, $(A82a)$ and $(A83a)$, can be solved by backward integration. The state as well as the optimal control values may be calculated by forward integration as described by equation $(A84)$, and by

$$\underbrace{\begin{pmatrix} u_{1t} \\ u_{2t} \\ \vdots \\ u_{Nt} \end{pmatrix}}_{u_t} = \underbrace{\begin{pmatrix} \tilde{u}_{11t} \\ \tilde{u}_{22t} \\ \vdots \\ \tilde{u}_{NNt} \end{pmatrix}}_{\tilde{u}_t} - \underbrace{\begin{pmatrix} \left(Q_{1t}^{u_1} \right)^{-1} B_{1t}' & 0 & \cdots & 0 \\ 0 & \left(Q_{2t}^{u_2} \right)^{-1} B_{2t}' & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \left(Q_{Nt}^{u_N} \right)^{-1} B_{Nt}' \end{pmatrix}}_{\overline{Q}_t^u} \left(H_t \hat{x}_t + h_t \right) . \quad (A85)$$