BLS WORKING PAPERS

U.S. DEPARTMENT OF LABOR
Bureau of Labor Statistics

OFFICE OF PRICES AND LIVING CONDITIONS

An Anticipative Feedback Solution for the Infinite-Horizon, Linear-Quadratic, Dynamic, Stackelberg Game

Baoline Chen, Rutgers University
Peter A. Zadrozny, U.S. Bureau of Labor Statistics

Working Paper 350
November 2001

AN ANTICIPATIVE FEEDBACK SOLUTION FOR THE INFINITE-HORIZON, LINEAR-QUADRATIC, DYNAMIC, STACKELBERG GAME*

Baoline Chen
Department of Economics
Rutgers University
Camden, NJ 08102
e-mail: baoline@crab.rutgers.edu

Peter A. Zadrozny
Bureau of Labor Statistics
2 Massachusetts Ave., NE
Washington, DC 20212
e-mail: zadrozny_p@bls.gov

September 17, 2001

JEL Classification: C61, C63, C73

Additional key words: noncooperative games, solving Riccati-type nonlinear algebraic equations.

ABSTRACT

This paper derives and illustrates a new suboptimal-consistent feedback solution for an infinite-horizon, linear-quadratic, dynamic, Stackelberg game. This solution lies in the same solution space as the infinite-horizon, dynamic-programming, feedback solution but puts the leader in a preferred equilibrium position. The idea comes from Kydland (1977) who suggested deriving a consistent feedback solution for an infinite-horizon, linear-quadratic, dynamic, Stackelberg game by varying the coefficients in the player's linear constant-coefficient decision rules. Here feedback is understood in the sense of setting a current control vector as a function of a predetermined state vector. The proposed solution is derived for discrete- and continuous-time games and is called the anticipative feedback solution. The solution is illustrated with a numerical example of a duopoly model.

*The paper represents the authors' views and does not represent any official positions of the Bureau of Labor Statistics. We thank the editors, referees, Jacek Krawczyk, and Randal Verbrugge for comments.

1. Introduction.

Dynamic Stackelberg (or leader-follower) games are useful tools for studying dynamic economic behavior in equilibrium settings in which some player is dominant. Because of their tractability, infinite-horizon, linear-quadratic, dynamic, Stackelberg (LQDS) games have received particular attention. LQDS games have been used to study non-competitive behavior in specific markets and to evaluate and design macroeconomic policies. For example, Sargent (1985) contains studies of energy markets based on LQDS games; Kydland and Prescott (1977) and Fischer (1980) studied optimal tax policy using DS games; Canzoneri and Gray (1985), Miller and Salmon (1985), and Turnovsky, Basar, and d'Orey (1988) studied international macroeconomic policy coordination using DS games; section 4 illustrates the present anticipative feedback solution in a LQDS game of a hypothetical industry. The anticipative feedback solution could be applied to the LQ approximation of any dynamic economic setting with a dominant agent.

Three decision spaces have been considered in dynamic games: open-loop, feedback, and closed-loop. In open-loop decisions, players set their control vectors as functions of time; in feedback decisions, players set their control vectors as functions of the current (or most recently determined or observed) state vector; in closed-loop decisions, players set their control vectors as functions of the history of the state vector, from the start of the game to the moment of decision. For example, Hansen, Epple, and Roberds (1985) considered open-loop solutions of discrete-time LQDS games, computed using Euler-type equations. Simaan and Cruz (1973) considered feedback solutions of DS games, computed using backwards recursions of dynamic programming. To emphasize the dynamic programming nature of these feedback solutions, we refer to them as dynamic programming feedback (DPF) solutions. Basar and Selbuz (1979), Basar and Olsder (1980), and Tolwinski (1981) considered classes of closed-loop solutions for discrete- and continuous-time, DS games, computed using non-standard (non-DP) recursions and differential equations. See Basar and Olsder (1995, ch. 7) for a comprehensive discussion of DS games.

A potential problem in DS games is that the solution which is optimal for the leader at the beginning of the game is time inconsistent. That is, it ceases to be optimal for the leader in subsequent periods. Consequently, the leader has an incentive to restart the game. In a rational-expectations setting, followers would recognize continual restarts. Such a succession of restarted leader-optimal solutions would be unsustainable and, hence, unappealing as a solution concept. The time inconsistency problem in DS games was first noted by Simaan and Cruz (1973), Kydland (1975, 1977), and Kydland and Prescott (1977) for open-loop solutions of DS games. In response, Simaan and Cruz (1973), Kydland (1975, 1977), and Kydland and Prescott (1977) considered DPF solutions of DS games. DPF solutions are time consistent by construction, but do not entirely solve the time consistency problem because in them the leader is continually tempted to switch to an optimal, open- or closed-loop, solution.

Basar and Selbuz (1979) and Basar and Olsder (1980) proposed closed-loop solutions for discrete- and continuous-time, DS games. The Basar-Selbuz-Olsder solutions require additional structural restrictions, beyond the usual concavity (or convexity), playability, and stability conditions (see section 2). However, whenever they are applicable, the Basar-Selbuz-Olsder solutions are time consistent. Nevertheless, even when applicable, the Basar-Selbuz-Olsder solutions are not subgame perfect. Tolwinski (1981) proposed a more general closed-loop solution for LQDS games (under weaker structural restrictions) which is nearly subgame perfect: if the follower deviates from the optimal solution path for some reason, the leader induces them to return to it after one period. The common feature of these closed-loop solutions is that they are incentive (or trigger) strategies in which the leader induces the follower to be on the solution path. In economics, the time consistency problem has similarly been addressed using incentive strategies (Barro and Gordon, 1983; Backus and Driffill, 1985; Rogoff, 1989).

The present paper introduces a new feedback solution for infinite-horizon LQDS games, called the anticipative feedback (AF) solution. The name is explained further in this section. Like the infinite-horizon DPF solution, the AF solution lies in the space of constant-coefficient, linear, feedback, decision rules and is subgame perfect, hence, is time consistent. However, the AF solution puts the leader closer to an open- or closed-loop optimal solution than the DPF solution. Thus, in the AF solution, the leader less tempted to switch to an optimal solution. The idea of the AF solution comes from Kydland

(1977, p.310), who suggested deriving a feedback solution for a discrete-time, infinite-horizon, LQDS game by varying the coefficients in players' linear, constant-coefficient, decision rules. The AF solution is derived for both discrete- and continuous-time versions of the LQDS games. Compared with the open-loop, DPF, and closed-loop solutions proposed by Hansen, Epple, and Roberds (1985), Simaan and Cruz (1973), and Tolwinski (1981), the AF solution has the following three merits:

1. The AF solution is in the same space as the infinite-horizon DPF solution, namely, the space of linear, constant-coefficient decision rules in which the current control vector feeds back only on the current state vector. This solution space is a product space of real-valued matrices of finite dimensions. As explained below, when anticipative effects are suppressed, the AF solution reduces to the infinite-horizon DPF solution. Compared to this DPF solution, the leader is better off in the AF solution and is, therefore, less tempted to switch to an optimal, time-inconsistent solution.

2. Like the DPF solution, the AF solution is subgame perfect by construction, hence, is time consistent.

3. Although open- and closed-loop solutions generally are preferred by the leader, their solution spaces of sequences of control vectors or decision functions are much more complicated. In infinite-horizon games, the sequences are infinite. To be practical as policy prescriptions, DS game solutions should involve simple and easily understood decision rules. The lower-dimensional AF solution is simpler and more easily understood.

Anticipative control is the leader's ability to influence the state vector's evolution by accounting for the follower's current reactions to changes in the leader's current and expected future control settings. The effect, manifested in the dependence of the optimal solution on the initial state vector (cf., Hansen, Epple, and Roberds, 1985), causes DS games to be time inconsistent. The AF solution is consistent by construction, through the assumption that coefficient matrices of decision rules are independent of the initial state vector. By virtue of the principle of optimality upon which it is based, the DPF solution cannot account for anticipative control effects.

In the present AF solution, anticipative control manifests itself through the matrix Ψ, that measures how the follower's optimal valuation matrix W_2 varies in response to variations in the leader's policy rule. Here $W_2 = \lim_{h\to\infty} W_2(h)$ is the limit of the follower's optimal valuation matrix as the planning horizon, h, goes to infinity. In the DPF solution at backwards recursion h, by taking $W_2(h-1)$ as given, the leader ignores anticipative control effects and, in effect, sets Ψ identically equal to zero. In the AF solution, the leader sets Ψ so as to induce the follower to make decisions which put the leader in a more favorable solution. Thus, Ψ is an incentive tool of the leader so that, like Basar-Selbuz-Olsder's and Tolwinski's solutions, the AF solution is an incentive-based enhancement of the DPF solution.

The AF solution relates more directly to the following previous work. Zadrozny (1988) reported the discrete-time AF solution in summary form, without giving any derivations or applications. Medanic (1978) derived a related continuous-time solution, using the maximum principle. Whereas Medanic randomized the initial state vector, here, as usual, it is taken as given. Otherwise, the AF solution appears not to have been reported before.

This paper is organized as follows: For simplicity, only the two-player game is treated. The extension to n-player games is conceptually straightforward but notationally tedious. Section 2 presents the discrete-time game. Section 3 derives nonlinear algebraic (or nonrecursive) Riccati-type solution equations for the discrete-time AF game. Appendix B derives analogous continuous-time solution equations. Section 4 describes a trust-region gradient method for solving the discrete-time AF equations and presents illustrative, numerical, DPF and AF, solutions of a duopoly model. Section 5 contains concluding remarks.

2. The Discrete-Time Game.

Two players are indexed by an ordered pair $(i,j) \in \{(1,2), (2,1)\}$, where i is the player on whom attention is focused and j is the opponent. Player 1 is the leader.

Let y(t) be an $n\times 1$ vector of "outputs" generated by the ARMAX process

$$(2.1)\qquad y(t) = A_1y(t-1) + \dots + A_py(t-p) + B_0u(t) + \dots + B_qu(t-q) + C_0\varepsilon(t) + \dots + C_r\varepsilon(t-r),$$

where p, q, and r are positive integers, u(t) is an $m\times 1$ vector of players' controls, and $\varepsilon(t)$ is an $n\times 1$ vector of independent (white-noise) disturbances with zero mean and constant covariance matrix, Σ_ε, i.e., $\varepsilon(t) \sim IID(0,\Sigma_\varepsilon)$.

Let $u(t) = (u_1(t)^T, u_2(t)^T)^T$, where $u_i(t)$ is the $m_i\times 1$ subvector controlled by player i and superscript T denotes vector or matrix transposition. Coefficients B_0, ..., B_q are partitioned conformably. There are no equality or inequality restrictions on u(t), so that it ranges over an m-dimensional Euclidean space. By appropriately zeroing out coefficients, a subvector of y(t) can be made exogenous to the game. For example, if $y(t) = (y_1(t)^T, y_2(t)^T)^T$ and conformable (2,1) blocks in A_1, ..., C_r are zero and Σ_ε is block diagonal, then, $y_2(t)$ is exogenous.

To write process (2.1) in state-space form, following Ansley and Kohn (1983), first, let x(t) be the $s\times 1$ state vector $x(t) = (x_1(t)^T, \dots, x_\nu(t)^T)^T$ where $\nu = \max(p,q+1,r+1)$ and each $x_1(t)$ is $n\times 1$, so that $s = \nu n$. Then, with $x_1(t) = y(t)$, equation (2.1) is equivalent to

$$(2.2)\qquad x(t) = Fx(t-1) + Gu(t) + H\varepsilon(t),$$

$$F = \begin{bmatrix} A_1 & I & 0 & \dots & 0 \\ \cdot & 0 & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & I \\ A_\nu & 0 & \cdot & \dots & 0 \end{bmatrix},\quad G = \begin{bmatrix} B_0 \\ \cdot \\ \cdot \\ \cdot \\ B_{\nu-1} \end{bmatrix},\quad H = \begin{bmatrix} C_0 \\ \cdot \\ \cdot \\ \cdot \\ C_{\nu-1} \end{bmatrix},$$

where $A_i = 0$ for $i > p$, $B_i = 0$ for $i > q$, and $C_i = 0$ for $i > r$.

The state-space representation of process (2.1), thus, comprises state equation (2.2) and observation equation $y(t) = Mx(t)$, where $M = [I, 0, \dots, 0]$ is an $n\times s$ selection matrix. Partition $G = [G_1, G_2]$ conformably with $u(t) = (u_1(t)^T, u_2(t)^T)^T$, so that $Gu(t) = G_1u_1(t) + G_2u_2(t)$.

In each period t, player i maximizes

(2.3) $$V_i(t,h) = E_{t-1}\sum_{k=0}^{h} \delta^k \pi_i(t+k)$$

with respect to linear feedback policies on $u_i(t)$, defined below, where $E_t[\cdot] = E[\cdot|\Omega(t)]$, $\Omega(t) = \{y(t-k), u(t-k) \mid k \geq 0\}$, δ is a real discount factor satisfying $0 < \delta < 1$, and

(2.4) $$\pi_i(t) = u(t)^T R_i u(t) + 2u(t)^T \bar{S}_i y(t-1) + y(t-1)^T \bar{Q}_i y(t-1).$$

The matrices R_i, $\bar{S}_i$, and $\bar{Q}_i$ determine players' profits (or losses) for different values of u(t) and y(t-1). For example, in the illustrative example of section 4, R_i, $\bar{S}_i$, and $\bar{Q}_i$ are derived from the output demand curve and the production function and, therefore, depend on the parameters of these structural components. Without loss of generality, we assume that R_i and $\bar{Q}_i$ are symmetric. Although we focus on $h = \infty$, initially we assume h is finite. Allowing some element of y(t) to be identically equal to one introduces a constant term into process (2.1) and linear terms into objective (2.3).

Defining $S_i = \bar{S}_i M$ and $Q_i = \bar{Q}_i M$, we write

(2.5) $$\pi_i(t) = u(t)^T R_i u(t) + 2u(t)^T S_i x(t-1) + x(t-1)^T Q_i x(t-1).$$

In accordance with the partition of u(t) into $u_i(t)$ and $u_j(t)$, partition R_i into R_i^{ii}, R_i^{ij}, and R_i^{jj}, $\bar{S}_i$ into $\bar{S}_i^1$ and $\bar{S}_i^j$, and, hence, S_i into S_i^i and S_i^j, to obtain

(2.6) $$\pi_i(t) = u_i(t)^T R_i^{ii} u_i(t) + 2u_i(t)^T R_i^{ij} u_j(t) + u_j(t)^T R_i^{jj} u_j(t)$$

$$+ 2u_i(t)^T S_i^i x(t-1) + + 2u_j(t)^T S_i^j x(t-1) + x(t-1)^T Q_i x(t-1),$$

for (i,j) = (1,2) and (2,1).

We assume that each player knows: (a) $\Omega(t-1)$ at the beginning of period t; (b) the game's structure, and (c) the game's parameters, i.e., the coefficients of process (2.2) and objective (2.3), for (i,j)

= (1,2) and (2,1). Each player, thus, has the same information about exogenously given quantities. The information set is complete except for knowledge of x(t), which might have to be inferred from Ω(t). Games in which different players have different information sets are substantially more complicated (Townsend, 1983) and are not considered here.

We assume that the players follow constant linear feedback policies $u_i(t+k) = P_i x(t+k-1|t+k-1)$, k = 0, ..., h, where $x(t+k-1|t+k-1) = E_{t+k-1}x(t+k-1)$. We make this assumption for reasons of simplicity and understandability, as discussed in the introduction. In particular, "constant" means that the policy coefficient matrices, P_i, are nonstochastic functions of the parameters and are independent of the initial state vector. We write the players' policy rules jointly as $u(t+k) = Px(t+k-1|t+k-1)$, where $P = [P_1^T, P_2^T]^T$.

Because the game has a linear-quadratic structure and the players have identical information sets, the principle of certainty equivalence (also called the separation principle) applies (Astrom, 1970, pp. 278-279). Certainty equivalence says that the equilibrium value of P is independent of the probability distributions of ε(t) and v(t), and hence can be computed independently of x(t+k-1|t+k-1). Because computation of x(t+k-1|t+k-1) in the present case where both players have the same information is a familiar Kalman filtering exercise (Anderson and Moore, 1979, pp. 165-192), we focus on computing P and set $\Sigma_\varepsilon = 0$, so that x(t+k-1|t+k-1) = x(t+k-1).

Infinite-horizon dynamic games generally require the following three types of assumptions: (i) second-order concavity assumptions to ensure that players' optimization problems have locally unique solutions, (ii) playability assumptions (the term comes from Lukes and Russell, 1971) to ensure that an equilibrium exists, i.e., that players' reaction functions "intersect;" and (iii) stability assumptions to ensure that players' objectives remain finite as $h \rightarrow \infty$.

Following standard practice (Basar and Selbuz, 1979; Basar and Olsder, 1980), we make a broad concavity assumption. The concavity assumption also serves to maintain playability. The stability assumption that we make is the familiar stabilizability condition in linear optimal control theory.

Thus, to cover concavity and playability, we assume that (A) $\pi_i(t)$ is concave in u(t) and the endogenous variables in y(t) and (B) $\pi_i(t)$ is strictly concave in $u_i(t)$. Assumption (A) is equivalent to assuming that the matrix which defines the purely endogenous part of $\pi_i(t)$ in terms of u(t) and y(t) is negative semi-definite and assumption (B) is equivalent to assuming that R_i^{ii} is negative definite.

In the case of stability, first, $\Phi = F + GP$ is the closed-loop matrix of the game. Second, we account for discounting by multiplying Φ by $\sqrt{\delta}$. That is, we define the discounted closed-loop matrix $\tilde{\Phi} = \sqrt{\delta}\Phi$ or, equivalently, $\tilde{\Phi} = \tilde{F} + \tilde{G}P$, where $\tilde{F} = \sqrt{\delta}F$ and $\tilde{G} = \sqrt{\delta}G$. $\tilde{\Phi}$ is said to be discrete-time stable if its eigenvalues are less than one in modulus. A sufficient (but not always necessary) condition for $V_i(t) = \lim_{h\to\infty} V_i(t,h)$ to be finite is that $\tilde{\Phi}$ is stable. We assume P is restricted to values that imply that $\tilde{\Phi}$ is stable. P's that imply stable $\tilde{\Phi}$'s are themselves called stable. To ensure that the set of stable P's is nonempty, we assume that (C) the ordered pair $[\tilde{F}, \tilde{G}]$ is stabilizable. Stabilizability ensures existence of stable, constant, linear, feedback policies (Wonham, 1967).

Defining stabilizability is somewhat involved (e.g., Kwakernaak and Sivan, 1972, pp. 53-65 and 459-462). However, let $y(t) = (y_1(t)^T, y_2(t)^T)^T$, where $y_1(t)$ is endogenous and $y_2(t)$ is exogenous. Then, first, stabilizability implies that, abstracting from exogenous variables and disturbances, for any initial value of $y_1(t)$, and for any target value y_1^*, there is a control sequence, $\{u^*(t+k)\}_{k=0}^{N-1}$ that takes y_1 from the initial value to y_1^* in a finite number of periods N. Second, stabilizability implies that the conditional mean of $y_2(t)$ is of exponential order $1/\sqrt{\delta}$, i.e., $||E_t y_2(t+k)|| < c||y_2(t)||/\sqrt{\delta^k}$, for k = 0, 1, ..., where c is a positive constant and $||\cdot||$ is a vector norm. Because $0 < \delta < 1$, the second condition of stabilizability implies that exogenous variables can be nonstationary within this limit.

Detectability, which is dual to stabilizability, is usually assumed for the underlying optimal-control problem (i.e., cooperative solution of the game) to help ensure that $\tilde{\Phi}$ is stable (Kwakernaak and Sivan, pp. 65-81, 247-283, 462-466, and 495-501). Because concavity assumptions (A) and (B) and detectability may be insufficient to ensure stability in the present game setting, detectability is not assumed formally. Instead, we assume directly that P is stable.

A Stackelberg equilibrium can now be defined. In addition to what each player has been assumed to know, player i has conjecture Ξ_{ij} about player j's policy, P_j, and conjectural variation $\partial\Xi_{ij}/\partial P_i$ about how P_j, reacts to infinitesimal changes in P_i. Given this information, conjectures, and conjectural variations, player i maximizes objective (2.3) with respect to P_i subject to state equation (2.2). Let $P_i = \Gamma_i(\Xi_{ij}, \partial\Xi_{ij}/\partial P_i)$ denote player i's resulting optimal feedback matrix as a function of Ξ_{ij} and $\partial\Xi_{ij}/\partial P_{ij}$ and let Γ_{21} denote first-partial derivatives of Γ_2 with respect to the first argument Ξ_{12}. Recall that player 1 is the leader. A Stackelberg equilibrium occurs when: (I) $P = [P_1^T, P_2^T]^T$ is such that the players optimize ($P_1 = \Gamma_1$ and $P_2 = \Gamma_2$), (II) players' conjectures are confirmed ($\Xi_{12} = P_2$ and $\Xi_{21} = P_1$), (III) the leader's conjectural variations are confirmed ($\partial\Xi_{12}/\partial P_1 = \Gamma_{21}$), and (IV) the follower's conjectural variations are null ($\partial\Xi_{21}/\partial P_2 = 0$). Section 3 derives computationally useful forms of these equilibrium conditions for the discrete-time infinite-horizon linear-quadratic game. Appendix B derives analogous continuous-time solution equations.

3. Derivation of Discrete-Time AF Solution Equations.

First, we state definitions and rules of matrix differentiation and, then, use the rules to derive discrete-time AF solution equations.

3.1. Definitions and Rules of Matrix Differentiation.

First, motivated by Magnus and Neudecker (1988), we define matrix derivatives in terms of matrix differentials. Then, we state two rules of matrix differentiation, a product rule and a trace rule, that we use to derive the AF solution equations.

Let $A(\theta) = \{A_{ij}(\theta)\}$ denote a real, differentiable, m×n, matrix function of a real, p×1 vector $\theta = (\theta_1, ..., \theta_p)^T$, where m, n, and p are any positive integers (previous uses of m, n, p, and q are temporarily suspended). Vector θ could be the vectorization of a matrix with a total of p elements. Let the $m\times n$ matrices $\partial_k A = \{\partial A_{ij}/\partial\theta_k\}$, for $k = 1, \ldots, p$, collect first-order derivatives of $A(\theta)$ in partial-derivative form. Let $dA_{ij} = \sum_{k=1}^{p} (\partial A_{ij}/\partial\theta_k)d\theta_k$, where $d\theta_k$ is an infinitesimal variation in θ_k. It suffices to consider $d\theta = (\hat{\theta}_1, \ldots, \hat{\theta}_p)^T d\varepsilon$, where $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_p)^T$ is a real vector of finite-valued elements and $d\varepsilon$ is a real, infinitesimal, scalar variation. Then, $dA = \{dA_{ij}\}$ is the $m\times m$ differential form of first-order derivatives of $A(\theta)$.

Product Rule. Let $A(\theta)$ and $B(\theta)$ be real, $m\times n$ and $n\times q$, differentiable, matrix functions of the real $p\times 1$ vector θ, where m, n, p, and q are any positive integers. Componentwise application of the scalar product rule of differentiation yields the matrix product rule of differentiation

(3.1) $$d(AB) = dA\cdot B + A\cdot dB$$

(Magnus and Neudecker, 1988, p. 148). For example, setting $\theta = P$, we use rule (3.1) to derive equation (3.5) from equation (3.4).

Trace Rule. Let A and B be real-valued, $m\times n$ and $n\times m$, matrices, so that $dA\cdot B$ is a meaningful matrix product, where dA is an infinitesimal variation in A. dA may be expressed as $\hat{A}d\varepsilon$, where $\hat{A}$ is a finite-valued, nonzero, $m\times n$ matrix and $d\varepsilon$ is an infinitesimal, nonzero, scalar variation. Let $tr(\cdot)$ denote the trace of a square matrix. The trace rule is:

(3.2) $$\text{If } tr(dA\cdot B) = 0, \text{ for all } dA = \hat{A}d\varepsilon \in \mathbf{R}^{m\times n}, \text{ then, } B = 0.$$

To prove trace rule (3.2), choose some $(i,j) \in \{(1,1), \ldots, (m,n)\}$. Consider $\hat{A}$ such that $\hat{A}_{ij} \neq 0$, $\hat{A}_{k\ell} = 0$, for $(k,\ell) \neq (i,j)$, and remember that $d\varepsilon \neq 0$. Then, $tr(dA\cdot B) = \sum_{k=1}^{m} \sum_{\ell=1}^{n} dA_{k\ell}B_{\ell k} = \hat{A}_{ij}B_{ji}d\varepsilon = 0$ implies $B_{ji} = 0$. Repeating this argument for all other $(i,j) \in \{(1,1), \ldots, (m,n)\}$, we complete the proof.

In subsection 3.2, we use rule (3.2) to convert first-order conditions from the unsolvable form $tr(dA\cdot B) = 0$ to the solvable form $B = 0$. In these applications, dA represents all possible variations in a policy rule coefficient matrix. For example, setting $dA = d_i P_i^T = d_2 P_2^T$, we use rule (3.2) to convert equation (3.5) to (3.6). Because the variations, $dA \neq 0$, are used to determine first-order conditions of LQ maximization problems, their $\hat{A}$'s conceptually assume any nonzero values.

3.2. Derivation of Solution Equations.

We now derive the discrete-time AF solution equations. Because $u(t+k) = Px(t+k-1)$ and the disturbance $\varepsilon(t)$ is being ignored, state equation (2.2) implies $x(t+k-1) = \tilde{\Phi}^k x(t-1)$, for $k = 0, \ldots, h$. Thus, objective (2.3) implies $V_i(t,h) = x(t-1)^T W_i(h) x(t-1)$, where $W_i(h)$ is generated recursively by

$$(3.3) \qquad W_i(k) = \tilde{\Phi}^T W_i(k-1)\tilde{\Phi} + P^T R_i P + P^T S_i + S_i^T P + Q_i,$$

for $k = 1, \ldots, h$, and $W_i(0) = P^T R_i P + P^T S_i + S_i^T P + Q_i$. Because $\tilde{\Phi}$ is stable, in the limit as $h \to \infty$, $V_i(t,\infty) = x(t-1)^T W_i x(t-1)$, where $W_i = \lim_{h\to\infty} W_i(h)$ satisfies

$$(3.4) \qquad W_i = \tilde{\Phi}^T W_i \tilde{\Phi} + P^T R_i P + P^T S_i + S_i^T P + Q_i.$$

Let d_i denote the differential induced by infinitesimal variations in P_i. Then, the immediate first-order necessary condition for maximizing $V_i(t,\infty)$ with respect to player i's policy rule P_i is $d_i V_i(t,\infty) = 0$. Because $x(t-1)$ is independent of P_i and can assume any value, $d_i V_i(t,\infty) = x(t-1)^T d_i W_i x(t-1) = 0$ implies $d_i W_i = 0$. To see what $d_i W_i = 0$ implies, we use product rule (3.1) to differentiate equation (3.4) with respect to P_i, impose $d_i W_i = 0$, and obtain

$$(3.5) \qquad d_i P_i^T [R_i^{ii} P_i + R_i^{ij} P_j + S_i^i + \tilde{G}_i^T W_i \tilde{\Phi}]$$

$$+ d_i P_j^T [(R_i^{ij})^T P_i + R_i^{jj} P_j + S_i^j + \tilde{G}_j^T W_i \tilde{\Phi}] + \ldots = 0,$$

where $(i,j) \in \{(1,2), (2,1)\}$, the dots here and below denote asymmetric terms repeated in transposed form, d_iP_i denotes player i's variation of their policy rule coefficient matrix, and d_iP_j denotes player i's conjectural variation about player j, i.e., player i's assumption about how P_j responds to infinitesimal variations in P_i.

Next, we specialize equation (3.5) as the first-order necessary condition for player 2, the follower in the game. Being a Nash player, player 2 has a null conjectural variation about the leader, i.e., $d_2P_1 = 0$. Thus, we set $(i,j) = (2,1)$ in (3.5), impose $d_2P_1 = 0$, take the trace of (3.5), use $tr(A) = tr(A^T)$ and $tr(AB) = tr(BA)$ to consolidate terms (including those represented by repeated dots), divide by 2, apply trace rule (3.2) with $dA = d_2P_2$, and obtain

$$(3.6) \qquad R_2^{22}P_2 + R_2^{21}P_1 + S_2^2 + \tilde{G}_2^T W_2 \tilde{\Phi} = 0.$$

To obtain the first-order condition of the leader, we express d_1P_2 in terms of d_1P_1. First, we use product rule (3.1) to differentiate equation (3.6) with respect to P_1 and obtain

$$(3.7) \qquad d_1P_2 = M_1d_1P_1 + M_2d_1W_2\tilde{\Phi},$$

where $M_1 = -[R_2^{22} + \tilde{G}_2^T W_2 \tilde{G}_2]^{-1}[R_2^{21} + \tilde{G}_2^T W_2 \tilde{G}_1],$

$$M_2 = -[R_2^{22} + \tilde{G}_2^T W_2 \tilde{G}_2]^{-1}\tilde{G}_2^T.$$

Concavity assumptions (A) and (B) stated in section 2 imply $[R_2^{22} + \tilde{G}_2^T W_2 \tilde{G}_2]$ is negative definite and, hence, is nonsingular.

Next, to express d_1W_2 in terms of d_1P_1, we differentiate equation (3.4) for i = 2 with respect to P_1, simplify the result using equation (3.6) (an envelope theorem), and obtain

$$(3.8) \qquad d_1W_2 = d_1P_1^T N_{12} + \tilde{\Phi}^T d_1W_2 \tilde{\Phi} + \ldots,$$

where $N_{12} = (R_2^{21})^T P_2 + R_2^{11}P_1 + S_2^1 + \tilde{G}_1^T W_2 \tilde{\Phi}.$

Because $\tilde{\Phi}$ is stable, equation (3.8) is equivalent to

(3.9) $$d_1W_2 = \sum_{j=0}^{\infty}(\tilde{\Phi}^T)^j[d_1P_1^T N_{12} + N_{12}^T d_1P_1]\tilde{\Phi}^j.$$

Next, we use equation (3.9) to eliminate d_1W_2 from equation (3.7), use the result to eliminate d_1P_2 from equation (3.5), for (i,j) = (1,2), and obtain

(3.10) $$d_1P_1^T[N_{11} + M_1^T N_{21}] + \sum_{j=0}^{\infty}(\tilde{\Phi}^T)^{j+1}[d_1P_1^T N_{12} + N_{12}^T d_1P_1]\tilde{\Phi}^j M_2^T N_{21} + \ldots = 0,$$

where $N_{11} = R_1^{11}P_1 + R_1^{12}P_2 + S_1^1 + \tilde{G}_1^T W_1\tilde{\Phi}$,

$N_{21} = (R_1^{12})^T P_1 + R_1^{22}P_2 + S_1^2 + \tilde{G}_2^T W_1\tilde{\Phi}$.

Next, we take the trace of equation (3.10), use $tr(A) = tr(A^T)$ and $tr(AB) = tr(BA)$ to consolidate terms, divide by 2, and obtain

(3.11) $$tr\{d_1P_1^T[N_{11} + M_1^T N_{21} + N_{12}\sum_{j=0}^{\infty}\tilde{\Phi}^j(\tilde{\Phi}N_{21}^T M_2 + M_2^T N_{21}\tilde{\Phi}^T)(\tilde{\Phi}^T)^j]\} = 0.$$

Because equation (3.11) is in the form of $tr(dA\cdot B) = 0$, where $dA = d_1P_1^T$ can assume any $n\times m_1$ value, trace rule (3.2) implies

(3.12) $$N_{11} + M_1^T N_{21} + N_{12}\Psi = 0,$$

where $$\Psi = \sum_{j=0}^{\infty}\tilde{\Phi}^j[\tilde{\Phi}N_{21}^T M_2 + M_2^T N_{21}\tilde{\Phi}^T](\tilde{\Phi}^T)^j.$$

Because $\tilde{\Phi}$ is a stable matrix,

(3.13) $$\Psi = \tilde{\Phi}\Psi\tilde{\Phi}^T + M_2^T N_{21}\tilde{\Phi}^T + \tilde{\Phi}N_{21}^T M_2.$$

The anticipative control effects manifest themselves through Ψ, which measures how the follower’s valuation matrix, W_2, varies in response to variations in the leader’s policy rule. In the DPF solution at backwards recursion h, by taking $W_2(h-1)$ as given, the leader ignores anticipative control effects and, in effect, sets $\Psi = 0$. Thus, by dropping equation (3.13) and setting $\Psi = 0$, the AF solution equations reduce to the DPF solution equations. Furthermore, if we set $M_1 = 0$ in equation (3.12), so that it reduces to $N_{11} = 0$, the DPF solution equations reduce to Nash equilibrium solution equations.

We have derived algebraic Riccati-type solution equations for the anticipative feedback solution of the discrete-time, linear-quadratic, infinite-horizon, Stackelberg game: equations (3.4), for i = 1 and 2, (3.6), (3.12), and (3.13). Equations (3.4), for i = 1, (3.12), and (3.13) are the leader's complete first-order conditions and equations (3.4), for i = 2, and (3.6) are the follower's complete first-order conditions. Let $\vartheta = (\vartheta_1^T, \vartheta_2^T)^T$, where $\vartheta_1 = (\text{vec}(P_1)^T, \text{vech}(W_1)^T, \text{vech}(\Psi)^T)^T$, $\vartheta_2 = (\text{vec}(P_2)^T, \text{vech}(W_2)^T)^T$, vec(·) denotes the columnwise vectorization of a matrix, and vech(·) denotes the columnwise vectorization of the nonredundant lower half of a symmetric matrix. Then, the AF solution equations comprise ms + (3/2)s(s+1) scalar-level equations for determining the same number of elements of ϑ.

4. Numerical Solution of Discrete-Time AF Equations.

First, we describe a numerical trust-region method for solving the AF equations and, then, illustrate the method with a duopoly model.

4.1. Numerical Solution Method.

The AF solution equations are nonlinear differentiable equations in ϑ and, therefore, are solvable using gradient methods such as the trust-region method (Moré et al., 1980). The trust-region method requires an initial value of ϑ, ϑ_0, which should be close to the AF solution value and, therefore, should satisfy its regularity conditions. Although the full AF regularity conditions are unknown, at a minimum ϑ_0 should imply that P is such that $\tilde{\Phi}$ is stable (otherwise W_1 and W_2 are likely to be undefined) and that W_1 and W_2 are negative semi-definite in endogenous state variables (otherwise second-order concavity conditions are violated).

A cautious approach in setting ϑ_0 for the AF solution is to compute a sequence of Pareto (or cooperative), Nash equilibrium (NE), and DPF solutions. The Pareto solution is a convenient starting point because its solution can be computed without initial values. The Pareto solution solves the problem of maximizing the weighted average of the players' expected present values, $\overline{V}(t,\infty) = \theta V_1(t,\infty) + (1-\theta)V_2(t,\infty)$, for some value of $\theta \in [0,1]$, with respect to the joint feedback matrix P, subject to the state law of motion (2.2). Computing the Pareto solution involves solving the standard discrete-time algebraic matrix Riccati equation, which can be done accurately and quickly using a Schur-decomposition method (Laub, 1979). Under the concavity and stabilizability conditions, the Pareto solution yields the desired stable P and the negative semi-definite endogenous part of $\overline{W} = \theta W_1 + (1-\theta)W_2$. We could use the Pareto solution as an initial value for computing the AF solution. A more cautious approach computes successive Pareto, NE, DPF, and AF solutions, using the Pareto, NE, and DPF solutions as initial values for the NE, DPF, and AF solutions. The idea here is that the Pareto, NE, DPF, and AF solutions should "line up" in the solution space because the leading player has ever greater dominance in the sequence of solutions. The last step of computing the AF solution is greatly simplified by using the fact that, given the first player's solution values ϑ_1, the second player solves a standard Riccati equation.

4.2. Illustrative Numerical Solutions of a Duopoly Model.

The model is a modification of the model in Chen and Zadrozny (2001). As before, subscripts i = 1 and 2 refer to the leading and following players, respectively. The players are firms that produce $q_1(t)$ and $q_2(t)$ amounts of a good. The demand for the good is given by

$$(4.1) \qquad p_q(t) = -\eta q(t) + d(t),$$

where $q(t) = q_1(t) + q_2(t)$, $\eta > 0$ is a slope parameter, and d(t) is the demand state generated by the AR(1) process

$$(4.2) \qquad d(t) = \phi_d d(t-1) + \zeta_d(t),$$

where $|\phi_d| < 1/\sqrt{\delta}$ and the disturbance $\zeta_d(t)$ is distributed IID with zero mean.

The firms use capital, $k_i(t)$, labor, $\ell_i(t)$, and materials, $m_i(t)$, to produce output and invest in capital. Investment, $n_i(t)$, has two stages: purchasing capital goods and installing them. Installing capital is an "output activity" because it uses resources that could otherwise be used to produce output. The output activities are restricted according to the production function

$$(4.3) \qquad h(q_i(t), n_i(t)) = g(k_i(t), \ell_i(t), m_i(t)),$$

where $g(\cdot)$ and $h(\cdot)$ are the constant elasticity functions,

$$(4.4) \qquad g(k_i(t), \ell_i(t), m_i(t)) = (\alpha_1 k_i(t)^\beta + \alpha_2 \ell_i(t)^\beta + \alpha_3 m_i(t)^\beta)^{1/\beta},$$

$$h(q_i(t), n_i(t)) = (\gamma_1 q_i(t)^\rho + \gamma_2 n_i(t)^\rho)^{1/\rho},$$

where $\alpha_i > 0$, $\alpha_1 + \alpha_2 + \alpha_3 = 1$, $\beta < 1$, $\gamma_i > 0$, $\gamma_1 + \gamma_2 = 1$, and $\rho > 1$. $|\beta-1|^{-1}$ is the constant elasticity of substitution among inputs and $|\rho-1|^{-1}$ is the constant elasticity of transformation between output activities. Including n_i in $h(\cdot)$ is a parsimonious way of specifying internal adjustment costs: for given input resources, ever more units of output must be forgone as investment increases. Adjustment costs arise only during the installation of capital goods. Mathematically, $\rho > 1$ is a necessary and sufficient condition for the output transformation curves to be concave. The transformation curves become more curved, and, hence, adjustment costs increase as ρ increases. Similarly, $\beta < 1$ is a necessary and sufficient condition for the input isoquants to be convex to the origin. The isoquants become more curved and input substitutability decreases as β decreases.

To obtain linear-quadratic optimization problems for the firms, we describe the production function in terms of the quadratic approximation of its dual variable production cost function (DVPCF). The variable production costs are $c_{qi}(t) = p_\ell(t)\ell_i(t) + p_m(t)m_i(t)$, where $p_\ell(t)$ and $p_m(t)$ are the hiring and purchase prices of labor and materials. The labor and materials costs are called variable because these inputs are free of adjustment costs. The DVPCF is denoted by $c_q(w_i(t))$ and defined as follows: given $w_i(t) = (w_{i1}(t), ..., w_{i5}(t))^T = (q_i(t), n_i(t), k_i(t), p_\ell(t), p_m(t))^T$, $c_q(w_i(t))$ = minimum of $p_\ell(t)\ell_i(t) + p_m(t)m_i(t)$ with respect to $\ell_i(t)$ and $m_i(t)$, subject to the production function (4.3)-(4.4).

The firms' remaining costs are the purchase costs of capital goods $c_{ni}(t) = p_n(t)n_i(t)$, where $p_n(t)$ is the purchase price of capital goods. Thus, the firms' profits are $\pi_i(t) = r_{qi}(t) - c_{qi}(t) - c_{ni}(t)$, where $r_{qi}(t) = p_q(t)q_i(t) = -\eta q(t)q_i(t) + d(t)q_i(t)$ is sales revenue. The quadratic approximation of $c_{qi}(t)$ is $(1/2)w_i(t)^T\nabla^2c_q(w_{i0})w_i(t)$, where $\nabla^2c_q(w_{i0})$ is the Hessian matrix of second-partial derivatives of $c_{qi}(t)$ evaluated at $w_{i0} = (1, 1, 1, \alpha_2, \alpha_3)^T$, a value that results in the simplest expression for $\nabla^2c_q(w_{i0})$. $\nabla^2c_q(w_{i0})$ is stated in appendix A in terms of the parameters of the production function. For simplicity, we write $\nabla^2c_q(w_{i0})$ as ∇^2c_q. Therefore,

(4.5) $$\pi_i(t) = -\eta q(t)q_i(t) + d(t)q_i(t) - (1/2)w_i(t)^T\nabla^2c_qw_i(t) - p_n(t)n_i(t).$$

The input prices are generated by the AR(1) processes

(4.6) $$p_n(t) = \phi_{pn}p_n(t\text{-}1) + \zeta_{pn}(t),$$

$$p_\ell(t) = \phi_{p\ell}p_\ell(t\text{-}1) + \zeta_{p\ell}(t),$$

$$p_m(t) = \phi_{pm}p_m(t\text{-}1) + \zeta_{pm}(t),$$

where the coefficients ϕ_{pn}, $\phi_{p\ell}$, and ϕ_{pm} are less than $1/\sqrt{\delta}$ in absolute value and the disturbances ζ_{pn}, $\zeta_{p\ell}$, and ζ_{pm} are distributed IID with zero means.

Each firm's capital accumulates according to the law of motion

(4.7) $$k_i(t) = \phi_kk_i(t\text{-}1) + n_i(t) + \zeta_k(t),$$

where $0 < \phi_k < 1$ and the disturbance ζ_k is distributed IID with mean zero.

The model's structural components have thus been specified.

Next, we simplify the firms' dynamic optimization problems by first solving for $\ell_i(t)$ and $m_i(t)$. We can do this because $\ell_i(t)$ and $m_i(t)$ are not control variables in the capital law of motion. Optimal values of $\ell_i(t)$ and $m_i(t)$, conditional on $q_i(t)$ and $n_i(t)$ being at their optimal values, are obtained using Shepard's lemma (an envelope theorem),

(4.8) $$\ell_i(t) = c_{41}q_i(t) + c_{42}n_i(t) + c_{43}k_i(t) + c_{44}p_\ell(t) + c_{45}p_m(t),$$

(4.9) $$m_i(t) = c_{51}q_i(t) + c_{52}n_i(t) + c_{53}k_i(t) + c_{54}p_\ell(t) + c_{55}p_m(t),$$

where $(c_{41}, ..., c_{45})$ and $(c_{51}, ..., c_{55})$ are the 4th and 5th rows of $\nabla^2 c_q$.

Then, to state the firms' remaining optimization problems in the general notation of sections 2 and 3, we define the 4×1 control vector $u(t) = (u_1(t), u_2(t))^T = (q_1(t), n_1(t), q_2(t), n_2(t))^T$ and the 6×1 state vector $x(t) = (k_1(t), k_2(t), p_i(t), p_\ell(t), p_m(t), d(t))^T$, and then assemble the dynamic equations (4.2), (4.6), and (4.7) as the state equation

(4.10) $$x_t = Fx_{t-1} + \begin{bmatrix} G_0 \\ 0_{4\times4} \end{bmatrix} u_t,$$

where $F = \text{diag}[\phi_k, \phi_k, \phi_{pi}, \phi_{p\ell}, \phi_{pm}, \phi_d]$, $G_0 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, and $0_{4\times4}$ is the 4×4 zero matrix. The matrices R_i, S_i, and Q_i, which define the profit function in the general notation, are stated in appendix A in terms of η and the elements of $\nabla^2 c_q$.

We computed examples of AF solutions for the model using a trust-region method described in section 4.1. The solutions were computed in less than 5 seconds on a personal computer using a 150-megahertz processor. A computed solution $\hat{\vartheta}$ satisfies (3.4) for i = 1 and 2, (3.6), (3.12), and (3.13) up to error matrices. The solution has k-digit precision if all elements of the error matrices are $\leq 10^{-k}$.

To compute the AF solution, first, we computed a Pareto solution for equally weighted player's objective functions, by solving a standard algebraic Riccati equation using a Schur-decomposition method. Second, we used the Pareto solution as an initial value for computing the DPF solution using the trust-region method. Finally, we used the DPF solution as an initial value for computing the AF solution using the trust-region method. We skipped the NE solution between the Pareto and DPF solutions. In all cases, $\tilde{\Phi}$ was stable and W_1 and W_2 were negative semi-definite in the two endogenous state variables, i.e., firms' capital stocks.

Table 1 reports Pareto, DPF, and AF solutions for five different pairs of production function elasticities. The remaining parameters are fixed at the following values: δ = .935, η = .5, $\gamma_1 = \gamma_2$ = .5, $\alpha_1 = \alpha_2 = \alpha_3$ = .333, and $\phi_k = \phi_{pi} = \phi_{p\ell} = \phi_{pm}$ = .9. In the Pareto solutions, players' objective functions are weighted equally, with θ = .5. The table reports the firms' optimized values, $V_i(t,\infty) = x(t-1)^T W_i x(t-1)$, for $x(t-1) = (1, 1, 1, 1, 1, 1)^T$, in the different solutions.

Table 1: Pareto, DPF, and AF Solutions of the Duopoly Model

Solution	**V_1**	**V_2**	**V_1**	**V_2**	**V_1**	**V_2**
	(β, ρ) = (-5, 5) (CES,CET)=(.17,.2)		**(β, ρ) = (-5, 7.5) (CES,CET)=(.17,.15)**		**(β, ρ) = (-5, 10) (CES,CET)=(.17,.11)**	
Pareto	.4292	.4292	.2611	.2611	.1417	.1417
AF	.4215	.4107	.2568	.2530	.1383	.1360
DPF	.4206	.4093	.2567	.2535	.1383	.1366
	(β, ρ) = (-5, 5) (CES,CET)=(.17,.2)		**(β, ρ) = (-3, 5) (CES,CET)=(.77,.2)**		**(β, ρ) = (-1, 5) (CES,CET)=(1.1,.2)**	
Pareto	.4292	.4292	2.6286	2.6286	5.0481	5.0481
AF	.4215	.4107	2.6199	2.6031	5.0368	4.9990
DPF	.4206	.4093	2.6188	2.6083	5.0345	5.0196

As expected, both players achieve the highest values in the Pareto solutions and the leader achieves higher values in the AF solutions than in the DPF solutions. At least in this model, for the parameter values considered, the values differ slightly among the five cases and between the players. As expected, the firms' values decline as β declines and ρ increases. This occurs because lower β's and higher ρ's make the firms' operations less flexible. First, lower β's imply lower input substitutability, which reduces firms' flexibility in changing input proportions when input prices change. Second, higher ρ's imply higher adjustment costs, which makes capital adjustments more costly. Generally, both firms' values increase as the leader shifts from DPF to AF decision rules.

5. Conclusion.

We have derived algebraic solution equations for a new solution, called the anticipative feedback (AF) solution, for discrete- and continuous-time, linear-quadratic, infinite-horizon, Stackelberg, two-player, dynamic games. The AF solution puts the leading player in a better position in comparison with the familiar dynamic programming feedback (DPF) solution. The paper illustrates discrete-time AF solutions for a duopoly model. The solutions are accurately and quickly computed using a trust-region method. The illustrations show that the leading firm indeed increases its value by switching from DPF to AF decisions, but that the value increases are small, at least in the duopoly model for the parameter values that are considered.

Appendix A: Coefficient Values of the Duopoly Model.

In this appendix, we state the R_i, S_i, Q_i matrices, for i = 1 and 2, that define the players' objective functions in the illustrative duopoly model. First, we state the elements of $\nabla^2 c_q$, denoted c_{ij}, in terms of the parameters of the production function. Then, we state the R_i, S_i, and Q_i matrices in terms of the output-demand slope η and the elements of $\nabla^2 c_q$.

The nonredundant upper-triangular elements of $\nabla^2 c_q$ are

$$c_{11} = \gamma_1(1-\gamma_1)(\rho-1) + \gamma_1^2\alpha_1(1-\beta)/(1-\alpha_1),$$

$$c_{12} = -\gamma_1\gamma_2(\rho-1) - \gamma_1^2\gamma_2(1-\beta)/(1-\alpha_1),$$

$$c_{13} = -\gamma_1\alpha_1(1-\beta)/(1-\alpha_1)],$$

$$c_{14} = c_{15} = \gamma_1/(1-\alpha_1),$$

$$c_{22} = \gamma_2(1-\gamma_2)(\rho-1)$$

$+ \gamma_2^2 \alpha_1(1-\beta)/(1-\alpha_1)$,

$c_{23} = -\gamma_2 \alpha_1(1-\beta)/(1-\alpha_1)$,

$c_{24} = c_{25} = \gamma_2/(1-\alpha_1)$,

$c_{33} = \alpha_1(1-\beta)[1+\alpha_1(2-\alpha_1)/(1-\alpha_1)]$,

$c_{34} = c_{35} = -\alpha_1/(1-\alpha_1)$,

$c_{44} = -\alpha_3/\alpha_2(1-\alpha_1)(1-\beta)$,

$c_{45} = 1/(1-\alpha_1)(1-\beta)$,

$c_{55} = -\alpha_2/\alpha_3(1-\alpha_1)(1-\beta)$.

Let $R_{i,jk}$ denote the (j,k) element of R_i and similarly for S_i and Q_i. For compactness, we state only nonzero elements. The nonzero, nonredundant, upper-triangular elements of R_1 and R_2 are

$R_{1,11} = -\eta - c_{11}$,

$R_{1,12} = -c_{12}$,

$R_{1,13} = -\eta$,

$R_{1,22} = -c_{22}$,

$R_{2,13} = -\eta$,

$R_{2,33} = -\eta - c_{11}$,

$R_{2,34} = -c_{12}$,

$R_{2,44} = -c_{22}$.

The nonzero elements of S_1 and S_2 are

$S_{1,11} = -c_{13}$,

$S_{1,14} = -c_{14}$,

$S_{1,15} = -c_{15}$,

$S_{1,21} = -c_{23}$,

$S_{1,23} = -1$,

$S_{1,24} = -c_{24}$,

$S_{1,25} = -c_{25}$,

$S_{2,32} = -c_{13}$,

$S_{2,34} = -c_{14}$,

$S_{2,35} = -c_{15}$,

$S_{2,36} = 1$,

$S_{2,42} = -c_{23}$,

$S_{2,43} = -1$,

$S_{2,44} = -c_{24}$,

$S_{2,45} = -c_{25}$.

The nonzero, nonredundant, upper-triangular elements of Q_1 and Q_2 are

$Q_{1,11} = -c_{33}$,

$Q_{1,14} = -c_{34}$,

$Q_{1,15} = -c_{35}$,

$Q_{1,44} = -c_{44}$,

$Q_{1,45} = -c_{45}$,

$Q_{1,55} = -c_{55}$,

$Q_{2,22} = -c_{33}$, $Q_{2,44} = -c_{44}$,

$Q_{2,24} = -c_{34}$, $Q_{2,45} = -c_{45}$,

$Q_{2,25} = -c_{35}$, $Q_{2,55} = -c_{55}$.

Appendix B: Continuous-Time Solution Equations.

In continuous time, analogously to the ARMAX process (2.1), y(t) is a $n\times 1$ vector generated by the stochastic differential equation

$$(B.1) \qquad D^{P}y(t) = A_1D^{P-1}y(t) + \ldots + A_py(t) + B_0D^{q}u(t) + \ldots + B_qu(t)$$

$$+ C_0D^{r}\varepsilon(t) + \ldots + C_r\varepsilon(t),$$

where D^j denotes the j-th mean-squared time derivative, p, q, and r are positive integers, $u(t) = (u_1(t)^T, u_2(t)^T)^T$ is an $m\times 1 = (m_1 + m_2)\times 1$ vector of players' controls, and $\varepsilon(t)$ is an $n\times 1$ vector of continuous-time white-noise disturbances.

To say that $\varepsilon(t)$ is continuous-time white noise means that it has independent increments, zero mean, and autocovariance function $\Sigma_{\varepsilon}\Delta(t_2-t_1)$, where Σ_{ε} is an $n\times n$, symmetric, positive definite, intensity matrix, $\Delta(t_2-t_1)$ is the Dirac delta function, and $t_2 \geq t_1$ are points in time. Strictly, $\varepsilon(t)$, its derivatives, and (B.1) are not well-defined, but for $p \geq 1 + \max(q,r)$, which is assumed, there is a well-defined (mean-squared) stochastic integral equation corresponding to (B.1) that gives it a rigorous foundation (Astrom 1970, pp. 13-90).

To put (B.1) in state-space form, we define a state vector x(t) exactly as in section 2 and, again, let $x_1(t) = y(t)$. Then, by a recursive-substitution argument similar to that yielding (2), we obtain

$$(B.2) \qquad Dx(t) = Fx(t) + Gu(t) + H\varepsilon(t),$$

where coefficient matrices F, G, and H are exactly as in the discrete-time state equation (2.2). Thus, the state-space representation of process (B.1) comprises state equation (B.2) and the observation equation $y(t) = Mx(t)$, where $M = [I, 0, ..., 0]$. As before, $u(t)$, G, and P partition as $u(t) = (u_1(t)^T, u_2(t)^T)^T$, $G = [G_1, G_2]$, and $P = [P_1^T, P_2^T]^T$.

Constant linear feedback policies are $u_i(t) = P_i x(t|t)$, where $x(t|t) = E_t[x(t)] = E\{x(t)|\Omega(t)]$ and $\Omega(t) = \{y(\tau), u(\tau)| \tau \leq t\}$. As in the discrete-time case, we suppress disturbances so that $x(t|t) = x(t)$.

At each moment t, given information $\Omega(t)$, assumptions about opponents' positions and reactions, and state equation (B.2), player i maximizes

$$\text{(B.3)} \qquad V_i(t,h) = E_t \int_{\tau=t}^{t+h} \exp[-\delta(\tau-t)]\pi_i(\tau)d\tau$$

with respect to P_i, subject to $u_i(t) = P_i x(t)$ and state equation (B.2), where $\delta > 0$ is a real discount factor and $\pi_i(\tau) = u(\tau)^T R_i u(\tau) + 2u(\tau)^T \overline{S}_i(\tau) + y(\tau)^T \overline{Q}_i y(\tau)$. As before, we assume that $\overline{Q}_i$ is symmetric, h is initially taken to be finite, $S_i = \overline{S}_i M$ and $Q_i = M^T \overline{Q}_i M$, and we consider $\pi_i(\tau)$ in the more detailed representation analogous to equation (2.4).

The discrete-time assumptions on parameters are retained, except for technical changes necessitated by the switch to continuous time. That is, as before, we assume (A) $\pi_i(\tau)$ is concave in $u(t)$ and in the endogenous variables in $y(t)$, (B) $\pi_i(t)$ is strictly concave in $u_i(t)$, and (C) after incorporating discounting, state equation (B.2) is stabilizable. Assumptions (A) and (B) impose the same restrictions as in the discrete-time case. However, whereas the assumption that R_i^{ii} is negative definite is generally not necessary in discrete time, it is necessary in continuous time.

In continuous time, the discounting and stability conditions are different. $\Phi = F + GP$ is still the closed-loop matrix, but in continuous time the discounted closed-loop matrix is $\tilde{\Phi} = \Phi - (\delta/2)I_n$, where I_n is the $n \times n$ identity matrix, or, equivalently, $\tilde{\Phi} = \tilde{F} + GP$, where $\tilde{F} = F - (\delta/2)I_n$. $\tilde{\Phi}$ is continuous-time stable if the real parts of its eigenvalues are negative. As before, P is restricted to values that imply that $\tilde{\Phi}$ is stable and the stabilizability of $[\tilde{F}, \tilde{G}]$ ensures that the set of stable P's is nonempty.

Constant linear feedback policies imply that $Dx(\tau) = \tilde{\Phi}\, x(\tau)$, hence, that $x(\tau) = \exp[\tilde{\Phi}(\tau-t)]x(t)$, for $\tau \geq t$ (Graham, 1981, pp. 108-110), where $\exp[\tilde{\Phi}(\tau-t)] = I + \Phi(\tau-t) + (1/2!)\tilde{\Phi}^2(\tau-t)^2 + \ldots$ is the matrix exponential. Thus, $V_i(t,h) = x(t)^T W_i(t,h)x(t)$, where

$$\text{(B.4)} \quad V_i(t,h) = \int_{\tau=t}^{t+h} \exp[\tilde{\Phi}^T(\tau-t)][P^TR_iP + P^TS_i + S_i^TP + Q_i]\exp[\tilde{\Phi}(\tau-t)]d\tau.$$

We premultiply equation (B.4) by $\tilde{\Phi}^T$, postmultiply it by $\tilde{\Phi}$, add the products together, and integrate the sum by parts. Because $\tilde{\Phi}$ is a stability matrix, in the limit as $h\rightarrow\infty$, we obtain

$$\text{(B.5)} \quad \tilde{\Phi}^TW_i + W_i\tilde{\Phi} + P^TR_iP + P^TS_i + S_i^TP + Q_i = 0,$$

where $W_i = \lim_{h\rightarrow\infty} W_i(t,h)$.

The immediate first-order necessary condition for maximizing $V_i(t,\infty)$ with respect to P_i is $d_iV_i(t,\infty) = 0$. Because $x(t)$ is given independently of P_i and can assume any value, $d_iV_i(t,\infty) = x(t)^Td_iW_ix(t) = 0$ implies $d_iW_i = 0$. To see what $d_iW_i = 0$ implies, we use product rule (3.1) to differentiate equation (B.5) with respect to P_i, impose $d_iV_i = 0$, and obtain

$$\text{(B.6)} \quad d_iP_i^T[R_i^{ii}P_i + R_i^{ij}P_j + S_i^i + G_i^TW_i]$$

$$+ d_iP_j^T[(R_i^{ij})^TP_i + R_i^{jj}P_j + S_i^j + G_j^TW_i] + \ldots = 0,$$

where $(i,j) \in \{(1,2,), (2,1)\}$, and the dots denote asymmetric terms repeated in transposed form.

To specialize equation (B.6) as the first-order necessary condition for the follower, we set $(i,j) = (2,1)$, impose $d_2P_1 = 0$, take the trace, use $tr(B) = tr(B^T)$ and $tr(AB)= tr(BA)$ to consolidate terms, divide by 2, apply trace rule (3.2) with $dA = d_2P_2$, and obtain

$$\text{(B.7)} \quad R_2^{22}P_2 + R_2^{21}P_1 + S_2^2 + G_2^TW_2 = 0.$$

To obtain the first-order necessary condition of the leader, we use product rule (3.1) to differentiate equation (B.7) with respect to P_1 and obtain

(B.8) $$d_1P_2 = M_1d_1P_1 + M_2d_1W_2,$$

where $M_1 = -(R_2^{22})^{-1}R_2^{21}$,

$M_2 = -(R_2^{22})^{-1}G_2^T$.

To express d_1W_2 in terms of d_1P_1, we differentiate equation (B.5), for i = 2, with respect to P_1, simplify the result using equation (B.7), and obtain

(B.9) $$\tilde{\Phi}^T d_1W_2 + d_1W_2\tilde{\Phi} + d_1P_1^T N_{12} + N_{12}^T d_1P_1 = 0,$$

where $N_{12} = (R_2^{21})^T P_2 + R_2^{ij}P_1 + S_2^1 + G_1^T W_2$.

Because $\tilde{\Phi}$ is a stable matrix, equation (B.9) is equivalent to

(B.10) $$d_1W_2 = \int_{\tau=0}^{\infty} \exp(\tilde{\Phi}^T\tau)[d_1P_1^T N_{12} + N_{12}^T d_1P_1]\exp(\tilde{\Phi}\tau)d\tau.$$

We use equation (B.10) to eliminate d_1W_2 from equation (B.8), then, use the result to eliminate d_1P_2 from equation (B.6), for (i,j) = (1,2), and obtain

(B.11) $$d_1P_1^T[N_{11} + M_1^T N_{21}]$$

$$+ \int_{\tau=0}^{\infty} \exp(\tilde{\Phi}^T\tau)[d_1P_1^T N_1 + N_{12}^T d_1P_1]\exp(\tilde{\Phi}\tau)M_2^T N_{21}d\tau = 0,$$

where $N_{11} = R_i^{ij}P_1 + R_1^{12}P_2 + S_1^1 + G_1^T W_1$,

$N_{21} = (R_1^{12})^T P_1 + R_1^{22}P_2 + S_1^2 + G_2^T W_1$.

Next, we take the trace of equation (B.11), use $tr(B) = tr(B^T)$ and $tr(AB) = tr(BA)$ to consolidate terms, divide by 2, and obtain

$$\text{(B.12)} \quad tr\{d_1P_1^T[N_{11} + M_1^T N_{21}$$

$$+ N_{12}\int_{\tau=0}^{\infty} \exp(\tilde{\Phi}\tau)[M_2^T N_{21} + N_{12}M_2]\exp(\tilde{\Phi}^T\tau)d\tau]\} = 0.$$

Because equation (B.12) is in the form of $tr(dA{\cdot}B) = 0$, where $dA = dP_1^T$ can assume any $n{\times}m_1$ value, trace rule (3.2) implies

$$\text{(B.13)} \quad N_{11} + M_1^T N_{21} + N_{12}\Psi = 0,$$

where $\Psi = \int_{\tau=0}^{\infty} \exp(\tilde{\Phi}\tau)[M_2^T N_{21} + N_{21}^T M_2]\exp(\tilde{\Phi}^T\tau)d\tau,$

or, equivalently,

$$\text{(B.14)} \quad \tilde{\Phi}\Psi + \Psi\tilde{\Phi}^T + M_2^T N_{21} + N_{21}^T M_2 = 0.$$

Thus, we have derived algebraic Riccati-type solution equations for the anticipative feedback solution of the continuous-time, linear-quadratic, infinite-horizon, Stackelberg, dynamic game: equations (B.5), for i = 1 and 2, (B.7), (B.13), and (B.14). Equations (B.5), for i = 1, (B.13), and (B.14) are the leader's complete first-order conditions and equations (B.5), for i = 2, and (B.7) are the follower's complete first-order conditions.

As in the discrete-time solution, if we set $\Psi \equiv 0$ and drop equation (B.14), the continuous-time Stackelberg AF solution reduces to the Stackelberg nonanticipative solution, which is analogous to the discrete-time DPF solution. In addition, if we set $M_1 \equiv 0$, the nonanticipative solution reduces to the Nash equilibrium solution.

REFERENCES

Anderson, B.D.O. and J.B. Moore (1979), Optimal Filtering, Englewood Cliffs, NJ: Prentice-Hall.

Ansley, C.F. and R. Kohn (1983), "Exact Likelihood of Vector Autoregressive Moving-Average Process with Missing or Aggregated Data," Biometrika 70: 275-278.

Astrom, K.J. (1970), Introduction to Stochastic Control Theory, Orlando, FL: Academic Press.

Backus, D. and J. Driffill (1985), "Rational Expectations and Policy Credibility Following a Change in Regime," Review of Economic Studies 52: 211-221.

Barro, R.J. and D.B. Gordon (1983), "Rules, Discretion and Reputation in a Model of Monetary Policy," Journal of Monetary Economics 12: 101-121.

Basar, T. and G.J. Olsder (1980), "Team-Optimal Closed-Loop Stackelberg Strategies in Hierarchical Control Problems," Automatica 16: 409-414.

Basar, T. and G.J. Olsder (1995), Dynamic Noncooperative Game Theory, second edition, London, UK: Academic Press.

Basar, T. and H. Selbuz (1979), "Closed-Loop Stackelberg Strategies with Applications in the Optimal Control of Multilevel Systems," IEEE Transactions on Automatic Control AC-24: 166-178.

Canzoneri, M. and J. Gray (1985), "Monetary Policy Games and the Consequences of Noncooperative Behavior," International Economic Review 26: 547-564.

Chen, B. and P.A. Zadrozny (2001), "Estimated U.S. Manufacturing Capital and Technology Based on an Estimated Dynamic Economic Model," typescript.

Fischer, S. (1980), "Dynamic Inconsistency, Cooperation, and the Benevolent Dissembling Government," Journal of Economic Dynamics and Control 2: 93-107.

Graham, A. (1981), Kronecker Products and Matrix Calculus with Applications, Chichester, UK: Ellis Horwood.

Hansen, L.P., D. Epple, and W. Roberds (1985), "Linear-Quadratic Duopoly Models of Resource Depletion," pp. 101-142 in Energy, Foresight, and Strategy, T.J. Sargent (ed.), Washington, DC: Resources for the Future.

Kwakernaak, H. and R. Sivan (1972), Linear Optimal Control Systems, New York, NY: Wiley-Interscience.

Kydland, F. (1975), "Noncooperative and Dominant Player Solutions in Discrete Dynamic Games," International Economic Review 16: 321-335.

Kydland, F. (1977), "Equilibrium Solutions in Dynamic Dominant-Player Models," Journal of Economic Theory 15: 307-324.

Kydland, F. and E.C. Prescott (1977), "Rules Rather than Discretion: The Inconsistency of Optimal Plans," Journal of Political Economy 85: 473-491.

Laub, A.J. (1979), "A Schur Method for Solving Algebraic Riccati Equations," IEEE Transactions on Automatic Control 24: 913-921.

Lukes, D.L. and D.L. Russell (1971), "A Global Theory for Linear-Quadratic Differential Games," Journal of Mathematical Analysis and Applications 33: 96-123.

Magnus, J.R. and H. Neudecker (1988), Matrix Differential Calculus with Applications in Statistics and Econometrics, New York, NY: John Wiley & Sons.

Medanic, J. (1978), "Closed-Loop Stackelberg Strategies in Linear-Quadratic Problems," IEEE Transactions on Automatic Control AC-23: 632-637.

Miller, M. and M. Salmon (1985), "Dynamic Games and the Time Inconsistency of Optimal Policy in Open Economies," Economic Journal 95: 124-137

Moré, J.J., B.S. Garbow, and K.E. Hillstrom (1980), "User Guide for MINPACK-1," Report ANL-80-74, Argonne National Laboratory, Argonne, IL.

Rogoff, K. (1989), "Reputation, Coordination, and Monetary Policy," pp. 236-264 in Modern Business Cycle Theory, R.J. Barro (ed.), Cambridge, MA: Harvard University Press.

Sargent T.J. (ed.) (1985), Energy, Foresight, and Strategy, Washington, DC: Resources for the Future.

Simaan, M. and J.B. Cruz, Jr. (1973), "Additional Aspects of the Stackelberg Strategy in Nonzero-Sum Games," Journal of Optimization Theory and Applications 11: 613-626.

Townsend, R.M. (1983), "Forecasting the Forecasts of Others," Journal of Political Economy 91: 546-588.

Tolwinski, B. (1981), "Closed-Loop Stackelberg Solution to a Multistage Linear-Quadratic Game," Journal of Optimization Theory and Applications 34: 485-501.

Turnovsky, S.J., T. Basar, and V. d'Orey (1988), "Dynamic Strategic Monetary Policies and Coordination in Interdependent Economies," American Economic Review 78: 341-361.

Wonham, W.M. (1967), "On Pole Assignment in Multi-Input Controllable Linear Systems," IEEE Transactions on Automatic Control AC-12: 660-665.

Zadrozny, P.A. (1988), "A Consistent Closed-Loop Solution for Infinite-Horizon Linear-Quadratic Dynamic Stackelberg Games," Journal of Economic Dynamics and Control 12: 155-159.