Oxford Economic Papers 43 (1991), 612–636

DIFFERENCE GAMES AND POLICY EVALUATION: A CONCEPTUAL FRAMEWORK

By A. J. de ZEEUW *and* F. van der PLOEG*

1. Introduction

THIS paper gives an overview of the various equilibrium concepts used in non-cooperative difference games and their economic interpretation. Difference games are dynamic games in discrete time. The state of the economy at time t, say y_t depends on the state of the economy at time $t-1$, y_{t-1}, and on the actions of the various players undertaken during this period. (Differential games are dynamic games in continuous time.) Difference games are unlike repeated games (supergames), because the latter refer to the repetition of a static game where the state of the economy in each game is independent of the state of the economy in previous games. History in repeated games matters only because players might condition their strategies on the history of play, but history in difference games matters also due to the dynamics of capital accumulation, wages, prices, etc.

To illustrate the various concepts employed in difference games, it is useful to discuss a classic example where actions can take on only one of two values. Figure 1 gives a simple example of such a dynamic game (due to Simaan and Cruz (1973)). The economy starts off in the state $y_0 = 0$. Subsequently each player can either take the action L or H. Each player minimizes a welfare loss function, which is time separable. The welfare losses incurred during the transition from the state at time 0 to the state at time 1 are given above the actions. From each state at time 1, each player can again take the actions L or H and at time 2 arrive at four possible states. The Nash solution concept represents the standard approach to non-cooperative games and is applicable when both players have equal strength. The actions in a Nash equilibrium must be the best response of player 1 to the action of player 2 and the best response of player 2 to the action of player 1. In dynamic games one distinguishes between the open-loop Nash equilibrium and the feedback Nash or subgame-perfect equilibrium. Two assumptions distinguish the feedback concept from the open-loop concept, namely information structure (Başar and Olsder (1982)) and period of commitment (Reinganum and Stokey (1985)). The open-loop Nash equilibrium presumes that the players at time 1 and 2 can only observe the initial state of the economy, y_0, i.e. have open-loop information patterns. The open-loop Nash equilibrium also presumes that at time 0 each player can make binding commitments about the actions he or she announces to undertake

* This paper arose from an earlier paper, entitled *Non-cooperative strategies for dynamic policy games and the problem of time inconsistency: A comment*, and has benefited from the comments of three anonymous referees, Andries Brandsma and the editorial advice of Peter Sinclair.

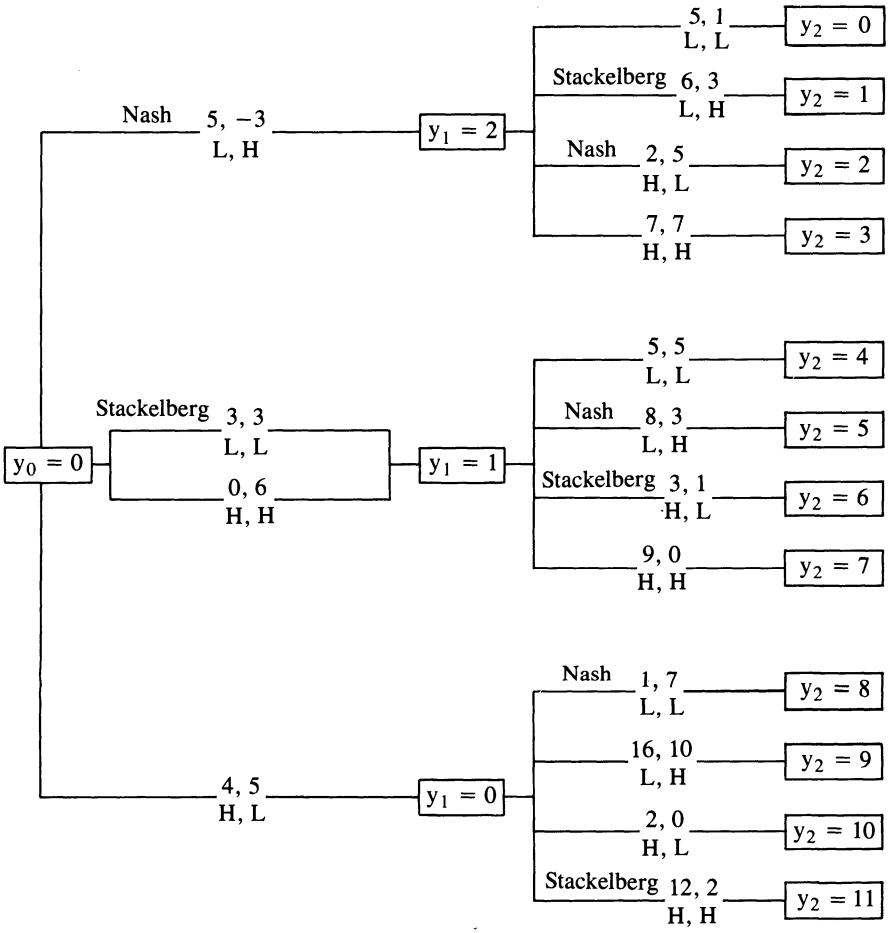

FIG. 1. Feedback equilibria in a dynamic game (extensive form)

in period 1 and 2, i.e. the period of commitment is equal to the entire planning period.

The normal form of the open-loop game associated with Figure 1 is presented in Table 1. The open-loop Nash equilibrium corresponds to the intersection of the reaction curve of player 1 (indicated by [a]) and the reaction curve of player 2 (indicated by [b]). The open-loop Nash equilibrium means that player 1 plays *H* in the first period and *L* in the second period, that player 2 plays *H* in both periods, and that the welfare losses are 8 to player 1 and 9 to player 2. The feedback Nash equilibrium presupposes that each player can observe the state of the economy at the beginning of the current period and therefore corresponds to a dynamic information structure and a period of commitment of one. It is constructed by imposing subgame perfectness, that is the Nash equilibrium for the whole game must remain a Nash equilibrium for every subgame starting

TABLE 1
Open-loop equilibria in a dynamic game (normal form)

Player 2 / *Player 1*	*LL*	*LH*	*HL*	*HH*
LL	8, 8	11, 6[a,d]	10, −2[b]	11, 0
LH	6, 4	12, 3	7, 2[b]	12, 4
HL	5, 12[a]	20, 15	5, 11	8, 9[a,b,c]
HH	6, 5[b]	16, 7	3, 7[a]	9,6

[a] Best response of player 1 to player 2 (player 1's reaction curve)
[b] Best response of player 2 to player 1 (player 2's reaction curve)
[c] Open-loop Nash equilibrium
[d] Open-loop Stackelberg equilibrium (with 2 as leader)

from an abritrary state at some point in time after the beginning of the whole game. One constructs the feedback Nash or subgame-perfect equilibrium by dynamic programming. This can be done with the aid of the extensive form of the game presented in Figure 1. First, one calculates the Nash equilibrium for each of the three subgames in the second period. One then adds on the resulting welfare losses to the welfare losses in the subgame of period 1, which results in 7, 2 for *L*, *H*, 11, 6 for *L*, *L*, 8, 9 for *H*, *H* and 5, 12 for *H*, *L*, and finally calculates the Nash equilibrium for the game starting from $y_0 = 0$ i.e. 7, 2 for *L*, *H*. In the resulting feedback Nash equilibrium player 1's actions are *L* and *H* whilst player 2's actions are *H* and *L* and both players are better off than in the open-loop Nash equilibrium (see Table 2). It is also possible to construct examples where the players are better off in the open-loop Nash equilibrium (see section 2.2). Hence, making use of information as it becomes available, can make players worse off in the context of a dynamic game, even though it always is profitable in a one-player context. There are many applications of open-loop and feedback Nash equilibria in dynamic games, e.g. conflict over the distribution of income in capitalist economies (Lancaster (1973)), conflict over the harvesting of a

TABLE 2
Welfare losses under the various outcomes

Solution concept	*Actions of player 1*	*Actions of player 2*	*Welfare losses*
Feedback Nash (subgame-perfect) equilibrium	LH	HL	7, 2
Feedback Stackelberg equilibrium (and reneging outcome)	LH	LL	6, 4
Open-loop Nash equilibrium	HL	HH	8, 9
Open-loop Stackelberg equilibrium	LL	LH	11, 6

common renewable resource (Reinganum and Stokey (1985); van der Ploeg (1986)), price sluggishness in duopolistic competition (Fershtman and Kamien (1987)), capacity investment in industrial organization (Reynolds (1987)), conflict over arms accumulation (van der Ploeg and de Zeeuw (1990)), and international pollution control (van der Ploeg and de Zeeuw (1991)).

The Lucas (1976) critique of econometric policy evaluation has increased the interest in applications of rational expectations and non-cooperative difference (or differential) game theory to dynamic economic or econometric models, because these techniques take explicit account of the reaction of the private sector (such as households and firms) to expected changes in government economic policy. Non-cooperative difference (or differential) games of the Stackelberg variety, with the government as leader and the private sector as follower, can provide a behavioural foundation of macroeconomic models with expectations of future government economic policy affecting the current state of the economy. Obviously, the advantage of difference (or differential) games of the Stackelberg variety over *ad hoc* rational expectations models is that they are immune to the Lucas (1976) critique policy evaluation as the behaviour of the private sector is no longer invariant to the policy rule adopted by the government.

To illustrate some counter-intuitive results and other problems found with Stackelberg equilibria, it is best to return to the example presented in Figure 1. The open-loop Stackelberg equilibrium with player 2 as leader and player 1 as follower assumes that both players have open-loop information structures (i.e. can only observe y_0) and that player 2 can make binding commitments about his or her future policy actions. Player 2 chooses the best action taking account of player 1's reaction curve (denoted by [a] in Table 1), so that he or she chooses to play L followed by H and obtains a welfare loss of 6. The follower plays L in both periods and gets a welfare loss of 11. Note that the leader is better off (which is always the case) whilst the follower is worse off (which is not always the case) than in the open-loop Nash equilibrium (see Table 2). It is well known that open-loop Stackelberg dynamic games (or economies with rational expectations) are characterized by forward-looking (as well as backward-looking) behaviour due to the anticipation of future actions of dominant players (such as the Treasury or the Central Bank). In such models the problem of time inconsistency arises, that is dominant players have an incentive to alter previous plans when they are called upon to execute those plans (Kydland and Prescott (1977)). For example, in the beginning of the planning period the leader finds it optimal to play L followed by H. However, once the first period has elapsed, the leader finds it optimal to renege and play L, inducing the follower to play H, as this reduces his or her welfare loss from 6 to 4. If there are no binding commitments, it is quite clear that such models are vulnerable to cheating from the side of the dominant player (e.g. the government) and, therefore, the initial plan of the dominant player will generally not be believed. It is important to note that cheating is only to be expected when the short-term gains of cheating outweigh the long-term losses of cheating,

which is more likely to happen when the rate of time preference used to discount punishments from cheating is large (Barro and Gordon (1983); Meijdam and de Zeeuw (1986)). In the case of incomplete information about preferences it is possible that the dominant player builds a reputation by being tough in the early stage of the games and blows its reputation in the final stages of the game (Kreps and Wilson, (1982); Backus and Driffill (1985)).

If the government cannot commit itself or does not have a strong reputation, the private sector cannot be expected to believe time-inconsistent announcements and therefore such policies are not credible. The feedback Stackelberg solution concept (Simaan and Cruz (1973)) assumes that the players can change their strategies at all points in time on the basis of observations on the evolution of the state of the economic system and is therefore by construction time-consistent. This solution concept can be seen as an extension of the principle of optimality (Bellman (1957)) to games. In terms of the example, one first calculates the Stackelberg equilibrium associated with every subgame in the second period, then adds on the resulting welfare losses to the welfare losses of the game starting with $y_0 = 0$, and finally calculates the Stackelberg equilbrium for this game. In the feedback Stackelberg equilibrium the follower plays L and H whilst the leader plays L and L. The associated welfare losses are 6 and 4, respectively. This provides a counter-example to the view that the leader is always better off in a Stackelberg than in a Nash game. In fact, the leader's welfare loss increases from 2 in the feedback Nash equilibrium to 4 in the feedback Stackelberg equilibrium whilst the follower's welfare loss improves from 7 to 6. It also provides a counter-example to the view that in the (time-consistent) feedback Stackelberg equilibrium the players are worse off than in the (time-inconsistent) open-loop Stackelberg equilibrium, because the leader reduces his or her welfare loss from 6 to 4 and the follower reduces his or her welfare loss from 11 to 6.

When the idea of a subgame is restricted to a game starting at some point of time from every possible state of the economy at that point of time, the feedback Stackelberg equilibrium may be called the subgame-perfect Stackelberg equilibrium, although Selten's (1975) original concept of subgame perfectness is only relevant for the Nash equilibrium concept. Because it is assumed that the players are ex ante given the opportunity to renege at each stage of the game, ex post they will not renege and therefore the feedback Stackelberg equilibrium leads to time-consistent policies by construction. However, subgame perfectness is stronger than time consistency, so that it is possible to formulate a time-consistent open-loop Stackelberg solution which is not subgame perfect (Meijdam and de Zeeuw (1986)). The point is that time consistency implies that there is no incentive to deviate from the equilibrium path and that subgame perfectness implies that there is no incentive to deviate from points that are off the equilibrium path either. The open-loop Nash equilibrium is another example of a time-consistent solution which is not subgame perfect, because the fact that there is no dominant player that can manipulate the current actions of the other players by making announcements

about its own future actions implies that as long as there are no unexpected deviations from the equilibrum path none of the players has an incentive to renege. When the information structure is such that players have information on all past states of the economy and the period of commitment is the planning period, it is common to refer to a closed-loop (with memory) dynamic game.

An alternative solution concept for non-cooperative games to the Nash or Stackelberg equilibrium is the consistent conjectural variations equilibrium, which was introduced in oligopoly theory by Bresnahan (1981) and was recently applied to an open-loop difference game (Hughes Hallett (1984); Brandsma and Hughes Hallett (1984)) and a feedback difference game (Başar, Turnovsky and d'Orey (1986)). Although it has been argued that the concept is logically inconsistent (Daughety (1985)); de Zeeuw and van der Ploeg (1987)), the main importance for the discussion in this paper is that the open-loop consistent conjectural variations equilibrium is time-inconsistent.

The conventional 'stacking' procedure is often applied to an economic system of difference games to obtain a final-form model (Theil (1964)), which is then used for policy evaluation purposes (e.g. Hughes Hallett, (1984)). A problem with this final-form approach is that in fact the open-loop model results, so that it is more difficult and cumbersome to discuss dynamic issues such as subgame perfectness and time inconsistency (even though it is possible to discuss with some difficulty the principle of multiperiod certainty equivalence in a one-player world).

This paper gives an overview of different solution concepts with their properties and derives the results for a standard class of linear-quadratic policy evaluation problems. The main difference between this framework and the example of Figure 1 is that the strategy space is continuous rather than discrete, which makes it much more appropriate for economic applications. A comparable survey can be found in Başar (1986), but this paper focusses on different issues and attempts to give more verbal explanation of the various solution concepts and associated issues of time consistency, subgame perfectness and credibility. Special attention is also given to the consistent conjectural variations approach, because it is felt that there are some problems with this approach. In section 2 a linear-quadratic class of difference games is formulated and different decision models or game-theoretic solution concepts are discussed. In section 3 properties such as time consistency, subgame perfectness and credibility are defined and evaluated. Section 4 concludes the paper.

2. Linear-quadratic difference games: An evaluation

2.1. *Model and solution concepts*

In this section some essential concepts for dynamic policy evaluation are discussed and a standard class of linear-quadratic difference games is formulated in order to elucidate the conceptual discussion.

The starting point is a linear dynamic economic model in state-space form:

$$y_t = A_t y_{t-1} + B_t^1 x_t^1 + B_t^2 x_t^2 + s_t, \; y_0 = \bar{y}_0. \tag{1}$$

The transition of the state y of the economy from period $t-1$ to period t is influenced by two players (such as the government and the private sector) who independently control the exogeneous variables x^1 and x^2, respectively. The non-controllable exogeneous variables are denoted by s. The objective of player i, $i = 1, 2$, is to minimise a quadratic welfare loss function over a finite horizon:[1]

$$w^i = \sum_{t=1}^{T} \tfrac{1}{2}\{y_t' Q_t^i y_t + x_t^{i'} R_t^i x_t^i\}, \qquad i = 1, 2, \tag{2}$$

where Q_t^i and R_t^i are symmetric and $Q_t^i \geqq 0$ and $R_t^i > 0$. An extension with linear terms in the welfare loss function is straightforward by redefining the state vector, y_t, and s_t in an appropriate way. The convex linear-quadratic structure is not essential for the discussion but facilitates analytical solutions. It can always be considered as an approximation to the real structure of a specific model. The problem is called an optimal control problem with two decision makers or a difference game.

The traditional approach (Theil (1964)) to an economic optimal control problem is to cast the economic model (1) into a *final-form* model:

$$y = B^1 x^1 + B^2 x^2 + s, \tag{3}$$

where y, x^1 and x^2 stack the state variables y_t and the policy instruments x_t^1 and x_t^2 for all periods of the finite planning horizon. Consequently, B^1 and B^2 are block-triangular matrices composed of A_t, B_t^1 and B_t^2, and s contains the non-controllable exogeneous variables s_t as well as the influence of the initial state vector $\bar{y}_0$.[2] The corresponding objective functionals become

$$w^i = \tfrac{1}{2}\{y' Q^i y + x^{i'} R^i x^i\}, \qquad i = 1, 2, \tag{4}$$

where the matrices Q^i and R^i are block-diagonal as the welfare loss functions (2) were assumed to be time separable. In this form the problem cannot be distinguished from a static problem, so that it corresponds to the normal form of the open-loop difference game. It explains why after this transformation into final form it is very difficult to discuss some dynamic issues such as the impact of new information or the absence of commitments to future actions. This will become clear in the sequel.

The by now standard approach to a difference game is to distinguish *information patterns* and *periods of commitment*. The decision makers or players announce strategies for the whole planning period but may or

[1] Attention is restricted to finite-horizon games, because this is analytically much more tractable. The infinite-horizon game can be viewed as the asymptotic case of the finite-horizon game as T tends to infinity.

[2] To be precise, $B^i = (B_{jk}^i)$ where $B_{jk}^i = 0$, $j < k$, $B_{jj}^i = B_j^i$, $B_{jk}^i = \Pi_{l=k}^{j-1}(A_l) B_l^i$, $j > k$, for $j = 1, \ldots T$, $k = 1, \ldots T$ and $i = 1, 2$.

may not be committed to stick to these strategies. A strategy is a mapping from the information set and time to the set of available actions. Considering the state of the economic system, this information set can in principle contain only the initial state (open-loop information), only the current state (closed-loop, no memory information) or all the states up to the current state (closed-loop, memory information). Memory information complicates matters considerably and is sometimes excluded on the grounds of bounded rationality (e.g. Rubinstein (1987)). The model with an open-loop information structure and a period of commitment equal to the planning horizon is called the *open-loop* model. The model with a closed-loop, (no) memory information structure and a period of commitment equal to the planning horizon is called the *closed-loop* model. The model with a closed-loop, no memory information structure and a period of commitment of one period is called the *feedback* model. In the feedback model the players have access to the current state of the economy and are ex-ante given the opportunity to renege on announced strategies at each stage of the game, so that in equilibrium they have no incentive to renege. The open-loop model is equivalent to the optimal control model based on a final-form economic model, which was described earlier in (3) and (4). It is also possible to have asymmetries between the two players such as different roles in the game (i.e. leader/follower), different information patterns and different periods of commitment (see also section 3 and Cohen and Michel (1988)).

The standard techniques to solve optimal control problems are Bellman's *dynamic programming* and Pontryagin's *minimum principle*. For an optimal deterministic control problem with one decision maker the two techniques yield the same optimal actions and performance.[3] For an optimal control problem with two or more decision makers these techniques lead in general to different solutions. The reason is that dynamic programming solves the feedback model and the minimum principle solves the open-loop model. To put it differently, dynamic programming presupposes information on the current state of the economic system and no commitments, whereas the minimum principle presupposes information on the initial state of the economic system and binding commitments. In the context of a game these assumptions have their influence, even when the world is deterministic. In the feedback model the players can observe the effects of the actions of their opponent and they can react to these observations, whereas in the open-loop model they cannot. Dynamic programming as a solution technique to a one-player optimal control problem is based on Bellman's principle of optimality. Dynamic programming as a solution framework for a difference game presupposes a generalization of the principle of optimality to dynamic games, which is also called subgame perfectness and which is treated in more detail in the next section.

[3] In a stochastic world, dynamic programming leads to policy feedback rules that take account of stochastic shocks and therefore leads to a lower expected welfare loss.

The two solution techniques have in common that they transform the dynamic optimization problem into a series of static optimization problems in a dynamic setting. When the minimum principle is applied, the optimization part of the solution is the static optimization of the Hamiltonian. When dynamic programming is applied the optimization part of the solution is the static optimization of the right-hand side of the Hamilton-Jacobi-Bellman equation. As a consequence the game theory involved can be reduced to static equilibrium concepts.

2.2. *Open-loop and feedback Nash equilibrium*

The standard non-cooperative equilibrium concept is the *Nash* concept which is based upon the idea that there should be no individual incentive for any player to deviate from the equilibrium. The Nash equilibrium assumes that strategy choice is simultaneous. Hence, the players choose their actions simultaneously and form expectations about each other's actions, which in equilibrium are fulfilled. This implies that the Nash equilibrium is the intersection of the hypothetical reaction curves which express the optimal decisions of each player conditional on the actions of the rival. For the prototype model (1), (2) the first-order conditions of the optimization problem,

$$R_t^i x_t^i + B_t^{i'}\{K_t^i y_t + g_t^i\} = 0, \qquad i = 1, 2, \tag{5}$$

where y_t is given by (1), lead to the hypothetical reaction functions

$$x_t^i = -(R_t^i + B_t^{i'} K_t^i B_t^i)^{-1} B_t^{i'}\{K_t^i(A_t y_{t-1} + B_t^j x_t^j + s_t) + g_t^i\},$$
$$j \neq i, \qquad i = 1, 2. \tag{5'}$$

For the *open-loop decision model* the terms $\{K_t^i y_t + g_t^i\}$ are the so-called co-states (also called shadow prices or adjoint variables) of Pontryagin's minimum principle. They show by how much the welfare loss is increased when there is a marginal increase in the state of the economy, y_t. The parameters K_t^i and g_t^i can be determined from the backward recursive equations:[4]

$$K_{t-1}^i = Q_{t-1}^i + A_t' K_t^i [E_t]^{-1} A_t, \tag{6}$$
$$K_T^i = Q_T^i,$$
$$g_{t-1}^i = A_t'\{K_t^i [E_t]^{-1}(s_t - B_t^1 [R_t^1]^{-1} B_t^{1'} g_t^1 - B_t^2 [R_t^2]^{-1} B_t^{2'} g_t^2) + g_t^i\},$$
$$g_T^i = 0, \qquad i = 1, 2, \tag{7}$$

where

$$E_t \equiv I + B_t^1 [R_t^1]^{-1} B_t^{1'} K_t^1 + B_t^2 [R_t^2]^{-1} B_t^{2'} K_t^2.$$

[4] An outline of the proofs of these and later results in this section can be found in the appendix.

For the *feedback decision model* the terms K_t^i and g_t^i are the parameters of the quadratic so-called value functions of dynamic programming:

$$\tfrac{1}{2}y'_{t-1}K^i_{t-1}y_{t-1} + g^{i'}_{t-1}y_{t-1} + c^i_{t-1}$$

$$= \min_{x^i}\{\tfrac{1}{2}y'_{t-1}Q^i_{t-1}y_{t-1} + \tfrac{1}{2}x^{i'}_t R^i_t x^i_t + \tfrac{1}{2}y'_t K^i_t y_t + g^{i'}_t y_t + c^i_t\}.$$

They follow from the backward recursive equations:

$$\begin{aligned}
K^i_{t-1} &= Q^i_{t-1} + A'_t[E'_t]^{-1}(I + K^i_t B^i_t[R^i_t]^{-1}B^{i'}_t)K^i_t[E_t]^{-1}A_t,\\
K^i_T &= Q^i_T,\\
g^i_{t-1} &= A'_t[E'_t]^{-1}(I + K^i_t B^i_t[R^i_t]^{-1}B^{i'}_t)\\
&\quad \{K^i_t[E_t]^{-1}(s_t - B^1_t[R^1_t]^{-1}B^{1'}_t g^1_t\\
&\quad - B^2_t[R^2_t]^{-1}B^{2'}_t g^2_t) + g^i_t\},\\
g^i_T &= 0, \qquad i = 1, 2.
\end{aligned} \tag{9}$$

The Nash equilibrium for both decision models is given by the intersection of the two hypothetical reaction functions, (5′):

$$x^i_t = G^i_t y_{t-1} + h^i_t, \qquad i = 1, 2, \tag{10}$$

where

$$G^i_t = -[R^i_t]^{-1}B^{i'}_t K^i_t[E_t]^{-1}A_t$$

and

$$h^i_t = -[R^i_t]^{-1}B^{i'}_t\{K^i_t[E_t]^{-1}(s_t - B^1_t[R^1_t]^{-1}B^{1'}_t g^1_t - B^2_t[R^2_t]^{-1}B^{2'}_t g^2_t) + g^i_t\}.$$

It is essential to note that the relationship between x^i_t and y_{t-1} in (10) is only a real functional relationship between actions and state of the economy in the feedback model; it does not represent the policy rule of player i in the open-loop model. Furthermore, the feedback equilibrium strategies $\{G^i_t, h^i_t\}$ are not binding; in the feedback model the strategies can be changed whenever one of the players wants to do so. However, because they form a feedback equilibrium, there will be no incentives to change the policy rule, even after unexpected events. The open-loop equilibrium consists of binding sequences of actions $\{x^i_t\}$ which result from (10) and (1) together with (6) and (7) and which only depend upon the initial state $\bar{y}_0$, so that unexpected state trajectories cannot have their influence. This open-loop outcome coincides with the Nash equilibrium of the static problem (3), (4). The transformation of the economic model into final form implies that the open-loop model with static information patterns and periods of commitment equal to the planning horizon is implicitly assumed. It is worth mentioning here that both open-loop and feedback policy rules can

be inferior to closed-loop memory policy rules where the players condition their strategies on information on current and past states of the economy (Başar and Olsder (1982), Section 6.3; de Zeeuw (1984), Section 4.3).

The open-loop and feedback Nash decision models can imply very different economic results. Consider as an example the problem of an oligopoly with restricted entry and exit harvesting a renewable resource with zero extraction costs, iso-elastic demand and serially uncorrelated shocks to the natural replenishment rate. It can then be shown that the open-loop extraction rates obey Hotelling-type arbitrage rules and are therefore efficient whilst the feedback Nash equilibrium leads to excessive extraction rates or even extinction of the resource (van der Ploeg (1986)). The reason is that when an individual firm decides to harvest an additional unit, it realizes that the lower stock increases harvesting costs to the other firms and therefore the other firms will in the feedback model react by harvesting less. This means that the marginal cost of harvesting an additional unit is less than in the absence of such a response from its rivals, hence the feedback model leads to excessive harvesting. (With free entry and exit, the harvesting rates in the feedback model become efficient.) To take another example, in a model of competitive arms accumulation between two countries, where each country has a 'guns versus butter' dilemma, the feedback Nash equilibrium proves to be more efficient and leads to less arms accumulation and thus to more consumption than the open-loop Nash equilibrium (van der Ploeg and de Zeeuw (1990)). The reason is that when one country decides to invest in an additional weapon, it realizes that the security of rival countries is threatened and therefore in the feedback model the rivals respond by investing more in weapons. Obviously, this increases the marginal cost of investment in an additional weapon and therefore the feedback model results in lower weapon stocks. (The policy recommendation is that countries should agree to monitor each other's weapon stocks.) This is an example where the feedback equilibrium proves to be better for both players in terms of utility than the open-loop one, which is in contrast with the usual implication in the literature.

2.3. *Open-loop and feedback Stackelberg equilibrium*

Another standard non-cooperative equilibrium concept is the *Stackelberg* concept. The difference with the Nash concept is the leader/follower structure which means that one of the players (the leader) acts first or, to put it differently, the action or strategy of the leader is part of the information set of the follower. This may be relevant when one of the players has a dominant position on the market. There are again two optimization problems. The first one determines the rational reaction of the follower to the action or strategy of the leader. This rational reaction, which is not a hypothetical reaction as in the Nash concept but a real reaction, is given by the reaction function for the follower (5′). The second optimization problem determines the optimal action or strategy of the leader given the rational reaction of the follower.

For the open-loop decision model this implies that the constraints of this optimization problem consist of the forward recursive system (1), equation (5′) for the follower and the backward recursive system for the co-states. The resulting open-loop Stackelberg equilibrium (Kydland (1975); Başar and Olsder (1982), Section 7.2; de Zeeuw (1984), Section 4.5) for the prototype model will not be given here, because it is not immediately relevant for this evaluation.[5] The backward recursiveness of the so-called adjoint system implies forward-looking behaviour of the follower, which leads to, what is called, time inconsistency of the optimal actions of the leader. The leader can, by making decisions about its future policy actions, manipulate the current policy actions of the follower. However, once the follower has implemented those actions, it might pay the leader (where it is possible) to renege and deviate from the previous decisions about its policies. These issues of time inconsistency will be dealt with in the next section.

For the feedback decision model the first-order conditions of the two optimization problems are

$$\begin{aligned} &R_t^i x_t^i + B_t^{i'}\{K_t^i y_t + g_t^i\} = 0 \\ &R_t^j x_t^j + \{B_t^{j'} + (\partial x_t^i/\partial x_t^j)' B_t^{i'}\}\{K_t^j y_t + g_t^j\} = 0 \end{aligned} \tag{11}$$

where i is the follower and j is the leader. This implies that the action x_t^i follows the action x_t^j, so that they are not simultaneous. The crucial difference with the Nash concept is the reaction coefficient $\partial x_t^i/\partial x_t^j = -(R_t^i + B_t^{i'} K_t^i B_t^i)^{-1} B_t^{i'} K_t^i B_t^j$. The feedback Stackelberg equilibrium is given by

$$\begin{aligned} x_t^i &= F_t^i\{A_t y_{t-1} + B_t^j x_t^j + s_t\} + F_t^{ii} g_t^i \\ x_t^j &= F_t^j\{A_t y_{t-1} + s_t\} + F_t^{ji} g_t^i + F_t^{jj} g_t^j \end{aligned} \tag{12}$$

where

$$\begin{aligned} F_t^{ii} &= -[R_t^i + B_t^{i'} K_t^i B_t^i]^{-1} B_t^{i'} \\ F_t^i &= F_t^{ii} K_t^i \\ F_t^{jj} &= -[R_t^j + B_t^{j'}(I + B_t^i F_t^i)' K_t^j (I + B_t^i F_t^i) B_t^j]^{-1} B_t^{j'}(I + B_t^i F_t^i)' \\ F_t^{ji} &= F_t^{jj} K_t^j B_t^i F_t^{ii} \\ F_t^j &= F_t^{jj} K_t^j (I + B_t^i F_t^i) \end{aligned}$$

and the backward recursions by

$$\begin{aligned} K_{t-1}^i = Q_{t-1}^i + A_t'\{&(I + B_t^j F_t^j)' F_t^{i'} R_t^i F_t^i (I + B_t^j F_t^j) \\ &+ (I + B_t^j F_t^j)'(I + B_t^i F_t^i)' K_t^i (I + B_t^i F_t^i)(I + B_t^j F_t^j)\} A_t \end{aligned}$$

[5] In any case, one could in principle obtain the open-loop Stackelberg equilibrium as the static Stackelberg equilibrium of the final-form model (3). That is, $x^i = -(R^i + B^{i'} Q^i B^i)^{-1}(B^i x^j + s)$ is the optimal reaction of the follower i to the actions of the leader j. The leader minimizes its welfare loss function subject to the reaction function of the follower, which gives $x^j = -(R^j + \bar{B}^{j'} Q^j \bar{B}^j)^{-1} s$ where $\bar{B}^j \equiv [I - B^i(R^i + B^{i'} Q^i B^i)^{-1}] B^j$.

$$K_T^i = Q_T^i$$

$$K_{t-1}^j = Q_{t-1}^j + A_t'\{F_t^{j'} R_t^j F_t^j + (I + B_t^j F_t^j)'(I + B_t^i F_t^i)' K_t^j (I + B_t^i F_t^i)(I + B_t^j F_t^j)\} A_t$$

$$K_T^j = Q_T^j$$

$$g_{t-1}^i = A_t'\{(I + B_t^j F_t^j)' F_t^{i'} R_t^i [F_t^i (I + B_t^j F_t^j) s_t + F_t^{ii} g_t^i + F_t^i B_t^j F_t^{ji} g_t^i + F_t^i B_t^j F_t^{jj} g_t^j] + (I + B_t^j F_t^j)'(I + B_t^i F_t^i)'\{K_t^i [(I + B_t^i F_t^i)(I + B_t^j F_t^j) s_t + B_t^i F_t^{ii} g_t^i + (I + B_t^i F_t^i) B_t^j (F_t^{ji} g_t^i + F_t^{jj} g_t^j)] + g_t^i\}\}$$

$$g_T^i = 0$$

$$g_{t-1}^j = A_t'\{F_t^{j'} R_t^j [F_t^j s_t + F_t^{ji} g_t^i + F_t^{jj} g_t^j] + (I + B_t^j F_t^j)'(I + B_t^i F_t^i)'\{K_t^j [(I + B_t^i F_t^i)(I + B_t^j F_t^j) s_t + B_t^i F_t^{ii} g_t^i + (I + B_t^i F_t^i) B_t^j (F_t^{ji} g_t^i + F_t^{jj} g_t^j)] + g_t^j\}\}$$

$$g_T^j = 0.$$

Given the parameters, the action x_t^i of the follower i is a function of the state y_{t-1} and the action x_t^j of the leader i which both belong to the follower's information set. The action x_t^j of the leader j is only a function of the state y_{t-1}. It is essential to note that for logical reasons the players have each other's action in their information set at the same time. Either player i acts first, so that the action x_j^i is part of the information set of player j, or it is the other way around. Otherwise, the equilibrium is not well-defined and one may end up with a multiplicity of 'solutions'. The follower just plays optimally given the state of the economy and the action of the leader. In the Stackelberg equilibrum the leader expects the follower to react rationally and the action is chosen accordingly. The rational reaction is determined by the first equation of (11) and influences the reaction coefficient $\partial x_t^i / \partial x_t^j$ as well as the state transition y_t in the second equation of (11). After substitution of this rational reaction, the second equation of (11) determines the optimal action of the leader and not an optimal reaction, because the leader is not reacting to the follower. These considerations are essentially of a static nature and they apply also to the open-loop decision model, especially when the final-form representation (3), (4) is used. The only difference is that in the feedback decision model the leader reacts indirectly to past actions of the follower through observations on the state of the economy. The feedback equilibrium is obtained from dynamic programming and therefore satisfies subgame perfectness in the sense described before. Subgame perfectness implies that the leader has no incentive to deviate in subsequent periods. These issues are treated in more detail in section 3.

2.4. *Consistent conjectural variations equilibrium*

Recently, a third non-cooperative equilibrium concept for difference games has been developed: the *consistent conjectural variations equilibrium* (for the open/closed-loop case from the final-form representation: Hughes Hallett (1984); Brandsma and Hughes Hallett (1984); for the feedback case: Başar, Turnovsky and d'Orey (1986). The equilibrium was originally introduced in the context of oligopoly theory (Bresnahan (1981)) and is based upon the concept of conjectural variation (Bowley (1924)). A conjectural variation in this context is a reaction coefficient $\partial x_t^i/\partial x_t^j$ as in (11), which comes from a conjecture of player j with respect to the reaction of player i. In the Stackelberg equilibrium the leader conjectures a rational reaction function of the follower and will be right in this conjecture. In the Nash equilibrium the two players conjecture the *action* of the other player and they are assumed to be right in their conjecture (consistency argument). The idea behind the consistent conjectural variations equilibrium is that the two players conjecture the *reaction* of the other player and that they are assumed to be right in their conjecture. In the literature up to now the equilibrium is determined by introducing conjectural variations for both players and requiring consistency of conjectural variations and hypothetical reaction coefficients. Since the Nash equilibrium requires correct conjectures of action for both players and the Stackelberg equilibrium requires that the leader has correct conjectures about the reactions of the follower, one could argue that the consistent conjectural variations equilibrium is a natural extension as it requires correct conjectures about reactions for both players. However, as it is done, the extension leads to logical inconsistencies (Daughety (1985); de Zeeuw and van der Ploeg (1987)). The reason is simply, that one cannot mix the idea of hypothetical reactions in notional time of the Nash concept with the idea of conjectured reactions, which actually degenerate the game into separate optimization problems. This will become more clear in the sequel. The consistent conjectural variations concept is not well-defined. It is certainly not true that the proposed equilibrium is a Nash equilibrium or, worse, a superior one (in contrast to the statements in Hughes Hallett (1984)[6] and Brandsma and Hughes Hallett (1984)).

The ideas of conjectures (about actions) and conjectural variations (about reactions) are alright, but the consistency argument should be different. There are two ways out. The first one is to formulate an infinite regress decision model of the type 'player i conjectures that player j conjectures that player i conjectures.... ad infinitum' (Daughety (1985)). The other way out is to start with conjectures and corresponding conjectural variations and to require

[6] In this and later papers an unfortunate mistake has slipped in. Apart from a type-setting error (the term $G_s^{(i)'}\{Q^{(i)}(I/0)$ should not appear at the end of the first line of equation (16)), there is also a more fundamental error (the '$-$' after $G_s^{(i)}$ in the first line of (16) should be a '$+$') which seems to lead a persistent life in later papers as well. However, Andries Brandsma has said in private communication that most computer algorithms are, in fact, based on equation (14) so that may of the empirical results may not be affected by this second error.

consistency of conjectures and actions. To stress the difference the resulting equilibrium will be called *consistent conjectures equilibrium.* Player i minimizes the welfare loss function (2) subject to (1) and subject to the conjectures about the reactions of player j, $x_t^j(x_t^i)$. Player j faces a similar problem. The conjectures $x_t^j(x_t^i)$ and $x_t^i(x_t^j)$ account for the conjectural variations $\partial x_t^j/\partial x_t^i$ and $\partial x_t^i/\partial x_t^j$. The first-order conditions with conjectures are for player i:

$$R_t^i x_t^i + \{B_t^{i'} + (\partial x_t^j/\partial x_t^i)' B_t^{j'}\}\{K_t^i y_t + g_t^i\} = 0$$
$$y_t = A_t y_{t-1} + B_t^i x_t^i + B_t^j x_t^j(x_t^i) + s_t$$

and for player j: (13)

$$R_t^j x_t^j + \{B_t^{j'} + (\partial x_t^i/\partial x_t^j)' B_t^{i'}\}\{K_t^j y_t + g_t^j\} = 0$$
$$y_t = A_t y_{t-1} + B_t^i x_t^i(x_t^j) + B_t^j x_t^j + s_t.$$

The first set of equations yields an equation in x_t^i determining the optimal action of player i and the second set yields an equation in x_t^j determining the optimal action of player j. In equilibrium consistency of conjectures and actions is required, that is these optimal actions have to fit the conjectures, and therefore yield restrictions on the parameters of the conjectures. This weaker concept of consistency is not enough to guarantee uniqueness and usually leads to multiple equilibria. How can the idea of consistency of conjectures and reactions arise? In that literature reaction functions are created by not substituting the conjectures in the state transition y_t, which does not seem very sensible. As first-order conditions with conjectures are then taken:

$$R_t^i x_t^i + \{B_t^{i'} + (\partial x_t^j/\partial x_t^i)' B_t^{j'}\}\{K_t^i y_t + g_t^i\} = 0$$
$$y_t = A_t y_{t-1} + B_t^i x_t^i + B_t^j x_t^j + s_t$$

and (13′)

$$R_t^j x_t^j + \{B_t^{j'} + (\partial x_t^i/\partial x_t^j)' B_t^{i'}\}\{K_t^j y_t + g_t^j\} = 0$$
$$y_t = A_t y_{t-1} + B_t^i x_t^i + B_t^j x_t^j + s_t.$$

The resulting 'reaction functions', $\hat{x}_t^i(x_t^j)$ and $\hat{x}_t^j(x_t^i)$, lead to reaction coefficients, $\partial \hat{x}_t^i/\partial x_t^j$ and $\partial \hat{x}_t^j(x_t^i)$, which must match the conjectural variations. There are, however, not only logical difficulties with this approach. The consistent conjectural variations equilibrium for the open-loop model in final form (3), (4) suffers from more problems. The result is time-inconsistent for the same reason as the Stackelberg open-loop equilibrium is (see section 3). Furthermore, the outcome is typically worse for the players than the Nash outcome and it suffers from non-uniqueness and instability. Two (static) examples will clarify these statements.

Example 1 (Hughes Hallett (1984) pp. 389–90)

Consider the game with objectives $w^i = y^2 + (x^i)^2$ where $y = x^1 + x^2 - 1$. The Nash equilibrium is $x^1 = 1/3$ with outcome $w^i = 2/9$.

Hughes Hallett argues that $x^i = 2/5$ is a better solution, because the associated outcome $w^i = 1/5$ implies an improvement for both players. This is not surprising, since it is well known that it is possible to find Pareto improvements over the Nash outcome even though there is no unilateral incentive for any player to deviate from the Nash equilibrium. In fact, $x^i = 2/5$ is what is generally called the Nash bargaining solution. However, this solution is not sustained as a consistent conjectural variations equilibrium. To find one, Hughes Hallett describes an iterative procedure and searches for a fixed point in the conjectured and actual 'reaction coefficients'. This procedure starts from an initial pair (d^1, d^2) of conjectural variations, where $d^i \equiv \partial x^i/\partial x^j$, and yields new pairs corresponding to the 'reaction coefficients' $(-(1 + d^2)/(2 + d^2), -(1 + d^1)/(2 + d^1))$. There are two fixed points here, namely $d^i = -3/2 \pm 1/2\sqrt{5}$ with corresponding actions $x^i = 1/2 \mp 1/10\sqrt{5}$ and outcome $w^i = 1/2 \mp 1/10\sqrt{5}$. Both consistent conjectural variations equilibria produce worse results for both players as compared to the Nash equilibrium. Futhermore, they do not satisfy the Nash property since each player can unilaterally improve by playing, for example, $x^i = 1/4 \pm 1/20\sqrt{5}$. Finally, it follows from the derivative of the fixed point mapping, $-1/(2 + d^i)^2 = -3/2 \pm 1/2\sqrt{5}$, that one of the fixed points ($d^i = -3/2 + 1/2\sqrt{5}$) is stable whilst the other is unstable.

The Nash bargaining solution $x^i = 2/5$ is, however, sustained as a consistent conjectures equilibrium. The conjectures 'my rival mimicks what I do', $x^i = x^j$, implies conjectural variations equal to 1 and leads to optimal actions $x^i = 2/5$, which are consistent with the conjectures. The non-cooperative Nash equilibrium is, as always, also sustained as a consistent conjectures equilibrium, since it corresponds to zero conjectural variations. A final example of such an equilibrium is the solution $x^i = 0$ with outcome $w^i = 1$, which results from the conjectures $x^i = -x^j$ with conjectural variations -1. However, this outcome is obviously unattractive for the players:

Example 2

Consider the game with objectives $w^i = 1/2(y'y + x^{i\prime}x^i)$ where $y = x^1 + x^2 + s$, $s = [1, 1]'$ and x^i are two-dimensional vectors. The non-cooperative Nash equilibrium is $x^i = [-1/3, -1/3]'$ with outcome $w^i = 2/9$.

The consistent conjectural variations (D^1, D^2) are characterized by

$$(D^i)^2 + 3D^i + I = 0. \tag{14}$$

There are an infinite number of solutions to (14) (see appendix), which can be found analytically after some tedious calculations. Hughes Hallett's iterative scheme corresponds to:

$$D^i_{s+1} = -(2I + D^j_s)^{-1}(I + D^j_s) = (2I + D^j_s)^{-1} - I.$$

The local stability of the iterative scheme in the neighbourhood of the fixed points follows if all of the eigenvalues of the Jacobian

$$\partial \operatorname{vec} D^i_{s+1}/\partial \operatorname{vec} D^j_s = -\{(2I + D^j_s)^{-1'} \otimes (2I + D^j_s)^{-1}\},$$

evaluated at D^j, are inside the unit circle. It can be shown after considerable manipulation that $D^i = (-3/2 + 1/2\sqrt{5})\, I$ is the only stable fixed point (see appendix). Again the corresponding welfare loss, $w^i = 1/2 - 1/10\sqrt{5}$, is higher than the welfare loss which can be obtained under the Nash concept. Finally, the Nash equilibrium is not sustained as a conjectural variations equilibrium, because $x^i = -1/3\, s$, $x^i = D^i(x^i + s)$ and (14) are inconsistent.

3. Time inconsistency, subgame perfectness and atomistic agents

In section 2 several decision models for dynamic policy evaluation problems have been discussed. This section discusses properties of these dynamic decision models, such as time inconsistency, subgame perfectness and credibility.

A decision model is *time inconsistent* if there is an incentive for one of the players to renege on the initially chosen strategy in the future (Kydland and Prescott (1977)). A decision model which typically is time-inconsistent is the open-loop Stackelberg equilibrium. It should be noted here that there is a semantic difficulty with this analysis. Strictly speaking it is assumed that it is impossible for players to renege in an open-loop decision model. However, in practice such commitments are difficult to enforce and time inconsistency may therefore be regarded as an undesirable property of a decision model. In the open-loop Stackelberg equilibrium the leader's strategy $\{x_1^j, \ldots, x_T^j\}$ is optimal given the follower's rational reaction $\{x_1^i, \ldots, x_T^i\}$. However, at time $s > 1$ the remaining strategy $\{x_s^j, \ldots, x_T^j\}$ of the leader is typically not optimal anymore, given the rational reaction of the follower at time $s > 1$. The reason is that at time s the actions $\{x_s^j, \ldots, x_T^j\}$ have done the job of influencing the past actions of the follower and can now be solely employed to influence the current and future actions of the follower. This is particularly so when the leader does not have sufficient instruments, in the face of private market failures, to achieve the first-best outcome in the first place. When the leader does have sufficient instruments to achieve the first-best outcome there is clearly no problem of time inconsistency (Hillier and Malcomson (1984)). The strategy announcement $\{x_s^j, \ldots, x_T^j\}$ can be considered as some sort of threat which helped to force the follower to play $\{x_1^i, \ldots, x_{s-1}^i\}$. The leader tries to influence the future expectations of the follower in order to get a better outcome by making such an announcement, irrespective of whether the leader will stick to the announcement or not.

For example, a benevolent government, who maximizes the gross consumers' surplus of the representative household, may announce taxation of the supply of labour rather than of capital tomorrow in order to induce agents to accumulate capital today. Once the capital stock is in place, it pays the government (improves economic welfare) to renege by taxing capital, instead of labour, tomorrow, despite the fact that the government has the same preferences as the representative household. The use of lump-sum taxation gives the first-best result in this case and avoids the problem of dynamic inconsistency (Fischer (1980)). The crucial point is the forward-looking behaviour of the

follower which looses its impact as soon as actions are performed. In the example, once the investment has occurred, the government can extract the quasi-rent on it. The same phenomenon occurs in other models with forward-looking variables such as models with rational expectations. For example, the optimal taxation of a monetary economy with a Cagan-type money demand function is time inconsistent (Calvo (1978)). The reason is that the government finds it optimal to announce a low monetary growth rate in order to induce large holdings of real money balances and low inflation, but once the real money balances have been accumulated it pays the government to renege and impose a surprise inflation tax. Strictly speaking, the incentives to renege in these two examples are only hypothetical as the commitment in the open-loop Stackelberg equilibrium extends over the entire planning period.

This type of forward-looking behaviour can easily be derived for the class of models defined in section 2. For the follower the strategy x_t^j of the leader has the same role as the exogeneous input s_t. The rational reaction x_t^i of the follower is given by equation (5′) which means that it is a function of x_t^j and g_t^i. According to (7) g_t^i is a function of all the future exogeneous inputs, so that x_t^i is a function of current and future actions of the leader:

$$x_t^i = \varphi(x_t^j, \ldots, x_T^j).$$

The feedback Stackelberg equilibrium, however, is time consistent by construction, because it is based on the idea that the players are *ex-ante* constantly given the opportunity to renege and therefore *ex-post* have no incentive to renege. But there is more. Feedback decision models have the stronger property of subgame perfectness. A game equilibrium is *subgame perfect* if it remains an equilibrium for any subgame. A subgame in this respect is a game with the same players, objectives and system dynamics, but starting from an arbitrary state y_s at time s, $1 \leqq s \leqq T$. This concept reveals precisely the structure of dynamic programming and thus of the feedback decision model. To avoid confusion with the original definition of subgame perfectness in extensive-form games (Selten (1975)), it might be better to call this concept subgame perfectness without perfect recall or Markov subgame perfectness (Fershtman (1989)). It can be said that subgame perfectness is time consistency on the equilibrium path as well as off the equilibrium path and is therefore a stronger concept. A subgame perfect equilibrium is robust against mistakes or other unexpected events (Selten (1975)). Because subgame perfectness is stronger than time consistency, it is possible to formulate a Stackelberg equilibrium which is time consistent but not subgame perfect and which can be called a consistent open-loop Stackelberg equilibrium (Buiter (1983)); Meijdam and de Zeeuw (1986)). Such an equilibrium ignores the effect of the leader's future actions on the follower's current actions. This was already proposed by Kydland and Prescott (1977, p. 476), but they did not recognize that this is only one possible consistent equilibrium. The consistent solution in Fischer (1980) is another example of a consistent open-loop Stackelberg equilibrium. It can be viewed as a subgame-perfect equilibrium with infinitely

TABLE 3
*Optimal dynamic taxation (Fischer, 1980)**

Solution	*Social welfare*	*Capital*	*Tax rates on capital*	*Tax rates on labour*
Command optimum	0.759	1.576		
Open-loop	0.706	1.274	0.334	0.332
Consistent open-loop	0.625	0.986	0.479	0.000
Subgame perfect	0.724	1.275	0.435	0.000

* The subgame-perfect solution is relevant under the assumption of a dominant government faced with only one household-producer. The consistent open-loop solution is relevant under the assumption of a dominant government faced with an infinite number of atomistic agents.

many identical consumers or atomistic agents (de Zeeuw, Groot and Withagen (1988)). This equilibrium leads to lower social welfare than the time-inconsistent open-loop Stackelberg equilibrium. Many authors have argued that this is the price one has to pay for a lack of credibility. However, it is a mistake to think that players are *always* worse off in a time-consistent equilibrium. Table 3 shows that for Fischer's (1980) example of optimal dynamic taxation of capital and labour the feedback Stackelberg equilibrium, which is based on the assumption of one dominant government and one large household-producer rather than on the assumption of atomistic agents, leads to higher social welfare than the open-loop Stackelberg equilibrium. The reason is that the losses from avoiding the problem of time inconsistency are off-set by the gains from additional information (see also the example of arms accumulation discussed in section 2.2).

It is also possible to achieve time consistency by requiring that the leader employs a feedback decision model, whereas the follower still employs the open-loop decision model (Cohen and Michel (1988)). In our framework the ideas of subgame perfectness and dynamic programming are equivalent and these ideas imply time consistency. However, there are time-consistent equilibria which are not subgame perfect and which can therefore not be found by dynamic programming.

The open-loop Nash equilibrium is time-consistent. As long as the state of the economic system follows the open-loop Nash equilibrium path none of the players has an incentive to renege. The open-loop consistent conjectural variations equilibrium, however, is time inconsistent for the same reason as the open-loop Stackelberg equilibrium is. Time inconsistency is a property of an equilibrium and can only be avoided by changing the equilibrium concept. There is no technical problem as is claimed by Hughes Hallett (1984) and Brandsma and Hughes Hallett (1984).[7]

A strategy is *credible* if it contains announcements on future actions and if

[7] Hillier (1987) points at the same mistake in his comment on Hughes Hallett (1987).

these announcements are believed by the other players. Announcements are believed if the announced actions are considered to be optimal at the time of action or, alternatively, if there will be no incentive, in the eyes of the other players, to act differently at the time the action has to be implemented. Time-inconsistent strategies are an example of strategies that are not credible. Strictly speaking, this applies only to decision models without commitments. The credibility problem can also occur in a static context where players are in principle of equal strength and act simultaneously. One of the players can try to become a Stackelberg leader by announcing the action beforehand. If the announcement has effect the Stackelberg equilibrium may result. In this case, however, the 'leader' can do even better, because generally there will be an incentive to deviate from the announcement under the assumption that the other player expects it to be true. In this simple framework the only credible announcement is the Nash action, because this is the only announcement that is at the same time the optimal reaction to the optimal reaction to the announcement (Meijdam and de Zeeuw (1986)). However, in a more advanced framework with imperfect (or incomplete) information, reputational effects in a sequential equilibrium can lead to credible strategies which differ from the complete-information Nash equilibrium (Kreps and Wilson (1982)). For example, with incomplete information it is possible to have the private sector believing announcements of the government that it will fight inflation, whereas with complete information the Nash announcement of high inflation is the only credible one (Backus and Driffill (1985)). Alternatively, reputational effects can occur when the game is repeated indefinitely (e.g. Barro and Gordon (1983)). When the discount rate is small enough, reneging results in punishments which are relatively large and therefore there may be no temptation to renege even though the policy actions may be time-inconsistent in the absence of such reputational effects.

4. Conclusion

In this paper several methods to analyze policy problems with two or more decision makers are evaluated. These methods employ decision models which are distinguished according to different non-cooperative game-theoretic solution concepts (Nash, Stackelberg, consistent conjectural variations), different information structures (open-loop, closed-loop) and different periods of commitment. The decision models are evaluated by considering properties such as time consistency, subgame perfectness and credibility, and the links with solution techniques like dynamic programming and Pontryagin's minimum principle are precisely described.

The formulation of the problem on the basis of economic models in final form is rejected for games, because it is difficult to employ crucial dynamic concepts and techniques in this formulation. The consistent conjectural variations equilibrium is rejected on principles of logic and the alternative consistent conjectures equilibrium is presented. The decision model with

open-loop information structure and binding commitments can suffer from time inconsistency. The assumption of binding commitments means that it is impossible to renege and therefore time inconsistency can, strictly speaking, not be a problem. However, time inconsistency is an unattractive feature for a decision model to display, because it puts great strain on the assumption of binding commitments. The decision models with closed-loop information structure and without commitments are subgame perfect and thus time-consistent, and therefore they deserve more attention. When the Stackelberg leader/follower structure is based upon announcements and not upon sequential actions, the requirement of credibility leads back to Nash. Credible strategies can also be based on reputational effects.

The class of difference games discussed in this paper assumes quadratic preferences and linear, discrete-time models, which keeps matters tractable. The extension of the methods discussed in this paper to continuous-time models is easy and leads to so-called differential games. The extension to models with non-quadratic preferences and non-linear models is not easy, which is a pity as many interesting economic problems fall into this category. Although it is relatively straightforward to develop iterative Gauss-Newton algorithms to derive *open-loop* Stackelberg or Nash equilibria, it is very difficult to calculate *feedback* Stackelberg or Nash equilibria for non-linear models and/or non-quadratic preferences. The reason is that it usually is impossible to find analytical expressions for the functional forms of the value functions. All that one can do in such cases is to discretize the space of control variables of the players and calculate the subgame-perfect solution numerically by dynamic programming (as was done in Figure 1). This procedure rapidly runs into combinatorial problems and is thus very expensive in terms of computer requirements of storage and time. (Discretization also means that it is not possible to calculate the consistent conjectural variations equilibrium, but this is not too serious as this equilibrium suffers from logical problems anyway.) This problem disappears when there are no externalities or market imperfections, because then it is possible to invoke the fundamental theorem of welfare economics which says that the market (read difference/differential game) outcome is the same as the outcome of a centrally planned or command economy. Kydland and Prescott (1982) use this to avoid the difficult derivation of value functions for the market outcome. Unfortunately, most interesting policy problems are real games and are therefore characterized by externalities and market imperfections so that the approach used by Kydland and Prescott (1982) cannot be applied. It follows that future research must be concerned with the technical difficulties of calculating value functions and subgame-perfect outcomes for non-linear models and non-quadratic preferences (cf. Lucas, 1987).

Tilburg University, The Netherlands
University of Amsterdam, The Netherlands

REFERENCES

BACKUS, D. and DRIFFILL, J. (1985). 'Inflation and reputation', *American Economic Review*, 75, 3, 530–8.

BARRO, R. J. and GORDON, D. M. (1983). 'Rules, discretion and reputaton in a model of monetary policy', *Journal of Monetary Economics*, 12, 101–21.

BASAR, T. and OLSDER, G. J. (1982). *Dynamic Noncooperative Game Theory*, Academic Press, New York.

BASAR, T. (1986). 'A tutorial on dynamic and differential games', in T. Başar (ed.), *Dynamic Games and Applications in Economics*, Springer-Verlag, Berlin.

BASAR, T., TURNOVSKY, S. J. and d'OREY, V. (1986). 'Optimal strategic monetary policies in dynamic interdependent economies', in T. Başar (ed.), *Dynamic Games and Applications in Economics*, Springer-Verlag, Berlin.

BELLMAN, R. (1957). *Dynamic Programming*, Princeton University Press, Princeton.

BOWLEY, A. L. (1924). *The Mathematical Groundwork of Economics*, Oxford University Press, Oxford.

BRANDSMA, A. S. and HUGHES HALLETT, A. J. (1984). 'Economic conflict and the solution of dynamic games', *European Economic Review*, 26, 13–32.

BRESNAHAN, T. F. (1981). 'Duopoly models with consistent conjectures', *American Economic Review*, 71, 934–45.

BUITER, W. H. (1983). 'Optimal and time-consistent policies in continuous time rational expectations models', discussion paper No. A.39, Econometrics Programme, London School of Economics.

CALVO, G. A. (1978). 'On the time consistency of optimal policy in a monetary economy', *Econometrica*, 46, 1411–28.

COHEN, D. and MICHEL, Ph. (1988). 'How should control theory be used to calculate a time-consistent government policy?', *The Review of Economic Studies*, 55, 182, 263–74.

DAUGHETY, A. F. (1985). 'Reconsidering Cournot: the Cournot equilibrium is consistent', *Rand Journal of Economics*, 16, 3, 368–79.

FERSHTMAN, Ch. and KAMIEN, M. I. (1987). 'Dynamic duopolistic competition with sticky prices', *Econometrica*, 55, 5, 1151–64.

FERSHTMAN, Ch. (1989). 'Fixed rules and decision rules: Time consistency and subgame perfection', *Economics Letters*, 30, 3, 185–91.

FISCHER, S. (1980). 'Dynamic inconsistency, cooperation and the benevolent dissembling government', *Journal of Economic Dynamics and Control*, 2, 1, 93–108.

HILLIER, B. and MALCOMSON, J. M. (1984). 'Dynamic inconsistency, rational expectations, and optimal government policy', *Econometrica*, 52, 6, 1437–51.

HILLIER, B. (1987). 'Forecasting and policy evaluation in economies with rational expectations: the discrete time case—a comment', *Bulletin of Economic Research*, 39, 1, 71–8.

HUGHES HALLETT, A. J. (1984). 'Non-cooperative strategies for dynamic policy games and the problem of time inconsistency', *Oxford Economic Papers*, 36, 381–99.

HUGHES HALLETT, A. J. (1987). 'Forecasting and policy evaluation in economies with rational expectations: the discrete time case', *Bulletin of Economic Research*, 39, 1, 49–70.

KREPS, D. M. and WILSON, R. (1982). 'Reputation and imperfect information', *Journal of Economic Theory*, 27, 253–79.

KYDLAND, F. E. (1975). 'Noncooperative and dominant player solutions in discrete dynamic games', *International Economic Review*, 16, 301–35.

KYDLAND, F. E. and PRESCOTT, E. C. (1977), 'Rules rather than discretion: the inconsistency of optimal plans', *Journal of Political Economy*, 85, 3, 473–91.

KYDLAND, F. E. and PRESCOTT, E. C. (1982). 'Time to build and aggregate fluctuations', *Econometrica*, 50, 6, 1345–70.

LANCASTER, K. (1973). 'The dynamic inefficiency of capitalism', *Journal of Political Economy*, 81, 5, 1092–109.

LUCAS, R. E. Jr. (1976). 'Econometric policy evaluation: a critique', in K. Brunner and A. H. Meltzer (eds.), *The Phillips Curve and Labour Markets*, North-Holland, Amsterdam.

LUCAS, R. E. Jr. (1987). *Models of Business Cycles*, Basil Blackwell, Oxford.

MEIJDAM, A. C. and de ZEEUW, A. J. (1986). 'On expectations, information and dynamic game equilibria', in T. Başar (ed.), *Dynamic Games and Applications in Economics*, Springer-Verlag, Berlin.

PLOEG, F. van der (1986). 'Inefficiency of credible strategies in stochastic renewable resource markets with iso-elastic demand and zero extraction costs', *Journal of Economic Dynamics and Control*, 10, 1/2, 309–14.

PLOEG, F. van der and de ZEEUW, A. J. (1990). 'Perfect equilibrium in a model of competitive arms accumulation', *International Economic Review*, 31, 1, 131–46.

PLOEG, F. van der and de ZEEUW, A. J. (1991). 'International aspects of pollution control', forthcoming in *Environmental and Resource Economics*.

REINGANUM, J. F. and STOKEY, N. L. (1985). 'Oligopoly extraction of a common property natural resource: the importance of the period of commitment in dynamic games', *International Economic Review*, 26, 1, 161–73.

REYNOLDS, S. S. (1987). 'Capacity investment, preemption and commitment in an infinite horizon model', *International Economic Review*, 28, 1, 69–88.

RUBINSTEIN, A. (1987). 'The complexity of strategies and the resolution of conflict: an introduction', *ST/ICERD*, discussion paper No. 87/150, London School of Economics.

SELTEN, R. (1975). 'Reexamination of the perfectness concept for equilibrium points in extensive games', *International Journal of Game Theory*, 4, 25–55.

SIMAAN, M. and CRUZ Jr., J. B. (1973). 'Additional aspects of the Stackelberg strategy in nonzero-sum games', *Journal of Optimization Theory and Applications*, 11, 6, 613–26.

THEIL, H. (1964). *Optimal Decision Rules for Government and Industry*, North-Holland, Amsterdam.

ZEEUW, A. J. de (1984). *Difference Games and Linked Econometric Policy Models*, Ph.D. Thesis, Tilburg University.

ZEEUW, A. J. de and PLOEG, F. van der (1987). 'Consistency of conjectures and reactions?', *Economics Letters*, 24, 2, 181–5.

ZEEUW, A. J. de, GROOT, F. and WITHAGEN, C. (1988). 'On credible optimal tax rate policies', mimeo., Tilburg University.

APPENDIX

1. Open-loop Nash equilibrium

Pontryagin's minimum principle is applied to the problem (2), (1) for each player, given the strategy of the other player. Because $Q_t^i \geqq 0$ and $R_t^i > 0$, the welfare loss functions (2) are strictly convex in x and convex in y. Because the state transition (1) is linear in y and x, the welfare loss functions (2) are in fact strictly convex on the strategy spaces. If follows, that the necessary conditions of the minimum principle are also sufficient and lead to a unique optimal solution.

The Hamiltonians are

$$H^i(y_{t-1}, x_t^i, t, p_t^i) = \tfrac{1}{2}\{y_{t-1}' Q_{t-1}^i y_{t-1} + x_t^{i\prime} R_t^i x_t^i\} + p_t^{i\prime}(A_t y_{t-1} + B_t^1 x_t^1 + B_t^2 x_t^2 + s_t), \qquad i = 1, 2, \tag{A.1}$$

where p_t^i is the so-called co-state or adjoint variable. Necessary and sufficient conditionx for minima are (1) and

$$R_t^i x_t^i + B_t^{i\prime} p_t^i = 0 \tag{A.2}$$

$$p_{t-1}^i = Q_{t-1}^i y_{t-1} + A_t' p_t^i,$$

$$p_T^i = Q_T^i y_T. \tag{A.3}$$

For an open-loop Nash equilibrium the two-point boundary value problems (1), (A.2), (A.3) for $i = 1, 2$ have to be solved simultaneously and can be solved analytically by postulating a linear

relationship between the state y and the co-state p:

$$p_t^i = K_t^i y_t + g_t^i. \tag{A.4}$$

Substitution of (A.4) into (A.2) leads to (5) and to

$$x_t^i = -[R_t^i]^{-1} B_t^{i'} \{K_t^i y_t + g_t^i\}. \tag{A.5}$$

Substitution of (A.5) into (1) and some rewriting yields

$$E_t y_t = A_t y_{t-1} + s_t - B_t^1 [R_t^1]^{-1} B_t^{1'} g_t^1 - B_t^2 [R_t^2]^{-1} B_t^{2'} g_t^2, \tag{A.6}$$

where

$$E_t \equiv I + B_t^1 [R_t^1]^{-1} B_t^{1'} K_t^1 + B_t^2 [R_t^2]^{-1} B_t^{2'} K_t^2,$$

which is assumed to be non-singular. Substitution of (A.6) into (A.4) and then of (A.4) into (A.3) leads to equations which have to hold for every y_{t-1} and which, therefore, lead to the backward recursive equations (6) and (7).

2. Feedback Nash equilibrium

The feedback Nash equilibrium is found by solving simultaneously the equations of Bellman's principle of dynamic programming

$$\begin{aligned} V^i(t, y_{t-1}) &= \min_{x^i} \{\tfrac{1}{2}\{y'_{t-1} Q_{t-1}^i y_{t-1} + x_t^{i'} R_t^i x_t^i\} \\ &\quad + V^i(t+1, y_t)\}, \\ V^i(T+1, y_T) &= \tfrac{1}{2} y_T' Q_T^i y_T. \end{aligned} \tag{A.7}$$

Because the welfare loss functions (2) are quadratic in y and x and the state transition (1) is linear in y and x, the optimal actions must be linear in y and hence the value functions must be quadratic in y:

$$V^i(t, y_{t-1}) = \tfrac{1}{2} y'_{t-1} K_{t-1}^i y_{t-1} + g_{t-1}^{i'} y_{t-1} + c_{t-1}^i, \tag{A.8}$$

where, without loss of generality, it is assumed that the matrices K are symmetric.

The minimizations of the right-hand sides of (A.7), where y_t is given by (1), lead to the conditions (5) which yield (A.5) and (A.6). These results give the values of the right-hand sides of (A.7). Because the equations of dynamic programming have to hold for every y_{t-1}, the quadratic terms in y_{t-1} of the left-hand sides and the right-hand sides of (A.7) and the linear terms in y_{t-1} can be compared separately. These comparisons yield the backward recursive equations (8) and (9).

3. Feedback Stackelberg equilibrium

The feedback Stackelberg equilibrium is found by solving the equations of dynamic programming (A.7), where now the action x_t^i of the follower is a function of the action x_t^j of the leader. The value functions are again given by (A.8). The minimizations of the right-hand sides of (A.7), where y_t is given by (1), lead to the conditions (11). The first equation of (11) and (1) lead to the rational reaction of the follower, which is given by the first equation of (12). The reaction coefficient becomes

$$\partial x_t^i / \partial x_t^j = F_t^i B_t^j. \tag{A.9}$$

Substitution of the first equation of (12) into (1) yields

$$y_t = (I + B_t^i F_t^i)(A_t y_{t-1} + B_t^j x_t^j + s_t) + B_t^i F_t^{ii} g_t^i. \tag{A.10}$$

Substitution of (A.9) and (A.10) into the second equation of (11) yields the second equation of (12). The matrices that are inverted are non-singular, because they are positive definite on the basis of the convexity assumptions with respect to the welfare loss functions.

Substitution of the second equation of (12) into the first equation of (12) and into (A.10) yields

$$x_t^i = F_t^i(I + B_t^j F_t^j)(A_t y_{t-1} + s_t) + F_t^{ii} g_t^i + F_t^i B_t^j F_t^{ji} g_t^i + F_t^i B_t^j F_t^{jj} g_t^j \tag{A.11}$$

and

$$y_t = (I + B_t^i F_t^i)(I + B_t^j F_t^j)(A_t y_{t-1} + s_t) + B_t^i F_t^{ii} g_t^i + (I + B_t^i F_t^i) B_t^j (F_t^{ji} g_t^i + F_t^{jj} g_t^j). \tag{A.12}$$

These results give the values of the right-hand sides of (A.7). Comparing quadratic terms and linear terms in y_{t-1} leads to the backward recursive equations of the feedback Stackelberg equilibrium.

4. Example 2

The solutions of equation (14) are:

$$\begin{bmatrix} -3/2 + 1/2\sqrt{5} & 0 \\ 0 & -3/2 + 1/2\sqrt{5} \end{bmatrix}, \begin{bmatrix} -3/2 - 1/2\sqrt{5} & 0 \\ 0 & -3/2 - 1/2\sqrt{5} \end{bmatrix},$$

$$\begin{bmatrix} -3/2 + 1/2\sqrt{5} & 0 \\ p & -3/2 - 1/2\sqrt{5} \end{bmatrix}, \begin{bmatrix} -3/2 - 1/2\sqrt{5} & 0 \\ q & -3/2 + 1/2\sqrt{5} \end{bmatrix},$$

$$\begin{bmatrix} r & s \\ \dfrac{-r^2 - 3r - 1}{s} & -3 - r \end{bmatrix}, p, q, r, s \in R, \quad s \neq 0.$$

The sets of eigenvalues of $\partial \operatorname{vec} D_{s+1}^i / \partial \operatorname{vec} D_s^j$ that have to be evaluated are $\{\text{four times } -3/2 + 1/2\sqrt{5}\}, \{\text{four times } -3/2 - 1/2\sqrt{5}\}, \{-3/2 + 1/2\sqrt{5}, -3/2 - 1/2\sqrt{5}, 1, 1\}, \{-3/2 + 1/2\sqrt{5}, -3/2 - 1/2\sqrt{5}, 1, 1\}$ and $\{-3/2 + 1/2\sqrt{5}, -3/2 - 1/2\sqrt{5}, 1, 1\}$, respectively, so that the first solution is the only one with all eigenvalues inside the unit circle and is therefore the only stable fixed point.