



*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

*No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.*

Dynamic Games and the Time  
Inconsistency of Optimal  
Policy in Open Economies

Marcus Miller and Mark Salmon

No. 232

**WARWICK ECONOMIC RESEARCH PAPERS**

DEPARTMENT OF ECONOMICS

UNIVERSITY OF WARWICK  
COVENTRY

Dynamic Games and the Time  
Inconsistency of Optimal  
Policy in Open Economies

Marcus Miller and Mark Salmon

No. 232

July 1983

The authors are grateful for constructive comments and suggestions from W.Buiter of L.S.E. and R.Hamalainen of the Helsinki University of Technology. They have also benefitted from discussions with C.Bean, P.Burridge and K.Roberts in particular, and with members of the Control Theory Centre at the University of Warwick.

This paper is a revision of an earlier one by the same authors (see References) and their thanks are due to Y.Slater for typing text and tables.

This paper is circulated for discussion purposes only and its contents should be considered preliminary.

NOTE: Annex 1 and 2 are available on request from Mrs.M.Perry,  
Department of Economics.

## Introduction

It has been argued that the Bretton Woods system (of pegged but adjustable exchange rates) set up after World War II was designed specifically to prevent the manipulation of exchange rates in pursuit of national macroeconomic objectives (see R.Cooper); so it is perhaps no coincidence that the ending of pegged rates has led to the re-emergence of theories and policies involving such "manipulation". Before the War, when inflation was low and unemployment high in the major industrialised nations, this involved "competitive depreciation" as countries tried to gain employment; in the late 1970's and early 1980's, however, countries like the UK and the US have embarked on policies of competitive appreciation in order to cut inflation, regarded as the first priority. Sharp movements in exchange rates, however, constitute a threat to orderly trade and stimulate protectionism, as Bergsten (1981) argues.

In this paper a framework for analysing the behaviour of rational policy makers in open economies with floating exchange rates is developed and applied to a relatively simple economic model. Such a framework should prove helpful in assessing (or constructing) plans for reforming current international monetary arrangements (as contained in the evidence and reports of the Treasury and Civil Service Committee on this topic, for example).

In the proposed framework, policy makers optimise in an environment with three main characteristics. First the economies are interdependent, so policy actions in any one country have significant impacts on other economies, second the economic processes are assumed to be

dynamic, so any optimisation involves plans over time, and third there is forward-looking behaviour on the part of private agents whose expectations of policy actions influence the effects of policy itself.

To characterise the behaviour of policy makers in an interdependent world, Koichi Hamada has applied ideas from game theory, specifically the notion of Nash behaviour and Stackelberg leadership, see for example his 1974 paper. Hamada's analysis is typically conducted in a static framework but, by using appropriate results from dynamic games, the same ideas can be applied to dynamic economies, as we show. (See also Buiter (1983) for a discussion of dynamic games and economic policy).

For the most part in what follows policy makers are assumed to take each others policy "paths" as given when choosing their own action (the Nash assumption). But, given the way in which they influence the expectations held by private agents, it appears that these policy makers are in the position of Stackelberg leaders vis a vis the private sector: specifically they can take into account the reaction of the private markets to policy announcements in choosing policy itself.

As is shown by the necessary conditions for the optimal open-loop behaviour of a Stackelberg leader, the policy is time inconsistent, as discussed in Kyland and Prescott (1977). And so too, for essentially the same reason, is optimal economic policy in an economy with perfect foresight, Calvo (1978) (cf. Driffill (1982) for a discussion of monetary policy under floating rates). The fact that in the presence of forward looking behaviour optimal economic policy is time inconsistent immediately implies, however, that it cannot be obtained by using dynamic programming techniques. The latter are based on Bellman's Principle of Optimality, so they are restricted to examining only time consistent policies which can be expressed in feedback form. Pontryagin's maximum principle, however, provides an open-loop solution

path which constitutes the optimal policy (although, when time inconsistent it cannot be meaningfully expressed in feedback form). For present purposes therefore the maximum principle is preferable, and by assuming an infinite time horizon provides analytically tractable solutions for optimum policy. The very fact that such solutions do not obey the Principle of Optimality must, however, raise some doubts about their credibility, so we investigate related time consistent policies on each occasion.

This paper is organised as follows. In Section 1 there is a brief review of the basic results of policy optimisation using the maximum principle for a single controller in an environment where there are no forward looking expectations influencing private behaviour. The open-loop optimal policy solutions are calculated (for the infinite time case) and shown to be time consistent.

In Section 2 the exercise is repeated in an N-country environment where each policy maker takes the policy path of the N-1 others as given in designing his/her own policy, so the system constitutes an N-person open-loop-Nash non zero-sum game. The solution paths (if they exist) are shown to be time consistent but are Pareto inferior to a centralised policy which is also calculated. In Section 3 the time inconsistent optimal solution of the open loop Stackelberg game is described, as are two closely related time consistent policy paths. The first of these is the "perfect cheating" solution where the leader changes the game by so using his freedom to make credible announcements as to deliberately mislead the follower; the second is the one where the leader "loses his leadership position" - and the game reverts to an open loop Nash solution. These game theoretic ideas are then used

to discuss optimal policy for a single policy maker in a general dynamic linear model with forward looking behaviour in Section 4 (and the contrast with the results of Section 1 is drawn).

As an Annex to Section 4 is an application to fiscal policy design in a small open economy with floating rates (Annex 1)\*. The effects on the roots of the system of applying time-inconsistent optimal policy are shown analytically, c.f. Livesey (1980), as are the consequences of "perfect cheating" and of "losing leadership". Numerical results obtained using "Saddlepoint" (Austin and Buiter (1982)) are also presented. The assumption of an infinite horizon helps to make the analysis more tractable than the finite time case examined by Driffill (1982).

In Section 5 the N-player open-loop Nash economic "game" of Section 2 is extended to include forward-looking behaviour by private agents. Optimal behaviour, on the assumption that policy makers act in Nash fashion towards each other but act as leaders vis a vis private agents, is derived and is seen to be neither time inconsistent nor Pareto optimal. Centralised, Pareto optimal policy, also calculated, is not time consistent either.

As an Annex to this section, the example earlier used is extended to cover two interdependent economies with a floating exchange rate (Annex 2)\*. The formal characterisation of optimal time-consistent Nash policy requires a high order differential equation; but the assumption that the two countries share a common economic structure (cf. Aoki (1981)) together with numerical solutions obtained using "Saddlepoint" make the analysis reasonably accessible.

In conclusion, some of the implications of the proposed framework are drawn out, and various further developments briefly outlined.

(1) A Review of policy optimisation with a single controller and a quadratic criterion.

Consider an economic system described by the linear state equation

$$(1.1) \quad Dx = Ax + Bu$$

where the state variables are represented by the vector  $x$  and the control variables by the vector  $u$ . In what follows all such vectors are implicitly time-indexed, so that  $x \equiv x(t)$ ; and the operator  $D$  indicates differentiation with respect to time.

The policy maker is assumed to select a path for  $u$  by minimising a quadratic cost function

$$(1.2) \quad V = \frac{1}{2} \int_{t_0}^{\infty} (x^T(s)Qx(s) + u^T(s)Ru(s))ds$$

where  $Q$  and  $R$  are symmetric matrices (positive semi-definite and positive definite respectively) representing the preferences of the policy maker, and the superscript  $T$  denotes transposition. The assumption of an infinite horizon is made for analytical convenience (as explained below) and the solution may be obtained by applying Pontryagin's "maximum principle" (see for example Intriligator (1971, Chapter 14) or Wiberg (1971, Chapter 10)).

The "Hamiltonian" relevant for this minimisation problem is



$$(1.3) \quad H = \frac{1}{2}x^T Qx + \frac{1}{2}u^T Ru + p^T (Ax + Bu)$$

where the vector  $p$  denotes the costate variables (or "shadow prices") attached to the state variables by the policy maker at each point in the plan. The necessary conditions for optimisation may be written

$$(1.4) \quad \frac{\partial H}{\partial u} = Ru + B^T p = 0$$

$$(1.5) \quad \frac{\partial H}{\partial x} = Qx + A^T p = -Dp$$

$$(1.6) \quad \frac{\partial H}{\partial p} = Ax + Bu = Dx, \text{ i.e. the model of (1.1) above.}$$

Equation (1.4) implies that the policy instruments are related to the costate variables as

$$(1.7) \quad u = -R^{-1}B^T p$$

and, on substituting this into equation (1.6), one obtains the "adjoint system" describing the dynamic behaviour of both the state and costate variables under optimal control:

$$(1.8) \quad \begin{bmatrix} Dx \\ Dp \end{bmatrix} = \begin{bmatrix} A & -BR^{-1}B^T \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix} \equiv M \begin{bmatrix} x \\ p \end{bmatrix}$$

If there are  $n$  state variables, equation (1.8) is a system of  $2n$  linear homogeneous differential equations, which will have a unique solution given  $2n$  "boundary" conditions. The first  $n$  of these are typically provided by the "predetermined" initial values of the state vector,  $x(t_0) = x_0$ ; the remaining  $n$  conditions are given by the "transversality" condition which, assuming controllability for the infinite time problem considered here, ensures that the solution depends only on the stable roots. Note that the adjoint system has a saddle point structure, indeed the

roots appear in pairs of opposite sign see Wiberg (1981, p.215), for example.

The dynamic adjustment of this system can be analysed easily with the aid of the usual canonical transformation. Let the column eigenvectors of  $M$  be denoted by the matrix  $C$  so that

$$(1.10) \quad MC = CA$$

where  $\Lambda$  is a diagonal matrix of eigenvalues. Then the variables  $x$  and  $p$  can in general be expressed as functions of canonical variables  $z$  (each associated with a single root) as follows:

$$(1.11) \quad \begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} z_s \\ z_u \end{bmatrix}$$

where

$$(1.12) \quad \begin{bmatrix} Dz_s \\ Dz_u \end{bmatrix} = \begin{bmatrix} \Lambda_s & \\ & \Lambda_u \end{bmatrix} \begin{bmatrix} z_s \\ z_u \end{bmatrix}$$

and  $z_s, z_u$  denote the canonical variables associated with the stable and unstable roots respectively.

As the transversality condition requires that  $z_u = 0$ , this implies  $x = C_{11}^{-1} z_s$  or  $z_s = C_{11}^{-1} x$ . Hence the open loop dynamics of the state variables under control are found from the initial condition  $x(t_0) = x_0$  and the equation

$$(1.13) \quad Dx = C_{11} \Lambda_s C_{11}^{-1} x.$$

which together imply

$$(1.14) \quad x = C_{11} e^{\Lambda_s(t-t_0)} C_{11}^{-1} x(t_0).$$

Since the transversality also requires

$$(1.15) \quad p = C_{21} z_s = C_{21} C_{11}^{-1} x$$

the open loop dynamics of the shadow prices are

$$(1.16) \quad p = C_{21} e^{\Lambda_s(t-t_0)} C_{11}^{-1} x(t_0)$$

and the optimal path planned for the instruments of policy is, given

(1.7) above, consequently

$$(1.17) \quad u = R^{-1} B^T C_{21} e^{\Lambda_s(t-t_0)} C_{11}^{-1} x(t_0).$$

It is easy to confirm that such a plan is "time consistent".

i.e. satisfies Bellman's Principle of Optimality which asserts

"an optimal policy has the property that whatever the initial state and decision [control] are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision". Bellman (1957).

Let us denote by a prefix to the state variable the date at which the open loop plan was formed and include explicitly the time index.

Then we find that reoptimising at  $t_1$  given the initial condition

$x(t_1)$  (achieved by following the optimal plan formed at  $t_0$ ) yields an optimal plan

$$\begin{aligned}
{}_{t_1}x(t) &= {}_{t_1}C_{11}e^{\Lambda_s(t-t_1)} {}_{t_1}C_{11}^{-1} {}_{t_0}x(t_1) \quad \text{for } t \geq t_1 \\
&= {}_{t_1}C_{11}e^{\Lambda_s(t-t_0)} {}_{t_1}C_{11}^{-1} x(t_0)
\end{aligned}$$

i.e. the optimal plan found at  $t_1$  is simply a continuation of the plan found at  $t_0$  provided the system was at  ${}_{t_0}x(t_1)$ . Hence that portion of the original plan from  $t_1$  onwards constitutes an optimal trajectory in its own right.

The same logic applies to the shadow prices and controls so for a time consistent plan

$${}_{t_1}p(t) = {}_{t_0}p(t), \quad {}_{t_1}u(t) = {}_{t_0}u(t), \quad \text{for } t_1 > t_0.$$

i.e. the shift of "origin" for optimisation from  $t_0$  to  $t_1$  leaves the planned values for costate and policy variables unaffected.

Since the plan of equations (1.15) and (1.16) is time consistent, so  ${}_{t_1}x(t) = {}_{t_0}x(t)$ ,  $t \geq t_1$ , as long as  $x(t_1) = {}_{t_0}x(t_1)$ . But if some extraneous displacement were to occur so that, for example,  $x(t_1) = {}_{t_0}x(t_1) + \delta$  then the optimal policy would not be to continue with the original plan. Instead the new plan would be based on shadow prices which differ from  ${}_{t_0}p(t)$  since

$$\begin{aligned}
{}_{t_1}p(t) &= {}_{t_1}C_{21}e^{\Lambda_s(t-t_1)} {}_{t_1}C_{11}^{-1} \begin{bmatrix} {}_{t_0}x(t_1) + \delta \end{bmatrix} \\
&= {}_{t_0}p(t) + {}_{t_1}C_{21}e^{\Lambda_s(t-t_1)} {}_{t_1}C_{11}^{-1} \delta, \quad t \geq t_1
\end{aligned}$$

which implies, given (1.7)

$${}_{t_1}u(t) = {}_{t_0}u(t) - R^{-1}B^TC_{21}e^{\Lambda_s(t_2-t_1)}C_{11}^{-1}\delta, \quad t \geq t_1$$

Thus the controls follow the original open loop trajectory chosen at  $t_0$  modified by an "innovation contingent response", starting at the time of the displacement and decaying ultimately to zero.

(2) Decentralised control: an N-country "Nash" differential game

Given a dynamic system with the state equation

$$(2.1) \quad \dot{x} = Ax + \sum_{i=1}^N B_i u_i$$

where the  $u_i$  denote the instrument(s) under the control of country  $i$ , one could characterise the behaviour of the system under decentralised optimal control as follows.

Let the  $i^{\text{th}}$  country minimise a cost functional of the form

$$V^i = \frac{1}{2} \int_{t_0}^{\infty} (x^T(s) Q_i x(s) + \sum_{j=1}^N u_j^T(s) R_{ij} u_j(s)) ds$$

where the weighting matrices are symmetric and satisfy  $Q_i \geq 0$ ,  $R_{ii} > 0$ ,  $R_{ij} \geq 0$ ,  $i \neq j$ ,  $i, j = 1, N$ .

Defining the Hamiltonian for the  $i^{\text{th}}$  country as

$$(2.2) \quad H^i = x^T Q_i x + \sum_{j=1}^N u_j^T R_{ij} u_j + p_i^T (Ax + \sum_{j=1}^N B_j u_j)$$

where  $p_i$  are the shadow prices relevant to the  $i^{\text{th}}$  country, the necessary conditions for cost minimisation become

$$(2.3) \quad \frac{\partial H^i}{\partial u_i} = R_{ii} u_i + B_i^T p_i = 0 \leftrightarrow u_i = -R_{ii}^{-1} B_i^T p_i$$

and (assuming that each country fails to realise that the other countries are also altering their controls in response to the environment)

$$(2.4) \quad Dp_i = -\frac{\partial H_i}{\partial x} = -Q_i x - A^T p_i$$

After substitution of the optimised values of the instruments, the dynamics of the system under decentralised control will for this "open loop" Nash game be

$$(2.5) \quad \begin{bmatrix} Dx \\ Dp_1 \\ \vdots \\ Dp_N \end{bmatrix} = \begin{bmatrix} A & -J_1 & \dots & -J_N \\ -Q_1 & -A^T & \dots & 0 \\ \vdots & \vdots & & \vdots \\ -Q_N & 0 & \dots & -A^T \end{bmatrix} \begin{bmatrix} x \\ p_1 \\ \vdots \\ p_N \end{bmatrix} \equiv M^N \begin{bmatrix} x \\ p_1 \\ \vdots \\ p_N \end{bmatrix}$$

where  $J_i \equiv B_i R_i^{-1} B_i^T$ ,  $i = 1, \dots, N$ . (See Basar and Olsder (1981, Ch.6.))

Assuming, as before, the dynamics of the system under such control should be restricted to the stable manifold, and that the state variables are predetermined at  $t_0$ , the open loop path for the state variables from  $x(t_0) = x_0$  is given by

$$(2.6) \quad Dx = C_{11} \Lambda_s C_{11}^{-1} x$$

where the  $C_{11}$ , and  $\Lambda_s$  are now defined with respect to the matrix  $M^N$  shown above in equation (2.5).

The "control rules" in operation must be of the form

$$(2.7) \quad u_i^1 = -R_{ii}^{-1} B_i^T p_i = -R_{ii}^{-1} B_i^T C_{i+1,1} C_{11}^{-1} x$$

where  $C_{i+1,1}$  refers to the appropriate block of the column eigenvectors

associated with  $\Lambda_s$ .

Note that despite the number of dynamic equations in the decentralised system, the number of distinct roots in (2.5) can only be thrice the number of state variables, since the dynamics of the state variables can be reduced to a second order differential equation in the state variables, and all the  $p_i$  have the same block,  $-A^T$ , on the diagonal adding a further  $N$  roots.

Pareto optimal policies may be derived by assuming a single controller and assigning weights to the individual countries cost functionals. The problem is then to minimise

$$V = \frac{1}{2} \int_{t_0}^{\infty} \left( \sum_{i=1}^N w_i (x(s))^T Q_i x(s) + \sum_{j=1}^N u_j^T(s) R_{ij} u_j(s) \right) ds$$

by choice of paths for the  $u_i$ .

Given the Hamiltonian

$$(2.8) \quad H = \frac{1}{2} \sum_{i=1}^N w_i (x^T Q_i x + \sum_{j=1}^N u_j^T R_{ij} u_j) + p^T (Ax + \sum_{i=1}^N B_i u_i)$$

the necessary conditions are

$$\left( \sum_{j=1}^N w_j R_{ji} \right) u_i = -B_i^T p$$

or

$$R_i^* u_i = -B_i^T p$$

$$(2.9) \quad u_i = -R_i^{*-1} B_i^T p$$

and

$$(2.10) \quad Dp = \frac{\partial H}{\partial x} = - \sum_{i=1}^N w_i Q_i x - A^T p$$



so the the system under "centralised" control would be

$$(2.11) \quad \begin{bmatrix} Dx \\ Dp \end{bmatrix} = \begin{bmatrix} A & \begin{matrix} N \\ -\sum_{i=1}^N J_i \end{matrix} \\ \begin{matrix} N \\ -\sum_{i=1}^N w_i Q_i \end{matrix} & \begin{matrix} J_i \\ -A^T \end{matrix} \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix}$$

where  $J_i \equiv B_i R_i^{*-1} B_i^T$   $i = 1, \dots, N$  as before.

The dynamic behaviour of the state variables is thus given by  $x(t_0) = x_0$  and

$$(2.12) \quad Dx = C_{11} \Lambda_s C_{11}^{-1} x,$$

for  $C_{11}$  and  $\Lambda_s$  defined by the adjoint system (2.11). The control rule is then given by

$$(2.13) \quad u_i = -R_i^{*-1} B_i^T C_{21} C_{11}^{-1} x.$$

In a system such as that under discussion, where there is no long-run inconsistency of targets (so each country can ultimately reach zero cost equilibrium) it is nevertheless true that the Nash-decentralised control will be inefficient relative to centralised control (which is Pareto-optimal). While one may in principle be able to solve for Pareto efficient policy, however, it is not obvious how such policies would be implemented and politically sustained.

As the optimal solutions for both the decentralised open-loop Nash game and for the centralised controller are "time consistent" in the sense defined in Section 1, they too could have been calculated by dynamic programming methods rather than by using the maximum principle.

(3) The time inconsistency of optimal policy in a two-person Stackelberg differential game

In Section 1 the unique solution for the optimal control problem for a single controller was obtained with the aid of initial conditions for all the state variables, whose starting values were all predetermined. By contrast the initial values for the co-state variables were determined by the "transversality" requirement that they lay on the stable manifold of an adjoint system with a saddlepoint structure.

It is not unusual in economics, however, to find that the state equations themselves have a saddle point structure, and that some of the state variables are determined by the restriction that they lie on the stable manifold of such a system (cf. the "rational expectations" models of floating exchange rates to be found in Dornbusch (1976), Buiter and Miller (1981) and Driffill (1982), for example). Since the application of control changes the dynamics of the system, as shown in Section 1, it must in general change the initial values of those state variables which are determined in this way. For such models therefore it is inappropriate to assume that the entire vector  $x$  is pre-determined at time  $t_0$ , as has been done in Sections 1 and 2 above.

the initial values of

In considering how to proceed where some of the state variables are determined endogenously to the control problem, it is illuminating first to consider in some detail the procedures used for analysing a two-person dynamic game with a Stackelberg leader, because the follower's behaviour is characterised by a saddle point structure. Specifically we will examine the open-loop solution of an infinite-horizon differential game with a Stackelberg leader, before returning to the general problem of control with "forward-looking" variables.

### A two-person Stackelberg Differential Game

For a Stackelberg game between two players with a linear structure, the state equation is

$$(3.1) \quad Dx = Ax + B_1 u_1 + B_2 u_2$$

where the policy variables of the leader are represented by  $u_1$  and those of the follower by  $u_2$ . The follower is assumed to minimise the cost function

$$(3.2) \quad V_2 = \frac{1}{2} \int_{t_0}^{\infty} (x(s)^T Q_2 x(s) + u_1^T(s) R_{21} u_1(s) + u_2^T(s) R_{22} u_2(s)) ds$$

where the weighting matrices are symmetric and satisfy  $Q_2 \geq 0$ ,

$R_{21} \geq 0$ ,  $R_{22} > 0$ , subject to (3.1) and the given value of the leaders announced path for  $u_1$  from time  $t_0$  onwards. The Hamiltonian for the follower is thus

$$H^2 = \frac{1}{2} (x^T Q_2 x + u_1^T R_{21} u_1 + u_2^T R_{22} u_2) + p_2^T (Ax + B_1 u_1 + B_2 u_2)$$

with corresponding first-order conditions

$$(3.3) \quad u_2 = -R_{22}^{-1} B_2^T p_2$$

$$(3.4) \quad Dp_2 = -\frac{\partial H^2}{\partial x} = -Q_2 x - A^T p_2.$$

where the follower's costate variable  $p_2$  implicitly depends on the leader's policy.

The leader calculates his optimal path for  $u_1$  taking into account the behaviour just ascribed to the follower by minimising

$$(3.5) \quad V_1 = \frac{1}{2} \int_{t_0}^{\infty} (x^T(s) Q_1 x(s) + u^T(s) R_{11} u(s) + u_2^T(s) R_{12} u_2(s)) ds$$

where the weighting matrices are symmetrical and satisfy  $Q_1 \geq 0$ ,  $R_{11} > 0$ ,  $R_{12} \geq 0$ , subject to the follower's first order conditions

$$(3.6) \quad Dx = Ax + B_1 u_1 + B_2 u_2$$

and

$$(3.7) \quad Dp_2 = -Q_2 x - A^T p_2$$

Since the control solution for the follower has a "saddle-point" structure, it is evident that the leader is faced by a set of state variables (in equations (3.6) and (3.7)) whose initial conditions are not all independent of his activity. Specifically  $p_2(t_0)$  will vary with the leader's policy.

For the leader the relevant Hamiltonian may be written

$$(3.8) \quad H^1 = \frac{1}{2} (x^T Q_1 x + u_1^T R_{11} u_1 + p_2^T J_{12} p_2) + p_1^T (Ax + B_1 u_1 - J_2 p_2) + p_2^{*T} (-Q_2 x - A^T p_2).$$

where  $p_2^*$  is the costate which the leader attaches to the follower's costate variable, where  $J_2 \equiv B_2 R_{22}^{-1} B_2^T$  and  $u_2 = -R_{11}^{-1} B_2^T p_2$  leads to  $u_2^T R_{12} u_2$  being written as  $p_2^T B_2 R_{22}^{-1} R_{12} R_{22}^{-1} B_2^T p_2 \equiv p_2^T J_{12} p_2$ .

The first order conditions for the leader are thus

$$(3.9) \quad u_1 = -R_{11}^{-1} B_1^T p_1 \equiv -J_1 p_1$$

$$(3.10) \quad Dp_1 = -Q_1 x + Q_2 p_2^* - A^T p_1$$

$$(3.11) \quad Dp_2^* = Ap_2^* + J_2 p_1 - J_{12} p_2.$$

So, collecting results, we obtain in matrix form (c.f. Simaan and Cruz (1973)) the necessary conditions for the open loop Stackelberg solution with player 1 acting as leader

$$(3.12) \quad \begin{bmatrix} Dx \\ Dp_2^* \\ Dp_1 \\ Dp_2 \end{bmatrix} = \begin{bmatrix} A & 0 & -J_1 & -J_2 \\ 0 & A & J_2 & -J_{12} \\ -Q_1 & Q_2 & -A^T & 0 \\ -Q_2 & 0 & 0 & -A^T \end{bmatrix} \begin{bmatrix} x \\ p_2^* \\ p_1 \\ p_2 \end{bmatrix} = M^S \begin{bmatrix} x \\ p_2^* \\ p_1 \\ p_2 \end{bmatrix}$$

where  $J_i = B_i R_{ii}^{-1} B_i^T$ ,  $i = 1, 2$

and  $J_{12} = B_2 R_{22}^{-1} R_{12} R_{22}^{-1} B_2^T$

with  $x(t_0)$  given,  $p_2^*(t_0) = 0$  and the transversality condition that the system be restricted to the stable manifold. Notice that with  $p_1$  and  $p_2$  free to adjust so as to satisfy the transversality conditions,

the leader will set  $p_2^*(t_0) = 0$  at the time of initial optimisation; as  $p_2^*$  is the shadow price of  $p_2$ , so doing reduces the "costs" the leader sees associated with  $p_2$  to a minimum.

The dynamics of adjustment of  $x$  and  $p_2^*$  under such conditions will be described by

$$(3.13) \quad \begin{bmatrix} \dot{x} \\ \dot{p}_2^* \end{bmatrix} = C_{11} \Lambda_s C_{11}^{-1} \begin{bmatrix} x \\ p_2^* \end{bmatrix}$$

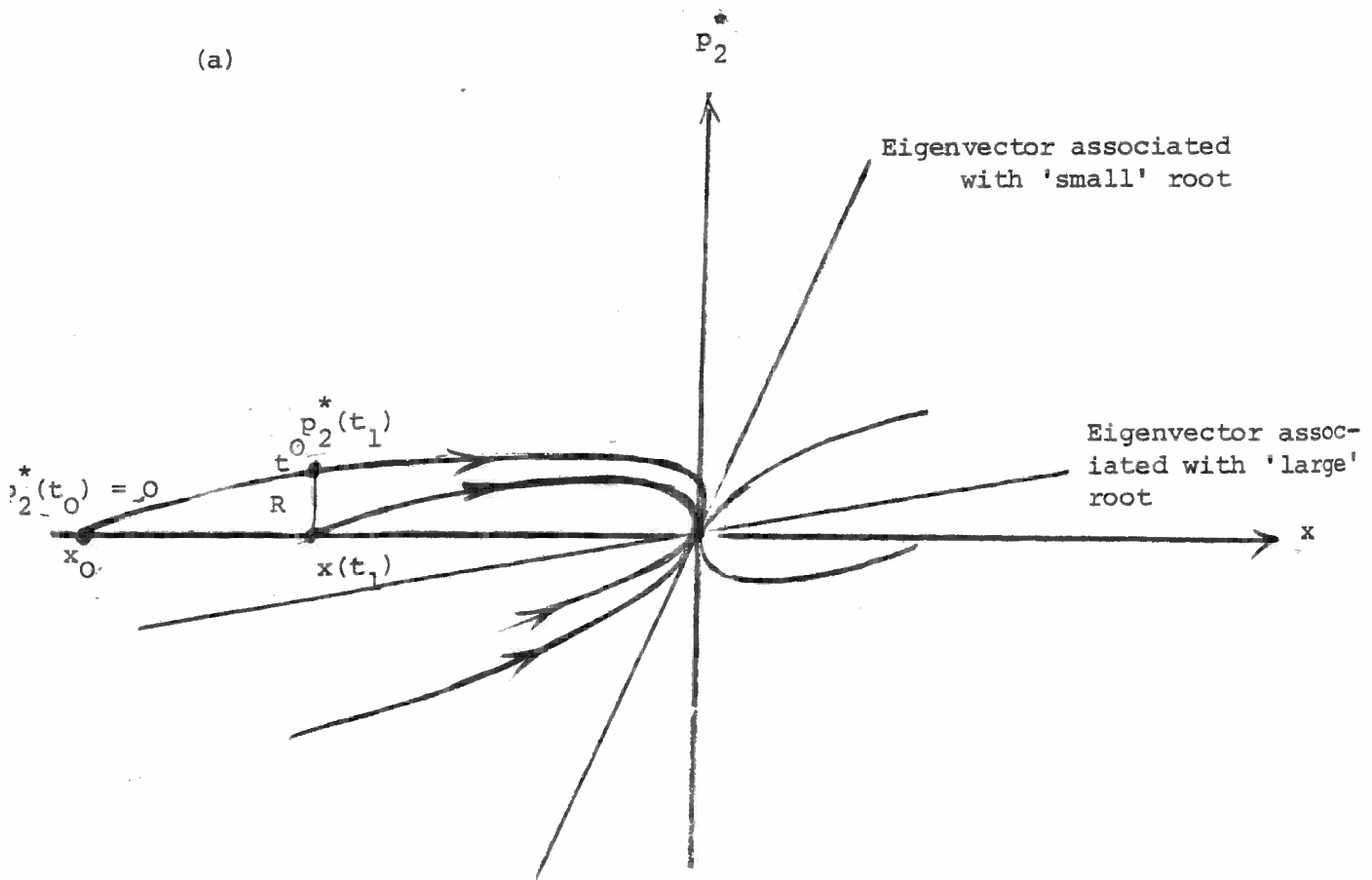
$$= \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} - \begin{bmatrix} J_1 & J_2 \\ -J_2 & J_{12} \end{bmatrix} C_{21} C_{11}^{-1} \begin{bmatrix} x \\ p_2^* \end{bmatrix}$$

Where  $\begin{bmatrix} C_{11} \\ C_{12} \end{bmatrix}$  are the stable column eigenvectors of  $M^S$

with the phase diagram of a stable node such as that shown in Figure 1a.

For the case when  $x$  is a scalar, equation (3.13) has a phase portrait with a stable node such as that shown in Figure 1a. As the leader is trying to minimise the costs of  $x$  being away from zero, one would expect the eigenvector associated with the (absolutely) larger stable root to lie close to the horizontal axis as shown - but not to coincide with the  $x$  axis except in degenerate cases. Various trajectories satisfying the differential equations and the transversality condition are shown converging to the origin, and the unique optimal trajectory associated with the initial condition  $x(t_0) = x_0$  is shown starting with  $p_2^*(t_0) = 0$ .

(a)



(b)

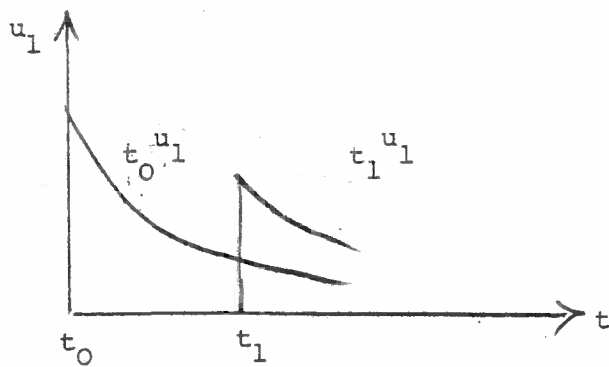


FIGURE 1. The time inconsistency of optimal policy in an open loop Stackelberg game.

As is evident from the phase portrait, this optimal path is not "time consistent". For if we consider the leader "reoptimising" at some later date  $t_1 > t_0$  when the state variable pursuing a deterministic path has reached  $x(t_1)$ , then his decision must be to reset  $p_2^*(t_1) = 0$ , which would lead to a new trajectory starting from point R in Figure 1a.

Resetting  $p_2^*$  to zero represents a change in the plan for future path for  $u_1$  originally announced by the leader, as is illustrated in Figure 1b, where the path for  $u_1$  announced at  $t_0$  is shown labelled as  ${}_{t_0}u_1$ . In this scalar example this is proportional to  $p_1$  as  $u_1 = R_{11}^{-1} B^T p_1$ , so with a change of units the same path represents  $p_1$  as well). At time  $t_1$  the announced path shifts to  ${}_{t_1}u_1$  (and this will be associated with a shift in  $p_1$  at time  $t_1$  for the reason just given).

As  ${}_{t_0}p_2^*(t_1) \neq {}_{t_1}p_2^*(t_1) = 0$ , the optimal plan formulated at  $t_0$  with the aid of the maximum principle is not time consistent, and we would not therefore expect to be able to determine it by using dynamic programming methods based on Bellman's Principle of Optimality. A necessary condition for Bellman's principle to apply in this context appears to be that  $p_2^*(t) = 0$  for all  $t > t_0$ , and we now consider when this time consistency constraint may be satisfied.

One optimal solution to the Stackelberg game which (if it exists) is time consistent is where the leader can carry out a policy of "perfect cheating" see Hamalainen (1981). It is assumed that the leader may announce any control path  $u_1^c$  which need not correspond to path actually chosen for  $u_1$ . Since the leader has, in virtue



of such announcements, a "costless" way of affecting the followers behaviour: (by manipulating  $p_2$ ) the perfect cheating solution involves setting  $p_2^*(t_0) = 0$  and  $Dp_2^* = 0$ , which from equation (3.11) above implies  $J_2 p_1 - J_{12} p_2 = 0$  or  $B_2 R_{22}^{-1} B_2^T p_1 = B_2 R_{22}^{-1} R_{12} R_{22}^{-1} B_2^T p_2$  so the shadow prices of the follower are linked to the shadow prices of the leader.

The dynamic behaviour of the state variables and the shadow prices under such a policy, also known in the literature as a "team Basar and Olsder (1980) optimal solution", can be obtained by simply setting equation (3.11) to zero and dropping equation (3.4) - because the follower is "cheated" into choosing those controls for  $u_2$  which suit the leader. This changes the time inconsistent system (3.12) to the lower-order time consistent system shown as

$$(3.14) \quad \begin{bmatrix} \dot{x} \\ Dp_1 \\ 0 \end{bmatrix} = \begin{bmatrix} A & -J_1 & -J_2 \\ -Q_1 & -A^T & 0 \\ 0 & J_2 & -J_{12} \end{bmatrix} \begin{bmatrix} x \\ p_1 \\ p_2 \end{bmatrix}$$

or alternatively

$$(3.15) \quad \begin{bmatrix} \dot{x} \\ Dp_1 \end{bmatrix} = \begin{bmatrix} A & -(J_1 + J_2 J_{12}^{-1} J_2) \\ -Q_1 & -A^T \end{bmatrix} \begin{bmatrix} x \\ p_1 \end{bmatrix}$$

where  $J_i = B_i R_{ii}^{-1} B_i^T$   $i = 1, 2$ .

$$J_{12} = B_2 R_{22}^{-1} R_{12} R_{22}^{-1} B_2^T$$

The existence and uniqueness of such a solution to a Stackelberg game is discussed in Basar and Olsder (1980). The necessary requirement discussed there (that the pair  $(A, (B_1, B_2))$  is stabilizable) appears to rule out the anomalous situation where the follower has control variables "which do not affect the state variable, and thereby cannot be detected by the leader" Basar and Olsder (1982, p.331).

An alternative approach which also avoids the problem of time inconsistency suggested to us by W.Buiter is to consider circumstances where the need for  $P_2^*$  does not arise at all, so that one simply deletes the 2nd row and 2nd column of 2.2. On doing this one finds that the resulting system is that of the open loop Nash game examined in the previous section! If the leader "loses his leadership role", because for example his announcements cease to be credible, the solution will become time consistent (though the game will also have changed its character).

In our subsequent treatment of the problem of policy design in an economy with forward-looking behaviour, both of the time consistent solutions referred to here, the "perfect cheating" proposed by Hamalainen and the "loss of leadership" suggested by Buiter will be used again, though the latter is typically labelled "Nash Policy" since the government becomes a Nash player rather than a Stackelberg leader. (The problem of time inconsistency would not arise if a single policy maker were to choose all instruments so as to minimise the weighted sum of the leader's and follower's costs; these Pareto optimal solutions are not investigated below as the cost functions of the follower are not typically available).

#### 4. Optimal control of an economy with forward looking behaviour

Having considered the issues to be faced by a Stackelberg leader in a dynamic game, we examine the problem for the policy maker in an economy with a saddle point dynamic structure. This involves some modification of the results of Section 1 as the state vector now includes some variables which are not historically predetermined but are free to "jump" (to satisfy transversality conditions).

The state variables which are not predetermined are typically "forward looking" prices which depend on the (expected) future path of the system, and there is a direct parallel between such variables and the shadow prices in an optimal control problem. Consider for example the equation describing the behaviour of these prices in the saddle point models of investment of Blanchard (1981) Summers (1981) and Hayashi (1982). The financial arbitrage equation between shares and other short-term financial assets is essentially:

$$\frac{Dv}{v} + \frac{d}{v} = \rho$$

where  $v$  is the (real) value of a share

$d$  is the real dividend paid on the share

$\rho$  is the instantaneous real rate of interest available elsewhere.

It is evident that such an equation is quite likely to introduce an unstable root as it implies  $Dv = \rho v - d$ , though one cannot of course say anything definite until the mechanisms determining  $\rho$

and  $d$  had been spelled out (as in the references).

It will be observed that such an arbitrage equation is almost identical to the first-order condition for the costate variable (shadow price) in a scalar case where there is a constant discount factor  $\rho$  in the criterion, namely

$$Dp - \rho p = - H_x \quad \text{or} \quad \frac{Dp}{p} + \frac{H_x}{p} = \rho$$

where  $p$  is the costate variable

$\rho$  is the discount factor

$$H_x \equiv \frac{\partial H}{\partial x} \quad \text{where } x \text{ is the state variable.}$$

In this paper there is no discount factor in the criterion function, but the analogy is nevertheless appropriate.

\*

In Annex 1 below we present a specific economic model containing a financial arbitrage equation (between domestic and foreign financial assets in an open economy) and possessing a saddlepoint structure. Here however, we apply the results just obtained for the Stackelberg game to the problem of policy design in a general linear model with a saddlepoint structure and some state variables not predetermined but free to "jump". The number of the latter, to be interpreted as "forward looking" prices (just like costate variables), is less than or equal to the number of unstable roots.

\* Not included in this copy: available on request

Assume the economic system with a saddle point structure is linear with a minimal state space representation

$$(4.1) \quad Dx = Ax + Bu$$

where  $x \equiv \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$  is the state vector (with  $x_1$  being predetermined at  $t_0$  and  $x_2$  "jump" variables which are free to adjust at the time of optimisation) and  $u$  is the vector of policy instruments to be chosen. Let the single policy maker minimise the quadratic cost functional

$$(4.2) \quad V = \frac{1}{2} \int_{t_0}^{\infty} (x^T(s)Qx(s) + u^T(s)Ru(s)) ds$$

by appropriate choice of  $u$ . The differential equations for the adjoint system of state and costate variables,  $p$ , described in Section 1 still apply so

$$(4.3) \quad \begin{bmatrix} Dx \\ Dp \end{bmatrix} = \begin{bmatrix} A & -J \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix}$$

where  $J = BR^{-1}B^T$  and  $p = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$  partitioned conformably with  $x$ .

The assumption made in Section 1, that the entire state vector is predetermined is, however, no longer appropriate. Instead, at  $t_0$  we take  $x_1$  to be predetermined "given by history" together with  $p_2$  (costates for  $x_2$ ) "given by optimality" - as discussed

below. It is convenient, therefore, to reorder the system so that as before the "predetermined" variables came first, thus

$$(4.4) \quad \begin{bmatrix} Dx_1 \\ Dp_2 \\ Dp_1 \\ Dx_2 \end{bmatrix} = \begin{bmatrix} A_{11} & -J_{12} & -J_{11} & A_{12} \\ -Q_{21} & -A_{22}^T & -A_{12}^T & -Q_{22} \\ -Q_{11} & -A_{21}^T & -A_{11}^T & -Q_{12} \\ A_{21} & -J_{12} & -J_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ p_2 \\ p_1 \\ x_2 \end{bmatrix} = M \begin{bmatrix} x_1 \\ p_2 \\ p_1 \\ x_2 \end{bmatrix}$$

where  $A, Q$  and  $J$  are partitioned to conform with  $x$  and  $p$ .

The path to be followed by  $x_1$  and  $p_2$  from their initial conditions can be described as a function of the stable eigenvalues and their associated eigenvectors in the usual way, so

$$(4.5) \quad \begin{bmatrix} Dx_1 \\ Dp_2 \end{bmatrix} = C_{11} \Lambda_s C_{11}^{-1} \begin{bmatrix} x_1 \\ p_2 \end{bmatrix}$$

$$\text{where } MC = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} \Lambda_s \\ \Lambda_u \end{bmatrix}$$

i.e.  $C$  denotes the column eigenvectors of  $M$  above. The remaining variables are linearly related to these by the expression

$$(4.6) \quad \begin{bmatrix} p_1 \\ x_2 \end{bmatrix} = C_{21} C_{11}^{-1} \begin{bmatrix} x_1 \\ p_2 \end{bmatrix} .$$

Apart from the reordering of the variables all is as in Section 1. The crucial question is of course the value at time  $t_0$  to be assigned to  $p_2$ , the shadow prices of the "jump" variables,  $x_2$ . From our consideration of the Stackelberg game it will come as no surprise that the cost minimising value for  $p_2$  is zero. As Driffil (1982, p.8) argues, such state variables

"can initially take on any value. They are not constrained by their past history. Since a costate variable indicates the (contribution) to the maximised value of the maximand of a marginal change in the starting point of the state variable, freedom to choose any starting point implies the marginal value of a change must be zero".

While  $x_2(t_0)$  is historically predetermined, therefore,  $p_2(t_0) = 0$ . Note, however, that in general  $p_2$  will cease to be zero once the system-under-control proceeds from the initial conditions established at  $t_0$ .

The optimal solution to the minimisation problem may be expressed in explicit 'open loop' form as

$$(4.7) \quad \begin{bmatrix} x_1(t_0) \\ p_2(t_0) \end{bmatrix} = C_{11} e^{A_s(t-t_0)} C_{11}^{-1} \begin{bmatrix} x_1(t_0) \\ 0 \end{bmatrix}$$

where the subscripts on the projected values of  $x_1$  and  $p_2$  indicate the date at which the plan was formed. This plan is not time consistent

however, however, because at a later date the optimal policy on the same criterion will not be to continue with the programme

$(t_0 x_1, t_0 p_2)$  shown in equation (4.7).

Faced with reoptimisation at some later date  $t_1 > t_0$ , the policymaker will proceed by resetting  $p_2$  to zero, so  $p_2(t_1) = 0$ . The new programme will therefore be

$$(4.8) \quad \begin{bmatrix} t_1 x_1 \\ t_1 p_2 \end{bmatrix} = C_{11} e^{\Lambda_s(t-t_1)} C_{11}^{-1} \begin{bmatrix} x(t_1) \\ 0 \end{bmatrix}$$

This will clearly involve a change in the path planned for  $p_2$ ; it will also involve a change in the other jump variables as

$$(4.9) \quad \begin{bmatrix} t_1 p_1 \\ t_1 x_2 \end{bmatrix} \equiv C_{21} e^{\Lambda_s(t-t_1)} C_{11}^{-1} \begin{bmatrix} t_0 x(t_1) \\ 0 \end{bmatrix} \neq C_{21} e^{\Lambda_s(t-t_0)} C_{11}^{-1} \begin{bmatrix} x(t_0) \\ 0 \end{bmatrix} \equiv \begin{bmatrix} t_0 p_1 \\ t_0 x_2 \end{bmatrix}$$

Since the optimal policy of (4.7) is not time consistent, it cannot be calculated using the standard techniques of dynamic programming, as Kydland and Prescott (1977) pointed out. Dynamic programming methods are based upon Bellman's Principle of Optimality which, as we have seen, means that only time consistent policies are considered. The open loop policy calculated using Pontryagin's Maximum Principle minimises the criterion function for a given initial condition (which is why we refer to it as optimal) even though it fails in this case to satisfy Bellman's Principle of Optimality.



Though the optimal plan is expressed as a time inconsistent open loop, this does not prevent a response to extraneous disturbances to the state vector so long as the response does not undermine the credibility of the initial trajectory. The optimal innovation contingent response which is associated with the trajectory of equation (4.7) is simply described. Let  $x(t_1) = {}_{t_0}x(t_1) + \delta$  where  $\delta$  is the disturbance, then the optimal policy thereafter will be

$$\begin{aligned}
 (4.10) \quad \begin{bmatrix} {}_{t_1}x_1(t) \\ {}_{t_1}p_2(t) \end{bmatrix} &= C_{11} e^{\Lambda_s(t-t_1)} C_{11}^{-1} \begin{bmatrix} {}_{t_0}x_1(t_1) + \delta \\ {}_{t_0}p_2(t_1) \end{bmatrix} \\
 &= \begin{bmatrix} {}_{t_0}x_1(t) \\ 0 \end{bmatrix} + C_{11} e^{\Lambda_s(t-t_1)} C_{11}^{-1} \begin{bmatrix} \delta \\ 0 \end{bmatrix}, \quad t \geq t_1
 \end{aligned}$$

i.e. it will be the original plan plus an "innovation contingent response" which is itself optimal, given the constraint that  ${}_{t_1}p_2(t_1) = {}_{t_0}p_2(t_1)$ . In responding to the disturbance, the policy maker is constrained only to take actions justified by the innovation itself; and clearly the second term in the last expression is an optimal response to such a disturbance from a position of equilibrium.

The "innovation contingent response" for  $p_2$  and  $x_2$  is thus

$$(4.11) \quad \begin{bmatrix} t_1 p_1(t) - t_0 p_1(t) \\ t_1 x_2(t) - t_0 x_2(t) \end{bmatrix} = C_{21} e^{\Lambda_s(t-t_1)} C_{11}^{-1} \begin{bmatrix} \delta \\ 0 \end{bmatrix}$$

and the movement of the controls can be calculated from (1.7)

(4.6) and (4.11).

As was the case for the Stackelberg game, we can construct a time consistent policy if we alter the basic assumptions made so far and assume instead that the policy maker is able to delude the forward looking markets so that  $p_2$  is always set to zero (and the equations describing the evolution of the forward looking markets are "inoperative"). Setting  $p_2 = Dp_2 = 0$  for the adjoint system (4.4) yields the dynamic equations for such a perfect cheating solution namely

$$(4.12) \quad \begin{bmatrix} Dx_1 \\ Dp_1 \\ 0 \end{bmatrix} = \begin{bmatrix} -A_{11} & -J_{11} & A_{12} \\ -Q_{11} & -A_{11}^T & -Q_{12} \\ -Q_{21} & -A_{12}^T & -Q_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ p_1 \\ x_2 \end{bmatrix}$$

which can of course be reduced by substitution for  $x_2$  so as to yield

$$(4.13) \quad \begin{bmatrix} Dx_1 \\ Dp_1 \end{bmatrix} = M_{pc} \begin{bmatrix} x_1 \\ p_1 \end{bmatrix}$$

where  $x_1$  is predetermined and  $p_1$  is given by the transversality requirement for this control problem, i.e.  $p_1 = C_{21} C_{11}^{-1} x_1$  where  $C$

is the column eigenvector associated with the stable roots of  $M_{pc}$ . It is evident that this perfect cheating solution is time consistent, since it involves an open loop path only in terms of  $x_1$ , and it must involve a smaller cost to the policy maker than the time inconsistent solution earlier considered as the policy maker now has one less constraint (the equation for  $Dx_2$  has been dropped).

As the economic example in Annex 1 shows, the policy maker is assumed to have the power to mislead the forward-looking markets by announcing a path for  $u$  (say  $t_0 u^c$ ) which bears no relation to the planned path  $t_0 u$ . It is unlikely that the policy maker would in the circumstances retain credibility; he may indeed lose all power to influence market expectations. In this case the policy maker "loses his leadership" and the situation reverts to one in which the policy makers takes the paths for the  $x_2$  variables as given in minimising his criterion.

To describe the system under this "Nash" policy assumption one simply deletes the row and column for  $p_2$  in equation (4.4). Thus  $p_2$  is eliminated from the system, (but the equation for  $Dx_2$  is retained as the policy maker has lost the power to mislead the market) so

$$(4.14) \quad \begin{bmatrix} Dx_1 \\ Dp_1 \\ Dx_2 \end{bmatrix} = \begin{bmatrix} A_{11} & -J_{11} & A_{12} \\ -Q_{11} & -A_{11}^T & -Q_{12} \\ A_{21} & -J_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ p_1 \\ p_2 \end{bmatrix} = M_{NP} \begin{bmatrix} x_1 \\ p_1 \\ p_2 \end{bmatrix}$$

As only  $x_1$  will be predetermined (with  $p_1$  and  $x_2$  "jumping" so as to remain on the stable manifold of the system with the transition matrix  $M_{NP}$ ) the solution will be time consistent.

These three solutions (the no-cheating time inconsistent optimal policy, "perfect cheating" and "Nash policy") are illustrated with a simple example of policy design in a small open economy under floating exchange rates in Annex 1.

## 5. Decentralised control - an N country Nash differential game with forward looking variables

We have considered how the standard time consistent results of optimal control in Section 1 can be modified to provide time-inconsistent optimal solutions for economic models with a saddle point structure arising from forward-looking variables (whose initial values were not historically predetermined). It is fairly straightforward to make analogous modifications to the results for the N-country open loop Nash game in Section 2 so as to allow for such forward looking variables there too. Even though the policy makers take each other's policy actions as predetermined, the optimal policies chosen will nevertheless be time inconsistent as the policy makers act as leaders vis a vis the financial markets.

The results obtained are then applied to a two country version of the open economy model for which time inconsistent optimal policies have already been derived see Annex 2.

### 5(a) Modifications to the open-loop Nash results of Section 2

It is easiest to begin by considering how Pareto optimal policy is to be determined if the state vector is partitioned into  $x_1$ , variables predetermined at  $t_0$ , and  $x_2$  "jump" variables. For Pareto-optimal policy there will only be one set of costates  $p$ , which are partitioned conformably. The dynamics of the system under centralised control, appearing earlier as equation (2.11), is now shown in

partitioned form at the top of Table A.

Before defining the eigenvectors which characterise the dynamics of the solution it is convenient, as before, to reorder the system so as to put first the variables  $x_1, p_2$  (which are predetermined at  $t_0$  and set equal to zero at  $t_0$  respectively), and this reordered system appears in the middle panel of Table A, where the composition of matrix  $M$  is shown in detail.

The open loop optimal solution formed at  $t_0$  will then simply be

$$(5.1) \quad \begin{bmatrix} x_1(t) \\ p_2(t) \end{bmatrix} = C_{11} e^{\Lambda_s(t-t_0)} C_{11}^{-1} \begin{bmatrix} x_1(0) \\ 0 \end{bmatrix}$$

and

$$(5.2) \quad \begin{bmatrix} p_1 \\ x_2 \end{bmatrix} = C_{21} C_{11}^{-1} \begin{bmatrix} x_1 \\ p_2 \end{bmatrix}$$

where  $\begin{bmatrix} C_{11} \\ C_{21} \end{bmatrix}$  are the column eigenvectors of the stable roots of

$M_{PO}$ . Such a solution is evidently time inconsistent. (Time consistent paths can be constructed, either for the "perfect cheating" case (where the national policy makers all collude to delude the financial markets) or for the "Nash policy" case where the policy makers collude but collectively fail to lead the financial markets; but this is not done here).

We come finally to the "time inconsistent optimal" (but not Pareto optimal) policy which arises when decision-making is decentralised and each country takes the open loop path of the other's policy as predetermined, but nevertheless acts as a leader vis a vis the financial markets. When there are  $N$  countries and  $N-1$  forward looking exchange rates, as in the economic example which follows, this is a game with  $N$  chiefs and  $N-1$  Indians!

Once again it is necessary to partition  $x$  into  $x_1$  and  $x_2$ , and also the shadow prices attached to these state variables by each of the separate policy makers. The coefficients of the partitioned system of this decentralised Nash game appear at the bottom of Table A, (cf. equation 2.5 above).

To obtain the paths for optimal (non cooperative) behaviour under the present assumptions, the system is reordered to

$$(5.3) \quad \begin{bmatrix} Dx_1 \\ Dp_2^1 \\ \vdots \\ Dp_2^N \\ Dp_1^1 \\ \vdots \\ Dp_1^N \\ Dx_2 \end{bmatrix} = M_{DN} \begin{bmatrix} x_1 \\ p_2^1 \\ \vdots \\ p_2^N \\ p_1^1 \\ \vdots \\ p_1^N \\ x_2 \end{bmatrix}$$

and the optimal time inconsistent policy expressed as

$$(5.4) \quad \begin{bmatrix} x_1 \\ t_0 \\ p_2^1 \\ \vdots \\ p_2^N \\ t_0 \end{bmatrix} = C_{11} e^{\Lambda_s(t-t_0)} C_{11}^{-1} \begin{bmatrix} x_1(0) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

on the assumption that all policy makers optimise at time  $t_0$ , where the  $C_{11}$  is the top block of the column eigenvector associated with the stable roots of  $M_{DN}$  in equation (5.3). As for time consistent policy paths, we can neglect "perfect cheating" as the policy makers cannot all succeed given their conflicting interests. If, however, all policy makers were to take the path for the financial asset prices as given, as well as the actions of other policy makers, (what we have called "Nash policy" vis-a-vis the financial markets), then the time consistent optimal policy is obtained by simply omitting the rows and columns relating to the  $p_2^i$  on  $p_2^i$  in panel (iii) of Table A so as to yield

$$(5.5) \quad \begin{bmatrix} Dx_1 \\ Dx_2 \\ p_1^1 \\ \vdots \\ p_1^N \end{bmatrix} = M_{NP} \begin{bmatrix} x_1 \\ x_2 \\ p_1^1 \\ \vdots \\ p_1^N \end{bmatrix}$$



TABLE A : Dynamic adjustment under Pareto-optimal and Decentralised policy

Pareto-optimal Control (in partitioned form)

$$(i) \begin{bmatrix} Dx_1 \\ Dx_2 \\ Dp_1 \\ Dp_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & -\Sigma J_{11}^i & -\Sigma J_{12}^i \\ A_{21} & A_{22} & -\Sigma J_{22}^i & -\Sigma J_{21}^i \\ -\Sigma w_i Q_{11}^i & -\Sigma w_i Q_{12}^i & -A_{12}^T & -A_{11}^T \\ -\Sigma w_i Q_{21}^i & -\Sigma w_i Q_{22}^i & -A_{21}^T & -A_{22}^T \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ p_1 \\ p_2 \end{bmatrix}$$

$$(ii) \begin{bmatrix} Dx_1 \\ Dp_2 \\ Dp_1 \\ Dx_2 \end{bmatrix} = \begin{bmatrix} A_{11} & -\Sigma J_{12}^i & -\Sigma J_{11}^i & A_{12} \\ -\Sigma w_i Q_{21}^i & -A_{22}^T & -A_{21}^T & -\Sigma w_i Q_{22}^i \\ -\Sigma w_i Q_{11}^i & -A_{12}^T & -A_{11}^T & -\Sigma w_i Q_{12}^i \\ A_{21} & -\Sigma J_{22}^i & -\Sigma J_{21}^i & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ p_1 \\ p_2 \end{bmatrix} \equiv M_{PO} \begin{bmatrix} x_1 \\ p_2 \\ p_1 \\ x_2 \end{bmatrix}$$

where  $x$ ,  $p$ ,  $A$ ,  $A^T$ ,  $Q^i$ ,  $J^i$  are conformably partitioned,  
 $\Sigma \equiv \sum_{i=1}^N$ , and the index showing the policy maker appears as a superscript.

Decentralised Control (in partitioned form)

$$(iii) \begin{bmatrix} Dx_1 \\ Dx_2 \\ Dp_1^1 \\ Dp_2^1 \\ \vdots \\ Dp_1^N \\ Dp_2^N \end{bmatrix} = \begin{bmatrix} A_{11} & -A_{12} & -J_{11} & \dots & -J_{11}^N & -J_{12}^N \\ A_{21} & -A_{22} & -J_{21} & & -J_{21}^N & -J_{22}^N \\ -Q_{11}^1 & -Q_{12}^1 & -A_{11}^T & & 0 & 0 \\ -Q_{21}^1 & -Q_{22}^1 & -A_{21}^T & & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -Q_{11}^N & -Q_{12}^N & 0 & \dots & -A_{11}^T & -A_{12}^T \\ -Q_{21}^N & -Q_{22}^N & 0 & & -A_{21}^T & -A_{22}^T \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ p_1^1 \\ p_1^1 \\ \vdots \\ p_1^N \\ p_2^N \end{bmatrix}$$

so the behaviour of  $x_1$  under such conditions is simply

$$(5.6) \quad x_1(t) = C_{11} e^{\Lambda_s(t-t_0)} C_{11}^{-1} x_1(0)$$

which is time consistent.

These various solutions are illustrated with an economic  
 \*\*  
 example in Annex 2.

\* Not included in this copy: available on request

\* not included: available on request

## Conclusions

The presence of forward-looking, rational-expectations behaviour means that cost-minimising macroeconomic policy is time inconsistent in the small open economy and also for both centralised and decentralised policy in an interdependent world. In such circumstances, some have argued for "pre-committment" to an open loop policy with no feedback; but it is easy to show that an open loop path for policy can with advantage be complemented by innovation contingent feedback, and the nature of this feedback was investigated.

Such innovation-contingent responses are designed to avoid the temptation to renege on past commitments; but given departures from the open loop path, there are undeniably temptations to "cheat". How great this temptation is depends very much on the parameters. In our specific example, zero weight on the current balance meant that the exchange rate would, under a policy of "perfect cheating", be treated as a costless control variable and moved wherever and whenever necessary to attain the output target; but with the current balance being weighted half as much as output, the perfect cheating policy hardly differed from time inconsistent optimal policy!

Given the temptation to cheat, it might be argued that there should be no policy variation, with instruments set so as to achieve targets in the long run but not depending in any way upon the current state of the system. This policy certainly avoids the temptation to cheat, but it appears to be a socially inefficient way of doing so. It was found in our example that if the policy-makers

were to adopt a time consistent approach, by simply eschewing their Stackelberg leadership position vis a vis the foreign exchange markets, this provided dynamics fairly similar to those available from the optimal policy, while both differed substantially from the dynamics with no policy variation.

Our example of a small open economy implied that it would be a great mistake for an individual policy maker to forego stabilisation policy "because optimal policy is time inconsistent in an open economy with floating rates". If the "first best" policy of an announced open-loop plus innovation contingent feedback (cf. Buiter (1980)) is rejected because it is time inconsistent, "second best" feedback strategies which approximate such a policy may nevertheless be available and should be considered.

On the assumption that the system was "controllable", the many country case exhibited the inefficiency of mutual Nash behaviour in forming policy, together with the time-inconsistency of leadership vis a vis private markets. The conditions under which such systems are indeed controllable are, in our view, one topic which merits further attention, as optimal decentralised behaviour is in itself no guarantee of controllability.

The potential gains from "co-ordination" were examined (with given country weights) using the maximum principle (where the value of the minimised criterion is easily represented analytically and calculated numerically); but an alternative approach to the design of international policy would be to elevate one of the countries

to the role of being a policy leader. Indeed some economists, C.P. Kindelberger in particular, have argued that the problems experienced by the international monetary system (both before the Second World War and after the Vietnam War) are the result of the failure of one country so to act. To analyse a two-level Stackelberg game (with the policy leader dominating other policy makers, and both affecting private expectations) would be a second useful extension of the framework proposed here, which can be handled by an elaboration of the same techniques.

A third extension to the techniques employed here would be to apply closed-loop Nash and Stackelberg games, rather than the open-loop forms used above, see Basar and Olsder (1982).

Even without these three extensions, we would hope that the above account has demonstrated both the usefulness of dynamic games as a vehicle for analysing important issues of policy formation in open economies, and the elegance of the maximum principle in computing time inconsistent optimal policy in the presence of forward looking behaviour.

## References:

- Aoki, M. (1981) Dynamic Analysis of Open Economies, London, Academic Press.
- Austin, G.P. and W.H.Buiter (1982) "Saddlepoint : a programme for solving continuous time linear rational expectations models", LSE Econometrics Programme, Discussion Paper A37 Department of Economics, London School of Economics.
- Basar, T. and G.T.Olsder (1980) "Team-optimal Closed-loop Stackelberg Strategies in Hierarchical Control Problems", Automatica, Vol.16, No.4, pp.
- (1982) Dynamic Non-Cooperative Game Theory, London, Academic Press.
- Bellman, R. (1957) Dynamic Programming, Princeton, N.J., Princeton University Press.
- Bergsten, C.F. (1982) "What to do about the US-Japan Economic Conflict", Foreign Affairs, Summer, pp.1059-1)75.
- Blanchard, O. (1981) "Output, the Stock Market and Interest Rates", American Economic Review, Vol.71, No.1, pp. 132-143.
- Buiter, W.H. (1980) "The Superiority of Contingent Rules over Fixed Rules in Models with Rational Expectations", Economic Journal, Vol. 91, pp.617-646.
- (1983) "Expectations and Control Theory", Economie Appliquée (forthcoming).
- Buiter W.H. and M.H.Miller (1981) "Monetary Policy and International Competitiveness", Oxford Economic Papers, Vol. 33, pp.143-175, Supplement.
- Calvo G.A. (1978) "On the Time Consistency of Optimal Policy in a Monetary Economy", Econometrica, Vol.46, No.6, pp.1411-1428.
- Cooper, Richard N. (forthcoming), "Economic Interdependence and Co-ordination of Economic Policies" in R.Jones and P.Kenen (eds.) Handbook on International Economics, Amsterdam, North Holland.
- Dornbusch, R. (1976) "Exchange Rate Dynamics", Journal of Political Economy, Vol. 84, December, pp.1161-1176.
- Driffill, J. (1982) "Optimal Money and Exchange Rate Policies", mimeo, Department of Economics University of Southampton.
- Hamada, K. (1974) "Alternative Exchange Rate Systems and the Interdependence of Monetary Policies" in National Monetary Policies and the International Financial System. R.Z.Aliber (ed.) London, University of Chicago Press.
- Hamalainen, R. (1981) "On the cheating problem in Stackelberg games", International Journal of Systems Science, Vol.12, No.6, pp.753-770.
- Hayashi, F. (1982) "Tobin's Marginal  $q$  and Average  $q$  : A Neoclassical Interpretation", Econometrica, Vol.50, No.1, pp.213-224, January.
- Intriligator, M.D. (1971) Mathematical Optimization and Economic Theory,

- Kindleberger, C.P. (1976) "Systems of International Organization" in D.P.Calleo (ed.), Money and the Coming World Order, New York, NY University Press.
- Kydland F.E. and E.C.Prescott (1977) "Rules Rather than Discretion : The Inconsistency of Optimal Plans", Journal of Political Economy, Vol.85, No.3.
- Livesey, D.A. (1980) "Stabilization Policy : A View from the Complex Plane", CARESS Working Paper No.80-9, pp.34, University of Pennsylvania.
- Miller, M.H. (1983) "Policy Co-ordination and Floating Exchange Rates", Memorandum submitted to the Treasury and Civil Service Committee, International Monetary Arrangements, Vol.III, pp.213-217, London, HMSO.
- Miller, M.H. and M.H.Salmon (1982) "Optimal Control and Economic Dynamics : An investigation of co-ordinated and decentralised policy formation in a multi-country setting with flexible exchange rates", mimeo, Department of Economics, University of Warwick.
- Mundell R.A. (1963) International Economics, N.Y. The Macmillan Company.
- Pontryagin L.S. et al (1962) The Mathematical Theory of Optimal Processes, translated by K.N.Tvirogoff, N.Y., International Publishers.
- Simaan, M. and J.B.Cruz (1973) "On the Stackelberg Strategy in Non zero-sum Games", Journal of Optimization Theory and Applications, Vol.11, No.5, pp.533-555.
- Summers, L.H. (1981) "Taxation and Corporate Investment : A q-Theory Approach", Brookings Papers on Economic Activity, No.1, pp.6--127.
- Treasury and Civil Service Committee (1982/3), International Monetary Arrangements Fourth Report of the 1982/83 Session (2 Reports, with Memoranda, Minutes of Evidence and Appendices, London, HMSO.
- Wiberg, D.M. (1971) State Space and Linear System. Schaum's Outline Series. N.Y., McGraw-Hill.