OPTIMAL CONTROL OF ECONOMETRIC MODELS

WITH AUTOCORRELATED DISTURBANCE TERMS*

Adrian Pagan

Econometric Research Program
PRINCETON UNIVERSITY
207 Dickinson Hall
Princeton, New Jersey

# OPTIMAL CONTROL OF ECONOMETRIC MODELS
## WITH AUTOCORRELATED DISTURBANCE TERMS*

Adrian Pagan

## 1. Introduction

Recently a spate of articles have appeared concentrating upon the application of control theory to econometric models and economic theory, e.g., [7], [13]. Perhaps the best known discrete formulation for a stochastic system is that proposed in a series of articles by Chow [4], [5], [6], [7], which concentrates upon the linear econometric model with reduced form

$$\underline{y}_t = \underline{A}\ \underline{y}_{t-1} + \underline{C}\ \underline{x}_t + \underline{\epsilon}_t \qquad (1.1)$$

where $\underline{y}_t$ is a $(p \times 1)$ vector, $\underline{x}_t$ a $(r \times 1)$ vector of control variables, $\underline{A}$ and $\underline{C}$ are non-stochastic matrices of

parameters and $E(\underline{\epsilon}_t) = 0$ , $E(\epsilon_t \epsilon_t') = V$ , $E(\underline{\epsilon}_t \underline{\epsilon}_s') = 0$ , $s \neq t$ . Given a quadratic welfare function whose expectation is to be maximized over T periods

$$W = E \sum_{t=0}^{T} \underline{Y}_t' \underline{K}_t \underline{Y}_t \qquad (1.2)$$

with $\underline{K}_t$ a symmetric, positive semi-definite matrix of weights reflecting the relative priorities of variables in the decision maker's welfare function, Chow shows that the optimal control rule obtained by maximizing (1.2) with respect to (1.1) is $\hat{\underline{x}}_t = \underline{G}_t E_{t-1}(\underline{Y}_{t-1})$ where $\underline{G}_t$ can be derived as the solution to the following set of equations:

$$\underline{G}_t = -(\underline{C}'\underline{H}_t \underline{C})^{-1} \underline{C}'\underline{H}_t \underline{A} \qquad (1.3a)$$

$$\underline{H}_{t-1} = \underline{K}_{t-1} + (\underline{A} + \underline{C}\,\underline{G}_t)' \underline{H}_t (\underline{A} + \underline{C}\,\underline{G}_t) \qquad (1.3b)$$

$$\underline{H}_T = \underline{K}_T \qquad (1.3c)$$

which come from a dynamic programming solution that involves the maximization of $\bar{W}_t = E_{t-1}(\underline{Y}'_t \underline{H}_t \underline{Y}_t)$ with respect to $\underline{x}_t$ for all $t = 1,\ldots,T$, where $E_{t-1}$ denotes that the

expectation is conditional on information prevailing in
period  t-1.[1,2]

A difficulty with  Chow's approach is the lack of
an explicit treatment of autocorrelation in the disturbance
term.  He suggests that the structural equations may be
adjusted with the Cochrane-Orcutt transformation if the dis-
turbances follow an autoregressive  (A.R.) format but this
will increase  p , i.e., the dimension of  $\underline{y}_t$ ,  with the
increase being particularly large if the fourth order pro-
cesses found by Wallis [16] to be present in a number of
quarterly models are encountered.  Furthermore, no such
simple solution is available if there are moving average
(M.A.)  errors entering the model; cases of this may be
found in [15] and [17].  Given the potential for discovery
of either form it seems apposite to extend Chow's framework
to encompass both disturbance formats with the stricture
that the dimension of the state variable  $\underline{y}_t$  should be no
larger than  p .

---

[1]Equations (1.3a)-(1.3c) may be derived from any of Chow's
articles by noting that there is no constant term in the
reduced form (1.1) and that there are no target levels
entering the welfare function, e.g., (1.3a) and (1.3b) cor-
respond to (9) and (14) of [7] with  $b_t = a_t = 0$ .
Section 5 extends the algorithm of this paper to the case
when  $b_t$ , $a_t \neq 0$ .

[2]Although  $E_{t-1}(\underline{y}_{t-1}) = \underline{y}_{t-1}$  in Chow's papers this is not
so in section 3.

Rather than present separate derivations for each type of disturbance it is useful to work with a representative that incorporates both, and this is defined in section 2. The general theory is then outlined in section 3 and given specific content in section 4. Section 5 modifies the resulting formulae to permit both a time-varying constant term to appear in the equations and targets for each of the variables, while section 6 discusses some complications arising in the implementation of the algorithm. Finally, section 7 provides an example based upon Klein's Model I of the U.S. economy.

## 2.  The Autocorrelation Representation

Attention will be focused upon the $q^{th}$ order AR.

$$u_t = \Phi_1 u_{t-1} + \ldots + \Phi_q u_{t-q} + e_t \qquad (2.1)$$

and MA

$$u_t = e_t + \Phi_1 e_{t-1} + \ldots + \Phi_q e_{t-q} \qquad (2.2)$$

processes, where $\Phi_1, \ldots, \Phi_q$ are non-stochastic matrices of parameters and $e_t$ has the properties $E(e_t) = 0$ , $E(e_t e_t') = \Omega$ , $E(e_t e_s') = 0$ for $t \neq s$ . Both (2.1) and (2.2)

may be written in the form

$$u_t = \Phi v_{t-1} + e_t \qquad (2.3)$$

$$v_t = B v_{t-1} + \xi_t \qquad (2.4)$$

by use of the following definitions

$$\Phi = [\Phi_1 \vdots \dots \vdots \Phi_q] \qquad (AR \text{ and } MA) \qquad (2.5)$$

$$v_t' = [u_t' \vdots \dots \vdots u_{t-q+1}'] \qquad (AR) \qquad (2.6)$$

$$v_t' = [e_t' \vdots \dots \vdots e_{t-q+1}'] \qquad (MA) \qquad (2.7)$$

$$B = \begin{bmatrix} \Phi_1 & \cdot & \cdot & \cdot & \Phi_q \\ 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & \cdot & \cdot & \cdot & 0 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ 0 & 0 & \cdot & \cdot & 1 & 0 \end{bmatrix} \qquad (AR) \qquad (2.8)$$

$$B = \begin{bmatrix} 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & \cdot & \cdot & \cdot & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdot & \cdot & 1 & 0 \end{bmatrix} \qquad (MA) \qquad (2.9)$$

$$\xi_t' = [e_t'\ 0\ \dots\ 0] \quad (\text{AR and MA}) \ . \qquad (2.10)$$

The structure of B assumes the crucial role, in section 4, of differentiating the control equations for each disturbance type. As it can also be shown that, if the errors were to follow an autoregressive-moving average structure [2], (2.3) and (2.4) would be appropriate, there is scope for a wider class of disturbances than considered explicitly in this paper.

3. General Theory

The reduced form is assumed to be

$$y_t = A\ y_{t-1} + C\ x_t + u_t \qquad (3.1)$$

where $u_t$ is given in (2.3) and (2.4). If $Y_t' = [y_t'\ v_t']$ the relations in (2.3), (2.4) and (3.1) may be summarized in the system

$$\begin{bmatrix} y_t \\ v_t \end{bmatrix} = \begin{bmatrix} A & \Phi \\ 0 & B \end{bmatrix} \begin{bmatrix} y_{t-1} \\ v_{t-1} \end{bmatrix} + \begin{bmatrix} C \\ 0 \end{bmatrix} x_t + \begin{bmatrix} e_t \\ \xi_t \end{bmatrix} \qquad (3.2)$$

or

$$\underline{Y}_t = \underline{A}\ \underline{Y}_{t-1} + \underline{C}\ \underline{x}_t + \underline{e}_t \ . \qquad (3.3)$$

As (3.3) and (1.1) are identical, the optimal controls may be found from (1.3a) - (1.3c). Simplification of these equations begins by partitioning $\underline{K}_t$ and $\underline{H}_t$ conformably with $\underline{y}_t$ viz.

$$\underline{K}_t = \begin{bmatrix} K_t & 0 \\ 0 & 0 \end{bmatrix} \qquad \underline{H}_t = \begin{bmatrix} \bar{H}_t & H_t^* \\ H_t^{*\,\prime} & H_t^{**} \end{bmatrix} \quad ,$$

using the knowledge that $\underline{H}_t$ is symmetric (this following from (1.3b) and (1.3c)). Then the equivalent of (1.3a) is

$$\underline{G}_t = - \left[ (C'\,0) \begin{pmatrix} H_t & H_t^* \\ H_t^{*\,\prime} & H_t^{**} \end{pmatrix} \begin{pmatrix} C \\ 0 \end{pmatrix} \right]^{-1} \left[ (C'\ 0) \begin{pmatrix} H_t & H_t^* \\ H_t^{*\,\prime} & H_t^{**} \end{pmatrix} \begin{pmatrix} A & \Phi \\ 0 & B \end{pmatrix} \right]$$

$$= - (C'H_t C)^{-1} [C'H_t A \vdots C'\ H_t\ \Phi + C'H_t^* B] \qquad (3.4)$$

$$= [G_t \vdots G_t^*] \quad . \qquad (3.5)$$

As evidenced by (3.4) it is unnecessary to evaluate all elements of $\underline{H}_t$ in order to compute the control rule contained in $\underline{G}_t$ , i.e. only $H_t$ and $H_t^*$ are required. As these both appear in the first "row" of the partitioned matrix $\underline{H}_t$ , they are obtained by multiplying the first "row" of

$(\underline{A} + \underline{C}\,\underline{G}_t)'\underline{H}_t$ by the matrix $(\underline{A} + \underline{C}\,\underline{G}_t)$ . As $(\underline{A} + \underline{C}\,\underline{G}_t)$ is

$$
\begin{bmatrix}
A + C\,G_t & \Phi + C\,G_t^* \\
\\
O & B
\end{bmatrix}
, \qquad (3.6)
$$

the first "row" of the product $(\underline{A} + \underline{C}\,\underline{G}_t)'\underline{H}_t$ becomes $[D'H_t \vdots D'H_t^*]$ where $D = A + C\,G_t$ . Performing the multiplication with the matrix $(\underline{A} + \underline{C}\,\underline{G}_t)$ yields

$$
[D'H_t(A + C\,G_t) \vdots D'H_t(\Phi + C\,G_t^*) + D'H_t^*\,B]
$$

or, in the recursive format of (1.3b),

$$
H_{t-1} = K_{t-1} + (A + C\,G_t)'\,H_t(A + C\,G_t) \qquad (3.7)
$$

$$
H_{t-1}^* = (A + C\,G_t)'\,[H_t(\Phi + C\,G_t^*) + H_t^*\,B] \qquad (3.8)
$$

$$
H_T = K_T , \; H_T^* = O \qquad . \qquad (3.9)
$$

The expressions for $G_t$ from (3.4) and $H_t$ in (3.7) reveal that the optimal controller $\hat{x}_t = G_t\,y_{t-1} + G_t^*\,E_{t-1}(v_{t-1})$ is composed of two terms; one accounting for the reduced form dynamics $(G_t y_{t-1})$ and one $(G_t^* E_{t-1} v_{t-1})$ for the disturbance

dynamics. Given such a division it is not surprising to find that the former is identical to the control rule found in Chow's articles.

## 4. Special Theory

To illustrate the methodology only the $q^{th}$ order AR is treated in detail. Using (2.5), (2.6) and (2.8) with a partitioning of $H_t^*$ as $[H_{1,t} \vdots \ldots \vdots H_{q,t}]$ allows $G_t^*$ in (3.4) to be written as

$$- (C'H_t C)^{-1} \left( C'H_t [\Phi_1 \vdots \ldots \vdots \Phi_q] + [C'H_{1,t} \vdots \ldots \vdots C'H_{q,t}] \begin{bmatrix} \Phi_1 \cdots \Phi_q \\ 1 \, 0 \ldots 0 \\ 0 \, 1 \ldots 0 \\ \vdots \qquad \vdots \\ 0 \, 0 .. 1 \, 0 \end{bmatrix} \right)$$

$$(4.1)$$

which, if $G_t^* = [G_{1,t} \vdots \ldots \vdots G_{q,t}]$, simplifies to

$$G_{j,t} = - (C'H_t C)^{-1}(C'H_t \Phi_j + C'H_{1,t} \Phi_j + C'H_{j+1,t}) \quad j=1,\ldots,q-1$$

$$(4.2)$$

$$G_{q,t} = - (C'H_t C)^{-1} (C'H_t \Phi_q + C'H_{1,t} \Phi_q) \quad . \quad (4.3)$$

The recursive equation (3.8) may be decomposed by substituting the following relations

$$(\Phi + C\,G_t^{*}) = (\Phi_1 \vdots \ldots \vdots \Phi_q) + (C\,G_{1,t} \vdots \ldots \vdots C\,G_{q,t})$$

$$= (\Phi_1 + C\,G_{1,t} \vdots \ldots \vdots \Phi_q + C\,G_{q,t})$$

$$H_t^{*}\,B = [H_{1,t} \vdots \ldots \vdots H_{q,t}] \begin{bmatrix} \Phi_1 & \ldots \Phi_q \\ 1 & 0 \ldots 0 \\ 0 & 1 \ldots 0 \\ 0 & 0 \ldots 1\ 0 \end{bmatrix}$$

$$= (H_{1,t}\Phi_1 + H_{2,t} \vdots H_{1,t}\Phi_2 + H_{3,t} \vdots \ldots \vdots H_{1,t}\Phi_q)$$

to yield

$$H_{j,t-1} = (A + C\,G_t)' \,[H_t(\Phi_j + C\,G_{j,t}) + H_{1,t}\Phi_j + H_{j+1,t}] \quad j=1,\ldots,q-1$$

$$(4.4)$$

$$H_{q,t-1} = (A + C\,G_t)' \,[H_t(\Phi_q + C\,G_{q,t}) + H_{1,t}\Phi_q] \qquad (4.5)$$

$$H_{j,t} = 0 \qquad\qquad j=1,\ldots,q \quad . \qquad\qquad (4.6)$$

Gathering the results from (3.4), (3.7), (3.9), (4.2), (4.3), (4.4), (4.5) and (4.6) leaves the optimal control equations for the $q^{th}$ order AR as

$$\hat{x}_t = G_t y_{t-1} + \sum_{j=1}^{q} G_{j,t}\, E_{t-1}(u_{t-j}) \qquad (4.7a)$$

$$G_t = -(C'H_t C)^{-1} C'H_t A \tag{4.7b}$$

$$G_{j,t} = -(C'H_t C)^{-1}(C'H_t \Phi_j + C'H_{1,t}\Phi_j + C'H_{j+1,t}) \quad j=1,\ldots,(q-1) \tag{4.7c}$$

$$G_{q,t} = -(C'H_t C)^{-1}(C'H_t \Phi_q + C'H_{1,t}\Phi_q) \tag{4.7d}$$

$$H_{t-1} = K_{t-1} + (A + C\,G_t)'\,H_t(A + C\,G_t) \tag{4.7e}$$

$$H_{j,t-1} = (A + C\,G_t)'[H_t(\Phi_j + C\,G_{j,t}) + H_{1,t}\Phi_j + H_{j+1,t}] \quad j=1,\ldots,(q-1) \tag{4.7f}$$

$$H_{q,t-1} = (A + C\,G_t)'\,[H_t(\Phi_q + C\,G_{q,t}) + H_{1,t}\Phi_q) \tag{4.7g}$$

$$H_T = K_T, \quad H_{j,T} = 0 \qquad j=1,\ldots,q \quad . \tag{4.7h}$$

The equations corresponding to $(4.7)$ for the MA case are identical except that $H_{1,t}$ is set equal to zero throughout, e.g., $(4.7c)$ would become

$$G_{j,t} = -(C'H_t C)^{-1}(C'H_t \Phi_j + C'H_{j+1,t}) \quad j=1,\ldots,(q-1) \quad .$$

Furthermore, in contrast to the AR case where $E_{t-1}(v_{t-1}) = v_{t-1}$, i.e., the values $u_{t-1},\ldots,u_{t-q}$ contained in $v_{t-1}$ are completely known at $t-1$, the components of

$E_{t-1}(v_{t-1})$ for an MA must be generated. To do this, set $z_t = v_{t-1}$ and observe that (2.3) and (2.4) constitute a state space representation from which $\hat{v}_{t-1} = E_{t-1}(v_{t-1}) = E_{t-1}(z_t) = \hat{z}_t$, i.e., the one-step predictor of $z_t$, must be extracted. From the theory underlying the Kalman-Bucy filter [ 10 ] this is given by $\hat{z}_t = B\hat{z}_{t-1} + K(u_{t-1} - \Phi\hat{z}_{t-1})$, where $K' = [1\ 0\ ...0]$, which results in a recursive relationship for the computation of $E_{t-1}(v_{t-1})$ .

Three comments may be made about (4.7) and its MA version. Firstly, the absence of $H_{1,t}$ for the MA case stems from the structure of B, i.e., the first row is zero for a MA but has $\Phi_j$ (j=1,...,q) as $j^{th}$ element for an AR. Secondly, there is a simple structure to (4.7) that allows for easy computer programming and enables the same program to be utilized for either disturbance format -- a symmetry that has been commented upon and exploited else-where in estimation [11]. Thirdly, the order of the matrices in (4.7) is never higher than that characteristic of the model if it had no autocorrelation; an original objective fulfilled.

## 5. Generalization of the Algorithm

It is now time to generalize the control problem to allow for specific targets in the welfare function and a constant term in the reduced form. There is little difficulty in so doing. In the expanded problem it is necessary to minimize

$$W = E[\sum_{t=0}^{T} (\underline{y}_t - \underline{a}_t)' \underline{K}_t (\underline{y}_t - \underline{a}_t)] \tag{5.1}$$

subject to $\underline{y}_t = \underline{A}\,\underline{y}_{t-1} + \underline{C}\,\underline{x}_t + \underline{b}_t + \underline{\epsilon}_t$ .

When $\underline{y}_t$ is the augmented vector of section 3 the vectors $\underline{a}_t$ and $\underline{b}_t$ will be $\underline{a}_t' = [a_t'\ 0]$, $\underline{b}_t' = [b_t'\ 0]$ where the correspondence is obvious. The optimal controller for (5.1) is now $\hat{\underline{x}}_t = \underline{G}_t\underline{y}_t + \underline{g}_t$ (Chow [7, eq. (8)]) where

$$\underline{g}_t = -(\underline{C}'\underline{H}_t\underline{C})^{-1}\underline{C}'(\underline{H}_t\underline{b}_t - \underline{h}_t) \tag{5.2}$$

$$\underline{h}_{t-1} = \underline{k}_{t-1} + (\underline{A} + \underline{C}\,\underline{G}_t)'(\underline{h}_t - \underline{H}_t\underline{b}_t) \tag{5.3}$$

$$\underline{k}_t = \underline{K}_t\,\underline{a}_t \tag{5.4}$$

$$\underline{h}_T = \underline{k}_T , \tag{5.5}$$

and all other matrices are as in (1.3a)-(1.3c) (eq. (5.3) is

(15) of [ 7 ]).  Substituting $\underline{h}_t' = [h_t'\ h_t^{*'}]$, $\underline{g}_t' = [g_t'\ g_t^{*'}]$ and the partitioned form of $\underline{H}_t$ into (5.2) gives

$$\begin{pmatrix} g_t \\ g_t^* \end{pmatrix} = - (C'H_t C)^{-1}(C'\ 0)\left[\begin{pmatrix} H_t & H_t^* \\ H_t^{*'} & H_t^{**} \end{pmatrix}\begin{pmatrix} b_t \\ 0 \end{pmatrix} - \begin{pmatrix} h_t \\ h_t^* \end{pmatrix}\right]$$

$$= - (C'H_t C)^{-1}\left[(C'0)\begin{pmatrix} H_t b_t & - h_t \\ H_t^{*'} b_t & - h_t^* \end{pmatrix}\right]$$

$$\therefore \quad g_t = - (C'H_t C)^{-1} C'(H_t b_t - h_t) \quad . \qquad (5.6)$$

Similarly, for (5.3),

$$\begin{bmatrix} h_{t-1} \\ h_{t-1}^* \end{bmatrix} = \begin{bmatrix} k_{t-1} \\ 0 \end{bmatrix} + \begin{bmatrix} (A+CG_t)' & 0 \\ (\Phi+CG_t^*)' & B' \end{bmatrix}\left[\begin{pmatrix} h_t \\ h_t^* \end{pmatrix} - \begin{pmatrix} H_t b_t \\ H_t^* b_t \end{pmatrix}\right].$$

As (5.6) reveals that only $h_t$ is required for the control computations, the appropriate recursion is

$$h_{t-1} = k_{t-1} + (A+C\ G_t)'\ (h_t - H_t b_t) \quad . \qquad (5.7)$$

(5.6) and (5.7) are identical to the adjustment that would be necessary if there were a constant term and/or targets but no autocorrelation. The influence of these equations on the controller is expressed by modifying (4.7a) to

$$\hat{x}_t = G_t y_{t-1} + \sum_{j=1}^{q} G_{j,t} E_{t-1}(u_{t-j}) + g_t \qquad (4.7a)'.$$

This completes the development of the optimal control law for the linear-quadratic problem with autocorrelation, and the benefits from viewing it in this framework may be enumerated.

  (i)   The dimensionality of the problem has been retained at $p$ , i.e., it corresponds to control with no autocorrelation.

  (ii)  By a suitable partitioning of the $\underline{H}_t$ matrix, it has emerged that not all elements of this matrix need be computed to derive the requisite law:  a result that has obvious benefits in terms of the reduced number of multiplications that must be performed.

  (iii) Because we have merely manipulated Chow's equations it is possible to extend the current results to other situations that he has examined. In particular, if $A, C$ and $\Phi$ are stochastic, one can employ the equations in [ 8 ] in an exactly analogous manner to the above to obtain the appropriate simplifications.

An Appendix considers the derivation of the expected welfare loss associated with the optimal control rule.

## 6. Implementation of the Algorithm

Under this heading we consider the conversion of a structural equation model (see [ 9 , p. 297] for nomenclature)

$$\Gamma\, y(t) + B\ x(t) + U(t) = 0 \tag{6.1}$$

to a reduced form $y(t) = \Pi\, x(t) + V(t)$ with an error format either of the type

$$V(t) = \Phi_1\, V(t-1) + \ldots + \Phi_q\, V(t-q) + E(t) \qquad (AR) \tag{6.2}$$

or

$$V(t) = E(t) + \Phi_1\, E(t-1) + \ldots + \Phi_q\, E(t-q) \qquad (MA) \tag{6.3}$$

where $E(t)$ has zero mean, finite contemporaneous covariance matrix $\Omega$ and no serial correlation. Once one of (6.2) or (6.3) is arrived at control can proceed with the algorithm described earlier.

(i) AR Case

Assume there is a $q^{th}$ order AR in the structure i.e.,

$$U(t) = A_1\, U(t-1) + \ldots + A_q\, U(t-q) + \tilde{E}(t) \ . \tag{6.4}$$

Pre-multiplying (6.4) by $-\Gamma^{-1}$ gives

$$- \Gamma^{-1} U(t) = - \Gamma^{-1} A_1 U(t-1) - \ldots - \Gamma^{-1} A_q U(t-q) - \Gamma^{-1} \tilde{E}(t) \qquad (6.5)$$

i.e.,

$$V(t) = \Phi_1 V(t-1) + \ldots + \Phi_q V(t-q) + E(t) \qquad (6.6)$$

where $V(t) = -\Gamma^{-1} U(t)$, $\Phi_j(t) = \Gamma^{-1} A_j \Gamma$ and $E(t) = -\Gamma^{-1}\tilde{E}(t)$. Now (6.6) has the requisite structure.

### (ii) MA Case

Assume there is a $q^{th}$ order MA in the structure, i.e.,

$$U(t) = \tilde{E}(t) + A_1 \tilde{E}(t-1) + \ldots + A_q \tilde{E}(t-q) \qquad . \qquad (6.7)$$

Pre-multiplying (6.7) by $-\Gamma^{-1}$ gives

$$- \Gamma^{-1} U(t) = -\Gamma^{-1} \tilde{E}(t) - \Gamma^{-1} A_1\tilde{E}(t-1) - \ldots - \Gamma^{-1}A_q \tilde{E}(t-q) \qquad (6.8)$$

or

$$V(t) = E(t) + \Phi_1 E(t-1) + \ldots + \Phi_q E(t-q) \qquad (6.9)$$

where $V(t) = -\Gamma^{-1} U(t)$, $\Phi_j = \Gamma^{-1} A_j \Gamma$ and $E(t) = -\Gamma^{-1} \tilde{E}(t)$ . Now (6.9) has the required format.

## 7. An Example

The chosen example involves the anti-depression experiments performed with Klein's model I by Theil [14]. This model has the following eight equations (where the nomenclature and definitions may be found in [9, p. 303]):

$$C_t = b_1 P_t + b_2 W_t + b_3 P_{t-1} + b_4$$

$$I_t = b_5 P_t + b_6 P_{t-1} + b_7 K_{t-1} + b_8$$

$$W_t^* = b_9 E_t + b_{10} E_{t-1} + b_{11} A_t + b_{12}$$

$$P_t = Y_t - W_t$$

$$E_t = Y_t + T_t - W_t^{**}$$

$$Y_t = C_t + I_t - T_t + G_t$$

$$K_t = K_{t-1} + I_t$$

$$W_t = W_t^* + W_t^{**} \quad .$$

Theil considers optimal policies for the years 1933-1936 with the strategy being, generally speaking, to restore the economy to the 1929 activity levels by 1936. Expressed

in numerical terms this implies the ideal values given in
Table 1 (see [ 14,        p. 79]), except that the desired
level of wages for 1936 was computed by applying the 1929
ratio of wages to income to the target income level for
1936 -- with the latter coming from substitution of the values
of Table 1 into  $Y_t = C_t + I_t - T_t + G_t$ -- and then, following
Theil, linearly interpolating values for 1933-5.  The division
of Table 1 is into the targets of consumption  (C) , invest-
ment  (I)  and total wage bill  (W)  and the instruments,
taxes  (T),  government expenditure  (G)  and the government
wage bill  (W**).

## TABLE 1

### Desired Values for Target Variables and Instruments

|      | 1933   | 1934   | 1935   | 1936   |
|------|--------|--------|--------|--------|
| C    | 49.69  | 53.78  | 57.88  | 61.97  |
| I    | -3.10  | 0.00   | 3.10   | 6.20   |
| W    | 37.86  | 41.42  | 44.98  | 48.53  |
| W**  | 5.038  | 5.254  | 5.469  | 5.685  |
| T    | 7.396  | 7.635  | 7.874  | 8.113  |
| G    | 10.438 | 10.868 | 11.298 | 11.727 |

Full Information Maximum likelihood estimates of
the parameters of Klein's Model I under the assumptions of
no autocorrelation and a first order moving average specifi-
cation are in Table 2.  An algorithm, described in [3 ],
was used to generate estimates for the  MA  specification,
under the assumption that pre-period values of the distur-
bances were zero.[3]

With a welfare function giving equal weight to each
of the variables in Table 1 and zero to all others, the
four-period feedback control matrices, when there is first
order M.A. autocorrelation, are in Table 3.  Table 4 contains
the expected values of the target and control variables for
1933-6 assuming that the optimal control rule is in use
(computed from equation (9) of the Appendix).  A comparison
of Table 4 with the desired values of Table 1 reveals two
outstanding features:

  (a)  a consistent failure to attain the wage target.

  (b)  the strange pattern of investment behaviour,
       probably induced by the negative effect of
       current profits but positive effect of lagged
       profits.

These points suggest that the welfare function would
need to be modified to give more weight to wages and

---

[3]My thanks to Ray Byron for providing his program.
Extensive tampering with it on my part makes it even more
urgent that the usual caveat on responsibility be noted.

## TABLE 2

### F.I.M.L. Estimates of Klein's Model I

| | | |
|---|---|---|
| $b_1$ | -0.232 | -0.019 |
| $b_2$ | 0.802 | 0.779 |
| $b_3$ | 0.386 | 0.267 |
| $b_4$ | 18.34 | 17.341 |
| $b_5$ | -0.801 | -0.514 |
| $b_6$ | 1.052 | 0.996 |
| $b_7$ | -0.148 | -0.182 |
| $b_8$ | 27.26 | 30.162 |
| $b_9$ | 0.234 | 0.294 |
| $b_{10}$ | 0.285 | 0.254 |
| $b_{11}$ | 0.235 | 0.201 |
| $b_{12}$ | 5.79 | 3.979 |
| MA(1) | | 0.224 |
| MA(2) | | 0.104 |
| MA(3) | | 0.132 |
| Likelihood | 1.782 | 0.537 |

investment if the control exercise was a serious one. However, as Klein's Model I is generally treated as a mere workhorse for experimentation, such a recomendation should not be taken too seriously.

To provide some index of the gains from recognizing autocorrelation patterns in the controller the expected welfare loss was computed twice; once with $G_t^*$ set equal to zero and

## TABLE 3

### Optimal Control Matrices

| | $T_t$ | $W_t^{**}$ | $G_t$ | $T_t$ | $W_t^{**}$ | $G_t$ |
|---|---|---|---|---|---|---|
| | | 1933 | | | 1934 | |
| $P_{t-1}$ | 0.0126 | -0.1447 | -0.4526 | 0.0108 | -0.1449 | -0.4510 |
| $E_{t-1}$ | -0.0919 | -0.1015 | -0.1044 | -0.0885 | -0.1014 | -0.1058 |
| $K_{t-1}$ | -0.0885 | -0.0076 | 0.1273 | -0.0685 | -0.0071 | 0.1200 |
| $e_1(t-1)$ | 0.0366 | -0.1029 | -0.1326 | 0.0343 | -0.1029 | -0.1317 |
| $e_2(t-1)$ | -0.0032 | -0.0023 | -0.0308 | -0.0031 | -0.0023 | -0.0307 |
| $e_3(t-1)$ | -0.0478 | -0.0527 | -0.0543 | -0.0460 | -0.0527 | -0.0550 |
| $e_4(t-1)$ | -0.0010 | -0.0031 | -0.0183 | -0.0010 | -0.0031 | -0.0183 |
| $e_5(t-1)$ | 0.0140 | 0.0155 | 0.0160 | 0.0135 | 0.0155 | 0.0162 |
| $e_8(t-1)$ | -0.0285 | 0.0801 | 0.1033 | -0.0267 | 0.0802 | 0.1026 |
| constant | -0.4326 | 12.421 | 28.312 | 4.2603 | 13.659 | 24.638 |
| | | 1935 | | | 1936 | |
| $P_{t-1}$ | -0.0910 | -0.1557 | -0.3790 | -0.2469 | -0.2394 | 0.0188 |
| $E_{t-1}$ | -0.0760 | -0.1002 | -0.1138 | -0.0704 | -0.1034 | -0.1016 |
| $K_{t-1}$ | -0.0084 | -0.0016 | 0.0815 | 0.0439 | 0.0183 | -0.0171 |
| $e_1(t-1)$ | 0.0174 | -0.1046 | -0.1202 | -0.0054 | -0.1167 | -0.0629 |
| $e_2(t-1)$ | -0.0117 | -0.0032 | -0.0246 | -0.0251 | -0.0105 | 0.0098 |
| $e_3(t-1)$ | -0.0395 | -0.0521 | -0.0592 | -0.0366 | -0.0537 | -0.0528 |
| $e_4(t-1)$ | -0.0057 | -0.0037 | -0.0149 | -0.0130 | -0.0076 | 0.0038 |
| $e_5(t-1)$ | 0.1161 | 0.0153 | 0.0174 | 0.0108 | 0.0158 | 0.0155 |
| $e_8(t-1)$ | -0.1353 | 0.0815 | 0.0936 | 0.0042 | 0.0909 | 0.0490 |
| constant | 15.851 | 14.244 | 13.067 | 10.598 | 14.450 | 22.587 |

## TABLE 4

### Expected Values of the Target
### and Endogenous Variables Under Control

|      | 1933   | 1934   | 1935  | 1936  |
|------|--------|--------|-------|-------|
| C    | 47.75  | 53.05  | 58.20 | 61.93 |
| I    | -7.854 | -1.231 | 2.808 | 5.024 |
| W    | 31.28  | 36.39  | 41.62 | 44.69 |
| W**  | 4.981  | 4.736  | 4.518 | 5.578 |
| T    | 5.991  | 6.519  | 7.588 | 9.049 |
| G    | 17.55  | 15.78  | 14.49 | 12.35 |

once with $G_t^*$ equal to the values of Table 3. The losses
were 385.13 and 383.98 respectively, signifying only a small
advantage to the employment of the autocorrelation-modified
controller -- a not unexpected result given the small estimated
M.A. parameters.

## 8. Conclusion

A method has been described for the optimal control
of linear econometric models with a quadratic performance
criterion when there is autocorrelation of either the auto-
regressive or moving average type. By augmenting the state
vector it was possible to convert the system to one with no auto-
correlation and hence to apply some well-known theorems. The
advantage of such a methodology is its potential for application
to a wider array of models than considered explicitly in this paper.

APPENDIX

## Computation of the Expected Welfare Loss Under Control

Given that a control rule has been established, it is of interest to compute the expected welfare loss under such a rule. This is most easily illustrated in the A.R. case where the task is to evaluate

$$E \sum_{t=1}^{T} (y_t - a_t)' K_t (y_t - a_t) \tag{1}$$

subject to

$$y_t = A y_{t-1} + C x_t + b_t + \Phi v_{t-1} + e_t \tag{2}$$

$$v_t = B v_{t-1} + \epsilon_t \tag{3}$$

and $E(e_t e_t') = \Omega$ , $E(\epsilon_t \epsilon_t') = \begin{bmatrix} \Omega & \vdots & 0 \\ 0 & \vdots & 0 \end{bmatrix} = R$ (since from (2.4) and (2.5) $\epsilon_t' = [e_t' \vdots 0]$).

(1) can be re-written as

$$\sum_{t=1}^{T} \text{tr } K_t [E(y_t y_t') - 2 a_t E(y_t') + a_t a_t'] .$$

The control rule is

$$x_t = G_t y_{t-1} + G_t^* v_{t-1} + g_t \quad , \tag{4}$$

which leaves the equation for $y_t$ , under control, as

$$y_t = (A + C G_t) y_{t-1} + (C G_t^* + \Phi) v_{t-1} + (C g_t + b_t) + e_t$$

or

$$y_t = D_t y_{t-1} + F_t v_{t-1} + d_t + e_t \quad . \qquad (5)$$

If $V_t = E(y_t y_t')$, $V_t^* = E(v_t y_t')$, $V_t^{**} = E(v_t v_t')$,
$P_t = E(y_t)$ and $P_t^* = E(v_t)$, the expected welfare loss is

$$\sum_{t=1}^{T} \text{tr } K_t[V_t - 2a_t P_t' + a_t a_t']$$

where $V_t$ and $P_t$ are the solutions of

$$V_t = D_t V_{t-1} D_t' + 2F_t V_{t-1}^* D_t' + 2F_t P_{t-1}^* d_t' + 2D_t P_{t-1} d_t' + F_t V_{t-1}^{**} F_t' +$$

$$d_t d_t' + \Omega \qquad (6)$$

$$V_t^* = B V_{t-1}^* D_t' + B V_{t-1}^{**} F_t' + B P_t^* d_t' + \binom{\Omega}{0} \qquad (7)$$

$$V_t^{**} = B V_{t-1}^{**} B' + R \qquad (8)$$

$$P_t = D_t P_{t-1} + F_t P_{t-1}^* + d_t \qquad (9)$$

$$P_t^* = B P_{t-1}^* \quad . \qquad (10)$$

Each of these equations emerges by utilizing (3) and (5) to get the various expectations, eg. $P_t$ is formed by taking the expectation of (5). Unfortunately, the M.A. situation is more complex since the control rule is now <u>linear in the expected value of $v(t-1)$ and not $v(t-1)$</u>, so that (5) is not

strictly appropriate, i.e., it should be $y_t = D_t y_{t-1} + CG_t^* \hat{v}_{t-1} + \Phi v_{t-1} + d_t + e_t$ where $\hat{v}_{t-1} = E_{t-1}(v_{t-1})$. When $\tilde{v}_{t-1} = \hat{v}_{t-1} - v_{t-1}$ it becomes $y_t = D_t y_{t-1} + F_t v_{t-1} + CG_t^* \tilde{v}_{t-1} + d_t + e_t$ . From the properties of the Kalman-Bucy filter [9] $E(\tilde{v}_{t-1}) = 0$ so that (9) and (10) hold, but the variance of $\tilde{v}_{t-1}$ is not zero and therefore (6) - (8) do not.

Can the use of (6) - (10) be justified for a M.A.? The answer is in the affirmative if $e(0),\ldots,e(-q+1)$ are known with certainty as $e(t)$ will then be computed exactly from $e(t) = u(t) - \sum_{j=1}^{q} \Phi_j \, e(t-j)$ for any $t$ in the control period. Because estimates of $e(0),\ldots,e(-q+1), A$, $\Phi$ and $C$ are likely to be obtained jointly by F.I.M.L. estimation, it seems logical to treat all estimates in the same fashion. In particular, if $A$, $\Phi$ and $C$ are regarded as non-stochastic, treating the estimated initial conditions as the correct values would seem justified. However, if the stochastic nature of $A$, $B$ and $C$ is recognized -- as in [8] -- new versions of (6)- (10) will need to be derived.

REFERENCES

[1]   Athans, M., "The Discrete Time Linear-Quadratic-Gaussian
       Stochastic Control Problem," _Annals of Economic and_
       _Social Measurement_, 1 (1972), 449-491.

[2]   Box, G.E.P. and G.M. Jenkins, _Time Series Analysis:_
       _Forecasting and Control_ (Holden-Day, San Francisco, 1970).

[3]   Byron, R.P., "The Computation of Maximum Likelihood Estimates
       for Linear Simultaneous Systems with Moving Average
       Disturbances," (Australian National University, 1973,
       unpublished M.S.).

[4]   Chow, G.C., "Optimal Stochastic Control of Linear Economic
       Systems," _Journal of Money Credit and Banking_, 2 (1970),
       291-302.

[5]   _____, "Optimal Control of Linear Econometric Systems
       with Finite Time Horizon," _International Economic_
       _Review_, 13 (1972), 16-25.

[6]   _____, "How Much Could be Gained by Optimal Stochastic
       Control Policies," _Annals of Economic and Social_
       _Measurement_, 1 (1972), 391-406.

[7]   _____, "Problems of Economic Policy from the Viewpoint
       of Optimal Control," Econometric Research Program
       Research Memorandum No. 139 (Nov. 1972), Princeton
       University.

[8]   _____, "Effect of Uncertainty on Optimal Control
       Policies," Econometric Research Program Memorandum No.
       145 (Sept. 1972), Princeton University.

[9]   Goldberger, A.S., _Econometric Theory_ (John Wiley, New York,
       1964).

[10]  Kalman, R.E. and R.S. Bucy, "New Results in Linear Filtering
       and Prediction Theory," _Journal of Basic Engineering_,
       83 (1961), 95-108.

[11]  Pagan, A.R., "A Generalization of the Treatment of Auto-
       correlation," Working Papers in Economics and Econometrics
       No. 10, Australian National University (Nov. 1972).

[12]  Pindyck, R.S., "An Application of the Linear Quadratic
       Tracking Problem to Economic Stabilization Policy,"
       _I.E.E.E. Transactions on Automatic Control_, 17 (1972),
       287-300.

[13]  Taylor, L., "The Existence of Optimal Distributed Lags,"
      The Review of Economic Studies, 37 (1970), 95-106.

[14]  Theil, H., Optimal Decision Rules for Government and
      Industry (North-Holland, Amsterdam, 1964).

[15]  Trivedi, P.K., "Retail Inventory Investment Behaviour,"
      Journal of Econometrics, 1 (1973), 61-80.

[16]  Wallis, K.F., "Testing for Fourth Order Autocorrelation
      in Quarterly Regression Equations," Econometrica,
      40 (1972), 617-636.

[17]  Williams, R.A., "Demand for Consumer Durables:  Stock
      Adjustment Models and Alternative Specifications of
      Stock Depletion," Review of Economic Studies, 39
      (1972), 281-295.