Multiple Regime Models with Switches in Exogeneity

Richard G. Pierse
Extended Essay for the MSc. Econometrics Examination
London School of Economics and Political Science

June 2nd 1979

1 Introduction

Mutiple regime models have been used by econometricians in several different fields, a major application being to models of market disequilibrium (see Maddala and Nelson [12]). One interesting field of application is the controlled economy where the policymaker switches between control instruments at different points in time, these switchpoints being generally known (unlike the situation in market disequlibrium models). A characteristic of such models is that the partitioning between "endogenous" and "exogenous" variables changes between regimes.

Multiple regime models with this characteristic have recently been analysed by Jean-François Richard in a forthcoming paper [15]. He shows that it is possible to test for the exogeneity of the control variables in the different regimes and hence to test whether or not the economy was being controlled.

This study is an attempt to implement the analysis of Richard for the FIML estimation and testing of multiple regime models and apply it to model the change in regime in the U.K. monetary sector of the introduction of 'Competition and Credit Control' in October 1971. An existing FIML computer program written by David Hendry was generalised to maximise the likelihood function of multiple regime models and to incorporate the tests proposed by Richard. The application involved the estimation of a money demand equation; in developing this equation a single equation approach was used and the general to simple methodology of Hendry and Mizon [11] was adopted. The resulting equation in seasonally unadjusted data provides direct comparison with the equations in [11].

The rest of this paper is organised as follows:

- In Section 2 some of the key concepts are developed for some simple models

- Section 3 derives Richard's exogeneity test for the case of a complete simultaneous model with a single regime
- Section 4 deals with the FIML estimation of multiple regime models and describes a testing procedure for these models
- Section 5 presents an application to the estimation of a simple two-equation model of the monetary sector. Results and conclusions are presented.

A note on exogeneity.

We define exogeneity for our purposes as follows: a variable is exogenous if we can run the analysis conditional on it without loss of information. This is a weak definition of exogeneity. In particular we do not require that lagged values of endogenous variables do not enter the determination of exogenous variables. Let x_t be an exogenous variable and y_t an endogenous variable. Then in the lag formulation

$$a(L)x_t = b(L)y_t + w_t$$

where L is the lag operator, we do not require that $b(L) = 0$ (although b_0 must be zero). Thus y_t may in fact be "causing" x_t in the Granger sense of causality (see Granger [7]).

2 Exogeneity in Simple Models

Suppose that we have two variables y_{1t} and y_{2t} which are jointly normally distributed with density

$$f(y_{1t}, y_{2t}) = N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{2.1}$$

or explicitly

$$f(y_{1t}, y_{2t}) = (2\pi)^{-1} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \begin{pmatrix} y_{1t} - \mu_1 \\ y_{2t} - \mu_2 \end{pmatrix}' \boldsymbol{\Sigma}^{-1} \begin{pmatrix} y_{1t} - \mu_1 \\ y_{2t} - \mu_2 \end{pmatrix} \right) \tag{2.2}$$

where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{21} & \sigma^{22} \end{pmatrix}.$$

The marginal distributions for y_{1t} and y_{2t} are given by

$$f(y_{1t}) = N(\mu_1, \sigma_{11}) \quad \text{and} \quad f(y_{2t}) = N(\mu_2, \sigma_{22}).$$

Defining $s_1 = \frac{\sigma_{11}\sigma_{22} - (\sigma_{12})^2}{\sigma_{11}}$ and $s_2 = \frac{\sigma_{11}\sigma_{22} - (\sigma_{12})^2}{\sigma_{22}}$ we have

$$\sigma^{11} = \frac{\sigma_{22}}{\sigma_{11}} s_1^{-1} \quad , \quad \sigma^{12} = -\frac{\sigma_{12}}{\sigma_{11}} s_1^{-1} \quad , \quad \sigma^{22} = s_1^{-1}$$

and the conditional distribution of y_{2t} given y_{1t} is

$$f(y_{2t}|y_{1t}) = f(y_{1t}, y_{2t})/f(y_{1t})$$

$$= \quad (2\pi)^{-\frac{1}{2}}\sigma_{11}^{\frac{1}{2}}(\sigma_{11}\sigma_{22} - (\sigma_{12})^2)^{-\frac{1}{2}} \exp(-\frac{1}{2s_1}[\frac{\sigma_{22}}{\sigma_{11}}(y_{1t} - \mu_1)^2 \tag{2.3}$$

$$-2(y_{1t} - \mu_1)(y_{2t} - \mu_2) + (y_{2t} - \mu_2)^2 - (\frac{\sigma_{11}\sigma_{22} - (\sigma_{12})^2}{\sigma_{11}^2})(y_{1t} - \mu_1)^2])$$

$$= (2\pi s_1)^{-\frac{1}{2}} \exp(-\frac{1}{2s_1}[(y_{2t} - \mu_2)^2 - 2\frac{\sigma_{12}}{\sigma_{11}}(y_{1t} - \mu_1)(y_{2t} - \mu_2) + \left(\frac{\sigma_{12}}{\sigma_{11}}\right)^2 (y_{1t} - \mu_1)^2])$$

which is a univariate distribution

$$f(y_{2t}|y_{1t}) = N(\mu_2 + \frac{\sigma_{12}}{\sigma_{11}}(y_{1t} - \mu_1), s_1). \tag{2.4}$$

Similarly we can derive the conditional density for y_{1t} given y_{2t} which is

$$f(y_{1t}|y_{2t}) = N(\mu_1 + \frac{\sigma_{12}}{\sigma_{22}}(y_{2t} - \mu_2), s_2). \tag{2.5}$$

From (2.4) we get the regression model

$$y_{2t} = a + by_{1t} + u_t \quad , \quad u_t \sim NI(0, \delta^2) \tag{2.6}$$

with $Cov(y_{1t}, u_t) = 0$, where $a = \mu_2 - b\mu_1$, $b = \frac{\sigma_{12}}{\sigma_{11}}$ and $\delta^2 = s_1$ and from (2.5) the regression model

$$y_{1t} = c + dy_{2t} + v_t \quad , \quad v_t \sim NI(0, \psi^2) \tag{2.7}$$

with $Cov(y_{2t}, v_t) = 0$, $c = \mu_1 - d\mu_2$, $d = \frac{\sigma_{12}}{\sigma_{22}}$ and $\psi^2 = s_2$.

As long as there are no prior cross-restrictions on the five parameters of the joint distribution ($\mu_1, \mu_2, \sigma_{11}, \sigma_{12}, \sigma_{22}$) we lose no information on (a, b, δ^2), the parameters of the conditional distribution of y_{2t}, by treating (μ_1, σ_{11}) as nuisance parameters and running the analysis conditional on y_{1t}. Equally, if we choose to treat (μ_2, σ_{22}) as nuisance parameters no information is lost by running the analysis conditional on y_{2t}. Thus (2.6) and (2.7) are equally valid parameterisations of the model (2.1) where we choose to treat (μ_1, σ_{11}) or (μ_2, σ_{22}) respectively as nuisance parameters.

Now consider the two-regime model

$$\begin{aligned} y_{1t} &= a_1 + b_1 y_{2t} + u_{1t}, & t \in I_1 = \{1, \cdots, T1 - 1\} \\ y_{2t} &= a_2 + b_{21} y_{1t} + u_{2t}, & t \in I_2 = \{T1, \cdots, T\} \end{aligned} \tag{2.8}$$

where $u_{1t} \sim NI(0,\delta_1^2)$, $u_{2t} \sim NI(0,\delta_2^2)$, $Cov(y_{2t},u_{1t}) = 0$ $t \in I_1$ and $Cov(y_{1t},u_{2t}) = 0$ $t \in I_2$. The switching time $T1$ is assumed to be known. There appears to be a switch in the exogeneity of the variables in this model at $T1$ with y_{2t} exogenous for $t \in I_1$ and y_{1t} exogenous for $t \in I_2$. However, it is clear from our previous analysis that this parameterisation is inadequate to describe such a switch in exogeneity. Let us assume that the joint distribution of (y_{1t}, y_{2t}) in regime i is

$$f(y_{1t}, y_{2t}) = N(\boldsymbol{\mu}^i, \boldsymbol{\Sigma}^i) \quad , \quad t \in I_1 \quad , \quad i = 1,2. \tag{2.9}$$

Then indeed (2.8) follows together with

$$(a_i, b_i, \delta_i^2) = (\mu_i^i - \frac{\sigma_{12}^i}{\sigma_{jj}^i}\mu_j^i, \frac{\sigma_{12}^i}{\sigma_{jj}^i}, \frac{\sigma_{11}^i\sigma_{22}^i - (\sigma_{12}^i)^2}{\sigma_{jj}^i}) \tag{2.10}$$

for $i = 1,2$, $j = 3-i$ where we treat (μ_j^i, σ_{jj}^i) as the nuisance parameters. Equally, however, we have the parameterisation

$$\begin{gathered} y_{1t} = a_3 + b_3 y_{2t} + u_t, \quad t = 1,\cdots,T \\ y_{2t} \sim NI(0,\phi_1^2) \quad t \in I_1, \quad y_{2t} \sim NI(0,\phi_2^2) \quad t \in I_2 \end{gathered} \tag{2.11}$$

with $u_t \sim NI(0,\delta_3^2)$ where

$$(a_3, b_3, \delta_3^2) = (\mu_1^i - \frac{\sigma_{12}^i}{\sigma_{22}^i}\mu_2^i, \frac{\sigma_{12}^i}{\sigma_{22}^i}, \frac{\sigma_{11}^i\sigma_{22}^i - (\sigma_{12}^i)^2}{\sigma_{22}^i}) \quad i = 1,2 \tag{2.12}$$

and $(\mu_2^i, \sigma_{22}^i) = (0, \phi_i^2)$ are treated as the nuisance parameters. (Note that we have the parameter correspondences $(a_1, b_1, \delta_1^2) = (a_3, b_3, \delta_3^2)$ for the first regime and $(a_2, b_2, \delta_2^2) = (-a_3 b_3 \lambda, b_3\lambda, \delta_3^2\lambda)$ for the second regime where $\lambda = \sigma_{22}^2/\sigma_{11}^2$.)

Equations (2.8) and (2.11) are both valid parameterisations of the data generation process (2.9) yet (2.8) exhibits a switch in exogeneity between regimes whereas in (2.11) the same variable y_{2t} is treated as exogenous over the whole period. The choice of exogenous variables in these models is quite arbitrary and the statement that y_{2t} is exogenous in (2.11) is no more than a statement that (μ_2, σ_{22}) are nuisance parameters. It is not subject to testing since it forces no restrictions on the joint distribution of the observable variables (2.9).

Suppose that, returning to the single regime model, we now do have a cross-restriction on the parameters of the joint distribution (2.1). We introduce a behavioural hypothesis

$$\mu_{2t} = c_0 + c_1\mu_{1t}. \tag{2.13}$$

This hypothesis says that the expectation of variable y_{2t} at time t depends on the expectation of y_{1t} at time t. We can interpret it as a behavioural rule followed by agents who take the joint distribution (2.1) as given. Looking at the conditional distribution of y_{2t} given y_{1t} we have from (2.4)

$$E(y_{2t}|y_{1t}) = \mu_2 + \frac{\sigma_{12}}{\sigma_{11}}(y_{1t} - \mu_1) \tag{2.14}$$

but from (2.13) this implies

$$E(y_{2t}|y_{1t}) = c_0 + \left(c_1 - \frac{\sigma_{12}}{\sigma_{11}}\right)\mu_{1t} + \frac{\sigma_{12}}{\sigma_{11}}y_{1t}. \tag{2.15}$$

In general this expression will now involve μ_{1t} and we will lose information on the parameters of the conditional distribution by running the regression model (2.6) treating y_{1t} as exogenous. Only if

$$c_1 = \frac{\sigma_{12}}{\sigma_{11}} \tag{2.16}$$

so that

$$E(y_{2t}|y_{1t}) = c_0 + \frac{\sigma_{12}}{\sigma_{11}}y_{1t} \tag{2.17}$$

do we lose no information by treating y_{1t} as exogenous. Condition (2.16) then is a direct test for the exogeneity of y_{1t}. An alternative intuitive explanation of this result may be helpful. Let us write

$$\mu_{1t} = E(y_{1t}|\ell_{t-1}) \quad , \quad \mu_{2t} = E(y_{2t}|\ell_{t-1})$$

where ℓ_{t-1} is the information set available at time t. Then the exogeneity condition (2.16) is equivalent to the condition that $y_{1t} \in \ell_{t-1}$ in which case $\mu_{1t} = y_{1t}$ and (2.15) becomes

$$E(y_{2t}|y_{1t}) = c_0 + c_1 y_{1t}. \tag{2.17$'$}$$

Similarly from the conditional distribution of y_{1t} given y_{2t} the hypothesis (2.13) gives

$$E(y_{1t}|y_{2t}) = \left(1 - \frac{\sigma_{12}}{\sigma_{22}}c_1\right)\mu_{1t} - \frac{\sigma_{12}}{\sigma_{22}}c_0 + + \frac{\sigma_{12}}{\sigma_{22}}y_{2t} \tag{2.18}$$

which leads to the exogeneity condition

$$c_1 = \frac{\sigma_{22}}{\sigma_{12}} \tag{2.19}$$

in which case

$$E(y_{1t}|y_{2t}) = -\frac{c_0}{c_1} + \frac{\sigma_{12}}{\sigma_{22}}y_{2t} \tag{2.20}$$

which valididates the parameterisation (2.7).

Treating y_{1t} (y_{2t}) as exogenous is only valid when the appropriate restriction ((2.16) or (2.19)) is satisfied and the two conditions cannot hold jointly (otherwise $\boldsymbol{\Sigma}$ would be singular). When y_{1t} is exogenous and if it is controlled then agents adjust their behaviour by revising their expectations according to the behavioural rule (2.13).

This framework provides a basis for examining multiple regime models with possible switches in exogeneity. We have seen that we need to specify the joint distribution of all the relevant variables in each regime, whether or not they are being controlled, so that the hypothesis of exogeneity can be tested. We now go on to develop this approach for the full information estimation of complete simultaneous equations models.

3 Exogeneity in a Single Regime

In the simple models considered in section 2 there were no predetermined variables. We now want to extend the analysis to allow for lagged dependent variables; we may also allow for some variables to be purely exogenous without proposing to test this. The general form of model for a single regime is

$$\boldsymbol{A}\boldsymbol{x}_t = \boldsymbol{B}\boldsymbol{y}_t + \boldsymbol{C}\boldsymbol{z}_t = \boldsymbol{u}_t \quad , \quad \boldsymbol{u}_t \sim NI(\boldsymbol{0}, \boldsymbol{\Sigma}) \tag{3.1}$$

with reduced form

$$\boldsymbol{y}_t = \boldsymbol{\Pi}\boldsymbol{z}_t + \boldsymbol{v}_t \quad , \quad \boldsymbol{v}_t \sim NI(\boldsymbol{0}, \boldsymbol{\Omega}) \tag{3.2}$$

where $\boldsymbol{y}_t$ is now an $(n \times 1)$ vector of observations on n endogenous variables and $\boldsymbol{z}_t$ is an $(m \times 1)$ vector of observations on predetermined variables. We assume that (3.1) is a complete model so that $\boldsymbol{B}$ is a square, non-singular $(n \times n)$ matrix and $\boldsymbol{\Pi} = -\boldsymbol{B}^{-1}\boldsymbol{C}$. The matrices $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{\Sigma}$ are functions of a vector of parameters $\boldsymbol{\theta}$.

The joint distribution of $\boldsymbol{y}_t$ is given by

$$f(\boldsymbol{y}_t|\boldsymbol{z}_t, \boldsymbol{\theta}) = N(\boldsymbol{\Pi}\boldsymbol{z}_t, \boldsymbol{\Omega}). \tag{3.3}$$

In formulating the data density of $\boldsymbol{Y}' = (\boldsymbol{y}_1 \cdots \boldsymbol{y}_T)$ we must take into account that $\boldsymbol{z}_t$ includes lagged $\boldsymbol{y}_t$. If $\boldsymbol{y}_t^0$ are the lagged endogenous variables in $\boldsymbol{z}_t$ and $\boldsymbol{z}_t^*$ are the purely exogenous variables then the data density, conditional with respect to initial conditions, is

$$f(\boldsymbol{y}_1 \cdots \boldsymbol{y}_T|\boldsymbol{y}_1^0, \boldsymbol{z}_1^* \cdots \boldsymbol{z}_T^*, \boldsymbol{\theta}) = \prod_{t=1}^{T} f(\boldsymbol{y}_t|\boldsymbol{y}_1^0, \boldsymbol{z}_t^*, \boldsymbol{\theta}). \tag{3.4}$$

From now on, for notational convenience, we will no longer make explicit the conditionality on $\boldsymbol{y}_1^0$, $\boldsymbol{z}_t^*$ but write

$$f(\boldsymbol{Y}|\boldsymbol{\theta}) = \prod_{t=1}^{T} f(\boldsymbol{y}_t|\boldsymbol{z}_t, \boldsymbol{\theta}). \tag{3.5}$$

Consider partitioning $\boldsymbol{y}_t$ into

$$\boldsymbol{y}_t' = (\boldsymbol{y}_{1t}' : \boldsymbol{y}_{2t}') \quad , \quad \boldsymbol{Y} = (\boldsymbol{Y}_1 : \boldsymbol{Y}_2) \tag{3.6}$$

where $\boldsymbol{y}_{1t} = (n_1 \times 1)$, $\boldsymbol{y}_{2t} = (n_2 \times 1)$ and $n_1 + n_2 = n$. $\boldsymbol{Y}_2$ is now defined to be exogenous if and only if there exists a partitioning of $\boldsymbol{\theta}$ into $\boldsymbol{\theta} = \boldsymbol{(}\boldsymbol{\theta}_1 : \boldsymbol{\theta}_2)$ such that

$$f(\boldsymbol{y}_t|\boldsymbol{z}_t, \boldsymbol{\theta}) = f(\boldsymbol{y}_{1t}|\boldsymbol{y}_{2t}, \boldsymbol{z}_t, \boldsymbol{\theta}_1) f(\boldsymbol{y}_{2t}|\boldsymbol{z}_t, \boldsymbol{\theta}_2) \tag{3.7}$$

where $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are not subject to any cross-restrictions and $\boldsymbol{\theta}_2$ may be regarded as nuisance parameters. Under these conditions the likelihood function $LL(\boldsymbol{\theta}|\boldsymbol{Y})$ factorises as

$$\begin{aligned} LL(\boldsymbol{\theta}|\boldsymbol{Y}) &= \prod_{t=1}^{T} f(\boldsymbol{y}_{1t}|\boldsymbol{y}_{2t}, \boldsymbol{z}_t, \boldsymbol{\theta}_1) \prod_{t=1}^{T} f(\boldsymbol{y}_{2t}|\boldsymbol{z}_t, \boldsymbol{\theta}_2) \\ &= LL_1(\boldsymbol{\theta}_1) LL_2(\boldsymbol{\theta}_2) \end{aligned} \tag{3.8}$$

and no relevant sample information is lost if we drop the factor $LL_2(\boldsymbol{\theta}_2)$. Note that in general the joint density (3.5) will not factorise since $\boldsymbol{y}_{2t}$ will not be independent of lagged values of $\boldsymbol{y}_{1t}$ in $\boldsymbol{z}_t$.

Let us partition the model (3.1) into two subsets of equations

$$\begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{y}_{1t} \\ \boldsymbol{y}_{2t} \end{pmatrix} + \begin{pmatrix} \boldsymbol{C}_1 \\ \boldsymbol{C}_2 \end{pmatrix} \boldsymbol{z}_t = \begin{pmatrix} \boldsymbol{u}_{1t} \\ \boldsymbol{u}_{2t} \end{pmatrix}. \tag{3.9}$$

We want to know the conditions under which $\boldsymbol{y}_{2t}$ is exogenous and the likelihood function factorises. Richard has proved [15, Theorem 3.1] that for the general case of models which may be incomplete, the conditions (in our notation) are

$$\boldsymbol{B}_{21} = \boldsymbol{0} \tag{3.10}$$

and

$$\boldsymbol{B}_{11}\boldsymbol{\Omega}_{12} + \boldsymbol{B}_{12}\boldsymbol{\Omega}_{22} = \boldsymbol{0} \tag{3.11}$$

where

$$\boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\Omega}_{11} & \boldsymbol{\Omega}_{12} \\ \boldsymbol{\Omega}_{21} & \boldsymbol{\Omega}_{22} \end{pmatrix}$$

is the reduced form covariance matrix. For the case of complete models we can translate the condition (3.11) into a condition on $\boldsymbol{\Sigma}$, the covariance matrix of the structural form since

$$\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{\Omega}\boldsymbol{B}' = \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Omega}_{11} & \boldsymbol{\Omega}_{12} \\ \boldsymbol{\Omega}_{12}' & \boldsymbol{\Omega}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{B}_{11}' & \boldsymbol{B}_{21}' \\ \boldsymbol{B}_{12}' & \boldsymbol{B}_{22}' \end{pmatrix} \tag{3.12}$$

so that

$$\begin{aligned} \boldsymbol{\Sigma}_{12} &= \boldsymbol{B}_{11}\boldsymbol{\Omega}_{11}\boldsymbol{B}_{21}' + \boldsymbol{B}_{12}\boldsymbol{\Omega}_{12}'\boldsymbol{B}_{21}' + \boldsymbol{B}_{11}\boldsymbol{\Omega}_{12}\boldsymbol{B}_{22}' + \boldsymbol{B}_{12}\boldsymbol{\Omega}_{22}\boldsymbol{B}_{22}' \\ &= (\boldsymbol{B}_{11}\boldsymbol{\Omega}_{12} + \boldsymbol{B}_{12}\boldsymbol{\Omega}_{22})\boldsymbol{B}_{22}' \end{aligned}$$

by (3.10). Thus (3.11) becomes for complete models

$$\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21}' = \mathbf{0}. \tag{3.13}$$

The conditions (3.10) and (3.13) are equivalent to the condition that the model (3.9) is block recursive.

The log-likelihood function for the model is

$$L(\boldsymbol{A}, \boldsymbol{\Sigma}) = -\frac{nT}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Sigma}| + T\log||\boldsymbol{B}|| - \frac{1}{2}tr(\boldsymbol{\Sigma}^{-1}\boldsymbol{U}'\boldsymbol{U}) \tag{3.14}$$

where

$$\boldsymbol{U}' = (\boldsymbol{u}_1 \cdots \boldsymbol{u}_T) = \begin{pmatrix} \boldsymbol{U}_1' \\ \boldsymbol{U}_2' \end{pmatrix}$$

$$\begin{aligned} L(\boldsymbol{A}, \boldsymbol{\Sigma}) = \; & -\frac{nT}{2}\log(2\pi) - \frac{T}{2}\log\begin{vmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22} \end{vmatrix} + T\log\left\|\begin{matrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \mathbf{0} & \boldsymbol{B}_{22} \end{matrix}\right\| \\ & -\frac{1}{2}tr\left(\begin{pmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}^{-1} \end{pmatrix}\begin{pmatrix} \boldsymbol{U}_1'\boldsymbol{U}_1 & \boldsymbol{U}_1'\boldsymbol{U}_2 \\ \boldsymbol{U}_2'\boldsymbol{U}_1 & \boldsymbol{U}_2'\boldsymbol{U}_2 \end{pmatrix}\right) \end{aligned} \tag{3.15}$$

$$\begin{aligned} = \; & \left\{\frac{-n_1T}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Sigma}_{11}| + T\log||\boldsymbol{B}_{11}|| - \frac{1}{2}tr(\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{U}_1'\boldsymbol{U}_1)\right\} \\ & +\left\{\frac{-n_2T}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Sigma}_{22}| + T\log||\boldsymbol{B}_{22}|| - \frac{1}{2}tr(\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{U}_2'\boldsymbol{U}_2)\right\} \end{aligned} \tag{3.16}$$

$$= L_1(\boldsymbol{B}_{11}, \boldsymbol{B}_{12}, \boldsymbol{C}_1, \boldsymbol{\Sigma}_{11}) + L_2(\boldsymbol{B}_{22}, \boldsymbol{C}_2, \boldsymbol{\Sigma}_{22}) \tag{3.17}$$

which does indeed factorise as in (3.8). It follows that the two factors can be estimated independently of each other, and if the parameters $(\boldsymbol{B}_{22}, \boldsymbol{C}_2, \boldsymbol{\Sigma}_{22})$ are nuisance parameters the factor L_2 can be dropped.

The conditions (3.10) and (3.13) give us a straightforward test for the exogeneity of the $\boldsymbol{y}_2$ variables. First we estimate the unrestricted model (3.14). We can concentrate out the $\boldsymbol{\Sigma}$ matrix in the usual way to give a concentrated log-likelihood function

$$L^*(\boldsymbol{A}, \boldsymbol{\Sigma}) = -\frac{nT}{2}\log(2\pi) - \frac{nT}{2} - \frac{T}{2}\log\left|\frac{\boldsymbol{A}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A}'}{T}\right| + T\log||\boldsymbol{B}|| \tag{3.18}$$

together with

$$\boldsymbol{\Sigma} = \frac{\boldsymbol{A}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A}'}{T} \tag{3.19}$$

and then maximise (3.18). Let the maximum be L^+. We then reestimate the model imposing the restrictions (3.10) amd (3.13). Note that we can concentrate $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\Sigma}_{22}$ out of (3.16)

$$\begin{aligned} \frac{\partial L}{\partial \boldsymbol{\Sigma}_{11}^{-1}} &= \frac{T}{2}\boldsymbol{\Sigma}_{11} - \frac{1}{2}\boldsymbol{A}_1\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A}_1' = \boldsymbol{0} \\ \Longrightarrow & \quad \boldsymbol{\Sigma}_{11} = \frac{1}{T}\boldsymbol{A}_1\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A}_1' \end{aligned} \tag{3.20}$$

$$\begin{aligned} \frac{\partial L}{\partial \boldsymbol{\Sigma}_{22}^{-1}} &= \frac{T}{2}\boldsymbol{\Sigma}_{22} - \frac{1}{2}\boldsymbol{A}_2\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A}_2' = \boldsymbol{0} \\ \Longrightarrow & \quad \boldsymbol{\Sigma}_{22} = \frac{1}{T}\boldsymbol{A}_2\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A}_2' \end{aligned} \tag{3.21}$$

giving the concentrated log-likelihood function

$$\begin{aligned} L^{**}(\boldsymbol{A}_1, \boldsymbol{A}_2; \boldsymbol{\Sigma}_{11}, \boldsymbol{\Sigma}_{22}) = & -\frac{nT}{2}\log(2\pi) - \frac{nT}{2} - \frac{T}{2}\log\left|\frac{\boldsymbol{A}_1\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A}_1'}{T}\right| \\ & + T\log||\boldsymbol{B}_{11}|| - \frac{T}{2}\log\left|\frac{\boldsymbol{A}_2\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A}_2'}{T}\right| + T\log||\boldsymbol{B}_{22}|| \end{aligned} \tag{3.22}$$

where

$$\begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \boldsymbol{C}_1 \\ \boldsymbol{0} & \boldsymbol{B}_{22} & \boldsymbol{C}_2 \end{bmatrix}.$$

Let L^{++} be the maximum of (3.22). Then on the null hypothesis that $\boldsymbol{y}_{2t}$ is exogenous

$$\lambda = 2(L^+ - L^{++}) \sim_a \chi^2(q_1 + q_2) \tag{3.23}$$

where q_1 is the number of unrestricted parameters in $\boldsymbol{B}_{21}$ in model (3.18) and q_2 is the number of parameter restrictions corresponding to the condition (3.13)

$= n1 \cdot n2$. (Not $2 \cdot n1 \cdot n2$ since $\boldsymbol{\Sigma}$ is already restricted to be symmetric). In fact often we may want to impose $\boldsymbol{B}_{21} = \mathbf{0}$ to identify the unrestricted model even before we test for exogeneity. In this case $q_1 = 0$ and we test only the $\boldsymbol{\Sigma}$ restrictions.

We can now go on to consider testing for switches in exogeneity in a multiple regime model using this test. If there are no constant parameters in different regimes then each may be estimated and tested separately. However, it is more reasonable to suppose that a subset of equations remain constant across regimes in which case we have cross-regime parameter restrictions. Then we must set up the joint likelihood function for all the regimes and maximise this subject to these parameter restrictions between regimes. Then we can test the exogeneity restrictions for each regime separately. This is the topic of the next section.

However, we want first to give an interpretation of the model we have been developing. We can consider it as an economy which is controlled by a policymaker. Then the first set of structural equations describes the economy on which control is forced and the second set the control process itself. The equations for $\boldsymbol{y}_{2t}$ are in fact reaction functions for the control instruments and by condition (3.10) the control authorities must make their plans solely on the basis of predetermined variables. Further, if the instruments are to be exogenous, condition (3.13) must also be satisfied. Switching between control instruments leads to switches in the exogeneity/endogeneity of variables.

The simplest illustration is an equilibrium model of a market where the supply of the commodity is under the control of the authorities. The authorities have two possible alternatives: they can either control the commodity price letting the market determine the quantity traded or they can control the supply and let the market determine the clearing price. We assume that the demand relationship is not changed by this regime switch. Then, taking a simple linear demand function we have, for the first regime, the demand equation

$$q_t = a_0 + a_1 p_t + v_t \quad , \quad E(v_t^2) = \sigma_v^2 \tag{3.24}$$

together with the authority reaction function for price

$$p_t = a_2 + a_3 z_t + w_t. \tag{3.25}$$

This reaction function depends solely on the single predetermined variable z_t.

In the second regime the authorities switch to controlling the commodity supply and so the demand relationship (3.24) now determines the price and we renormalise it as

$$p_t = b_0 + b_1 q_t + \varepsilon_t \quad , \quad E(\varepsilon_t^2) = \sigma_\varepsilon^2 \tag{3.26}$$

and suppose for simplicity that the same predetermined variable z_t continues to determine the new reaction function which is then

$$q_t = b_2 + b_3 z_t + \eta_t. \tag{3.27}$$

The constancy of the demand relation across regimes means that we have restrictions between the parameters in each regime which are

$$(b_0, b_1, \sigma_\varepsilon^2) = (-\frac{a_0}{a_1}, \frac{1}{a_1}, \frac{1}{(a_1)^2}\sigma_v^2). \tag{3.28}$$

Full information maximum likelihood estimation requires maximising the joint likelihood function of the two regimes subject to all prior restrictions on the parameters which include these non-linear cross-regime restrictions.

4 FIML Estimation

In this section we consider the FIML estimation of multiple regime models in which we want to be able to test for switches in exogeneity. We have r regimes which we assume, without loss of generality, to have been operating sequentially. For regime i the model is

$$A^i(\theta)x_t = u_t \quad , \quad t \in I_i. \tag{4.1}$$

The switchpoints between regimes are assumed to be known. the log-likelihood for regime i is

$$\begin{aligned} L_i(\boldsymbol{\theta}) = \quad & -\frac{nT_i}{2}\log(2\pi) - \frac{T_i}{2}\log\left|\boldsymbol{\Sigma}^i(\boldsymbol{\theta})\right| + T_i \log\left|\left|\boldsymbol{B}^i(\boldsymbol{\theta})\right|\right| \\ & -\frac{1}{2}tr(\boldsymbol{\Sigma}^i(\boldsymbol{\theta})^{-1}\boldsymbol{A}^i(\boldsymbol{\theta})\boldsymbol{X}_i'\boldsymbol{X}_i\boldsymbol{A}^i(\boldsymbol{\theta})') \end{aligned} \tag{4.2}$$

where

$$\boldsymbol{X}' = (\boldsymbol{x}_1 \cdots \boldsymbol{x}_T) = (\boldsymbol{X}_1' : \cdots : \boldsymbol{X}_r'),$$

and the joint log-likelihood for all r regimes is

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{r} L_i(\boldsymbol{\theta}). \tag{4.3}$$

The elements of the coefficient matrix $\boldsymbol{A}^i$ and the covariance matrix $\boldsymbol{\Sigma}^i$ of each regime are functions of a common vector of parameters $\boldsymbol{\theta}$. This is because we hypothesise that some relationships are common to more than one regime so that there are cross-restrictions between some elements of $\boldsymbol{A}^i$ and $\boldsymbol{\Sigma}^i$ for different regimes. When there are no cross-restrictions we can partition $\boldsymbol{\theta}$ so that

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{r} L_i(\boldsymbol{\theta}_i) \tag{4.4}$$

and the r vectors $\boldsymbol{\theta}_i$ have no elements in common. It follows that each regime can be estimated separately by standard FIML programs so these models are not of interest to us.

From (4.3) we have that

$$\frac{\partial L}{\partial \boldsymbol{\theta}} = \sum_{i=1}^{r} \frac{\partial L_i}{\partial \boldsymbol{\theta}} \quad , \quad \text{and} \quad \frac{\partial^2 L}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \sum_{i=1}^{r} \frac{\partial^2 L_i}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}. \tag{4.5}$$

If the $\boldsymbol{\Sigma}^i$ are unrestricted we can concentrate them out of the log-likelihood as usual to get

$$L^*(\boldsymbol{\theta}^*) = \sum_{i=1}^{r} L_i^*(\boldsymbol{\theta}^*) \tag{4.6}$$

where

$$\begin{aligned} L_i^*(\boldsymbol{\theta}^*) = \quad & -\frac{nT_i}{2}\log(2\pi) - \frac{nT_i}{2} - \frac{T_i}{2}\log\left|\frac{1}{T_i}\boldsymbol{A}^i(\boldsymbol{\theta}^*)\boldsymbol{X}_i'\boldsymbol{X}_i\boldsymbol{A}^i(\boldsymbol{\theta}^*)'\right| \\ & +T_i \log\left|\left|\boldsymbol{B}^i(\boldsymbol{\theta}^*)\right|\right| \end{aligned} \tag{4.7}$$

and

$$\widetilde{\boldsymbol{\Sigma}}^i = \frac{1}{T_i}\boldsymbol{A}^i\boldsymbol{X}_i'\boldsymbol{X}_i\boldsymbol{A}^{i\prime}. \tag{4.8}$$

We have specified nothing, so far, about the nature of the mappings from $\boldsymbol{\theta}$ to the elements of $\boldsymbol{A}^i$ and $\boldsymbol{\Sigma}^i$. If these mappings are non-linear then we say that the model is non-linear in parameters. We now show that by choosing the normalisation of the model in each regime we can keep the model linear in parameters.

We want to ensure that the parameters of each regime are identified before we consider imposing cross-regime restrictions. This is because we will later want to test these extra restrictions. We will consider only exclusion restrictions on $\boldsymbol{A}^i$ (the only available form of restriction in many standard FIML packages) where some elements of $\boldsymbol{A}^i$ are set a priori to zero. Let $\boldsymbol{\Phi}^k$ be a diagonal matrix defined by

$$\begin{aligned} \phi_{jj}^k &= \quad 1 \text{ if } a_{kj}^i = 0 \\ &= \quad 0 \text{ otherwise.} \end{aligned} \tag{4.9}$$

Then a necessary and sufficient condition for the kth equation to be identified is that

$$rank(\boldsymbol{A}^i\boldsymbol{\Phi}^k) = n - 1 \tag{4.10}$$

where n is the number of equations in the system. For the complete model to be identified condition (4.10) must be satisfied for each equation ($k = 1, \cdots, n$) of each regime ($i = 1, \cdots, r$).

Condition (4.10) uniquely identifies the parameters of each equation only up to multiplication by a scalar (see for example Schmidt [16]). To eliminate the remaining indeterminacy we need to specify in addition some normalisation rule. Conventionally we normalise along the diagonal setting

$$B^i_{kk} = -1 \quad (k = 1, \cdots, n). \tag{4.11}$$

However, this particular normalisation rule is quite arbitrary and we are free to normalise each equation on any variable with a non-zero coefficient. FIML procedures are invariant to the rule chosen.

In fact by choosing a different model normalisation in different regimes we can make each regime linear in parameters $\boldsymbol{\theta}$. We can illustrate this with the two equation model developed at the end of the last section. The two regimes with the conventional normalisation can be written as

$$\begin{pmatrix} -1 & a_1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} q_t \\ p_t \end{pmatrix} + \begin{pmatrix} a_0 & 0 \\ a_2 & a_3 \end{pmatrix} \begin{pmatrix} c \\ z_t \end{pmatrix} = \begin{pmatrix} \boldsymbol{u}^1_{1t} \\ \boldsymbol{u}^1_{2t} \end{pmatrix} \tag{4.12}$$

$$\begin{pmatrix} -1 & 0 \\ b_1 & -1 \end{pmatrix} \begin{pmatrix} q_t \\ p_t \end{pmatrix} + \begin{pmatrix} b_2 & b_3 \\ b_0 & 0 \end{pmatrix} \begin{pmatrix} c \\ z_t \end{pmatrix} = \begin{pmatrix} \boldsymbol{u}^2_{1t} \\ \boldsymbol{u}^2_{2t} \end{pmatrix} \tag{4.13}$$

with covariance matrices

$$\boldsymbol{\Sigma}^1 = \begin{pmatrix} \sigma^1_{11} & \sigma^1_{12} \\ \sigma^1_{21} & \sigma^1_{22} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma}^2 = \begin{pmatrix} \sigma^2_{11} & \sigma^2_{12} \\ \sigma^2_{21} & \sigma^2_{22} \end{pmatrix} \quad \text{respectively.}$$

The cross-restrictions between parameters are then

$$(b_0, b_1, \sigma^2_{11}) = (-\frac{a_0}{a_1}, \frac{1}{a_1}, \frac{1}{(a_1)^2}\sigma^1_{11}) \tag{4.14}$$

which are non-linear. However, if we keep the demand relation normalised on q_t in both regimes, instead of (4.13) we have

$$\begin{pmatrix} -1 & b_1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} q_t \\ p_t \end{pmatrix} + \begin{pmatrix} b_0 & 0 \\ b_2 & b_3 \end{pmatrix} \begin{pmatrix} c \\ z_t \end{pmatrix} = \begin{pmatrix} \boldsymbol{u}^2_{1t} \\ \boldsymbol{u}^2_{2t} \end{pmatrix} \tag{4.15}$$

and the cross-restrictions are simply

$$(b_0, b_1, \sigma^2_{11}) = (a_0, a_1, \sigma^1_{11}). \tag{4.16}$$

In general, by preserving the normalisation of equations which remain constant across regimes, which means changing the normalisation of the system from the conventional diagonal normalisation, we avoid non-linear restrictions between the regimes. This proves to be convenient. There is no inherent difficulty in treating non-linearity in parameters if we wish to do so (and we might wish to allow for other non-linear restrictions apart from those arising from the constancy of some relationships across regimes) but this is not necessary. From now on we will assume the mapping from $\boldsymbol{\theta}$ to A^i and $\boldsymbol{\Sigma}^i$ to be linear for each regime.

The numerical problem to be solved, then, is the maximisation of (4.3) with $L_i(\boldsymbol{\theta})$ defined by (4.2) and a linear mapping from $\boldsymbol{\theta}$ to A^i and $\boldsymbol{\Sigma}^i$. Clearly the simplest approach would have been to write a special program for the problem. However, it seemed more useful to attempt to generalise an existing FIML program, to find out how difficult it might prove to implement the necessary modifications within an existing framework. We have seen that cross-regime restrictions may involve the $\boldsymbol{\Sigma}^i$ matrices and in any case we need to impose block diagonality of $\boldsymbol{\Sigma}^i$ for Richard's test for exogeneity. For these reasons the program GENRAM, written by David Hendry [9], was chosen because it already included the option of some restrictions on the covariance matrix.

When we have cross-restrictions on $\boldsymbol{\Sigma}^i$ we cannot in general concentrate the matrix out of the likelihood function and we have to deal with the problem that the elements of $\boldsymbol{\Sigma}^i$ are already constrained by the restrictions of symmetry and positive-definiteness. For the numerical optimisation routine the parameters must be unconstrained so that we cannot work directly with the elements of $\boldsymbol{\Sigma}^i$. We make use of the fact that we can factorise $\boldsymbol{\Sigma}^i$ as

$$\boldsymbol{\Sigma}^i = \boldsymbol{H}^i \boldsymbol{H}^{i\prime} \tag{4.17}$$

where $\boldsymbol{H}^i$ is a lower triangular matrix with $n(\frac{n+1}{2})$ non-zero elements which are unconstrained. Now if we wish to impose restrictions on $\boldsymbol{\Sigma}^i$ these must be translated into restrictions on $\boldsymbol{H}^i$.

We wish to consider two types of restriction. Firstly we have cross-regime covariance restrictions arising from the constancy of some relationships across regime shifts. We assume that the first block of n_1^* equations remain constant over all regimes. Let us partition $\boldsymbol{\Sigma}^i$ and $\boldsymbol{H}^i$ for each regime as

$$\boldsymbol{\Sigma}^i = \begin{pmatrix} \boldsymbol{\Sigma}^i_{11} & \boldsymbol{\Sigma}^i_{12} \\ \boldsymbol{\Sigma}^i_{21} & \boldsymbol{\Sigma}^i_{22} \end{pmatrix} \quad , \quad \boldsymbol{H}^i = \begin{pmatrix} \boldsymbol{H}^i_{11} & \mathbf{0} \\ \boldsymbol{H}^i_{21} & \boldsymbol{H}^i_{22} \end{pmatrix} \tag{4.18}$$

where we note that $\boldsymbol{H}^i_{11}$ and $\boldsymbol{H}^i_{22}$ are square lower triangular matrices. Then the covariance restriction is that $\boldsymbol{\Sigma}^i_{11} = \boldsymbol{\Sigma}^j_{11}, \forall i,j$ which implies $\boldsymbol{H}^i_{11} = \boldsymbol{H}^j_{11}, \forall i,j$.

Secondly we want to apply Richard's test for the exogeneity of the control process for each regime. We saw in section 3 that this implies block diagonality of $\boldsymbol{\Sigma}^i$. There for a single regime we were able to analytically concentrate the likelihood but when we also have to deal with cross-regime restrictions this is no longer possible and we must deal with the restrictions explicitly. Let (4.18) now correspond to a partitioning into blocks of n_1 and n_2 equations ($n_1^* \leq n_1$) where the last block represents the equations of the control process, Then Richard's covariance restriction $\boldsymbol{\Sigma}_{12}^i = \boldsymbol{\Sigma}_{21}^i = \mathbf{0}$, $\forall i$ becomes $\boldsymbol{H}_{21}^i = 0$, $\forall i$.

When there are no restrictions the number of unconstrained $\boldsymbol{\Sigma}$ parameters to be estimated is $rn(\frac{n+1}{2})$. Cross-regime restrictions subtract $(r-1)n_1^*(\frac{n_1^*+1}{2})$ parameters and exogeneity restrictions another rn_1n_2 parameters. However, from the point of view of the order of the numerical problem to be solved, imposing $\boldsymbol{\Sigma}$ restrictions increases the size by

$$\frac{r}{2}\left[n_1(n_1+1)+n_2(n_2+1)\right]-\frac{r-1}{2}\left[n_1^*(n_1^*+1)\right] \tag{4.19}$$

extra parameters. This will become prohibitively expensive for all but small values of n and r.

We can now summarise the extra information needed by the modified program. Let

$$\boldsymbol{\psi}^i = vec\boldsymbol{A}^i \quad \text{and} \quad \boldsymbol{\phi}^i = \boldsymbol{S}^i\boldsymbol{\psi}^i \tag{4.20}$$

where $\boldsymbol{S}^i$ denotes a selection matrix picking out unrestricted elements only. Note that since there are no restrictions between the parameters of $\boldsymbol{A}^i$ and $\boldsymbol{\Sigma}^i$ we can partition $\boldsymbol{\theta}$ into $\boldsymbol{\theta}' = (\boldsymbol{\theta}_1' : \boldsymbol{\theta}_2')$ and write $\boldsymbol{A}^i$ $(\boldsymbol{\theta}_1)$ and $\boldsymbol{\Sigma}^i(\boldsymbol{\theta}_2)$. Then the additional information is:

1. (i) The number of regimes r.
   (ii) Their periods of operation I_i, $i = 1, \cdots, r$.
   (iii) The size of the block of constant equations n_1^*.
   (iv) The size of the block of control equations n_2.

2. For each regime:

   (i) The mapping from $\boldsymbol{\psi}^i$ to $\boldsymbol{\phi}^i$ (i.e. which elements of $\boldsymbol{A}^i$ are unrestricted)
   (ii) The mapping from $\boldsymbol{\phi}^i$ to $\boldsymbol{\theta}_1$ (i.e. which element of $\boldsymbol{\theta}_1$ corresponds to each unrestricted element of $\boldsymbol{A}^i$)
   (iii) The model normalisation (i.e. which elements of $\boldsymbol{A}^i$ are preset to -1)

3. Initial values for the parameter vector $\boldsymbol{\theta}_1$.

The program generates its own initial values for $\boldsymbol{\theta}_2$, the parameters of $\boldsymbol{\Sigma}^i$ and also defines the mapping from $\boldsymbol{\Sigma}^i$ to $\boldsymbol{\theta}_2$ on the basis of 1(iii) and 1(iv).

Now we want to develop tests for the validity of the restrictions we impose on the model. The most general model we consider is a multiple regime model with no cross-regime restrictions and no restrictions on the covariance matrices. We can then find a sequential testing procdure of two nested hypotheses. First we test the hypothesis of the constancy of some chosen sub-block of equations over the regime shifts. Then if this test does not reject the null hypothesis we can go on to test the hypothesis of the exogeneity of the control equations in each regime using Richard's test developed in section 3. This second test then becomes a test of whether or not the economy was controlled.

If, however, the first test rejects the hypothesis of a constant sub-block of equations then each regime is a completely separate model. Of course we can still test for exogeneity in each regime but even if this test fails to reject the null in each case we cannot really interpret this as the authorities switching between instruments in their control of the economy, since the economy being controlled changes with the regime.

The general unrestricted model can be estimated easily using the concentrated log-likelihood function defined by (4.6) – (4.8). Let the maximum of this log-likelihood be L^+. Then we reestimate the model imposing all the cross-regime restrictions. Initial values for the parameters of the covariance matrices can be derived from the solution of the unrestricted model. Let L^{++} be the maximum of this log-likelihood and let p be the number of unrestricted parameters in the first n_1^* rows of the $\boldsymbol{A}^i$ matrix. The first test is then

$$2(L^+ - L^{++}) \sim_a \chi^2_{((r-1)(p+n_1^*(\frac{n_1^*+1}{2})))}. \tag{4.21}$$

Lastly we reestimate the second model imposing the block diagonality on the $\boldsymbol{\Sigma}^i$ matrices for the exogeneity test. The joint test for exogeneity in each regime is

$$2(L^{++} - L^{+++}) \sim_a \chi^2_{(rn_1n_2)} \tag{4.22}$$

where L^{+++} is the maximum log-likelihood of the third model. We can also apply the exogeneity test separately to each regime. Let L_i^{++} and L_i^{+++} be the corresponding maxima for regime i. Then

$$2(L_i^{++} - L_i^{+++}) \sim_a \chi^2_{(n_1n_2)}. \tag{4.23}$$

5 A Model of the Money Market

As an illustration of the general approach already outlined it was decided to estimate a simple two-equation model of the money market to try and model the change in regime of the introduction of Competition and Credit Control (CCC) in October 1971.

Firstly we must explain the economic background to the change. (For a good account of British monetary policy before 1971 see the chapter by Goodhart in

[5]. Goodhart [6] gives a very interesting inside view of the Bank of England thinking behind the change and early experience of operating the new regime.) Before 1971 the control of short term interest rates was the main concern of the monetary authorities. The key to their control was the Bank Rate which was an institutional rate set directly by the Bank of England. Other short term interest rates moved closely in line with it; in particular a Clearing banks cartel existed which linked the interest rates on Advances and time deposits directly to Bank Rate. Changes in the Bank Rate came to have a strong "announcement effect" signalling the direction of monetary policy. On the other hand there was no attempt to directly control the main monetary aggregates (which only came to be regarded by the Bank as important indicators towards the end of the 1960s) although starting in 1964 there was some attempt to reduce bank lending to the private sector, especially persons, in a series of formal "Requests" to the Clearing banks. However, these were quantitative ceilings imposed on the market rather than operating through it. We can distinguish four main targets of monetary policy: the balance of payments, the rate of unemployment, the rate of inflation and, in the longer term, the growth rate. Of these the balance of payments was perhaps the most important. (For a statement of the aims of monetary policy by the authorities themselves see the Radcliffe Report [14]).

The effect of the 1971 changes was to switch emphasis to control of the monetary aggregates. In 1972 the Bank Rate was replaced by a Minimum Lending Rate (MLR) which is determined by the market (The Bank sets it at 0.5 above the current Treasury bill rate). The Clearing bank cartel was abolished and the banks encouraged to compete for funds, thus allowing interest rates to move freely in response to market presures. In this way control over interest rates was relaxed in favour of control over the monetary aggregates, although the mechanism of control was viewed as being through portfolio adjustment in response to changes in asset relative prices. (In the new spirit of free competition the formal ceilings on lending were abandoned although in 1973 they were reintroduced in a new guise as the Supplementary Special Deposit scheme or "corset".) The targets of monetary policy since 1971 have been essentially the same as before, with perhaps less emphasis on the balance of payments, although the managed floating of the exchange rate in 1972 has not alleviated all external balance problems.

We attempted to model this complicated policy change in a simple two regime model of two equations: the demand for money equation and a reaction function for the control variable (interest rate in the first regime and money supply in the second) although clearly, with such an oversimplified model, our results can only be very tentative. Previous work on reaction functions for the instruments of monetary policy for the U.K. has been done by Douglas Fisher [3], [4] and also by Goodhart (in his appendix to [5]). The reaction function treats a policy control variable as a function of the target variables which are the objects of control.

The coefficients then measure the response of the control variable to changes in the target variables. For our policies, in order to test for the exogeneity of the control variables, our reaction functions must be purely predetermined and we may interpret this as reflecting the necessary lags in the reaction of the authorities.

Rather than try to estimate the model from the start by FIML it was decided to do preliminary single equation estimation to choose the dynamic specification of the equations. This allowed access to the much fuller diagnostic tests available in programs such as GIVE (Hendry [10]), especially tests for serial correlation, which was important since we wanted to avoid dealing with this problem in the full simultaneous estimation of the two regime model. Also the preliminary estimates obtained provided us with initial values to use in the FIML program. In determining the dynamic structure the methodology of Hendry and Mizon [11] was adopted. Economic theory gives us little help in choosing dynamic specification and, rather than impose the dynamics a priori, we follow a consistent data-based approach. Starting from the most general hypothesis we are prepared to accept (generally a maximum lag length on all variables) we can then test for parameter restrictions in a series of Wald tests until we reach the most parsimonious description consistent with the data. There is no unique sequence for testing and the particular restrictions we test may often be governed by what seems intuitively plausible or appealing. However, this procedure should prevent us from imposing invalid restrictions.

One objection sometimes raised against this approach is the problem of multicollinearity in parameter estimates in the most general equation forms. In practice this did not greatly hinder the simplification process since, although t-ratios in these cases were in general quite small using conventional significance levels, the relative sizes of t-ratios were found to be still quite a good indication of the relative importance of lags. A more critical practical problem was degrees of freedom (especially for the reaction functions) which restricted the maximum lag length which could reasonably be estimated.

Seasonally unadjusted data was used. This avoided the problem of possible distortion of the true dynamic strucure arising from using series separately seasonally adjusted using different filters (see Wallis [17]). For the money demand equation it also gave us an interesting comparison with the results of Hendry and Mizon who used adjusted data.

The two endogenous variables used in the analysis were personal sector M3 and the Local Authority short term interest rate. It is sometimes argued that the appropriate monetary indicator is narrow money M1 rather than M3 which includes an interest bearing asset, time deposits. However, the view of the Bank at the time (see for example Goodhart [6]) was that, because of the ease of shifting between time deposit and current accounts, M1 could not be controlled, so that M3 was the appropriate indicator for control. Total M3 was tried but this was not very successful (even in an unrestricted specification with up to 8 quarter lags on all variables there was significant residual autocorrelation) and personal sector M3 was found to perform much better. This accords with the results of studies by the Bank of England (see Hacche [8]). Finding a single interest rate with the dual property of being both monetary control variable and a measure of the true opportunity cost of holding money proved difficult. Several short rates were tried but finally the Local Authority rate was chosen as the best proxy.

The general unrestricted money demand equation took the form

$$\ln M_t = \sum_{j=0}^{J} (\alpha_j \ln Y_{t-j} + \beta_j \ln P_{t-j} + \gamma_j \ln r_{t-j} + \delta_j \ln M_{t-j-1}) + c_0 + \sum_{i=1}^{3} c_i q_{t-i} + e_t \qquad (5.1)$$

where

$M_t =$ personal sector M3

$Y_t =$ personal disposable income at constant 1975 prices

$P_t =$ the deflator for Y_t

$r_t =$ the Local Authority interest rate as a percentage

and q_{t-i}, $i = 1, \cdots, 3$ are the usual seasonal dummies. Data was available quarterly over the period 5501-7802. Estimating (5.1) by OLS over the full data period with $J = 5$ gave the results in Table 1. The three χ^2 statistics quoted are tests for residual autocorrelation: χ^2_{12} is the Box-Pierce random residual correlogram test (see Pierce [13]), χ^2_6 is a Lagrange multiplier test for serial correlation up to the 6th order, χ^2_1 is the squared Durbin h-statistic for a first order AR process (Durbin [2]). These statistics show no evidence of dynamic misspecification or serial correlation in the residuals, confirming that choice of $J = 5$ is adequately general (as might be expected for quarterly data). Perhaps the most striking feature of the equation is the strong affirmation of unit price elasticity suggesting a reparameterisation with $\Delta \ln(M/P)$ as the dependent variable. Comparison of Table 1 with Table 1 in Hendry and Mizon (estimated with $J = 4$ over a much shorter period 6301-7503) shows much similarity in the lag patterns suggesting that the seasonal bias in using adjusted data may not be very important in this case. In particular there is the same $\Delta \ln r_t$ term with perverse sign, although in general the interest rate terms in our equation are much less significant (Hendry and Mizon used the consol yield)

To guard against possible simultaneous equation bias in using OLS, equation (5.1) was reestimated using instrumental variables for r_t. Note that in the complete two equation model if $\boldsymbol{\Sigma}$ is not diagonal then OLS on M_t will be inconsistent

Table 1: Equation 5.1 with $J = 5$. OLS 5603–7802

$j \backslash Vble$	$\ln M_{t-j-1}$	$\ln r_{t-j}$	$\ln Y_{t-j}$	$\ln P_{t-j}$
0	1.07 (.13)	.011 (.013)	.33 (.11)	.85 (.33)
1	−.28 (.18)	−.019 (.019)	−.15 (.13)	−.95 (.45)
2	.06 (.18)	.015 (.020)	.06 (.13)	−.03 (.47)
3	.20 (.19)	−.017 (.019)	−.13 (.13)	−.08 (.45)
4	−.26(.18)	.027 (.020)	.04 (.13)	.59 (.45)
5	.06 (.13)	−.012 (.015)	−.01 (.12)	−.29 (.29)
$\sum_j$	.85	.005	.14	.09

$c_0 = -.53\,(.25)$ $\quad c_1 = -.001\,(.011)$ $\quad c_2 = -.003\,(.010)$ $\quad c_3 = .008\,(.011)$

$R^2 = .9992$ $\quad \widehat{\sigma} = .0159$ $\quad \chi^2_{12} = 3.84$ $\quad \chi^2_6 = 8.18$ $\quad \chi^2_1 = 1.28$

(standard errors in parentheses)

but if $\mathbf{\Sigma}$ is diagonal (Richard's condition satisfied) then OLS is maximum likelihood. Full 2SLS using as instruments all predetermined variables in the most general specification of both equations (i.e. including all lags up to $J = 5$) would have exceeded the program dimension limitations but it was possible to use $j = 1, \cdots, 4$ on all the additional instruments (giving 39 instruments in all). However, the results were virtually identical to Table 1 with the interest rate terms coming out even less significant.

Before attempting to simplify the unrestricted equation (5.1) it was decided to test its stability over the regime change in October 1971. It is well known that the introduction of CCC was followed by a very large expansion in M3 (most especially in company holdings but also to a lesser extent in personal sector holdings), an expansion not matched in the narrow money definition M1, and Hacche [8] and Goodhart [6] reported a complete breakdown in the Bank of England's own forecasting equations. Since one of the effects of CCC was to encourage banks to compete for finance through the rates paid on deposit accounts, Hacche argued convincingly that this breakdown could be explained by the omission of own-rate terms from the Bank's demand equations, terms which would not have mattered before 1971 when the differential between deposit rates and the Bank Rate was constant. However, own-rate terms were not found to be significant either by Hacche or Artis and Lewis [1] who explained the expansion of M3 as a disequilibrium excess supply of money. (In an earlier equation for total M3 we tested the importance of own-rates by including interest differential terms of the form $\ln(r^*/r)_{t-j}$ where r^* was the time deposit rate. However, these terms came up with perverse negative signs).

A Chow test was carried out on equation 5.1 reestimated over the period ending

Table 2: Solved Coefficients from Equation 5.2. OLS 5603-7802

$j\backslash Vble$	$\ln M_{t-j-1}$	$\ln r_{t-j}$	$\ln Y_{t-j}$	$\ln P_{t-j}$
0	1.10	.011	.27	.65
1	$-.14$	$-.011$	$-.23$	$-.75$
2	0.0	0.0	0.0	$-.41$
3	.26	0.0	0.0	.15
4	$-.26$	.009	0.0	.54
5	0.0	0.0	0.0	$-.14$
$\sum_j$	.96	.009	.04	.04

7103. The statistic was $F_{(27,60)} = 2.62$, the 5 significance level for $F_{(24,60)}$ being 1.70, so that the hypothesis of parameter stability was rejected. This probably reflects the clearly inadequate proxying of interest effects in the equation.

Simplifying equation (5.1) we obtained

$$
\begin{array}{llll}
\Delta \ln(M/P)_t = & 27\Delta \ln Y_t & +.14\Delta \ln(M/P)_{t-1} & -.35\Delta \ln P_t \\
 & (.08) & (.10) & (.20) \\
 & -.55\Delta \ln P_{t-2} & -.40\Delta^2 \ln P_{t-2} & +.26\Delta \ln(M/P)_{t-4} \\
 & (.20) & (.19) & (.10) \\
 & +.011\Delta \ln r_t & +.009 \ln r_{t-4} & -.04 \ln(M/PY)_{t-1} \\
 & (.011) & (.008) & (.02)
\end{array}
$$

$$c_0 = -.17\ (.07) \quad c_1 = -.01\ (.01) \quad c_2 = -.01\ (.01) \quad c_3 = -.003\ (.005)$$

$$R^2 = .644 \quad \widehat{\sigma} = .0151 \quad \chi^2_{12} = 6.86 \quad \chi^2_6 = 4.51 \quad \chi^2_1 = .002. \tag{5.2}$$

The χ^2 tests still reject the hypothesis of serial correlation in the residuals and the F-ratio test for the restrictions imposed on (5.1) is $F_{(15,60)} = .533$. (The 5 significance level for $F_{(15,60)}$ is 1.84). The solved coefficients from (5.2) are recorded in Table 2 and these quite closely correspond with the coefficients on the unrestricted equation. However, like Hendry and Mizon's equation 21, the specification (5.2) was primarily chosen so as to include decision variables with sensible economic interpretations. In the short-run inflation has a negative influence on money demand through three inflation terms , including a term for the rate of acceleration of inflation $\Delta^2 \ln P_{t-2}$. However, in long run equilibrium when all variables are growing along steady-state growth paths the equation exhibits unit elasticity with respect to both prices and income through the inverse velocity levels term $\ln(M/PY)_{t-1}$. This term also serves as an error feeedback correction mechanism in the short run, allowing agents to adjust from previous disequilibrium in the relationship betwen their real income and money holdings.

In the long run we would expect some levels interest rate effect. However, the coefficient on $\ln r_{t-4}$ in (5.2) has the "wrong" sign and is statistically insignificant, its value being approximately equal to the sum of the interest rate coefficients on the unrestricted equation. This again seems to reflect the unsatisfactory nature of the Local Authority rate (perhaps any single rate) as a proxy for the opportunity cost to the personal sector of holding money. It is clear that to have gone on to try and produce a fully satisfactory equation would have required much further work experimenting with various interest rates (possibly including several different rates) in an attempt to model the portfolio choice properly. For the purpose of this exercise it was decided to accept (5.2), however unsatisfactory, as the best equation we had been able to find.

In approaching the estimation of the reaction functions we had immediately to face the problem that with only 27 observations for the second regime (7104-7802) it was quite impossible to estimate an unrestricted function. Instead we chose the dynamic specification for the first regime on the basis of an unrestricted equation and then were forced to assume that this specification remained the same after the regime switch. We note that this assumption is a very strong one since we might expect the reaction lags of different control instruments to be very different.

The general hypothesis for the first regime was

$$\begin{aligned}\ln r_t = &\sum_{j=0}^{J}(a_j \ln(B/P)_{t-j} + b_j \ln U_{t-j} + d_j \ln P_{t-j} + e_j \ln E_{t-j}) \\ &+ \sum_{k=1}^{K}(f_k \ln r_{t-k} + g_k \ln M_{t-k}) + c_0 + \sum_{i=1}^{3} c_i q_{t-i} + \varepsilon_t \end{aligned} \tag{5.3}$$

where

U_t = the percentage rate of unemployment
B_t = the level of reserves
E_t = the Eurodollar rate.

The Eurodollar rate was included on the hypothesis that a long run objective of interest rate control would be to keep U.K. interest rates in line with foreign rates. Equation (5.3) was estimated over the period 5602-7103 taking $J = K = 4$ giving the results in Table 3. The χ^2 tests show no evidence of residual autocorrelation. Having little a priori theory to guide us to a sensible reparameterisation the equation was simplified by choosing the most important lags from Table 3. The resulting equation was

Table 3: Equation 5.3 with $J = K = 4$. OLS 5602–7103

$j \backslash Vble$	$\ln r_{t-j-1}$	$\ln M_{t-j-1}$	$\ln P_{t-j}$	$\ln U_{t-j}$	$\ln(B/P)_{t-j}$	$\ln E_{t-j}$
0	.83 (.18)	−.22 (1.19)	6.01 (3.2)	−.03 (.24)	−.09 (.19)	.17 (.18)
1	−.35 (.22)	.89 (1.36)	−7.51 (4, 4)	−.35 (.39)	−.42 (.25)	−.10 (.29)
2	.19 (.23)	.70 (1.39)	−.07 (4.2)	.49 (.41)	.01 (.23)	.21 (.31)
3	.27 (.17)	−1.20 (1.05)	−5.24 (4.0)	−.72 (.43)	−.07 (.25)	−.37 (.29)
4	—	—	6.22 (3.3)	.48 (.26)	.33 (.24)	.24 (.19)
$\sum_j$	.94	.17	−.59	−.13	−.24	.15

$c_0 = 1.40\,(2.47)$ $\quad c_1 = -.11\,(.13)$ $\quad c_2 = -.14\,(.12)$ $\quad c_3 = -.02\,(.14)$

$R^2 = .916$ $\quad \widehat{\sigma} = .1006$ $\quad \chi^2_{12} = 4.79$ $\quad \chi^2_6 = 6.14$ $\quad \chi^2_1 = .03$

Table 4: Solved Coefficients from Equation 5.4. OLS 5602–7103

$j \backslash Vble$	$\ln r_{t-j-1}$	$\ln M_{t-j-1}$	$\ln P_{t-j}$	$\ln U_{t-j}$	$\ln(B/P)_{t-j}$	$\ln E_{t-j}$
0	1.008	−.19	2.17	0.0	0.0	0.0
1	−.47	.19	−1.98	−.53	−.41	−.008
2	.47	0.0	−.19	.53	0.0	0.0
3	0.0	0.0	−3.78	−.37	0.0	0.0
4	0.0	0.0	3.78	.37	.41	0.0
$\sum_j$	1.008	0.0	0.0	0.0	0.0	−.008

$$\begin{array}{llll}
\Delta \ln r_t = & \underset{(2.10)}{2.17 \Delta \ln P_t} & \underset{(2.35)}{-3.78 \Delta \ln P_{t-3}} & \underset{(.16)}{-.53 \Delta \ln U_{t-1}} \\
 & \underset{(.17)}{-.37 \Delta \ln U_{t-3}} & \underset{(.13)}{-.47 \Delta \ln r_{t-2}} & \underset{(.10)}{-.41 \Delta_3 \ln(B/P)_{t-1}} \\
 & \underset{(.81)}{-.19 \Delta \ln(M/P)_{t-1}} & \underset{(.008)}{+.008 \ln(r/E_{-1})_{t-1}} &
\end{array}$$

$$c_0 = .07\ (.04) \quad c_1 = -.05\ (.06) \quad c_2 = -.05\ (.04) \quad c_3 = -.07\ (.07)$$

$$R^2 = .526 \quad \widehat{\sigma} = .0954 \quad \chi^2_{12} = 12.28 \quad \chi^2_6 = 2.96 \quad \chi^2_1 = .19. \tag{5.4}$$

The F-test on the imposed restrictions gave $F_{(20,30)} = .746$ (5 significance level $F_{(20,30)} = .1.93$). Although this test failed to reject the restrictions imposed in (5.4) the solved coefficients in Table 4 are rather different from those in Table 3. This is especially true of the inflation terms which have changed considerably in magnitude. The last term in (5.4) was a levels effect intended to capture a long run relationship between U.K. and foreign interest rates. However, it was found to be insignificant.

The specification (5.4) was reestimated over the period 7104-7802 with $\Delta \ln(M/P)$ as the dependent variable. This gave the equation

$$\begin{array}{llll}
\Delta \ln(M/P)_t = & -.95\Delta \ln P_t & -.29\Delta \ln P_{t-3} & -.04\Delta \ln U_{t-1} \\
 & (.18) & (.19) & (.03) \\
 & -.05\Delta \ln U_{t-3} & -.004\Delta \ln r_{t-2} & -.01\Delta_3 \ln(B/P)_{t-1} \\
 & (.03) & (.01) & (.005) \\
 & +.11\Delta \ln(M/P)_{t-1} & -.02 \ln(r/E_{-1})_{t-1} & \\
 & (.16) & (.008) &
\end{array}$$
$$c_0 = .06\ (.01) \quad c_1 = -.03\ (.008) \quad c_2 = .0004\ (.007) \quad c_3 = -.02\ (.009)$$
$$R^2 = .912 \quad \widehat{\sigma} = .0098 \quad \chi^2_{12} = 16.72. \tag{5.5}$$

Caution must be exercised in the interpretation of this equation because of the number of degrees of freedom (only 15). This means that the very good fit (high R^2, small $\widehat{\sigma}$) is probably spurious. Also the χ^2_{12} statistic is large compared with our previous equations, although still not significant. (The other χ^2 tests were unfortunately not available for this equation). Comparing with (5.4) we see that, although most of the coefficients have changed sign (as expected) a few (as with equation (5.4)) seem to have perverse signs.

These reaction functions are not very satisfactory, in particular because of the ad hoc way in which the dynamic specification was chosen from the unrestricted equation (5.3). Nevertheless, our purpose being illustrative only, the specifications (5.4) and (5.5) were used, together with the demand equation specification (5.2) for the purpose of a FIML estimation of the complete model to illustrate the application of Richard's test for exogeneity.

The complete model was estimated three times to carry out the two likelihood ratio tests described at the end of section 4. Firstly, the log-likelihood was maximised without any cross-restrictions, allowing the coefficients on the money demand equation to take different values in the two regimes. (This corresponds to L^+ in section 4). Then the model was reestimated constraining the coefficients on the demand equation to be the same for both regimes. (This corresponds to L^{++}). Finally, the second model was reestimated imposing Richard's exogeneity restrictions which in this model implies constraining the two covariance matrices $\mathbf{\Sigma}^i$ to be diagonal. (This corresponds to L^{+++}). The exact log-likelihoods obtained were

$$L^+ = 292.06 \quad , \quad L^{++} = 276.69 \quad , \quad L^{+++} = 267.23 \tag{5.6}$$

so that the two χ^2 tests corresponding to (4.21) and (4.22) were

$$\chi^2_{(9)} = 30.74 \quad \text{and} \quad \chi^2_{(2)} = 18.92 \tag{5.7}$$

respectively.

The first test convincingly rejects the hypothesis of a stable money demand function, confirming the result of the Chow test done on the single equation. (This test has only 9 degrees of freedom because the additional cross-restriction on the covariance terms that $\sigma^1_{11} = \sigma^2_{11}$ was not imposed). We have argued above that rejection of the null on this first test means that we cannot regard Richard's joint exogeneity test as a test of whether or not the economy was being controlled. Nevertheless, it is still interesting that Richard's test strongly rejects the hypothesis that the instruments of control were exogenous for this model imposing a stable money demand function. The individual χ^2 tests on each regime (corresponding to (4.23)) were

$$\chi^2_{(1)} = 13.52 \quad \text{and} \quad \chi^2_{(1)} = 5.40 \tag{5.8}$$

respectively, showing that the rejection of the null was much stronger for the first regime.

We also applied Richard's test for exogeneity to each regime estimately separately (allowing a shift in the money demand equation between the two regimes). This gave the statistics

$$\chi^2_{(1)} = 2.46 \quad \text{and} \quad \chi^2_{(1)} = .114 \tag{5.9}$$

showing remarkably different results from the previous tests. Taking these results at face value the implication is clear: in both regimes separately the control instruments appear to have been exogenous but the regime change itself was accompanied by a shift in the money demand function. We should be cautious about putting too much weight on this conclusion without further evidence (in particular it would have been interesting to test the exogeneity of the money supply in the first regime and the interest rate in the second). Before accepting the hypothesis of a shift in the money demand function we would like to have done more work with different relative interest rates to model the portfolio decision more carefully. Nevertheless, the hypothesis of a change in the behaviour of the personal sector cannot be ruled out, especially since one explicit purpose of CCC was to bring about a change in the behaviour of the banks.

The FIML estimates for the second model (with the demand equation constrained to be the same in both regimes) were as follows:

$$\begin{array}{llll} \Delta \ln(M/P)_t = & 31\Delta \ln Y_t & -.003\Delta \ln(M/P)_{t-1} & +.17\Delta \ln P_t \\ & (.08) & (.02) & (.22) \\ & +.015\Delta \ln P_{t-2} & -.02\Delta^2 \ln P_{t-2} & +.28\Delta \ln(M/P)_{t-4} \\ & (.027) & (.03) & (.11) \\ & +.11\Delta \ln r_t & +.003 \ln r_{t-4} & +.03 \ln(M/PY)_{t-1} \\ & (.03) & (.01) & (.03) \end{array} \tag{5.10}$$

$$
\begin{array}{llll}
\Delta \ln r_t = & .18\Delta \ln P_t & +.24\Delta \ln P_{t-3} & -.50\Delta \ln U_{t-1} \\
& (.17) & (.19) & (.13) \\
& +.01\Delta \ln U_{t-3} & -.24\Delta \ln r_{t-2} & -.12\Delta_3 \ln(B/P)_{t-1} \\
& (.17) & (.11) & (.08) \\
& +.23\Delta \ln(M/P)_{t-1} & -.02 \ln(r/E_{-1})_{t-1} & \\
& (.12) & (.05) &
\end{array}
\tag{5.11}
$$

for regime 1 and

$$
\begin{array}{llll}
\Delta \ln(M/P)_t = & 1.83\Delta \ln P_t & +.10\Delta \ln P_{t-3} & -.17\Delta \ln U_{t-1} \\
& (.14) & (.11) & (.11) \\
& -.03\Delta \ln U_{t-3} & +.006\Delta \ln r_{t-2} & +.007\Delta_3 \ln(B/P)_{t-1} \\
& (.08) & (.04) & (.02) \\
& +.20\Delta \ln(M/P)_{t-1} & -.06 \ln(r/E_{-1})_{t-1} & \\
& (.07) & (.03) &
\end{array}
\tag{5.12}
$$

for regime 2, with covariance matrices

$$
\mathbf{\Sigma}^1 = \begin{pmatrix} .0005 & -.0014 \\ -.0014 & .0103 \end{pmatrix} \quad , \quad \mathbf{\Sigma}^2 = \begin{pmatrix} .0005 & .0004 \\ .0004 & .0014 \end{pmatrix} . \tag{5.13}
$$

Comparison with the single equation estimates ((5.2), (5.4) and (5.5)) shows considerable differences in the significance as well as the magnitude of individual coefficients.

Experience with the FIML program highlighted the importance of good initial values in the iteration routine if the speed of convergence was to be increased. Two algorithms were used: Gill-Murray-Pitfield which was fast from good initial values but sometimes would fail to converge if started at all far away from a maximum, and the Powell conjugate directions algorithm which was much more robust but considerably slower (sometimes taking up to 40 iterations or 15-20 seconds on the CDC 7600 computer).

References

[1] Artis, M. J., and Lewis, M. K. The demand for money: stable or unstable. *The Banker 124* (1974), 239–247.

[2] Durbin, J. Testing for serial correlation in least squares regression when some of the regressors are lagged dependent variables. *Econometrica 38* (1970), 410–421.

[3] Fisher, D. The demand for money in britain: quarterly results 1951 to 1967. *Manchester School 36* (1968), 329–344.

[4] Fisher, D. The instruments of monetary policy and the generalised trade-off function for Britain, 1955–1968. *Manchester School 38* (1970), 209–222.

[5] Goodhart, C. A. E. Monetary policy in the United Kingdom. In *Monetary Policy in Twelve Industrial Countries*, K. Holbik, Ed. Federal Reserve Bank of Boston, Boston, MA, 1980, ch. 12.

[6] Goodhart, C. A. E. Problems of monetary management: the U.K. experience. In *Inflation, Depression and Economic Policy in the West, Lessons from the 1970's*, S. C. A, Ed. Basil Blackwell, Oxford, 1980.

[7] Granger, C. W. J. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica 37* (1969), 424–438.

[8] Hacche, G. The demand for money in the United Kingdom: experience since 1971. *Bank of England Quarterly Bulletin 14* (1974), 284–305.

[9] Hendry, D. F. *GENRAM. A generalised maximum likelihood estimation program for economertic systems*, 1978. London School of Economics Computer Unit.

[10] Hendry, D. F. *Technical manual for GIVE*, 1978. London School of Economics Computer Unit.

[11] Hendry, D. F., and Mizon, G. E. Serial correlation as a convenient simplification, not a nuisance. *Economic Journal 88* (1979), 549–563.

[12] Maddala, G. S., and Nelson, F. D. Maximum likelihood methods for models of markets in disequilibrium. *Econometrica 42* (1974), 1013–1030.

[13] Pierce, D. A. Distribution of residual autocorrelations in the regression model with autoregressive-moving average errors. *Journal of the Royal Statistical Society, Series B 33* (1971), 140–146.

[14] Report, R. *The Report of the Radcliffe Committee on the Working of the Monetary System.* Her Majesty's Stationary Office, London, 1959. Cmnd. 827.

[15] Richard, J.-F. Models with several regimes and changes in exogeneity. *Review of Economic Studies* (1979). (forthcoming).

[16] Schmidt, P. *Econometrics.* Marcel Dekker Inc., New York, 1976. Statistics: textbooks and monographs Volume 18.

[17] Wallis, K. F. Seasonal adjustment and relations between variables. *Journal of the American Statistical Association 69* (1974), 18–31.