OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 52, 2 (1990)
0305-9049 $3.00

MAXIMUM LIKELIHOOD ESTIMATION AND INFERENCE ON COINTEGRATION – WITH APPLICATIONS TO THE DEMAND FOR MONEY

Søren Johansen, Katarina Juselius

I. INTRODUCTION

1.1. Background

Many papers have over the last few years been devoted to the estimation and testing of long-run relations under the heading of *cointegration*, Granger (1981), Granger and Weiss (1983), Engle and Granger (1987), Stock (1987), Phillips and Ouliaris (1986), (1987), Johansen (1988b), (1989), Johansen and Juselius (1988), *canonical analysis*, Box and Tiao (1981), Velu, Wichern and Reinsel (1987), Pena and Box (1987), *reduced rank regression*, Velu, Reinsel and Wichern (1986), and Ahn and Reinsel (1987), *common trends*, Stock and Watson (1987), *regression with integrated regressors*, Phillips (1987), Phillips and Park (1986a), (1988b), (1989), as well as under the heading *testing for unit roots*, see for instance Sims, Stock, and Watson (1986). There is a special issue of this *BULLETIN* (1986) dealing mainly with cointegration and a special issue of the Journal of Economic Dynamics and Control (1988) dealing with the same problems.

We start with a vector autoregressive model (cf. (1.1) below) and formulate the hypothesis of cointegration as the hypothesis of reduced rank of the long-run impact matrix $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$. The main purpose of this paper is to demonstrate the method of maximum likelihood in connection with two examples. The results concern the calculation of the maximum likelihood estimators and likelihood ratio tests in the model for cointegration under linear restrictions on the cointegration vectors $\boldsymbol{\beta}$ and weights $\boldsymbol{\alpha}$. These results are modifications of the procedure given in Johansen (1988b) and apply the multivariate technique of partial canonical correlations, see Anderson (1984) or Tso (1981).

For inference we apply the results of Johansen (1989) on the asymptotic distribution of the likelihood ratio test. These distributions are given in terms of a multivariate Brownian motion process and are tabulated in the Appendix. Inferences on $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ under linear restrictions can be conducted using the usual χ^2 distribution as an approximation to the distribution of likelihood ratio test. We also apply the limiting distribution of the maximum likelihood estimator to a Wald test for hypotheses about $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

1.2. The Statistical Model

Consider the model

$$H_1: \mathbf{X}_t = \mathbf{\Pi}_1 \mathbf{X}_{t-1} + \ldots + \mathbf{\Pi}_k \mathbf{X}_{t-k} + \boldsymbol{\mu} + \mathbf{\Phi} \mathbf{D}_t + \boldsymbol{\varepsilon}_t, (t = 1, \ldots, T), \tag{1.1}$$

where $\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_T$ are $\mathrm{IIN}_p(\mathbf{0}, \mathbf{\Lambda})$ and $\mathbf{X}_{-k+1}, \ldots, \mathbf{X}_0$ are fixed. Here the variables $\mathbf{D}_t$ are centered seasonal dummies which sum to zero over a full year. We assume that we have quarterly data, such that we include three dummies and a constant term. The unrestricted parameters $(\boldsymbol{\mu}, \mathbf{\Phi}, \mathbf{\Pi}_1, \ldots, \mathbf{\Pi}_k, \mathbf{\Lambda})$ are estimated on the basis of T observations from a vector autoregressive process. For a p-dimensional process with quarterly data this gives Tp observations and $p + 3p + kp^2 + p(p+1)/2$ parameters.

In general, economic time series are non-stationary processes, and VAR-systems like (1.1) have usually been expressed in first differenced form. Unless the difference operator is also applied to the error process and explicitly taken account of, differencing implies loss of information in the data. Using $\Delta = 1 - L$, where L is the lag operator, it is convenient to rewrite the model (1.1) as

$$\Delta \mathbf{X}_t = \mathbf{\Gamma}_1 \Delta \mathbf{X}_{t-1} + \ldots + \mathbf{\Gamma}_{k-1} \Delta \mathbf{X}_{t-k+1} + \mathbf{\Pi} \mathbf{X}_{t-k} + \boldsymbol{\mu} + \mathbf{\Phi} \mathbf{D}_t + \boldsymbol{\varepsilon}_t, \tag{1.2}$$

where

$$\mathbf{\Gamma}_i = -(\mathbf{I} - \mathbf{\Pi}_1 - \ldots - \mathbf{\Pi}_i), \qquad (i = 1, \ldots, k-1),$$

and

$$\mathbf{\Pi} = -(\mathbf{I} - \mathbf{\Pi}_1 - \ldots - \mathbf{\Pi}_k). \tag{1.3}$$

Notice that model (1.2) is expressed as a traditional first difference VAR-model except for the term $\mathbf{\Pi X}_{t-k}$. It is the main purpose of this paper to investigate whether the coefficient matrix $\mathbf{\Pi}$ contains information about long-run relationships between the variables in the data vector. There are three possible cases:

(i) Rank($\mathbf{\Pi}$) $= p$, i.e. the matrix $\mathbf{\Pi}$ has full rank, indicating that the vector process $\mathbf{X}_t$ is stationary.
(ii) Rank($\mathbf{\Pi}$) $= 0$, i.e. the matrix $\mathbf{\Pi}$ is the null matrix and (1.2) corresponds to a traditional differenced vector time series model.
(iii) $0 < \mathrm{rank}(\mathbf{\Pi}) = r < p$ implying that there are $p \times r$ matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ such that $\mathbf{\Pi} = \boldsymbol{\alpha\beta}'$.

The cointegration vectors $\boldsymbol{\beta}$ have the property that $\boldsymbol{\beta}'\mathbf{X}_t$ is stationary even though $\mathbf{X}_t$ itself is non-stationary. In this case (1.2) can be interpreted as an error correction model, see Engle and Granger (1987), Davidson (1986) or Johansen (1988a). Thus the main hypothesis we shall consider here is the hypothesis of r cointegration vectors

$$H_2: \mathbf{\Pi} = \boldsymbol{\alpha\beta}', \tag{1.4}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $p \times r$ matrices.

We further investigate linear hypotheses expressed in terms of the coefficients $\boldsymbol{\mu}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, and in particular the relation between the constant term and the reduced rank matrix $\boldsymbol{\Pi}$. If $\boldsymbol{\Pi}$ is restricted as in H_2, see (1.4) and $\boldsymbol{\mu} \neq \mathbf{0}$ the non-stationary process $\mathbf{X}_t$ has linear trends with coefficients which are functions of $\boldsymbol{\mu}$ only through $\boldsymbol{\alpha}'_\perp \boldsymbol{\mu}$, where $\boldsymbol{\alpha}_\perp$ is a $p \times (p-r)$ matrix of vectors chosen orthogonal to $\boldsymbol{\alpha}$. Thus the hypothesis $\boldsymbol{\mu} = \boldsymbol{\alpha}\boldsymbol{\beta}'_0$, or alternatively $\boldsymbol{\alpha}'_\perp \boldsymbol{\mu} = \mathbf{0}$, is the hypothesis about the absence of a linear trend in the process. Note that when $\boldsymbol{\mu} = \boldsymbol{\alpha}\boldsymbol{\beta}'_0$ we can write

$$\boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{X}_{t-k} + \boldsymbol{\mu} = \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{X}_{t-k} + \boldsymbol{\alpha}\boldsymbol{\beta}'_0 = \boldsymbol{\alpha}\boldsymbol{\beta}^{*\prime}\mathbf{X}^*_{t-k},$$

where $\boldsymbol{\beta}^* = (\boldsymbol{\beta}', \boldsymbol{\beta}'_0)'$ and $\mathbf{X}^*_{t-k} = (\mathbf{X}'_{t-k}, 1)'$. This is useful for the calculations. Since the asymptotic distributions of the test statistics and estimators depend on which assumption is maintained, it is important to choose the appropriate model formulation. This has been pointed out for instance by West (1989), Dolado and Jenkinson (1988). The mathematical results for the multivariate model (1.2) are given in Johansen (1989).

1.3. The Data

We have chosen to illustrate the procedures by data from the Danish and Finnish economy on the demand for money.[1] The relation $m = f(y, p, c)$ expresses money demand m as a function of real income y, price level p and the cost of holding money c. Price homogeneity was first tested and since it was clearly accepted by data the empirical analysis here will be for real money, real income and some proxies measuring the cost of holding money. Money, income and prices were measured in logarithms, since multiplicative effects are assumed.

The two data sets differ both as to which variables are included and the length of the sample. More interestingly, however, the institutional relations in the two economies have been quite different in the sample period. In Denmark, financial markets have been much less restricted than in Finland, where both interest rates and prices have been subject to regulation for most of the sample period. One would expect this to show up in the empirical results and it does.

For the Danish data the sample is 1974.1–1987.3. As a proxy for money demand $m2$ was chosen because the data available on a quarterly basis are based on more homogeneous definitions for $m2$ than for $m1$. The cost of holding money was assumed to be approximately measured by the difference between the bank deposit rate, i^d, for interest bearing deposits (which are the main part of $m2$) and the bond rate, i^b, which plays an important role in the Danish economy. The two interest rates were included unrestrictedly in the analysis, but subsequently tested for equal coefficients with opposite signs. The inflation rate, Δp, was also included as a possible proxy for the cost of

[1] For a general review of theoretical and empirical results on the demand for money, see for instance Laidler (1985).

holding money, but since it did not enter significantly into the cointegration relation for money demand it was omitted from the present analysis.

For the Finnish data the sample is 1958.1–1984.3. In this case $m1$ was chosen since the $m1$ cointegration relation was found to enter the demand for money equation more significantly and hence illustrated the methodology better. Since interest rates have been regulated, a good proxy for the actual costs of holding money is difficult to find. The inflation rate, Δp, is a natural candidate and therefore is included in the data set. Moreover, the marginal rate of interest, i^m, of the Bank of Finland is included in spite of the fact that the marginal rate measures restrictedness of money rather than the cost of holding money. It has, however, been chosen as a determinant of Finnish money demand in other studies and therefore is also included here. All series are graphed in Figure 3 and Figure 4 in Section IV. The data are available from the authors on request.

The paper is structured as follows: Section II discusses the various hypotheses we shall investigate and in Section III the notation is introduced for the maximum likelihood procedure. The next section derives the estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ under the assumption of cointegration and the last two sections investigate estimates and tests for $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ under linear restrictions.

Throughout, the two examples are used to motivate the statistical analysis and to illustrate the mathematically derived concepts.

II. A CLASSIFICATION OF THE VARIOUS HYPOTHESES

The hypotheses we consider consists of the hypothesis H_2 on the existence of cointegrating relations combined with linear restrictions on either the cointegrating relations or their weights:

$$H_2: \boldsymbol{\Pi} = \boldsymbol{\alpha\beta}',$$

$$H_3: \boldsymbol{\Pi} = \boldsymbol{\alpha\varphi}'\mathbf{H}' \text{ (or } \boldsymbol{\beta} = \mathbf{H}\boldsymbol{\varphi}),$$

$$H_4: \boldsymbol{\Pi} = \mathbf{A}\boldsymbol{\psi\beta}' \text{ (or } \boldsymbol{\alpha} = \mathbf{A}\boldsymbol{\psi}),$$

$$H_5: \boldsymbol{\Pi} = \mathbf{A}\boldsymbol{\psi\varphi}'\mathbf{H}' \text{ (or } \boldsymbol{\beta} = \mathbf{H}\boldsymbol{\varphi} \text{ and } \boldsymbol{\alpha} = \mathbf{A}\boldsymbol{\psi}),$$

and H_j^* is H_j augmented by $\boldsymbol{\mu} = \boldsymbol{\alpha\beta}_0'$ for $j = 2,\ldots,5$.

Note that the hypothesis H_1, where $\boldsymbol{\Pi}$ is unrestricted, can be written as H_2 with $r = p$. Hence, in this case the restriction $\boldsymbol{\mu} = \boldsymbol{\alpha\beta}_0'$ is the same as having $\boldsymbol{\mu}$ unrestricted. When we estimate model (1.2) under the hypothesis $\boldsymbol{\Pi} = \boldsymbol{\alpha\beta}'$ the choice of hypothesis about $\boldsymbol{\mu}$ becomes important. For the Danish data there does not seem to be any linear trend in the non-stationary processes (cf. Figure 3) and we will estimate models of the form H_j^*. For the Finnish data, however, there seems to be a linear trend in the non-stationary processes (cf. Figure 4) and models of the form H_j will be estimated.

The matrices $\mathbf{A}(p \times m)$ and $\mathbf{H}(p \times s)$ are known and define linear restrictions on the parameters $\boldsymbol{\alpha}(p \times r)$ and $\boldsymbol{\beta}(p \times r)$. The restrictions reduce the parameters to $\boldsymbol{\varphi}(s \times r)$ and $\boldsymbol{\psi}(m \times r)$, where $r \leq s \leq p$ and $r \leq m \leq p$. The important distinction between the H and the H^* hypotheses is that H^* define a restriction on $\boldsymbol{\mu}$, namely that it lies in the space spanned by $\boldsymbol{\alpha}$ or that $\boldsymbol{\alpha}'_\perp \boldsymbol{\mu} = \mathbf{0}$, hence that no trend is present. In the following the discussion will be concerned with the H hypotheses, but can easily be extended to include the case of H^*. In the above scheme note that $H_5 = H_4 \cap H_3$ and that $H_3 \subset H_2$ and $H_4 \subset H_2$. In fact all hypotheses are special cases of H_5 if we choose either $\mathbf{A}$ or $\mathbf{H}$ as the identity matrix.

The relations between the various hypotheses are illustrated in Figure 1.

All these hypotheses are restrictions of the matrix $\boldsymbol{\Pi}$ which under H_1 contains p^2 parameters. Under the hypothesis H_2 there are $pr + (p - r)r$ parameters which are further restricted to $sr + (p - r)r$ under H_4. Finally $mr + (s - r)r$ parameters remain under H_5. Note also that the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are not identified in the sense that given any choice of the matrix $\xi(r \times r)$, the choice $\boldsymbol{\alpha\xi}$ and $\boldsymbol{\beta}(\boldsymbol{\xi}')^{-1}$ will give the same matrix $\boldsymbol{\Pi}$, and hence determine the same probability distribution for the variables. One way of expressing this is to say that what the data can determine is the space spanned by the columns in $\boldsymbol{\beta}$, the cointegration space, and the space spanned by $\boldsymbol{\alpha}$. In general we present the results normalized by the coefficient of some of the variables, usually $m1$ and $m2$ respectively.

Note also that for each value of $r (0 \leq r \leq p)$ there is a corresponding hypothesis $H_2(r)$ of r or fewer cointegrating relations. The analysis makes it

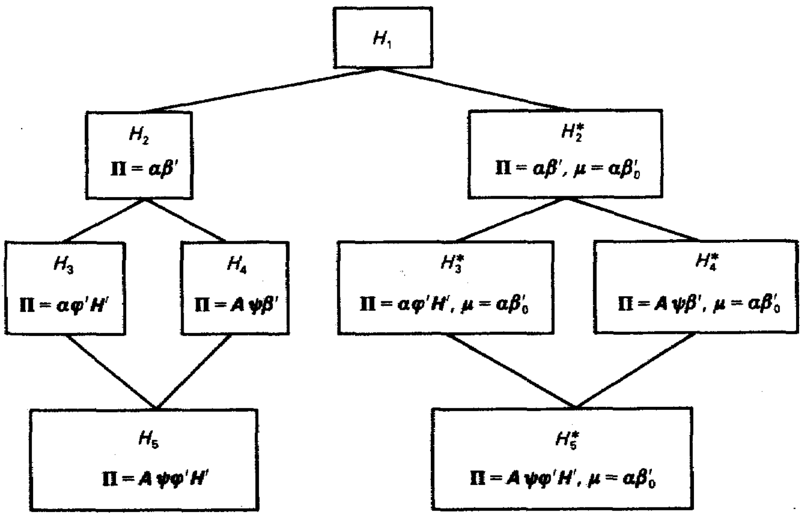


Fig. 1. The relation between the various hypothesis studied, starting with the most general VAR model (H_1) and introducing the restriction of cointegration (H_2) as well as linear restrictions on the cointegration vectors ($\boldsymbol{\beta}$) and the weights ($\boldsymbol{\alpha}$) in H_3 and H_4. The assumption of no trend $\boldsymbol{\mu} = \boldsymbol{\alpha\beta}_0'$ is indicated by a^*.

possible to conduct inferences about the value of r by testing $H_2(r)$ in H_1 or by testing $H_2(r)$ in $H_2(r+1)$.

III. THE MAXIMUM LIKELIHOOD PROCEDURE

In the following we will use the parameterization (1.2). The reason for this is that the parameters

$$(\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_{k-1}, \mathbf{\Phi}, \boldsymbol{\mu}, \mathbf{\Pi}, \mathbf{\Lambda})$$

are variation independent and, since all the models we are interested in are expressed as restrictions on $\boldsymbol{\mu}$ and $\mathbf{\Pi}$, it is possible to maximize over all the other parameters once and for all. We shall give details in the case of hypotheses about $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ without restricting $\boldsymbol{\mu}$, but mention how results are modified when $\boldsymbol{\alpha}'_\perp \boldsymbol{\mu} = \mathbf{0}$. Generally we use a superscript * to indicate that we are analysing a model where $\boldsymbol{\alpha}'_\perp \boldsymbol{\mu} = \mathbf{0}$. The model with $\boldsymbol{\mu} = \mathbf{0}$ and $\mathbf{\Phi} = \mathbf{0}$ was analysed in Johansen (1988b) and Johansen and Juselius (1988).

We now consider maximum likelihood estimation of the parameters in the unrestricted model:

$$\Delta \mathbf{X}_t = \mathbf{\Gamma}_1 \Delta \mathbf{X}_{t-1} + \ldots + \mathbf{\Gamma}_{k-1} \Delta \mathbf{x}_{t-k+1} + \mathbf{\Pi} \mathbf{X}_{t-k} + \boldsymbol{\mu} + \mathbf{\Phi} \mathbf{D}_t + \boldsymbol{\varepsilon}_t. \tag{3.1}$$

The results (3.2)–(3.9) are well-known but reproduced to establish the notation. This will be useful for discussing the estimators and tests later.

We first introduce the notation $\mathbf{Z}_{0t} = \Delta \mathbf{X}_t$, $\mathbf{Z}_{1t}$ denotes the stacked variables $\Delta \mathbf{X}_{t-1}, \ldots, \Delta \mathbf{X}_{t-k+1}, \mathbf{D}_t$, and 1, and $\mathbf{Z}_{kt} = \mathbf{X}_{t-k}$. Similarly, $\mathbf{\Gamma}$ is the matrix of parameters corresponding to $\mathbf{Z}_{1t}$, i.e. the matrix consisting of $\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_{k-1}, \mathbf{\Phi}$ and $\boldsymbol{\mu}$. Thus $\mathbf{Z}_{1t}$ is a vector of dimension $p(k-1)+3+1$ and $\mathbf{\Gamma}$ is a matrix of dimension $p \times (p(k-1)+3+1)$.

The model expressed in these variables becomes

$$\mathbf{Z}_{0t} = \mathbf{\Gamma} \mathbf{Z}_{1t} + \mathbf{\Pi} \mathbf{Z}_{kt} + \boldsymbol{\varepsilon}_t, \qquad (t = 1, \ldots, T). \tag{3.2}$$

For a fixed value of $\mathbf{\Pi}$, maximum likelihood estimation consists of a regression of $\mathbf{Z}_{0t} - \mathbf{\Pi} \mathbf{Z}_{kt}$ on $\mathbf{Z}_{1t}$ giving the normal equations

$$\sum_{t=1}^{T} \mathbf{Z}_{0t} \mathbf{Z}'_{1t} = \mathbf{\Gamma} \sum_{t=1}^{T} \mathbf{Z}_{1t} \mathbf{Z}'_{1t} + \mathbf{\Pi} \sum_{t=1}^{T} \mathbf{Z}_{kt} \mathbf{Z}'_{1t}. \tag{3.3}$$

The product moment matrices are denoted:

$$\mathbf{M}_{ij} = T^{-1} \sum_{t=1}^{T} \mathbf{Z}_{it} \mathbf{Z}'_{jt}, \qquad (i, j = 0, 1, k). \tag{3.4}$$

Then (3.3) can be written as:

$$\mathbf{M}_{01} = \mathbf{\Gamma} \mathbf{M}_{11} + \mathbf{\Pi} \mathbf{M}_{k1},$$

or

$$\mathbf{\Gamma} = \mathbf{M}_{01} \mathbf{M}_{11}^{-1} - \mathbf{\Pi} \mathbf{M}_{k1} \mathbf{M}_{11}^{-1}. \tag{3.5}$$

This leads to the definition of the residuals

$$\mathbf{R}_{0t} = \mathbf{Z}_{0t} - \mathbf{M}_{01}\mathbf{M}_{11}^{-1}\mathbf{Z}_{1t}, \tag{3.6}$$

$$\mathbf{R}_{kt} = \mathbf{Z}_{kt} - \mathbf{M}_{k1}\mathbf{M}_{11}^{-1}\mathbf{Z}_{1t}, \tag{3.7}$$

i.e. the residuals we would obtain by regressing $\Delta\mathbf{X}_t$ and $\mathbf{X}_{t-k}$ on $\Delta\mathbf{X}_{t-1}, \dots, \Delta\mathbf{X}_{t-k+1}, \mathbf{D}_t$, and 1.

The concentrated likelihood function becomes:

$$|\mathbf{\Lambda}|^{-T/2}\exp\left\{-\sum_{t=1}^{T}(\mathbf{R}_{0t}-\mathbf{\Pi}\mathbf{R}_{kt})'\mathbf{\Lambda}^{-1}(\mathbf{R}_{0t}-\mathbf{\Pi}\mathbf{R}_{kt})/2\right\}. \tag{3.8}$$

We express the estimates under the model H_1 by introducing the notation

$$\mathbf{S}_{ij} = T^{-1}\sum_{t=1}^{T}\mathbf{R}_{it}\mathbf{R}_{jt}' = \mathbf{M}_{ij} - \mathbf{M}_{i1}\mathbf{M}_{11}^{-1}\mathbf{M}_{1j}, \qquad (i,j=0,k), \tag{3.9}$$

and formulate these well-known results in

THEOREM 3.1: In the model:

$$H_1: \Delta\mathbf{X}_t = \sum_{j=1}^{k-1}\mathbf{\Gamma}_j\Delta\mathbf{X}_{t-j} + \mathbf{\Pi}\mathbf{X}_{t-k} + \boldsymbol{\mu} + \mathbf{\Phi}\mathbf{D}_t + \boldsymbol{\varepsilon}_t,$$

the parameters are estimated by ordinary least squares and we have:

$$\hat{\mathbf{\Pi}} = \mathbf{S}_{0k}\mathbf{S}_{kk}^{-1}, \tag{3.10}$$

and

$$\hat{\mathbf{\Lambda}} = \mathbf{S}_{00} - \mathbf{S}_{0k}\mathbf{S}_{kk}^{-1}\mathbf{S}_{k0}, \tag{3.11}$$

$$L_{\max}^{-2/T}(H_1) = |\hat{\mathbf{\Lambda}}|. \tag{3.12}$$

The estimate of $\mathbf{\Pi}$ inserted into (3.5) gives the estimate of $\mathbf{\Gamma}$.

Under the hypothesis $H_2^*: \mathbf{\Pi} = \boldsymbol{\alpha\beta}'$ and $\boldsymbol{\mu} = \boldsymbol{\alpha\beta}_0'$, which will be investigated for the Danish data, it is convenient to define $\mathbf{Z}_{0t}^* = \mathbf{Z}_{0t} = \Delta\mathbf{X}_t$ and let $\mathbf{Z}_{1t}^*$ be the stacked variables $\Delta\mathbf{X}_{t-1}, \dots, \Delta\mathbf{X}_{t-k+1}, \mathbf{D}_t$, whereas $\mathbf{Z}_{kt}^{*\prime} = \mathbf{X}_{t-k}^{*\prime} = (\mathbf{X}_{t-k}', 1)$. Thus we have moved the constant from the regressors into the vector $\mathbf{X}_{t-k}^*$. Further, we define $\mathbf{\Gamma}^*$ as the matrix of the relevant parameters $\mathbf{\Gamma}_1, \dots, \mathbf{\Gamma}_{k-1}, \mathbf{\Phi}$. Similarly we define $\mathbf{M}_{ij}^*$ and $\mathbf{S}_{ij}^*$. Note in particular that $\mathbf{S}_{kk}^*$ is $(p+1)\times(p+1)$.

3.1. The Empirical Analysis of the Unrestricted Model H_1

Model (3.1) including a constant term and seasonal dummies is fitted to the Danish and Finnish money demand data described in Section 1.3. For $k=2$, the residuals for the Danish data passed the test for being uncorrelated (see Table 1 below). For the Finnish data, the test statistic for the residuals in the equation for Δy is almost significant. The autocorrelogram suggests that there is some seasonality left in the residuals, but since the seasonal autocorrelation is rather small we have chosen to ignore this. Accordingly, model (3.1) with $k=2$ was fitted to both data sets. After conditioning on the first two data realizations, the number of observations left for estimation was 53 in the Danish and 104 in the Finnish data.

Since the parameter estimates of $\mathbf{\Gamma}_1$, $\mathbf{\Phi}$, $\boldsymbol{\mu}$ and $\mathbf{\Lambda}$ are not of particular interest in this paper, they are not reported. The estimates of $\mathbf{\Pi}$ are reported in Table 7 Section VI, and the standard error of regression estimates in Table 1 below. The normality assumption is tested by the Jarque & Bera test (Jarque and Bera, 1980), and reported below. For the Finnish data the residuals from the Δi^m and Δp equations do not pass the test. The deviations from normality are mainly due to too many large residuals. They are, however, approximately symmetrically distributed around zero, which probably is less serious than a skewed distribution. The robustness of the ML cointegration procedure for deviations from normality has not been investigated so far.

TABLE 1

Some Test Statistics for the niid Assumption for the Residuals in the Model (1.2) with $k=2$

	The Danish data				*The Finnish data*			
	$\Delta m2$	Δy	Δi^b	Δi^d	$\Delta m1$	Δy	Δi^m	Δp
τ_1	7.15	11.48	10.57	7.34	11.30	19.21	4.30	6.99
τ_2	2.12	1.93	1.06	1.61	1.61	1.88	10.86	28.02
$\hat{\sigma}_\varepsilon$	0.019	0.019	0.007	0.005	0.045	0.029	0.034	0.011

where $\tau_1 = T\Sigma r_i^2 (i=1,\ldots,10) \sim \chi^2(10)$,

$$\tau_2 = \frac{T-m}{6}\left(SK^2 + \frac{EK^2}{4}\right) \overset{as}{\sim} \chi^2(2),$$

m is the number of regressors, SK is skewness and EK is excess kurtosis.
$\hat{\sigma}_\varepsilon$ is the standard error of regression estimate.

IV. DERIVATION OF THE ESTIMATES OF $\boldsymbol{\alpha}$ AND $\boldsymbol{\beta}$ UNDER THE HYPOTHESIS $\mathbf{\Pi}=\boldsymbol{\alpha\beta}'$ AND THE LIKELIHOOD RATIO TEST FOR THIS HYPOTHESIS

Consider the model H_2, where $\mathbf{\Pi}=\boldsymbol{\alpha\beta}'$. The estimation of $\mathbf{\Gamma}_1,\ldots,\mathbf{\Gamma}_{k-1}$, $\mathbf{\Phi}$ and $\boldsymbol{\mu}$ is the same as before leading to (3.8). For fixed $\boldsymbol{\beta}$, it is easy to estimate $\boldsymbol{\alpha}$ and $\mathbf{\Lambda}$ by regressing $\mathbf{R}_{0t}$ on $\boldsymbol{\beta}'\mathbf{R}_{t-k}$ to obtain:

$$\boldsymbol{\alpha}(\boldsymbol{\beta}) = \mathbf{S}_{0k}\boldsymbol{\beta}(\boldsymbol{\beta}'\mathbf{S}_{kk}\boldsymbol{\beta})^{-1}, \tag{4.1}$$

$$\hat{\mathbf{\Lambda}}(\boldsymbol{\beta}) = \mathbf{S}_{00} - \mathbf{S}_{0k}\boldsymbol{\beta}(\boldsymbol{\beta}'\mathbf{S}_{kk}\boldsymbol{\beta})^{-1}\boldsymbol{\beta}'\mathbf{S}_{k0} = \mathbf{S}_{00} - \hat{\boldsymbol{\alpha}}(\boldsymbol{\beta})(\boldsymbol{\beta}'\mathbf{S}_{kk}\boldsymbol{\beta})\,\hat{\boldsymbol{\alpha}}(\boldsymbol{\beta})', \tag{4.2}$$

and

$$L_{\max}^{-2/T}(\boldsymbol{\beta}) = |\hat{\mathbf{\Lambda}}(\boldsymbol{\beta})| = |\mathbf{S}_{00} - \mathbf{S}_{0k}\boldsymbol{\beta}(\boldsymbol{\beta}'\mathbf{S}_{kk}\boldsymbol{\beta})^{-1}\boldsymbol{\beta}'\mathbf{S}_{k0}|. \tag{4.3}$$

As shown in Johansen (1988b) (see also Tso (1981)), one proceeds to estimate $\boldsymbol{\beta}$ by applying the identity.[2]

$$|\mathbf{S}_{00} - \mathbf{S}_{0k}\boldsymbol{\beta}(\boldsymbol{\beta}'\mathbf{S}_{kk}\boldsymbol{\beta})^{-1}\boldsymbol{\beta}'\mathbf{S}_{k0}| = |\mathbf{S}_{00}|\,|\boldsymbol{\beta}'\mathbf{S}_{kk}\boldsymbol{\beta} - \boldsymbol{\beta}'\mathbf{S}_{k0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0k}\boldsymbol{\beta}|/|\boldsymbol{\beta}'\mathbf{S}_{kk}\boldsymbol{\beta}|$$
$$= |\mathbf{S}_{00}|\,|\boldsymbol{\beta}'(\mathbf{S}_{kk} - \mathbf{S}_{k0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0k})\boldsymbol{\beta}|/|\boldsymbol{\beta}'\mathbf{S}_{kk}\boldsymbol{\beta}|. \tag{4.4}$$

This is minimized by noting that an expression like $|\boldsymbol{\beta}'(\mathbf{M}_1 - \mathbf{M}_2)\boldsymbol{\beta}|/|\boldsymbol{\beta}'\mathbf{M}_1\boldsymbol{\beta}|$ can be minimized by solving the equation $|\lambda\mathbf{M}_1 - \mathbf{M}_2| = 0$, where $\mathbf{M}_1 = \mathbf{S}_{kk}$, $\mathbf{M}_2 = \mathbf{S}_{k0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0k}$.

The results will be summarized in

THEOREM 4.1: Under the hypothesis

$$H_2: \mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}',$$

the maximum likelihood estimator of $\boldsymbol{\beta}$ is found by the following procedure: First solve the equation

$$|\lambda\mathbf{S}_{kk} - \mathbf{S}_{k0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0k}| = 0, \tag{4.5}$$

giving the eigenvalues $\hat{\lambda}_1 > \ldots > \hat{\lambda}_p$ and eigenvectors $\hat{V} = (\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_p)$ normalized such that[3] $\hat{\mathbf{V}}'\mathbf{S}_{kk}\hat{\mathbf{V}} = \mathbf{I}$.

The choice of $\boldsymbol{\beta}$ is now

$$\hat{\boldsymbol{\beta}} = (\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_r), \tag{4.6}$$

which gives

$$L_{\max}^{-2/T}(H_2) = |\mathbf{S}_{00}| \prod_{i=1}^{r} (1 - \hat{\lambda}_i). \tag{4.7}$$

The estimates of the other parameters are found by inserting $\hat{\boldsymbol{\beta}}$ into the above equations. The likelihood ratio test statistic for the hypothesis H_2 in H_1, since H_1 is a special case of H_2 for the choice $r = p$, is:

$$-2\ln(Q; H_2 \mid H_1) = -T \sum_{i=r+1}^{p} \ln(1 - \hat{\lambda}_i). \tag{4.8}$$

[2] This is based on the general result

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}' & \mathbf{C} \end{vmatrix} = |\mathbf{A}|\,|\mathbf{C} - \mathbf{B}'\mathbf{A}^{-1}\mathbf{B}| = |\mathbf{C}|\,|\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}'|$$

where in this application $\mathbf{A} = \mathbf{S}_{00}$, $\mathbf{B} = \mathbf{S}_{0k}\boldsymbol{\beta}$ and $\mathbf{C} = \boldsymbol{\beta}'\mathbf{S}_{kk}\boldsymbol{\beta}$.

[3] Many computer packages contain procedures for solving the eigenvalue problem $|\lambda\mathbf{I} - \mathbf{A}| = 0$, where $\mathbf{A}$ is symmetric. One can easily reduce (4.5) to this problem by first decomposing $\mathbf{S}_{kk} = \mathbf{C}\mathbf{C}'$, for some non-singular $p \times p$ matrix $\mathbf{C}$. Now (4.5) is equivalent to

$$|\lambda\mathbf{I} - \mathbf{C}^{-1}\mathbf{S}_{k0}\mathbf{S}_{00}\mathbf{S}_{0k}\mathbf{C}'^{-1}| = 0,$$

which has the same eigenvalues $\hat{\lambda}_1 > \ldots > \hat{\lambda}_p$ but eigenvectors $e_1, \ldots, e_p$. The eigenvectors of (4.5) are then found as $\hat{\mathbf{v}}_i = \mathbf{C}'^{-1}\mathbf{e}_i$.

Similarly the likelihood ratio test statistic for testing $H_2(r)$ in $H_2(r+1)$ is given by

$$-2\ln(Q;r|r+1) = -T\ln(1-\hat{\lambda}_{r+1}). \tag{4.9}$$

Under the hypothesis $H_2^*: \mathbf{\Pi} = \boldsymbol{\alpha\beta}'$ and $\boldsymbol{\mu} = \boldsymbol{\alpha\beta}_0'$ the same results hold but derived from $\mathbf{S}_{ij}^*$ rather than $\mathbf{S}_{ij}$.

The asymptotic distributions of the likelihood ratio test statistics (4.8) and (4.9) are found in Johansen (1989), and are not given by the usual χ^2 distributions, but as multivariate versions of the Dickey-Fuller distribution. These distributions are conveniently described by certain stochastic integrals, and can be tabulated by simulation, see the Appendix.

Consider first the case $\boldsymbol{\alpha}'_\perp \boldsymbol{\mu} = \mathbf{0}$. By suitably normalizing the equation (4.5) and letting $T \to \infty$ one can show that $T(\hat{\lambda}_{r+1}, \ldots, \hat{\lambda}_p)$ converge to the roots of the equation

$$\left|\rho \int_0^1 \mathbf{F}\mathbf{F}' \mathrm{d}t - \int_0^1 \mathbf{F}(\mathrm{d}\mathbf{U}') \int_0^1 (\mathrm{d}\mathbf{U})\mathbf{F}'\right| = 0, \tag{4.10}$$

where $\mathbf{U}(t) = \{U_1(t), \ldots, U_{p-r}(t)\}$ is a $(p-r)$-dimensional Brownian motion and the $(p-r)$-dimensional stochastic process $\mathbf{F}(t) = \{F_1(t), \ldots, F_{p-r}(t)\}$ is defined by

$$F_i(t) = U_i(t) - \int_0^1 U_i(s)\,\mathrm{d}s, \qquad (i = 1, \ldots, p-r). \tag{4.11}$$

Further $\int_0^1 \mathbf{F}\mathbf{F}'\mathrm{d}t$ is a $(p-r)\times(p-r)$ matrix of random variables defined by the ordinary Riemann integrals

$$\int_0^1 F_i(t)\,F_j(t)\,\mathrm{d}t, \qquad (i, j = 1, \ldots, p-r), \tag{4.12}$$

and

$$\int_0^1 \mathbf{F}(\mathrm{d}\mathbf{U}') = \left[\int_0^1 (\mathrm{d}\mathbf{U})\mathbf{F}'\right]'$$

is defined as the matrix of stochastic integrals[4]

$$\int_0^1 F_i \mathrm{d}U_j \qquad (i, j = 1, \ldots, p-r). \tag{4.13}$$

With this notation it follows that the statistic (4.8) satisfies

$$-2\ln(Q;H_2|H_1)=-T\sum_{i=r+1}^{p}\ln(1-\hat{\lambda}_i)\simeq\sum_{i=r+1}^{p}T\hat{\lambda}_i$$

which converges weakly to

$$\sum_{i=1}^{p-r}\hat{\rho}_i=tr\left\{\int_0^1(\mathrm{d}\mathbf{U})\mathbf{F}'\left[\int_0^1\mathbf{F}\mathbf{F}'\mathrm{d}t\right]^{-1}\int_0^1\mathbf{F}(\mathrm{d}\mathbf{U}')\right\},$$

where $tr\{\mathbf{M}\}$ denotes the trace of the matrix $\mathbf{M}$. The statistic (4.8) is therefore called the trace statistic (trace). Similarly

$$-2\ln(Q;r|r+1)=-T\ln(1-\hat{\lambda}_{r+1})\simeq T\hat{\lambda}_{r+1},$$

converges weakly to

$$\hat{\rho}_1=\lambda_{\max}\left\{\int_0^1(\mathrm{d}\mathbf{U})\mathbf{F}'\left[\int_0^1\mathbf{F}\mathbf{F}'\mathrm{d}t\right]^{-1}\int_0^1\mathbf{F}(\mathrm{d}\mathbf{U}')\right\},$$

where $\lambda_{\max}\{\mathbf{M}\}$ denotes the marginal eigenvalue of the matrix $\mathbf{M}$. The statistic (4.9) is called the maximal eigenvalue statistic ($\lambda_{\max}$).

If $r=p-1$, then both U and F are one dimensional. Then the test statistics are equal, since the trace equals the (maximal) eigenvalue, and the asymptotic distribution of the statistic can be expressed as

$$\left\{\int_0^1[U-\bar{U}]\mathrm{d}U\right\}^2\Big/\int_0^1[U(s)-\bar{U}]^2\mathrm{d}s$$

where

$$\bar{U}=\int_0^1 U(u)\mathrm{d}u.$$

This statistic is the square of the statistic $\hat{\tau}_\mu$ tabulated in Fuller (1976) p. 373.

The distribution of the trace and the maximal eigenvalue of the roots of (4.10) depend only on the dimension $p-r$, i.e. the number of non-stationary

[4] The definition of a stochastic integral is analogous to the definition of a Riemann integral. We let U and F be two continuous stochastic processes on the unit interval like the Brownian motions. Then we consider a partion of the unit interval and the Riemann sum

$$R=\Sigma_k F(t_{k-1})(U(t_k)-U(t_{k-1}))$$

The function $U(\cdot)$ has infinite variation but finite quadratic variation, i.e. $\sup(t_k)\Sigma_k(U(t_k)-U(t_{k-1}))^2\le c<\infty$. This can be used to prove the existence of the limit of R in L_2, i.e. there exists a random variable, which we shall call $\int F\,\mathrm{d}U$, such that $E(R-\int F\,\mathrm{d}U)^2$ converges to zero.

components under the hypothesis. The distributions are tabulated by simulation and are given in the Appendix in Table A2.

Next consider the case where $\boldsymbol{\alpha}'_{\perp}\boldsymbol{\mu} \neq \mathbf{0}$, i.e. the trend is present under the null hypothesis. We can express the results in this case by choosing a different definition of $\mathbf{F}$:

$$F_i(t) = U_i(t) - \int_0^1 U_i(s)\,\mathrm{d}s, \qquad (i=1,\dots,p-r-1), \tag{4.14}$$

$$F_i(t) = t - 1/2, \qquad (i=p-r). \tag{4.15}$$

It is instructive to consider again the case of $p-r=1$, where the statistic reduces to

$$\left[\int_0^1 (t-\tfrac{1}{2})\,\mathrm{d}U\right]^2 \Big/ \int_0^1 (t-\tfrac{1}{2})^2\mathrm{d}t$$

which is distributed as $\chi^2(1)$. This is the well-known result (West (1989)) that if the linear trend is present under the hypothesis of non-stationarity then the usual asymptotics hold for the likelihood ratio test. The distribution of the trace and maximal eigenvalue of the equation (4.10) with this choice of $\mathbf{F}$ is tabulated by simulation in the Appendix and given in Table A1.

Since the distribution with $\boldsymbol{\alpha}'_{\perp}\boldsymbol{\mu} = \mathbf{0}$ has broader tails, cf. Tables A1 and A2 in the Appendix, the p-value should be calculated from the latter distribution.

Under H_2^* (i.e. $\boldsymbol{\mu} = \boldsymbol{\alpha}\boldsymbol{\beta}'_0$) the asymptotic distribution of the test statistics (4.8) and (4.9) can be shown to be distributed as above but with $\mathbf{F}$ defined by:

$$F_i(U) = U_i(u) \qquad (i=1,\dots,p-r),$$

$$F_i(u) = 1 \qquad (i=p-r+1).$$

These distributions are tabulated by simulation in Table A3. The relation between the applications of the three distributions is illustrated in Figure 2 below:

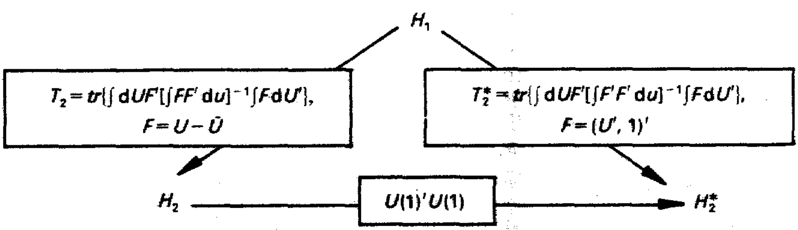


Fig. 2. The relation between the hypotheses H_1, H_2 and H_2^*, and the test statistics used to test them. Note that $T_2^* = T_2 + U(1)'U(1)$.

4.1. *The Empirical Analysis*[5]

In Table 2 the estimated eigenvalues $\hat{\lambda}$, the normalized eigenvectors $\hat{\mathbf{V}}$, and weights $\hat{\mathbf{W}} = \mathbf{S}_{0k}\hat{\mathbf{V}}$ are reported, for the two data sets. The graphs corresponding to the eigenvectors and the original data are presented in Figures 3 and 4. Note that the eigenvectors $\hat{v}_i^*$ for the Danish data are of dimension 5, where the last coefficient is the estimated intercept. For the Finnish data μ is assumed to contain effects both from the intercept and from the linear trend (see discussion in Section I).

In Table 3 the likelihood ratio test statistics are calculated and compared to the 95 percent quantiles of the appropriate limiting distribution. Two versions of the test procedure are reported in Table 3. The first is based on the trace and the second on the maximum eigenvalue, see Theorem 4.1, (4.8) and (4.9).

The Danish Data

On the basis of the plots of the series (see Figure 3) a model without a linear trend in the non-stationary part of the process was assumed. Thus a constant 1 was appended to the vector $\mathbf{X}_{t-2}$, and the calculations were performed as described in Section I, giving the matrices $\mathbf{S}_{ij}^*$. The results of the eigenvalue and eigenvector calculations are given in Table 2. First we consider the number of cointegration vectors, beginning with the hypothesis $r \leq 1$ versus the general alternative H_1. Using the trace test procedure gives $-2\ln(Q) = -T\Sigma_{i=2}^4 \ln(1-\hat{\lambda}_i^*) = 19.06$.

The 95 percent quantile, 35.07, in the asymptotic distribution, see Table A3, is not significant. Hence there is no evidence in the Danish data for more than one cointegration relation. If we test the hypothesis that $r=0$ in H_1 we get a test value of 49.14, which is found to have a p-value of appr. 10 percent. If instead we apply the maximum eigenvalue test, and test $H_2(r=0)$ in $H_2(r \leq 1)$ we find $-2\ln(Q; r=0 \mid r \leq 1) = 30.08$ which is in the upper tail of the distribution of $\lambda_{\max}$ for $r=0$ with a p-value of 2.5 percent. We conclude that there is only one cointegration vector in the Danish data. This hypothesis will be maintained below. It must be noted that since the $T\hat{\lambda}_i$ are ordered they cannot be independent, not even in the limit.

Thus all the tests performed in Table 3 are highly dependent on one another.

Finally, as a check that the maintained assumption about the absence of trend is data consistent, the test for $H_2^*(r \leq 1)$ in $H_2(r \leq 1)$ (see Figure 2) was performed:

$$-2\ln(Q; H_2^*(1) \mid H_2(1)) = T \ln\{|\mathbf{S}_{00}^*|(1-\hat{\lambda}_i^*)/|\mathbf{S}_{00}|(1-\hat{\lambda}_1)\}$$

$$= -T\sum_{i=2}^{4} \ln\{(1-\hat{\lambda}_i^*)/(1-\hat{\lambda}_i)\} = 1.99.$$

[5] The calculations have been performed in the computer package RATS, VAR Econometrics, Inc/Doan Associates.

TABLE 2

The Eigenvalues $\hat{\boldsymbol{\lambda}}$ and Eigenvectors $\hat{\mathbf{V}}$ as well as the Weights $\hat{\mathbf{W}}$ for the Danish and Finnish Data

	The Danish data						The finnish data			
	Eigenvalues $\hat{\lambda}^*$						Eigenvalues $\hat{\lambda}$			
	(0.4332	0.1776	0.1128	0.0434	0)		(0.3093	0.2260	0.0731	0.0295)
	Eigenvectors $\hat{V}^*$						Eigenvectors $\hat{V}$			
$m2$	1.00	1.00	1.00	1.00	1.00	$m1$	1.00	0.04	1.00	1.00
y	−1.03	−1.37	−3.23	−1.88	−0.63	y	−0.98	−0.06	−0.92	1.61
i^b	5.21	0.24	0.54	24.40	1.70	i^m	−7.09	−0.09	−0.26	−1.38
i^d	−4.22	6.84	−5.65	−14.30	−1.90	Δp	−7.02	1.00	−1.82	−15.69
1	−6.06	−4.27	7.90	−2.26	−8.03					
	Weights $\hat{W}^*$						Weights $\hat{W}$			
$m2$	−0.21	0.00	0.04	0.00	0.00	$m1$	0.03	−0.46	−0.13	0.00
y	0.12	0.02	0.05	0.00	0.00	y	0.02	−0.13	0.01	−0.01
i^b	0.02	−0.01	0.00	0.00	0.00	i^m	0.05	1.07	−0.01	0.00
i^d	0.03	−0.03	0.00	0.00	0.00	Δp	0.01	−0.39	0.01	0.00

TABLE 3

Test statistics for the hypothesis H_2^* and H_2 for various values of r versus r + 1 (λ_{max}) and versus the general alternative H_1 (trace) for the Danish and Finnish data. The 95% quantiles are taken from Table A2 (H_2) and A3 (H_2^*)

	The Danish data					The Finnish data			
H_2^*	trace	trace (0.95)	λ_{max}	λ_{max} (0.95)	H_2	trace	trace (0.95)	λ_{max}	λ_{max} (0.95)
$r \leq 3$	2.35	9.09	2.35	9.09	$r \leq 3$	3.11	8.08	3.11	8.08
$r \leq 2$	8.69	20.17	6.34	15.75	$r \leq 2$	11.01	17.84	7.90	14.60
$r \leq 1$	19.06	35.07	10.37	21.89	$r \leq 1$	37.65	31.26	26.64	21.28
$r = 0$	49.14	53.35	30.08	28.17	$r = 0$	76.14	48.42	38.49	27.34

This relation holds, since $|\mathbf{S}_{00}|\Pi_1^4(1-\hat{\lambda}_i) = |\mathbf{S}_{00}^*|\Pi_1^5(1-\hat{\lambda}_i^*) = |\mathbf{S}_{00} - \mathbf{S}_{0k}\mathbf{S}_{kk}^{-1}\mathbf{S}_{k0}|$. The asymptotic distribution of the test statistic is $\chi^2(3)$ and thus not significant.

The coefficient estimates of the cointegrating relation are found in Table 2 as the first column in $\hat{\mathbf{V}}^*$. The interpretation of the cointegration vector as an error correction mechanism measuring the excess demand for money is straightforward, with the estimate of the equilibrium relation given by

$$m2 = 1.03y - 5.21i^b + 4.22^d + 6.06.$$

Similarly $\hat{\boldsymbol{\alpha}}$ is found as the first column in the matrix $\hat{\mathbf{W}}^* = \mathbf{S}_{02}\hat{\mathbf{V}}^*$:

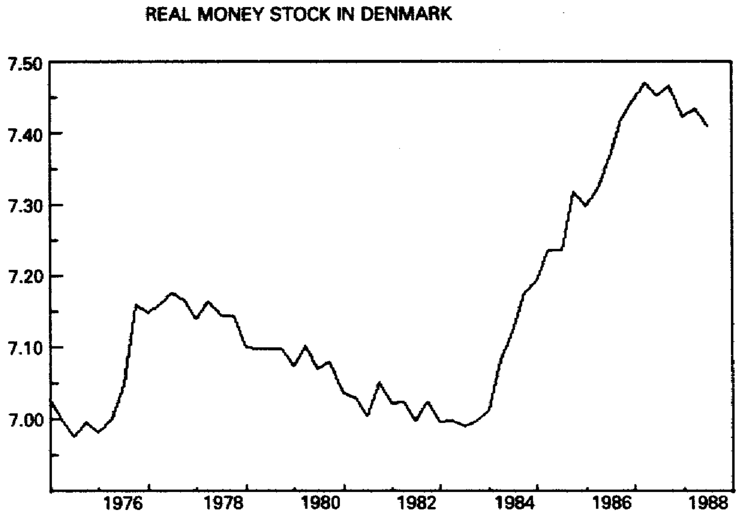
$$\hat{\boldsymbol{\alpha}}' = (-0.213, 0.115, 0.023, 0.029).$$

The coefficients of $\hat{\boldsymbol{\alpha}}$ can be interpreted as the weights with which excess demand for money enters the four equations of our system, and it is natural to give them an economic meaning in terms of the average speed of adjustment towards the estimated equilibrium state, such that a low coefficient indicates slow adjustment and a high coefficient indicates rapid adjustment. In the first equation, which measures the changes in money balances, the average speed of adjustment is approximately 0.213, whereas in the remaining three equations the adjustment coefficients are lower though of the 'correct' sign. In particular the last two adjustment coefficients are low, and the hypothesis that some subset of the adjustment coefficients is zero will be formally tested in Section VI.

The Finnish Data

As discussed earlier a model that allows for linear trends is fitted to the Finnish data. The estimated eigenvalues, vectors and weights are given in Table 2 and the test statistics in Table 3, which indicate that at least 2 but possibly 3 cointegration vectors are present.

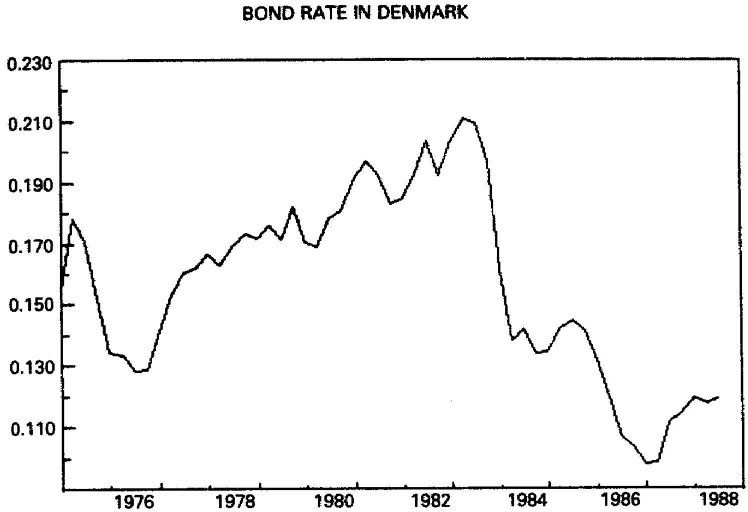
The acceptance of the third relies on a p-value of approximately 20 percent, which usually would be considered too high. But since the power of



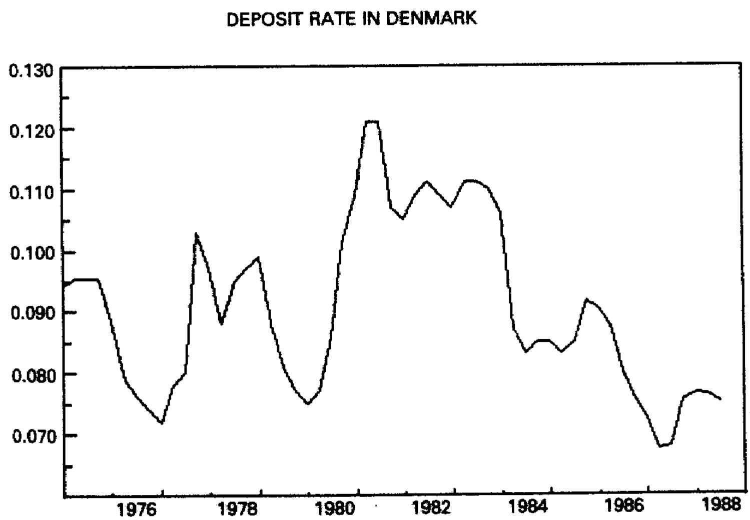
(a)

(b)

Fig. 3. The graphs of the cointegration relation $\hat{\mathbf{V}}_i'\mathbf{X}_t$ for $i=1,4$ and the original Danish data. The sample is: 1974.1–1987.3.



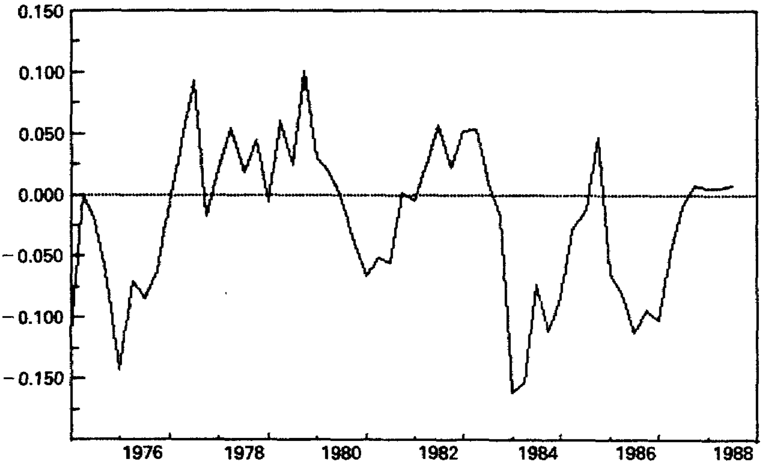
(c)



(d)

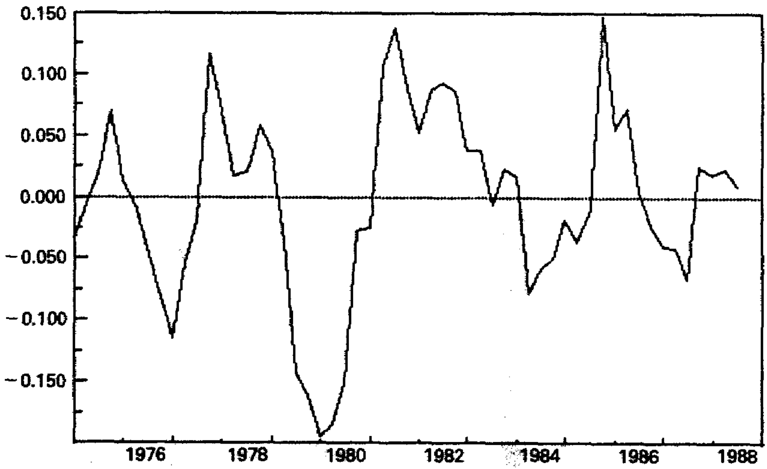
Fig. 3. Continued

COINTEGRATION RELATION 1



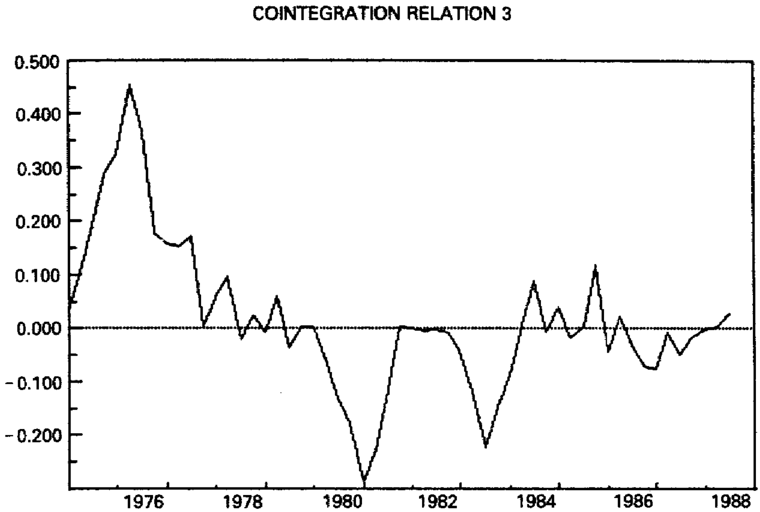
(e)

COINTEGRATION RELATION 2

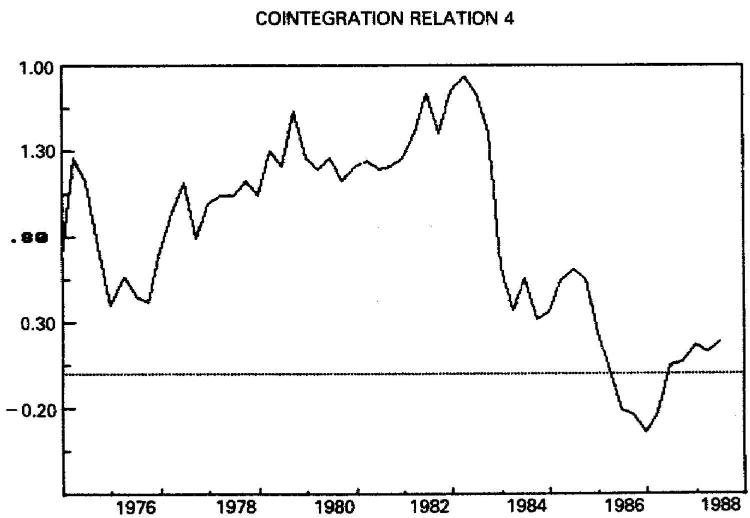


(f)

Fig. 3. Continued



(g)



(h)

Fig. 3. Continued

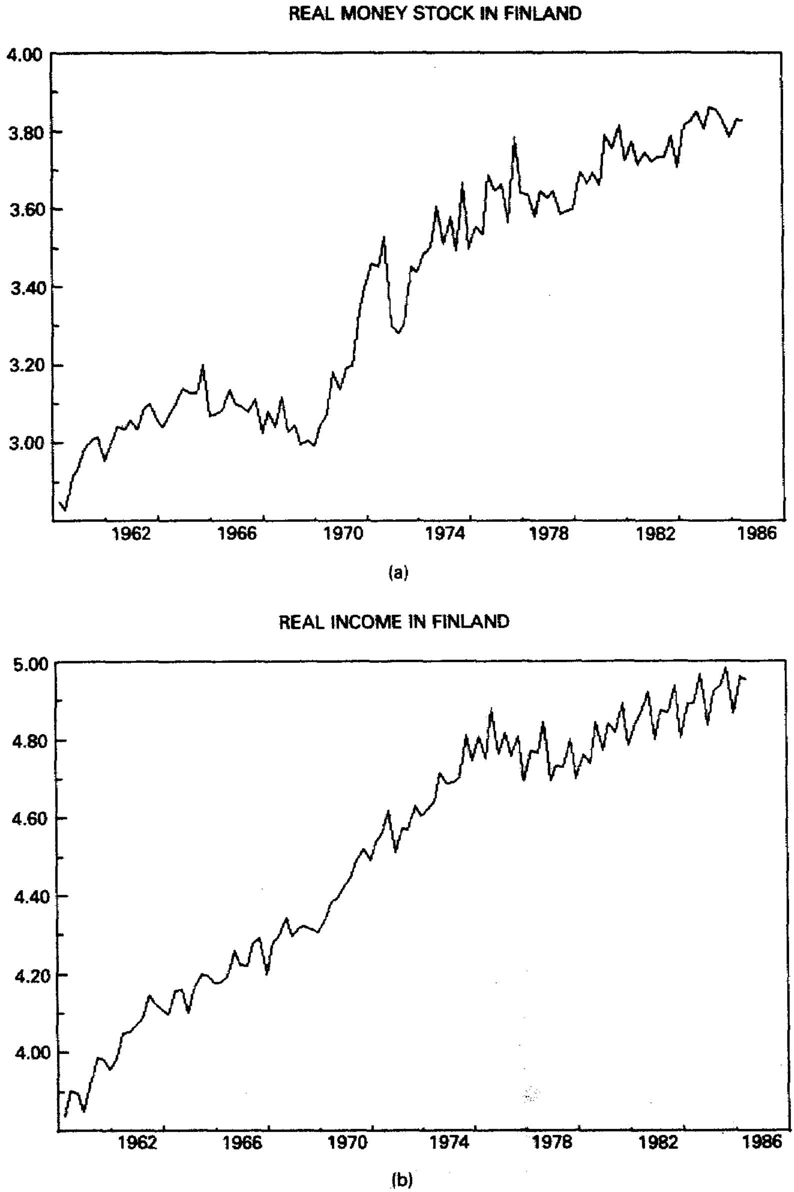
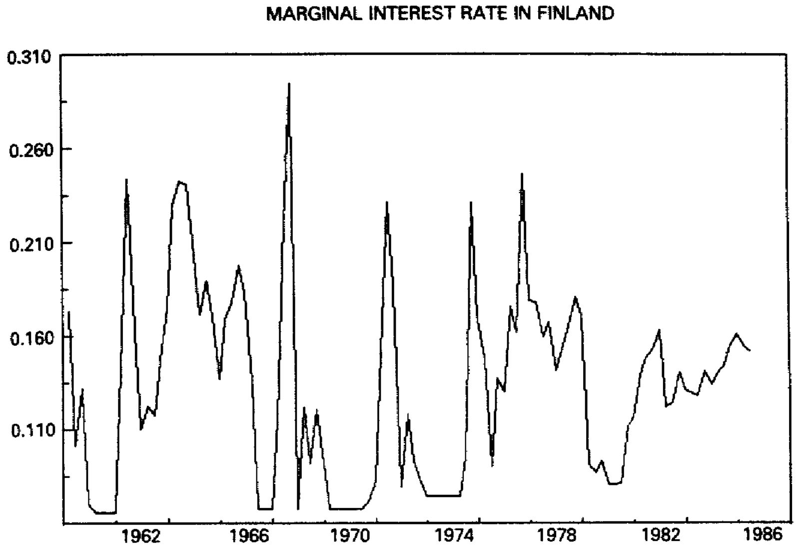
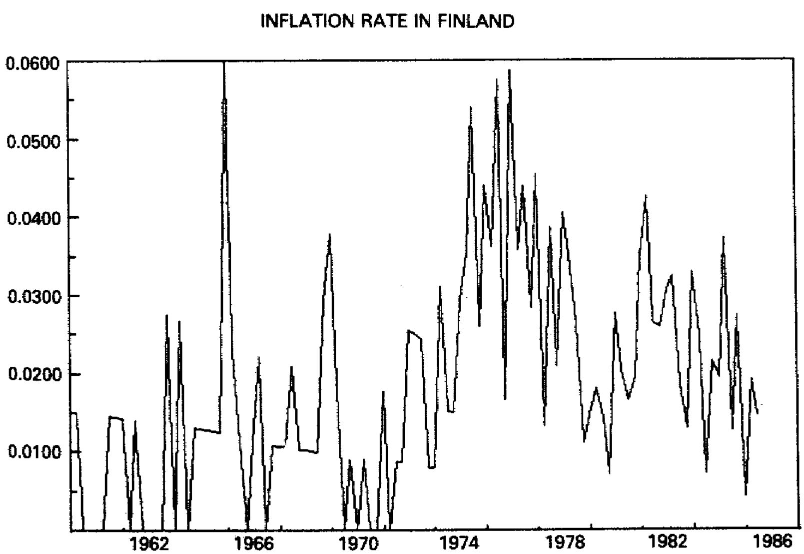


Fig. 4. The graphs of the cointegration relation $\hat{V}_i'X_t$ for $i=1,4$ and the original Finnish data. The sample is: 1958.1–1984.3.

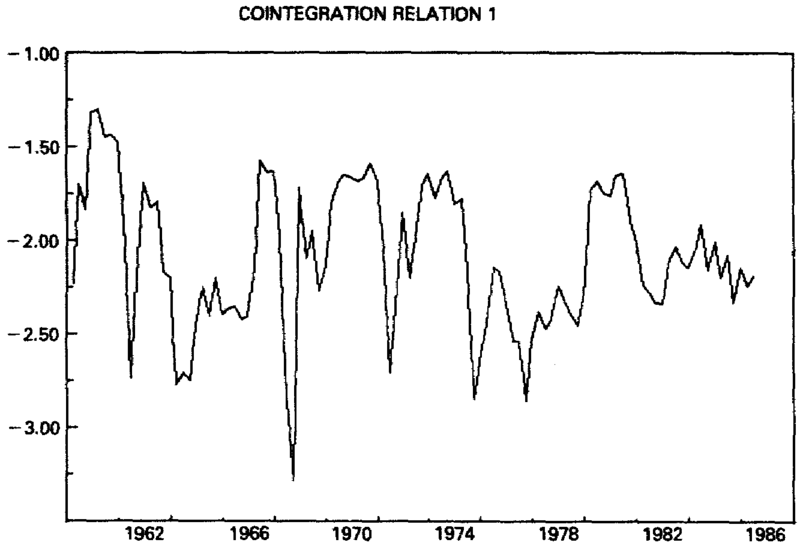


(c)

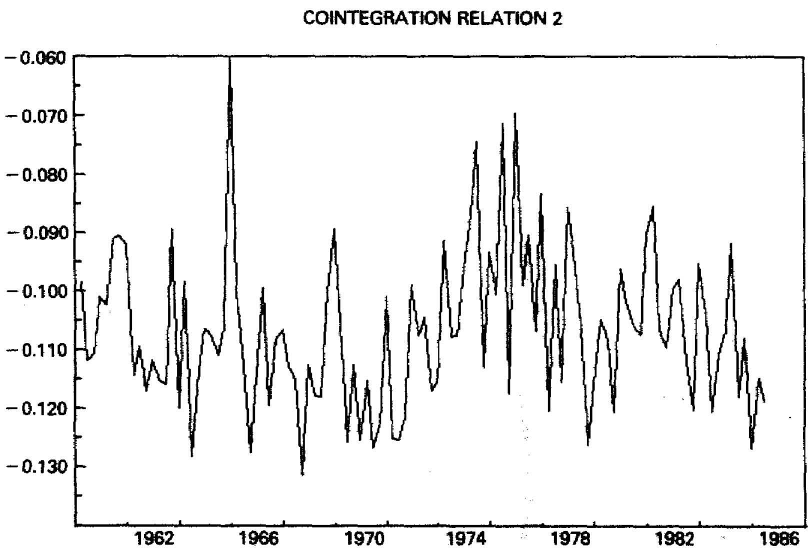


(d)

Fig. 4. Continued



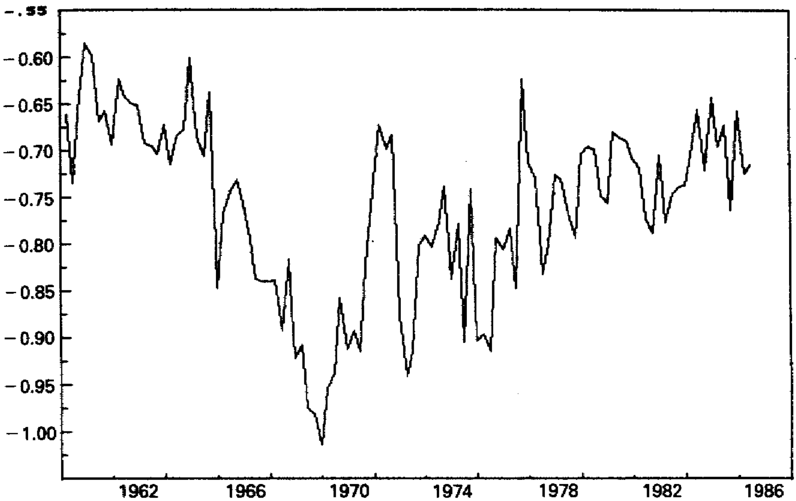
(e)



(f)

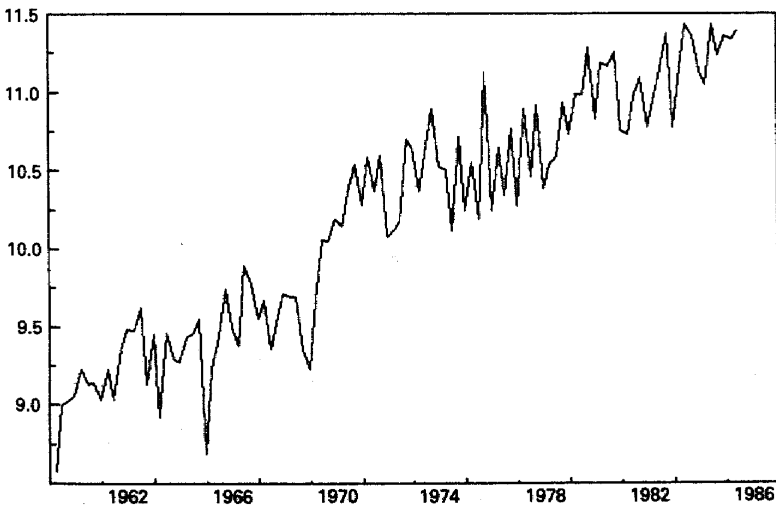
Fig. 4. Continued

COINTEGRATION RELATION 3



(g)

COINTEGRATION RELATION 4



(h)

Fig. 4. Continued

the tests are likely to be low for cointegration vectors with roots close to but outside the unit circle, it seems reasonable in certain cases to follow a test procedure which rejects for higher p-values than the usual 5 percent. One reason why we have kept $r=3$ in this case is that the hypothesis of proportionality between money and income, i.e. $\beta_{i.1} = -\beta_{i.2}$ seems consistent with the data for the 3 eigenvectors.[5] If $m1$ and y appear in a cointegration vector with equal coefficients of opposite sign, they should do so in all cointegration vectors, see Section V.

Next we test the hypothesis that the linear trend is absent, i.e. $H_2^*(3)$ in $H_2(3)$. We found that

$$-2\ln(Q; H_2^*(3) \mid H_2(3)) = T\ln\left\{|\mathbf{S}_{00}^*| \prod_{i=1}^{3}(1-\hat{\lambda}_i^*)/|\mathbf{S}_{00}| \prod_{i=1}^{3}(1-\hat{\lambda}_i)\right\}$$

$$= -T\ln\{(1-\hat{\lambda}_4^*)/(1-\hat{\lambda}_4)\} = 4.78.$$

Since the asymptotic distribution of this statistic is $\chi^2(1)$, it is significant, and the hypothesis $H_2(3)$ is maintained. We find $\hat{\boldsymbol{\beta}}$ as the first three columns of $\hat{\mathbf{V}}$ from Table 2 and $\hat{\boldsymbol{\alpha}}$ as the corresponding columns of the weights $\hat{\mathbf{W}}$. Note that given the full matrices $\hat{\mathbf{V}}$ and $\hat{\mathbf{W}}$ one can estimate $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ for any value of r.

For the case $r>1$, the interpretation of $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\alpha}}$ is not straightforward. A heuristic interpretation is however possible by considering the estimates in Table 2. Note that $\hat{\beta}_{i.2} \approx -\hat{\beta}_{i.1}$, $i=1,2,3$, and that $\hat{\boldsymbol{\beta}}_2$ is approximately proportional to $(0,0,0,1)$. Thus, $\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2$ and $\hat{\boldsymbol{\beta}}_3$ can be approximately represented as linear combinations of the vectors $(-1,1,0,0)$, $(0,0,0,1)$, and $(0,0,1,0)$, implying that $m1-y$, i^m and Δp are stationary. This means that the only interesting cointegration relation found is between $m1$ and y. However, a linear combination between these three vectors might be more stable (in terms of the roots of the characteristic polynomial) than the individual vectors themselves and this linear combination could in fact be the economically interesting relation. In particular, one would expect that the linear combination which is most correlated with the stationary part of the model, namely the first eigenvector, is of special interest. Although there is some arbitrariness in the case $r>1$, the ordering of the eigenvectors provided by the estimation procedure is likely to be useful.

The estimates reported in Table 2 indicate that $\hat{\boldsymbol{\beta}}_2$ is approximately measuring the inflation rate, whereas $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_3$ seem to contain information about $m1-y$. Note also that $\hat{\alpha}_{11}$ and $\hat{\alpha}_{13}$ have opposite sign. The sign to be expected for 'excess demand for money' should be negative, but $\hat{\alpha}_{13}$ dominates $\hat{\alpha}_{11}$, so that the 'excess demand for money' enters with a negative sign in the first equation. The value of $\hat{\alpha}_{i2}$ can be interpreted as the weight with which the inflation rate enters equation i. In Table 7 Section VI, the

[5] It seems reasonable to denote the first coordinate of the cointegration vector $\boldsymbol{\beta}_i$, say, by $\beta_{i.1}$. In ordinary matrix notation we then have $\beta_{i.1} = \beta_{1i}$.

estimate of $\mathbf{\Pi}=\boldsymbol{\alpha\beta}'$, i.e. the estimate of the combined effects of all three cointegration vectors, is reported. It is striking how well the proportionality hypothesis between money and income is maintained in all equations of the system.

This completes the investigation of the model H_2 and H_2^* in H_1 and we turn now to the models H_3 and H_3^* in H_2.

V. ESTIMATION AND TESTING UNDER LINEAR RESTRICTIONS ON $\boldsymbol{\beta}$

Model $H_3:\boldsymbol{\beta}=\mathbf{H}\boldsymbol{\varphi}$ is a formulation of a linear restriction on the cointegration vectors. The hypothesis specifies the same restriction on all the cointegration vectors. The reason for this is the following: If we have two cointegration vectors in which m and y, say, enter then any linear combination of these relations will also be a cointegrating relation. Thus it will in general be possible to find some relation which has, say, equal coefficients with opposite sign to m and y, corresponding to a long-run unit elasticity. This is clearly not interesting, and only if the proportionality restriction is present in all $\boldsymbol{\beta}$ vectors, is it meaningful to say that we have found a unit elasticity.

5.1. Likelihood Ratio Tests

Under H_3 we have the restriction $\boldsymbol{\beta}=\mathbf{H}\boldsymbol{\varphi}$ where $\mathbf{H}$ is $(p\times s)$, but that means that the estimation of $\mathbf{\Gamma}_1,\ldots,\mathbf{\Gamma}_{k-1}$, $\mathbf{\Phi}$, $\boldsymbol{\mu}$, $\boldsymbol{\alpha}$ and $\mathbf{\Lambda}$ is given as for fixed $\boldsymbol{\beta}=\mathbf{H}\boldsymbol{\varphi}$, and $\boldsymbol{\varphi}$ has to be chosen to minimize

$$|\boldsymbol{\varphi}'(\mathbf{H}'\mathbf{S}_{kk}\mathbf{H}-\mathbf{H}'\mathbf{S}_{k0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0k}\mathbf{H})\boldsymbol{\varphi}|/|\boldsymbol{\varphi}'(\mathbf{H}'\mathbf{S}_{kk}\mathbf{H})\boldsymbol{\varphi}| \tag{5.1}$$

over the set of all $s\times r$ matrices $\boldsymbol{\varphi}$. This problem has the same kind of solution as above and we formulate the results in Theorem 5.1 below. A subscript indicates which hypothesis we are currently working with. Throughout, the estimator without subscript will be the estimator under H_2 or H_2^*.

THEOREM 5.1: Under the hypothesis

$$H_3:\boldsymbol{\beta}=\mathbf{H}\boldsymbol{\varphi},$$

we find the maximum likelihood estimator of $\boldsymbol{\beta}$ as follows: First solve

$$|\lambda\mathbf{H}'\mathbf{S}_{kk}\mathbf{H}-\mathbf{H}'\mathbf{S}_{k0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0k}\mathbf{H}|=0, \tag{5.2}$$

for $\hat{\lambda}_{3.1}>\ldots>\hat{\lambda}_{3.s}$ and $\hat{\mathbf{V}}_3=(\hat{\mathbf{v}}_{3.1},\ldots,\hat{\mathbf{v}}_{3.s})$ normalized by $\hat{\mathbf{V}}_3'(\mathbf{H}'\mathbf{S}_{kk}\mathbf{H})\hat{\mathbf{V}}_3=\mathbf{I}$.

Choose

$$\hat{\boldsymbol{\varphi}}=(\hat{\mathbf{v}}_{3.1},\ldots,\hat{\mathbf{v}}_{3.r})\text{ and }\hat{\boldsymbol{\beta}}_3=\mathbf{H}\hat{\boldsymbol{\varphi}}, \tag{5.3}$$

and find the estimates of $\boldsymbol{\alpha}$, $\mathbf{\Lambda}$ and $\mathbf{\Gamma}$ from (4.1), (4.2), and (3.5). The maximized likelihood becomes

$$L_{\max}^{-2/T}(H_3)=|\mathbf{S}_{00}|\prod_{i=1}^{r}(1-\hat{\lambda}_{3.i}), \tag{5.4}$$

which gives the likelihood ratio test of the hypothesis H_3 in H_2 as

$$-2\ln(Q;H_3\,|\,H_2)=T\sum_{i=1}^{r}\ln\{(1-\hat{\lambda}_{3,i})/(1-\hat{\lambda}_i)\}. \tag{5.5}$$

The asymptotic distribution of this statistic is shown in Johansen (1989) to be χ^2 with $r(p-s)$ degrees of freedom.

Under the hypothesis $H_3^*:\boldsymbol{\beta}=\mathbf{H}\boldsymbol{\varphi}$ and $\boldsymbol{\mu}=\boldsymbol{\alpha}\boldsymbol{\beta}_0'$, the same results hold.

5.1.1. The Empirical Analysis

The Finnish Data

We consider the hypothesis that there is proportionality between money and income, so that the coefficients of money and income are equal with opposite sign, i.e.

$$H_3:\beta_{i.1}=-\beta_{i.2},\qquad (i=1,2,3)$$

In matrix notation the hypothesis can be formulated as:

$$\boldsymbol{\beta}=\begin{pmatrix}-1&0&0\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}\boldsymbol{\varphi},$$

where $\boldsymbol{\varphi}$ is a 3×3 matrix. Solving (5.2) gives the eigenvalues in Table 4. These are compared to the eigenvalues of the unrestricted model H_2. The test statistic is calculated as $-2\ln(Q)=0.02+3.51+0.29=3.82$ which is compared to $\chi^2_{0.95}(r(p-s))=\chi^2(3(4-3))=7.81$. Thus the hypothesis of equal coefficients with opposite sign for $m1$ and y, is clearly accepted. The corresponding restricted $\boldsymbol{\beta}$-estimates hardly change at all compared to the unrestricted estimates of Table 2 and they are therefore not reported here.

With the imposed proportionality restriction we now have three cointegration vectors restricted to a three dimensional space defined by the restriction that $m1$ and y have equal coefficients with opposite sign. Thus the hypothesis H_3 is really the hypothesis of a complete specification of $sp(\boldsymbol{\beta})$. In this space we can choose to present the results in any basis we want and it seems natural to consider the three variables $m1-y$, i^m and Δp. Thus the conclusion about the Finnish data is that the last two variables i^m and Δp are already stationary, and the first two, y and $m1$, are cointegrated.

The Danish Data

In the Danish data we found $r=1$. Based on the unrestricted estimates in the previous section it seems natural to formulate two linear hypotheses in this case, both of which are economically meaningful:

$$H_{3.1}^*:\beta_{1.1}=-\beta_{1.2},$$

and

$$H^*_{3.2}: \beta_{1.3} = -\beta_{1.4}.$$

In matrix formulation the first hypothesis is expressed as

$$\boldsymbol{\beta} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{\varphi},$$

where $\boldsymbol{\varphi}$ is a 4×1 vector. Solving (5.2) gives the eigenvalues in Table 4. These are compared to the eigenvalues of the unrestricted H^*_2 model.

The test of $H^*_{3.1}$ in H^*_2 consists of comparing $\hat{\lambda}^*_{3.1}$ and $\hat{\lambda}^*_1$ by the test

$$-2\ln(Q; H^*_{3.1} \mid H^*_2) = T\{\ln(1-\hat{\lambda}^*_{3.1}) - \ln(1-\hat{\lambda}^*_1)\} = 0.05$$

The asymptotic distribution of this quantity is given by the χ^2 distribution with degrees of freedom $r(p-s) = 1(4-3) = 1$. The test statistic is clearly not significant, and we can thus accept the hypothesis that for the Danish data the coefficients of $m2$ and y are equal with opposite sign.

The second hypothesis that the coefficients for the bond rate and the deposit rate are equal with opposite sign is now tested. This hypothesis implies that the cost of holding money can be measured as the difference between the bond yield and the yield from holding money in bank deposits. Since $H^*_{3.1}$ was strongly supported by the data, we will test $H^*_{3.2}$ within $H^*_{3.1}$. This will now be formulated in matrix notation as

$$\boldsymbol{\beta}^* = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{\varphi},$$

where $\boldsymbol{\varphi}$ is a 3×1 vector. Solving (5.2) we get the eigenvalues reported in Table 4. The test for the hypothesis is given by

$$-2\ln(Q) = 53\{\ln(1-\hat{\lambda}^*_{3.2}) - \ln(1-\hat{\lambda}^*_{3.1})\} = 0.88$$

which should be compared with the χ^2 quantiles with $r(s_1 - s_2) = 1(4-3) = 1$ degree of freedom. It is not significant and we conclude the analysis of the cointegration vectors for the Danish demand for money by the restricted estimate

$$\hat{\boldsymbol{\beta}}^* = (1.00, -1.00, 5.88, -5.88, -6.21),$$

The corresponding estimate of α is given by

$$\hat{\boldsymbol{\alpha}}^* = (-0.177, 0.095, 0.023, 0.032).$$

TABLE 4

The Eigenvalues and the Corresponding Test Statistics for Testing Restrictions on $\boldsymbol{\beta}$

			The Finnish data							
	Eigenvalues $\hat{\lambda}_i$					$-T\ \ln(1-\hat{\lambda}_i)$				
H_2:	0.309	0.226	0.073	0.030		38.49	26.64	7.89	3.11	
H_3:	0.309	0.199	0.070			38.47	23.13	7.60		
			The Danish data							
	Eigenvalues $\hat{\lambda}_i^*$					$-T\ \ln(1-\hat{\lambda}_i^*)$				
H_2^*:	0.433	0.178	0.113	0.043	0	30.09	10.36	6.34	2.35	0
$H_{3.1}^*$:	0.433	0.172	0.044	0.006		30.04	10.01	2.36	0.32	
$H_{3.2}^*$:	0.423	0.045	0.006			29.16	2.44	0.32		

5.2. *The Wald Test*

Instead of the likelihood ratio tests which require estimation under the model H_2 and H_3, one can directly apply the results of model H_2 given in Table 2 to calculate some Wald tests. The idea is to express the restrictions on $\boldsymbol{\beta}$ as $\mathbf{K}'\boldsymbol{\beta}=\mathbf{0}$ and then normalize $\mathbf{K}'\hat{\boldsymbol{\beta}}$ by its 'standard deviation'.

It is shown in Johansen (1989) that if $\hat{\mathbf{v}}^*$ denotes the eigenvectors corresponding to $\hat{\lambda}_2^*,\dots,\hat{\lambda}_5^*$, (see the Danish data in Table 5) then, in case $r=1$, the quantity

$$\omega = T^{1/2}\mathbf{K}^{*\prime}\hat{\boldsymbol{\beta}}_1^* \Bigg/ \left\{(\hat{\lambda}_1^{*-1}-1)\left(\sum_{i=2}^{5}(\mathbf{K}^{*\prime}\hat{\mathbf{v}}_i^*)^2\right)\right\}^{1/2}$$

is asymptotically Gaussian with mean 0 and variance 1. Hence $\mathbf{K}^*=(\mathbf{K}',0)'$, such that $\mathbf{K}^{*\prime}\hat{\boldsymbol{\beta}}^*=\mathbf{K}'\hat{\boldsymbol{\beta}}$, i.e. the contrast involves only the coefficients of the variables, not the constant term. This statistic is easily calculated from Table 5.

If more than one cointegration vector is present, as in the Finnish data, then the Wald statistic is given by

$$\omega^2 = T\otimes tr\{(\mathbf{K}'\hat{\boldsymbol{\beta}}(\hat{\mathbf{D}}^{-1}-\mathbf{I})^{-1}\hat{\boldsymbol{\beta}}'\mathbf{K})(\mathbf{K}'\hat{\mathbf{v}}\hat{\mathbf{v}}'\mathbf{K})^{-1}\},$$

where $\hat{\mathbf{v}}$ is the eigenvector corresponding to $\hat{\lambda}_4$, and $\hat{\mathbf{D}}=\text{diag}(\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3)$ (see the Finnish data in Table 5). The asymptotic distribution of this statistic is χ^2 with $r(p-s)$ degrees of freedom, where $\mathbf{K}$ is $p\times(p-s)$. In this case $r=p-1=3$, and since $r\le s\le p=4$ and, since $s=p$ is no restriction, we can only test a hypothesis with $s=r=p-1=3$, corresponding to a completely specified $\boldsymbol{\beta}$.

The above test statistics require the normalization of $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{v}}$ as in (4.5). An alternative expression for this statistic which can be applied for any normalization is

$$\omega^2 = T\otimes tr\{(\mathbf{K}'\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\Lambda}}^{-1}\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\beta}}'\mathbf{K})(\mathbf{K}'\hat{\mathbf{v}}(\hat{\mathbf{v}}'\mathbf{S}_{kk}\hat{\mathbf{v}})^{-1}\hat{\mathbf{v}}'\mathbf{K})\}.$$

5.2.1. The Empirical Analysis

Since the calculations are numerically simpler for the normalization $\hat{\mathbf{V}}'\mathbf{S}_{kk}\hat{\mathbf{V}}=\mathbf{I}$, it will be used to illustrate the Wald tests. In Table 5 the eigenvalues and eigenvectors for this normalization are reported.

The Danish Data

We start by the hypothesis

$$H_{3.1}^{*}:\beta_{1.1}=-\beta_{1.2}$$

expressed as $\mathbf{K}'\boldsymbol{\beta}=(1,1,0,0)\boldsymbol{\beta}=0$. The Wald statistic is then calculated as follows:

First, $T^{1/2}\mathbf{K}'\hat{\boldsymbol{\beta}}=53^{1/2}(-21.97+22.70)=5.31$, and

$$\sum_{i=2}^{5}(\mathbf{K}^{*\prime}\hat{\mathbf{v}}_i^{*})^2=(14.66-20.05)^2+(7.95-25.64)^2 + (1.02-1.93)^2+(11.36-7.20)^2=360.13.$$

Then the test statistic becomes

$$\omega=5.31/(1/0.4332-1)\times 360.13)^{1/2}=0.24.$$

The second hypothesis

$$H_{3.2}^{*}:\beta_{1.3}=-\beta_{1.4}$$

is tested in a similar way. Note however, that $H_{3.2}^{*}$ is now tested within H_2^{*} and not within $H_{3.1}^{*}$. The test statistics becomes 1.32. Both these statistics are asymptotically normalized Gaussian and the values found are hence not significant.

The Finnish Data

For the Finnish data we only test the hypothesis:

$$H_3:\beta_{i.1}=-\beta_{i.2},\qquad (i=1,2,3).$$

This can be formulated as $\mathbf{K}'\boldsymbol{\beta}=(1,1,0,0)\boldsymbol{\beta}=0$.

First we find from Table 5 that $\mathbf{K}'\hat{\mathbf{w}}\hat{\mathbf{w}}'\mathbf{K}=(1.38+2.22)^2=12.96$ and

$$\mathbf{K}'\hat{\boldsymbol{\beta}}(\hat{\mathbf{D}}^{-1}-I)^{-1}\hat{\boldsymbol{\beta}}'\mathbf{K}=\frac{(-2.93+2.86)^2}{0.3093^{-1}-1}+\frac{(4.58-6.06)^2}{0.2260^{-1}-1} +\frac{(-11.13+10.24)^2}{0.0731^{-1}-1}=0.83.$$

The test statistic becomes $\omega^2=104\times 0.83/12.96=5.66$, which is not significant in the χ^2 distribution with 3 degrees of freedom.

Notice that the Wald test in all cases gives a value of the test statistic which is larger than the value for the likelihood ratio test statistic. This just

TABLE 5

The Eigenvalues $\hat{\lambda}$ and Eigenvectors $\hat{\mathbf{V}}$ Based on the Normalization $\hat{\mathbf{V}}'\mathbf{S}_{kk}\hat{\mathbf{V}} = \mathbf{I}$

	The Danish data						The Finnish data			
	Eigenvalues $\hat{\lambda}^*$						Eigenvalues $\hat{\lambda}$			
	(0.4332	0.1776	0.1128	0.0434	0)		(0.3093	0.2260	0.0731	0.0295)
	Eigenvectors $\hat{\mathbf{V}}^*$						Eigenvectors $\hat{\mathbf{V}}$			
$m2$	−21.97	14.66	7.95	1.02	11.36	$m1$	−2.93	4.58	−11.13	1.38
y	22.70	−20.05	−25.64	−1.93	−7.20	y	2.86	−6.06	10.24	2.22
i^b	−114.42	3.56	4.28	25.00	19.28	i^m	20.79	−9.23	−2.99	−1.90
i^d	92.64	100.26	−44.88	−14.65	−21.53	Δp	20.58	104.17	20.28	−21.62
1	113.16	−62.59	62.75	−2.32	−91.28					

emphasizes the fact that we are relying on asymptotic results and a careful study of the small sample properties is needed.

VI. ESTIMATION AND TESTING UNDER RESTRICTIONS ON α

Let us now turn to the hypothesis H_4 where α is restricted by $\alpha = \mathbf{A}\psi$ in the model H_2. Here $\mathbf{A}$ is a $(p \times m)$ matrix. It is convenient to introduce $\mathbf{B}(p \times (p-m)) = \mathbf{A}_\perp$, such that $\mathbf{B}'\mathbf{A} = \mathbf{0}$. Then the hypothesis H_4 can be expressed as $\mathbf{B}'\alpha = \mathbf{0}$.

The concentrated likelihood function (3.8) can be expressed in the variables given by

$$\mathbf{A}'(\mathbf{R}_{0t} - \alpha\beta'\mathbf{R}_{kt}) = \mathbf{A}'\mathbf{R}_{0t} - \mathbf{A}'\mathbf{A}\psi\beta'\mathbf{R}_{kt}, \tag{6.1}$$

$$\mathbf{B}'(\mathbf{R}_{0t} - \alpha\beta'\mathbf{R}_{kt}) = \mathbf{B}'\mathbf{R}_{0t}. \tag{6.2}$$

In the following, we factor out that part of the likelihood function which depends on $\mathbf{B}'\mathbf{R}_{0t}$, since it does not contain the parameters ψ and β. To save notation, we define:

$$\Lambda_{aa} = \mathbf{A}'\Lambda\mathbf{A}, \ \Lambda_{ab} = \mathbf{A}'\Lambda\mathbf{B}, \ \mathbf{S}_{ak.b} = \mathbf{S}_{ak} - \mathbf{S}_{ab}\mathbf{S}_{bb}^{-1}\mathbf{S}_{bk}$$
$$= \mathbf{A}'\mathbf{S}_{0k} - \mathbf{A}'\mathbf{S}_{00}\mathbf{B}(\mathbf{B}'\mathbf{S}_{00}\mathbf{B})^{-1}\mathbf{B}'\mathbf{S}_{0k}, \text{ etc.}$$

The factor corresponding to the marginal distribution of $\mathbf{B}'\mathbf{R}_{0t}$ is given by

$$|\mathbf{B}'\mathbf{B}|^{T/2}|\Lambda_{bb}|^{-T/2}\exp\left\{-\sum_{t=1}^{T}(\mathbf{B}'\mathbf{R}_{0t})'\Lambda_{bb}^{-1}(\mathbf{B}'\mathbf{R}_{0t})/2\right\}, \tag{6.3}$$

and gives the estimate

$$\hat{\Lambda}_{bb} = \mathbf{S}_{bb} = \mathbf{B}'\mathbf{S}_{00}\mathbf{B}, \tag{6.4}$$

and the maximized likelihood function from the marginal distribution

$$L_{\max}^{-2/T} = |\mathbf{S}_{bb}|/|\mathbf{B}'\mathbf{B}|. \tag{6.5}$$

The other factor corresponds to the conditional distribution of $\mathbf{A}'\mathbf{R}_{0t}$ and $\mathbf{R}_{kt}$ conditional on $\mathbf{B}'\mathbf{R}_{0t}$ and is given by

$$|\mathbf{A}'\mathbf{A}|^{T/2}|\Lambda_{aa.b}|^{-T/2}\exp\left\{-\sum_{t=1}^{T}(\mathbf{A}'\mathbf{R}_{0t} - \mathbf{A}'\mathbf{A}\psi\beta'\mathbf{R}_{kt} - \Lambda_{ab}\Lambda_{bb}^{-1}\mathbf{B}'\mathbf{R}_{0t})'\right.$$
$$\left.\times \Lambda_{aa.b}^{-1}(\mathbf{A}'\mathbf{R}_{0t} - \mathbf{A}'\mathbf{A}\psi\beta'\mathbf{R}_{kt} - \Lambda_{ab}\Lambda_{bb}^{-1}\mathbf{B}'\mathbf{R}_{0t})/2\right\}. \tag{6.6}$$

It is a well-known result from the theory of the multivariate normal distribution that the parameters Λ_{bb}, $\Lambda_{ab}\Lambda_{bb}^{-1}$ and $\Lambda_{aa.b}$ are variation independent and hence that the estimate of $\Lambda_{ab}\Lambda_{bb}^{-1}$ is found by regression for fixed ψ and

$\boldsymbol{\beta}$ giving

$$\hat{\mathbf{\Lambda}}_{ab}\hat{\mathbf{\Lambda}}_{bb}^{-1}(\boldsymbol{\psi},\boldsymbol{\beta})=(\mathbf{S}_{ab}-\mathbf{A}'\mathbf{A}\boldsymbol{\psi}\boldsymbol{\beta}'\mathbf{S}_{kb})\mathbf{S}_{bb}^{-1}, \tag{6.7}$$

and new residuals defined by

$$\tilde{\mathbf{R}}_{at}=\mathbf{A}'\mathbf{R}_{0t}-\mathbf{S}_{ab}\mathbf{S}_{bb}^{-1}\mathbf{B}'\mathbf{R}_{0t},$$

$$\tilde{\mathbf{R}}_{kt}=\mathbf{R}_{kt}-\mathbf{S}_{kb}\mathbf{S}_{bb}^{-1}\mathbf{B}'\mathbf{R}_{0t}.$$

In terms of $\tilde{\mathbf{R}}_{at}$ and $\tilde{\mathbf{R}}_{kt}$ the concentrated likelihood function has the form (3.8) which means that the estimation of $\boldsymbol{\beta}$ follows as before.[6]

THEOREM 6.1: Under the hypothesis

$$H_4\colon \boldsymbol{\alpha}=\mathbf{A}\boldsymbol{\psi},$$

the maximum likelihood estimator of $\boldsymbol{\beta}$ is found as follows: First solve the equation

$$|\lambda\mathbf{S}_{kk.b}-\mathbf{S}_{ka.b}\mathbf{S}_{aa.b}^{-1}\mathbf{S}_{ak.b}|=0, \tag{6.8}$$

giving $\hat{\lambda}_{4.1}>\ldots>\hat{\lambda}_{4.m}>\hat{\lambda}_{4.m+1}=\ldots=\hat{\lambda}_{4.p}=0$ and $\hat{\mathbf{V}}_4=(\hat{\mathbf{v}}_{4.1},\ldots,\hat{\mathbf{v}}_{4.p})$ normalized such that $\hat{\mathbf{V}}_4'\mathbf{S}_{kk.b}\hat{\mathbf{V}}_4=\mathbf{I}$.

Now take

$$\hat{\boldsymbol{\beta}}_4=(\hat{\mathbf{v}}_{4.1},\ldots,\hat{\mathbf{v}}_{4.r}), \tag{6.9}$$

which gives the estimates

$$\boldsymbol{\psi}=(\mathbf{A}'\mathbf{A})^{-1}\mathbf{S}_{ak.b}\hat{\boldsymbol{\beta}}_4 \tag{6.10}$$

and

$$\hat{\boldsymbol{\alpha}}_4=\mathbf{A}\hat{\boldsymbol{\psi}}=\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'(\mathbf{S}_{0k}-\mathbf{S}_{00}\mathbf{B}(\mathbf{B}'\mathbf{S}_{00}\mathbf{B})^{-1}\mathbf{B}'\mathbf{S}_{0k})\hat{\boldsymbol{\beta}}_4, \tag{6.11}$$

$$\hat{\mathbf{\Lambda}}_{aa.b}=\mathbf{S}_{aa.b}-\mathbf{A}'\mathbf{A}\hat{\boldsymbol{\psi}}\hat{\boldsymbol{\psi}}'\mathbf{A}'\mathbf{A}=\mathbf{S}_{aa.b}-\mathbf{A}'\hat{\boldsymbol{\alpha}}_4\hat{\boldsymbol{\alpha}}_4'\mathbf{A}, \tag{6.12}$$

and the maximized likelihood function

$$L_{\max}^{-2/T}(H_4)=|\mathbf{B}'\mathbf{B}|^{-1}|\mathbf{A}'\mathbf{A}|^{-1}|\mathbf{S}_{bb}||\mathbf{S}_{aa.b}|\prod_{i=1}^{r}(1-\hat{\lambda}_{4.i})=|\mathbf{S}_{00}|\prod_{i=1}^{r}(1-\hat{\lambda}_{4.i}). \tag{6.13}$$

The estimate of $\mathbf{\Lambda}$ can be found from (6.4), (6.7) and (6.12), and $\mathbf{\Gamma}$ is estimated from (3.5).

The likelihood ratio test statisic of H_4 in H_2 is

$$-2\ln(Q;H_4|H_2)=T\sum_{i=1}^{r}\ln\{(1-\hat{\lambda}_{4.i})/(1-\hat{\lambda}_i)\}. \tag{6.14}$$

The asymptotic distribution of this test statistic is given by a χ^2 distribution with $r(p-m)$ degrees of freedom, see Johansen (1989). The same result holds for testing $H_4^*\colon \boldsymbol{\alpha}=\mathbf{A}\boldsymbol{\psi}$ in H_2^*.

[6] It is convenient to calculate the relevant product moment matrices as

$$S_{ij.b}=T^{-1}\sum_{t=1}^{T}\tilde{R}_{it}\tilde{R}_{jt}' \qquad (i,j=a,k)$$

The following very simple Corollary is useful for explaining the role of single equation analysis:

COROLLARY 6.2: If $m = r = 1$ then the maximum likelihood estimate of $\boldsymbol{\beta}$ is found as the coefficients of $\mathbf{X}_{t-k}$ in the regression of $\mathbf{A}'\Delta\mathbf{X}_t$ on $\mathbf{X}_{t-k}$, $\mathbf{B}'\Delta\mathbf{X}_t$, and $\Delta\mathbf{X}_{t-1}, \ldots, \Delta\mathbf{X}_{t-k+1}$, $\mathbf{D}_t$ and the constant.

PROOF: It suffices to notice that when $m = r = 1$, then only one cointegration vector has to estimated. It is seen from (6.8) that since the matrix $\mathbf{S}_{ka.b}\mathbf{S}_{aa.b}^{-1}\mathbf{S}_{ak.b}$ is singular and in fact of rank 1, then only one eigenvalue is non-zero, and the corresponding eigenvector is proportional to $\mathbf{S}_{kk.b}^{-1}\mathbf{S}_{ka.b}$, which is exactly the regression coefficient of $\mathbf{R}_{kt}$ obtained by regressing $\mathbf{A}'\mathbf{R}_{0t}$ on $\mathbf{B}'\mathbf{R}_{0t}$ and $\mathbf{R}_{kt}$. This can of course be seen directly from (6.6) since $\mathbf{A}'\mathbf{A}\boldsymbol{\psi}$ is 1×1 and can be absorbed into $\boldsymbol{\beta}$, which shows that $\boldsymbol{\beta}$ is given by the regression as described. If, in particular, $\boldsymbol{\alpha}$ is proportional to $(1,0,\ldots,0)$, then ordinary least squares analysis of the first equation will give the maximum likelihood estimation for the cointegration vector. An empirical illustration of this will be presented below.

Finally we just state briefly how one solves the estimation and testing of the model $H_5: \boldsymbol{\beta} = \mathbf{H}\boldsymbol{\varphi}$ and $\boldsymbol{\alpha} = \mathbf{A}\boldsymbol{\varphi}$. In this case we note that $\boldsymbol{\beta}'\mathbf{R}_{kt} = \boldsymbol{\varphi}'\mathbf{H}'\mathbf{R}_{kt}$ which leads to solving (5.2) where $\mathbf{R}_{kt}$ has been replaced by $\mathbf{H}'\mathbf{R}_{kt}$. Thus restricting $\boldsymbol{\beta}$ to lie in $sp(\mathbf{H})$ implies that the levels of the process should be transformed by $\mathbf{H}'$.

Since $\boldsymbol{\alpha} = \mathbf{A}\boldsymbol{\psi}$ we solve (6.8), where we have conditioned on $\mathbf{B}'\mathbf{R}_{0t}$. In other words if we assume that the equations for $\mathbf{B}'\mathbf{R}_{0t}$ do not contain the parameter $\boldsymbol{\alpha}$, i.e. $\mathbf{B}'\boldsymbol{\alpha} = \mathbf{0}$, then we can correct for these before solving the eigenvalue problem. It is now clear how one should solve the model $H_5 = H_4 \cap H_3$, where restrictions have been imposed on $\boldsymbol{\beta}$ as well as on $\boldsymbol{\alpha}$, namely by solving the eigenvalue problem

$$|\lambda \mathbf{H}'\mathbf{S}_{kk.b}\mathbf{H} - \mathbf{H}'\mathbf{S}_{ka.b}\mathbf{S}_{aa.b}^{-1}\mathbf{S}_{ka.b}\mathbf{H}| = 0. \tag{6.15}$$

This gives the final solution to the estimation problem of H_5. Notice how (6.15) contains the previous problems by choosing either $\mathbf{H} = \mathbf{I}$ or $\mathbf{A} = \mathbf{I}$ or both. We have, however, chosen to present the analysis of restrictions on $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ separately in order to simplify the notation.

Finally, note that a linear restriction on $\boldsymbol{\beta}$ implies a transformation of the process, and that a linear restriction on $\boldsymbol{\alpha}$ implies a conditioning. Thus all the calculations can easily be performed starting with the product moment matrices $\mathbf{S}_{ij}$ and applying the usual operations of finding marginal (transformed) and conditional variances followed by an eigenvalue routine.

6.1. The Empirical Analysis

In Section V it was shown that the hypothesis about proportionality between money and income, $\beta_{i1} = -\beta_{i2}$, was accepted both for the Danish and the Finnish data, and that the hypothesis $\beta_{13} = -\beta_{14}$ was accepted for the Danish data. Thus it seems natural to move directly to the H_5 and the H_5^*

hypothesis, see Section II, and test hypotheses about $\boldsymbol{\alpha}$ in the $\boldsymbol{\beta}$-restricted models. For illustrative purposes we will, however, also present the empirical results for just one H_4 hypothesis, i.e. a restriction on $\boldsymbol{\alpha}$ for unrestricted $\boldsymbol{\beta}$.

The Danish Data

We denote the cointegration vector by $(\beta_1,\ldots,\beta_5)$ and the weights by $(\alpha_1,\ldots,\alpha_4)$. Since we have no *a priori* hypothesis about the α's except that $\alpha_1 \neq 0$, we have at most three hypotheses about zero restrictions on α to test. We have chosen to demonstrate the hypothesis: $H_{5.1}: \alpha_3 = 0$, in the text, and report the results of other tests in Table 6.

The test results are summarized in the upper part of the table. The estimates of $\boldsymbol{\beta}_1$ and $\boldsymbol{\alpha}_1$ are presented below the test statistics. To facilitate comparison with the previous results the first column of the table gives the estimates under the unrestricted model H_2^*, the second and third under one and two $\boldsymbol{\beta}$ restrictions, the next three under three $\boldsymbol{\alpha}$ restrictions and finally the last column gives the estimates under three $\boldsymbol{\alpha}$ restrictions but for unrestricted $\boldsymbol{\beta}$.

Based on the calculated values of $-T\ln(1-\lambda_{i.1}^*)$ in Table 6 it is now possible to test any of the three α hypotheses $H_{5.i}^*$, $i=1,2,3$ against $H_{3.2}^*$, or any of the $H_{5.i}^*$ hypotheses with fewer restrictions on α. The likelihood ratio test statistic for $H_{5.i}^*$ versus $H_{3.2}^*$ is calculated as

$$-2\ln(Q; H_{5.i}^* | H_{3.2}^*) = T\{\ln(1-\hat{\lambda}^*(H_{5.i}^*)) - \ln(1-\hat{\lambda}_{3.2}^*)\}$$

which is asymptotically distributed as χ^2 with $(p-(p-i))r = i$ degrees of freedom, when $r=1$. For instance we consider $H_{5.1}^*$ versus $H_{3.2}^*$ and find

$$-2\ln(Q; H_{5.1}^* | H_{3.2}^*) = 29.15 - 27.96 = 1.19 < 3.84 = \chi_{0.95}^2(1).$$

The other tests have been linearly ordered in Table 6, and we can choose any hypothesis of interest and test against hypotheses with fewer restrictions. Since we had no strong *a priori* hypotheses about α_2, α_3 and α_4, the various tests we have performed can be seen as a form for data exploration rather than as specification testing in the strict sense.

We proceed to the H_4^* hypotheses described in the last column of Table 6, i.e.

$$H_{4.1}^*: \alpha_2 = \alpha_3 = \alpha_4 = 0$$

for unrestricted $\boldsymbol{\beta}$. As shown in Corollary 6.2, the acceptance of this hypothesis would legitimize the use of single equation estimation of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and is therefore of particular interest.

The hypothesis $H_{4.1}^*$ is first tested by the likelihood ratio test and the corresponding estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ derived. We then give the corresponding ordinary least squares estimates and show that the two procedures give the same result.

TABLE 6

Estimates for the Danish Data of the First Eigenvalue and the Corresponding $\boldsymbol{\beta}$- and $\boldsymbol{\alpha}$-vector under various hypotheses about $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$

	H_2^*	$H_{3.1}^*$	$H_{3.2}^*$	$H_{5.1}^*$	$H_{5.2}^*$	$H_{5.3}^*$	$H_{4.1}^*$
β-restrictions	—	$\beta_1 = -\beta_2$	$\beta_1 = -\beta_2$ and $\beta_3 = -\beta_4$	$\beta_1 = -\beta_2$ and $\beta_3 = -\beta_4$	$\beta_1 = -\beta_2$ and $\beta_3 = -\beta_4$	$\beta_1 = -\beta_2$ and $\beta_3 = -\beta_4$	—
α-restrictions	—	—	—	$\alpha_3 = 0$	$\alpha_3 = \alpha_4 = 0$	$\alpha_3 = \alpha_4 = \alpha_2 = 0$	$\alpha_3 = \alpha_4 = \alpha_2 = 0$
$\hat{\lambda}_1^*$	0.433	0.432	0.423	0.410	0.356	0.286	0.357
$-T \ln(1 - \hat{\lambda}_1^*)$	30.09	30.04	29.15	27.96	23.34	17.91	23.42
β_1	1.00	1.00	1.00	1.00	1.00	1.00	1.00
β_2	−1.03	−1.00	−1.00	−1.00	−1.00	−1.00	−0.96
β_3	5.21	5.30	5.88	5.95	5.81	5.88	4.76
β_4	−4.22	−4.29	−5.88	−5.95	−5.81	−5.88	−2.57
β_5	−6.06	−6.26	−6.21	−6.22	−6.21	−6.21	−6.58
α_1	−0.213	−0.212	−0.177	−0.152	−0.137	−0.197	−0.254
α_2	0.115	0.108	0.095	0.099	0.139	0	0
α_3	0.023	0.022	0.023	0	0	0	0
α_4	0.029	0.030	0.032	0.029	0	0	0

The appropriate $\mathbf{A}$ and $\mathbf{B}$ matrices are now

$$\mathbf{A}=\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \mathbf{B}=\begin{pmatrix}0&0&0\\1&0&0\\0&1&0\\0&0&1\end{pmatrix},$$

on the basis of which the matrices $\mathbf{S}^*_{kk.b}$, $\mathbf{S}^*_{ka.b}$, $\mathbf{S}^*_{aa.b}$ and $\mathbf{S}^*_{ak.b}$ can be calculated. Solving the eigenvalue problem (6.8) gives one eigenvalue 0.357 and consequently one eigenvector $\hat{\boldsymbol{\beta}}$, as well as one estimate $\hat{\boldsymbol{\alpha}}$. Normalized by the coefficient of $m2$ the estimates are

$$\hat{\boldsymbol{\beta}}(H^*_{4.1})=(1.00, -0.96, 4.76, -2.57, -6.58)$$
$$\hat{\boldsymbol{\alpha}}(H^*_{4.1})=(-0.25, 0, 0, 0).$$

The test statistic for this hypothesis about $\boldsymbol{\alpha}$ is then given by

$$\begin{aligned}-2\ln(Q; H^*_{4.1}\,|\,H^*_2) &= T\{\ln(1-\hat{\lambda}^*_{4.1})-\ln(1-\hat{\lambda}^*_1)\}\\ &= 30.09-23.42=7.67<7.81=\chi^2_{0.95}(3).\end{aligned}$$

Although the test statistic is not significant at the 5 percent level it would be so at a slightly higher level. On the basis of this we conclude that there is no strong support for restricting α_2, α_3 and α_4 to zero.

The single equation estimation corresponds to the calculation of the static long-run solution of general autoregressive model

$$A_1(L)\,m_t=A_2(L)\,y_t+A_3(L)\,i^b_t+A_4(L)\,i^d_t+a_5+\mathbf{A}_6'\mathbf{D}_t+\varepsilon_t, \tag{6.16}$$

where ε_t are independent Gaussian variables with mean zero and variance σ^2_ε, and $A_i(L)$, $i=1,\ldots,4$, is a lag polynomial of order 2, normalized at $A_1(0)=1$.

The static long-run solution is obtained by evaluating (6.16) at $L=1$, which gives the estimate of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ normalized by the coefficient of m_t as

$$\hat{\boldsymbol{\beta}}=A_1(1)^{-1}[-A_1(1), A_2(1), A_3(1), A_4(1), a_5]$$
$$\hat{\alpha}=A_1(1).$$

The OLS estimation of (6.16) evaluated at $L=1$ gives:

$$(A_1(1),\ldots, A_4(1), a_5)=(0.254, 0.244, -1.211, 0.654, 1.698),$$

from which the static long-run solution can be calculated as

$$\begin{aligned}m=\ &0.96y-4.76i^b+2.57i^d+6.58,\\ &(0.19)\quad(0.83)\quad(1.46)\quad(2.06)\end{aligned}$$

i.e. exactly the same estimates as in the restricted maximum likelihood procedure, see Corollary 6.2.

We conclude the empirical analysis by a comparison of the estimated $\mathbf{\Pi}$-matrices under the full unrestricted H_1-model and the final version $\mathbf{\Pi}=\boldsymbol{\alpha}\boldsymbol{\beta}'$ with data consistent restrictions on $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ (see Table 7). For the Danish data

TABLE 7

The Unrestricted **Π**-*Matrix Compared to the Reduced Rank Restricted* **Π**

	The Danish data						The Finnish Data			
	−0.18	0.11	−1.04	0.64	1.58		−0.12	0.11	−0.23	−0.42
H_1^*:	0.19	−0.31	0.66	−0.65	−0.39	H_1:	0.02	−0.04	−0.14	−0.21
	0.01	−0.02	0.08	−0.17	−0.06		0.10	−0.10	−0.48	0.67
	0.00	0.02	0.14	−0.31	−0.07		0.00	0.01	0.00	−0.48
	−0.15	0.15	−0.89	0.89	0.93		−0.11	0.11	−0.22	−0.57
H_5^*:	0.10	−0.10	0.60	−0.60	−0.62	H_3:	0.03	−0.03	−0.13	−0.42
	0	0	0	0	0		0.10	−0.10	−0.47	0.53
	0.03	−0.03	0.17	−0.17	−0.18		−0.00	0.00	−0.01	−0.31

the number of parameters (excluding the constant) has been reduced from 16 to 4, whereas for the Finnish data the reduction is from 16 parameters to 6.

VII. SUMMARY

In this paper we have addressed the estimation and testing problem of long-run relations in economic modelling. The solution we propose is to start with a relatively simple model specifying a vector valued autoregressive process (VAR) including a constant term and seasonal dummies, and with independent Gaussian errors. The hypothesis of the existence of cointegration vectors is formulated as the hypothesis of reduced rank of the long-run impact matrix. This is given a simple parametric form which allows the application of the method of maximum likelihood and likelihood ratio tests. In this way we can derive estimates and test statistics for the hypothesis of a given number of cointegration vectors, as well as estimates and tests for linear hypotheses about the cointegration vectors and their weights. The asymptotic inferences concerning the number of cointegrating vectors involve non-standard distributions, see Johansen (1989), and these are tabulated by simulation. Inference concerning linear restrictions on the cointegration vectors and their weights can be performed using the usual χ^2 methods. The test procedures are in general likelihood ratio tests, but in the case of linear restrictions on β a Wald test procedure is suggested as an alternative to the likelihood ratio test procedure.

It is shown that the inclusion of the constant term in the general VAR-model has significant effects on the statistical properties of the described tests for the reduced rank model. The role of the constant term is closely related to the question of whether there are linear trends or not in the levels of the data, and it is demonstrated that the estimation procedure as well as the distribution of the test statistics of the reduced rank model is strongly affected by the assumption of how the constant term is related to the stationary and the non-stationary part of the model.

The proposed methods are illustrated by money demand data from the Danish and the Finnish economy. The applications were chosen to illustrate various aspects of the cointegration method. The model for the Danish demand for money is specified without assuming a linear trend in the data, whereas the Finnish model allows for linear trends in the non-stationary part of the model. The order of cointegration was one for the Danish version, which simplified the interpretation of the cointegration vectors as a long-run relation in the levels of the process. For the Finnish data there were three cointegration vectors which served to illustrate the interpretational problems when there are several cointegration vectors in the data.

Institute of Mathematical Statistics and Institute of Economics,
University of Copenhagen

Date of Receipt of Final Manuscript: December 1989

APPENDIX. SIMULATION OF THE LIMITING DISTRIBUTIONS

The limit distributions are expressed as functions of the stochastic matrix

$$\left\{\int_0^1 (\mathrm{d}\mathbf{U})\mathbf{F}'\left[\int_0^1 \mathbf{F}\mathbf{F}'\mathrm{d}t\right]^{-1}\int_0^1 \mathbf{F}(\mathrm{d}\mathbf{U})'\right\},$$

see Section IV.

The $(p-r)$-dimensional Brownian motion $\mathbf{U}(t)=\{U_1(t),\ldots,U_{p-r}(t)\}$ is approximated by a random walk with $T=400$ steps. Thus we generate a $T\times(p-r)$ array of i.i.d. Gaussian variables

$$\varepsilon_{ti}, \qquad (t=1,\ldots,T,\ i=1,\ldots,p-r)$$

and calculate X_t from

$$X_{ti}=\sum_{s=1}^{t}\varepsilon_{si}, \qquad (t=1,\ldots,T,\ i=1,\ldots,p-r),$$

with $X_{0i}=0$, $i=1,\ldots,p-r$. In case the process $\mathbf{F}$ is given by $\mathbf{U}-\bar{\mathbf{U}}$, see (4.11) the stochastic matrix $\int\mathbf{F}\mathbf{F}'\mathrm{d}t$ and $\int\mathbf{F}\,\mathrm{d}\mathbf{U}'$ are approximated by

$$T^{-2}\sum_{t=1}^{T}(\mathbf{X}_{t-1}-\bar{\mathbf{X}}_{-1})(\mathbf{X}_{t-1}-\bar{\mathbf{X}}_{-1})'$$

and

$$T^{-1}\sum_{t=1}^{T}(\mathbf{X}_{t-1}-\bar{\mathbf{X}}_{-1})\,\varepsilon_t'$$

respectively, where $\bar{\mathbf{X}}_{-1}=T^{-1}\sum_{t=1}^{T}\mathbf{X}_{t-1}$. From these expressions we calculate

$$\left\{\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t(\mathbf{X}_{t-1}-\bar{\mathbf{X}}_{-1})'\left[\sum_{t=1}^{T}(\mathbf{X}_{t-1}-\bar{\mathbf{X}}_{-1})(\mathbf{X}_{t-1}-\bar{\mathbf{X}}_{-1})\right]^{-1}\sum_{t=1}^{T}(\mathbf{X}_{t-1}-\bar{\mathbf{X}}_{-1})\,\boldsymbol{\varepsilon}_t'\right\}.$$

From this matrix the trace and the maximum eigenvalue are calculated. On the basis of 6,000 simulations the quantiles are found as the appropriate order statistics.

If instead $\mathbf{F}$ is given by (4.14) and (4.15), we replace in the above calculation the last component of $\mathbf{X}_{t-1}-\bar{\mathbf{X}}_{-1}$ by $t-T/2$, and if $\mathbf{F}$ is given by (4.16) and (4.17) then $\bar{\mathbf{X}}_{-1}$ is dropped and $\mathbf{X}_{t-1}$ is extended by an extra component 1.

TABLE A1

Distribution of the Maximal Eigenvalue and Trace of the Stochastic Matrix

$$\int(d\mathbf{U})\,\mathbf{F}'[\int\mathbf{F}\mathbf{F}'du]^{-1}\int\mathbf{F}(d\mathbf{U}')$$

where $\mathbf{U}$ is an m-dimensional Brownian motion and $\mathbf{F}$ is $\mathbf{U}-\bar{\mathbf{U}}$, except that the last component is replaced by $t-1/2$, see Theorem 4.1

dim	*50%*	*80%*	*90%*	*95%*	*97.5%*	*99%*	*mean*	*var*
				Maximal eigenvalue				
1.	0.447	1.699	2.816	3.962	5.332	6.936	1.030	2.192
2.	6.852	10.125	12.099	14.036	15.810	17.936	7.455	12.132
3.	12.381	16.324	18.697	20.778	23.002	25.521	12.951	18.549
4.	17.719	22.113	24.712	27.169	29.335	31.943	18.275	23.837
5.	23.211	27.899	30.774	33.178	35.546	38.341	23.658	28.330
				Trace				
1.	0.447	1.699	2.816	3.962	5.332	6.936	1.030	2.192
2.	7.638	11.164	13.338	15.197	17.299	19.310	8.250	14.065
3.	18.759	23.868	26.791	29.509	32.313	35.397	19.342	32.103
4.	33.672	40.250	43.964	47.181	50.424	53.792	34.184	55.249
5.	52.588	60.215	65.063	68.905	72.140	76.955	52.998	82.106

Simulations are performed replacing the Brownian motion by a Gaussian random walk with 400 steps and the process is stimulated 6,000 times.

TABLE A2

Distribution of the Maximal Eigenvalue and Trace of the Stochastic Matrix

$$\int(d\mathbf{U})\,\mathbf{F}'[\int\mathbf{F}\mathbf{F}'du]^{-1}\int\mathbf{F}(d\mathbf{U}')$$

where $\mathbf{U}$ is an m-dimensional Brownian motion and $\mathbf{F}=\mathbf{U}-\bar{\mathbf{U}}$

dim	*50%*	*80%*	*90%*	*95%*	*97.5%*	*99%*	*mean*	*var*
				Maximal eigenvalue				
1.	2.415	4.905	6.691	8.083	9.658	11.576	3.030	7.024
2.	7.474	10.666	12.783	14.595	16.403	18.782	8.030	12.568
3.	12.707	16.521	18.959	21.279	23.362	26.154	13.278	18.518
4.	17.875	22.341	24.917	27.341	29.599	32.616	18.451	24.163
5.	23.132	27.953	30.818	33.262	35.700	38.858	23.680	29.000
				Trace				
1.	2.415	4.905	6.691	8.083	9.658	11.576	3.030	7.024
2.	9.335	13.038	15.583	17.844	19.611	21.962	9.879	18.017
3.	20.188	25.445	28.436	31.256	34.062	37.291	20.809	34.159
4.	34.873	41.623	45.248	48.419	51.801	55.551	35.475	56.880
5.	53.373	61.566	65.956	69.977	73.031	77.911	53.949	84.092

Simulations are performed by replacing the Brownian motion by a Gaussian random walk with 400 steps and the process is simulated 6,000 times.

TABLE A3

Distribution of the Maximal Eigenvalue and Trace of the Stochastic Matrix

$$\int (\mathbf{dU})\,\mathbf{F}'[\int \mathbf{FF}'\mathrm{d}u]^{-1}\int \mathbf{F}(\mathbf{dU}')$$

where **U** is an m-dimensional Brownian motion and **F** is an $(m+1)$-dimensional process equal to **U** extended by 1, see Theorem 4.1

dim	*50%*	*80%*	*90%*	*95%*	*97.5%*	*99%*	*mean*	*var*
				Maximal eigenvalue				
1.	3.474	5.877	7.563	9.094	10.709	12.740	4.068	6.738
2.	8.337	11.628	13.781	15.752	17.622	19.834	8.917	13.021
3.	13.494	17.474	19.796	21.894	23.836	26.409	14.050	18.698
4.	18.592	22.938	25.611	28.167	30.262	33.121	19.172	23.607
5.	23.817	28.643	31.592	34.397	36.625	39.672	24.433	28.954
				Trace				
1.	3.474	5.877	7.563	9.094	10.709	12.741	4.068	6.738
2.	11.381	15.359	17.957	20.168	22.202	24.988	12.017	19.192
3.	23.243	28.768	32.093	35.068	37.603	40.198	23.868	37.529
4.	38.844	45.635	49.925	53.347	56.449	60.054	39.431	59.854
5.	58.361	66.624	71.472	75.328	78.857	82.969	58.954	89.072

Simulations are performed by replacing the Brownian motion by a Gaussian random walk with 400 steps and the process is simulated 6,000 times.

REFERENCES

Ahn, S. K. and Reinsel, G. C. (1987). 'Estimation for Partially Non-stationary Multivariate Autoregressive Models', University of Wisconsin.

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis.* New York, Wiley.

Box, G. E. P. and Tiao, G. C. (1981). 'A Canonical Analysis of Multiple Time Series with Applications', *Biometrika*, Vol. 64, pp. 355–65.

Davidson, J. (1986). 'Cointegration in Linear Dynamic Systems', Discussion paper, LSE.

Dolado, J. J. and Jenkinson, T. (1988). 'Cointegration: A Survey of Recent Developments', Mimeo, Institute of Economics and Statistics, Oxford University.

Engle, R. F. and Granger, C. W. J. (1987). 'Co-integration and Error Correction: Representation, Estimation and Testing', *Econometrica*, Vol. 55, pp. 251–76.

Fuller, W. A. (1976). *Introduction to Statistical Time Series*, New York, Wiley.

Granger, C. W. J. (1981). 'Some Properties of Time Series Data and their Use in Econometric Model Specification', *Journal of Econometrics*, Vol. 16, pp. 121–30.

Granger, C. W. J. and Weiss, A. A. (1983). 'Time Series Analysis of Error Correction Models', in *Studies in Econometrics, Time Series and Multivariate Statistics*, in Karlin, S., Amemiya, T. and Goodman, L. A. (eds.), New York, Academic Press, pp. 255–78.

Jarque, C. M. and Bera, A. K. (1980). 'Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals', *Economic Letters*, Vol. 6, pp. 255–59.

Jeganathan, P. (1988). 'Some Aspects of Asymptotic Theory with Applications to Time Series Models', Technical Report The University of Michigan.

Johansen, S. (1988a). 'The Mathematical Structure of Error Correction Models', *Contemporary Mathematics*, American Mathematical Society, Vol. 80, pp. 359–86.

Johansen, S. (1988b). 'Statistical Analysis of Cointegration Vectors', *Journal of Economic Dynamics and Control*, Vol. 12, pp. 231–54.

Johansen, S. (1989). 'Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models', forthcoming, *Econometrica*.

Johansen, S. and Juselius, K. (1988). 'Hypothesis Testing for Cointegration Vectors — With an Application to the Demand for Money in Denmark and Finland', Preprint University of Copenhagen.

Laidler, D. E. W. (1985). 'The Demand for Money', Oxford, Philip Allan.

Pena, D. and Box, G. E. P. (1987). 'Identifying a Simplifying Structure in Time Series', *Journal of the American Statistical Association*, Vol. 82, pp. 836–43.

Phillips, P. C. B. (1987). 'Multiple Regression with Integrated Time Series', Cowles Foundation discussion paper No. 852.

Phillips, P. C. B. (1988). 'Optimal Inference in Cointegrated Systems', Cowles Foundation discussion paper No. 866.

Phillips, P. C. B. and Durlauf, S. N. (1986). 'Multiple Time Series Regression with Integrated Processes', *Review of Economic Studies*, Vol. 53, pp. 473–95.

Phillips, P. C. B. and Ouliaris, S. (1986). 'Testing for Cointegration', Cowles Foundation discussion paper No. 809.

Phillips, P. C. B. and Ouliaris, S. (1987). 'Asymptotic Properties of Residual Based Tests for Cointegration', Cowles Foundation discussion paper No. 847.

Phillips, P. C. B. and Park, J. Y. (1986a). 'Asymptotic Equivalence of OLS and GLS in Regression with Integrated Regressors', Cowles Foundation discussion paper No. 802.

Phillips, P. C. B. and Park, J. Y. (1988). 'Statistical Inference in Regressions with Integrated Processes. Part 1', Econometric Theory, Vol. 4, pp. 468–97.

Phillips, P. C. B. and Park, J. Y. (1989). 'Statistical Inference in Regressions with Integrated Processes. Part 2', Econometric Theory, Vol. 3, pp. 95–131.

Sims, A., Stock, J. H. and Watson, M. W. (1986). 'Inference in Linear Time Series Models with some Unit Roots', Preprint.

Stock, J. H. (1987). 'Asymptotic Properties of Least Squares Estimates of Cointegration Vectors', *Econometrica*, Vol. 55, pp. 1035–56.

Stock, J. H. and Watson, M. W. (1988). 'Testing for Common Trends', *Journal of the American Statistical Association*, Vol. 83, pp. 1097–1107.

Tso, M. K.-S. (1981). 'Reduced-rank Regression and Canonical Analysis', *Journal of the Royal Statisical Society B*, Vol. 43, pp. 183–89.

Velu, R. P., Reinsel, G. C. and Wichern, D. W. (1986). 'Reduced Rank Models for Multiple Time Series', *Biometrika*, Vol. 73, pp. 105–18.

Velu, R. P., Wichern, D. W. and Reinsel, G. C. (1987). 'A Note on Non-stationary and Canonical Analysis of Multiple Time Series Models', *Journal of Time Series Analysis*, Vol. 8, pp. 479–87.

West, D. D. (1989). 'Asymptotic Normality when Regressors Have a Unit Root', *forthcoming Econometrica.*