Søren Johansen

The Role of the Constant Term in Cointegration Analysis of Nonstationary Variables

Preprint
January
1992

1

Institute of Mathematical Statistics
University of Copenhagen

ISSN 0902-8846

Søren Johansen

THE ROLE OF THE CONSTANT TERM IN COINTEGRATION ANALYSIS OF NONSTATIONARY VARIABLES

Preprint 1992 No. 1

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

January 1992

THE ROLE OF THE CONSTANT TERM IN COINTEGRATION ANALYSIS OF NONSTATIONARY VARIABLES

Søren Johansen

Institute of Mathematical Statistics
Universitetsparken 5
2100 Copenhagen Ø
Denmark

Keywords and Phrases: cointegration; autoregressive models; linear trend; quadratic trend; reduced rank regression;

ABSTRACT

The autoregressive model for cointegrated variables is analyzed with respect the role of the constant term. A number of models for I(1) variables defined by restrictions on the constant term is discussed, and it is shown that statistical inference can be performed by reduced rank regression. The asymptotic distribution of the test statistics and estimators are found. A similar analysis is given of models for I(2) variables.

1. INTRODUCTION

The constant term in an autoregressive model for nonstationary variables gives rise to a trend. Similarly a linear term in the model gives rise to a polynomial trend of degree determined by the coefficients of the autoregressive model. It is the purpose of this paper to discuss in detail how the assumption of cointegration leads to various interpretations of the constant term and the linear term in an autoregressive model.

If we assume about the model that it defines I(1) variables, then the linear term implies a quadratic trend in the variables, and we discuss below some hypotheses expressed as nonlinear restrictions on the adjustment coefficients and the deterministic terms with the purpose of testing that there is no quadratic trend, that the linear trend only appears in the nonstationary components of the process, that the variables are trend stationary etc. It turns out that estimation of parameters under various hypotheses can be performed by reduced rank regression, but the asymptotic distribution of the test for cointegration depends on the assumption about the deterministic part of the model, that is, on the presence of trends in the variables. These problems have been discussed for the constant term in Johansen and Juselius (1990) and in Johansen (1992a), and the new results here are the results for the quadratic trend and the results for trend stationarity. The results are collected for comparison, and some new tables provided.

If the model allows for I(2) variables, the constant term will imply a quadratic trend in the process and there are many different models that can be expressed as restrictions on the various adjustment coefficients and the constant. We give a brief survey of the I(2) model and the properties of the process and discuss how the statistical analysis of the I(2) model can be performed by repeated application of reduced rank analysis.

2. THE I(1) MODELS AND THEIR INTERPRETATION

We first define some models for I(1) variables derived by restricting the deterministic term in the defining equations. Then we discuss the interpretation of these models and investigate how the various restrictions on the deterministic terms influence the behavior of the process.

2.1 The I(1) models

We therefore consider the vector autoregressive model for the p–dimensional process written as a reduced form error correction model

$$\Delta X_t = \Pi X_{t-1} + \Sigma_1^{k-1}\Gamma_i \Delta X_{t-i} + \mu_0 + \mu_1 t + \epsilon_t, \; t = 1,\dots,T, \qquad (2.1)$$

with initial values $X_{-k+1},\dots,X_0$ kept fixed and errors that are independent Gaussian with mean zero and variance matrix Ω. The hypothesis of cointegration is formulated as the hypothesis of reduced rank of Π, that is

$$\Pi = \alpha\beta' \qquad (2.2)$$

where α and β are $p \times r$ matrices, or in other words, that the rank of Π is less than or equal to r.

We define $\Gamma = I + \Pi - \Sigma_1^{k-1}\Gamma_i$, and apply Granger's theorem, see Engle and Granger (1987), in the form given in Johansen (1989), which states that if α and β have rank r, and $\alpha_\perp'\Gamma\beta_\perp$ has full rank, where $\alpha_\perp$ and $\beta_\perp$ is a $p \times (p-r)$ matrix of full rank orthogonal to α and β respectively, then X_t has the representation

$$X_t = X_0 + C\Sigma_1^t \epsilon_i + \tfrac{1}{2}\tau_2 t^2 + \tau_1 t + \tau_0 + Y_t - Y_0 \qquad (2.3)$$

where Y_t is a stationary process with $\beta' X_0 = \beta' Y_0$. Here $C = \beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp'$, and $\tau_2 = C\mu_1$. Note that if $r = 0$ then $\alpha = \beta = 0$, and we can take $\alpha_\perp = \beta_\perp = I$. If $r = p$ we define $\alpha_\perp = \beta_\perp = 0$ and $C = 0$.

It follows that in general X_t has a quadratic trend, $\tfrac{1}{2}\tau_2 t$, and that the stochastic part of X_t is a nonstationary process with stationary differences, a so called I(1) process. Note also that the reduced rank of Π implies that the stochastic part of $\beta' X_t$ is stationary. The space $sp(\beta)$ is called the cointegrating space and the vectors in $sp(\beta)$ the cointegrating vectors. The vectors α are the adjustment vectors since they measure the rate of adjustment of the process X_t to the disequilibrium error $\beta' X_{t-1}$. The coefficients τ_0 and τ_1 are rather complicated, but can be found by inserting (2.3) into (2.1) and identifying coefficients to 1 and t, as functions of the parameters and the initial values of the model (2.1).

The purpose of this paper is to discuss the role of the deterministic term $\mu_t = \mu_0 + \mu_1 t$ under the assumption of reduced rank and various restrictions on μ_0 and μ_1. By the role of the deterministic term we mean that the behavior of the deterministic trend of the process X_t depends critically on the relation between μ_t and the adjustment coefficients α. To analyze this we therefore decompose the parameters μ_i in the directions of α and $\alpha_\perp$ as follows:

$$\mu_i = \alpha\beta_i + \alpha_\perp\gamma_i, \qquad i = 0,1.$$

Thus $\beta_i = (\alpha'\alpha)^{-1}\alpha'\mu_i$ and $\gamma_i = (\alpha'_\perp\alpha_\perp)^{-1}\alpha'_\perp\mu_i$. We then define a number of nested submodels of the general model (2.1).

$$H_0(r): \quad \mu_t = \alpha\beta_0 + \alpha_\perp\gamma_0 + (\alpha\beta_1 + \alpha_\perp\gamma_1)t, \tag{2.4}$$

$$H_0^*(r): \quad \mu_t = \alpha\beta_0 + \alpha_\perp\gamma_0 + \alpha\beta_1 t, \tag{2.5}$$

$$H_1(r): \quad \mu_t = \alpha\beta_0 + \alpha_\perp\gamma_0, \tag{2.6}$$

$$H_1^*(r): \quad \mu_t = \alpha\beta_0, \tag{2.7}$$

$$H_2(r): \quad \mu_t = 0. \tag{2.8}$$

TABLE I

The relation between the hypotheses discussed for the I(1) model.

$$\begin{array}{ccccc}
H_0(0) & \subset\ldots\subset & H_0(r) & \subset\ldots\subset & H_0(p) \\
\cup & & \cup & & \cup \\
H_0^*(0) & \subset\ldots\subset & H_0^*(r) & \subset\ldots\subset & H_0^*(p) \\
\| & & \cup & & \cup \\
H_1(0) & \subset\ldots\subset & H_1(r) & \subset\ldots\subset & H_1(p) \\
\cup & & \cup & & \cup \\
H_1^*(0) & \subset\ldots\subset & H_1^*(r) & \subset\ldots\subset & H_1^*(p) \\
\| & & \cup & & \cup \\
H_2(0) & \subset\ldots\subset & H_2(r) & \subset\ldots\subset & H_2(p)
\end{array}$$

The next section discusses the properties of the process X_t under the various models defined by (2.4) – (2.8).

2.2 The interpretation of the I(1) models

All the models above have the property that the process is I(1) and that the stochastic part of $\beta' X_t$ is stationary. We focus here on the interpretation of the models for different restrictions of the constant and linear terms. The general model $H_0(r)$ with unrestricted deterministic term allows for a quadratic trend in the process X_t determined by the slope coefficients

$$\tau_2 = C\mu_1 = \beta_\perp(\alpha'_\perp\Gamma\beta_\perp)^{-1}\alpha'_\perp\mu_1 = \beta_\perp(\alpha'_\perp\Gamma\beta_\perp)^{-1}\alpha'_\perp\alpha_\perp\gamma_1.$$

Note, however, that the linear combinations $\beta' X_t$ have no quadratic trend, since $\beta'\tau_2 = 0$. Thus the quadratic trend is eliminated by the linear combinations β, but the process $\beta' X_t$ still has a linear trend.

The model $H_0^*(r)$ is characterized by the absence of the quadratic trend since $\alpha'_\perp\mu_1 = 0$ or $\gamma_1 = 0$ and hence $\tau_2 = 0$, but the model allows for the possibility of a linear trend in all components of the process, a trend which can not be eliminated by the cointegrating relations β. Thus a linear trend is allowed even in the cointegrating relations, each of which therefore represents a stationary process plus a linear trend or a trend stationary process. In particular if a unit vector is cointegrating then the corresponding component of X_t is trend stationary. The expression for the linear trend τ_1 when $\mu_1 = \alpha\beta_1$ is given by

$$\tau_1 = \bar{\beta}_\perp\Gamma^{-1}_{\alpha_\perp\beta_\perp}\gamma_0 + \bar{\beta}_\perp\Gamma^{-1}_{\alpha_\perp\beta_\perp}\Gamma_{\alpha_\perp\beta}\beta_1 - \bar{\beta}\beta_1.$$

Here we have used the notation $\bar{\beta} = \beta(\beta'\beta)^{-1}$, and $\Gamma_{\alpha_\perp\beta_\perp} = \bar{\alpha}'_\perp\Gamma\bar{\beta}_\perp$. This expression shows how the contributions from μ_1 and μ_0 enter into the slope of the linear trend. Note that the cointegrating relations $\beta' X_t$ have a trend given by $\beta'\tau_1 t = -\beta_1 t$.

For the present purpose we define a trend stationary process as a process that can be decomposed into a stationary process and a linear trend. A problem that often faces the econometrician is to make a choice between describing a given time series as a trend stationary process or an I(1) process. Since the sample paths of such two processes observed over a short interval can easily be mistaken, one will expect that the real decision to choose between the two descriptions should be based on economic insight. In some cases, however, it is of interest to conduct a statistical test to see if one can make the distinction on the basis of the data.

Model $H_0^*(r)$ allows for r trend stationary variables, and p–r variables that are composed of I(1) variables and a linear trend. Thus if one wants to test that a given variable, X_{1t}, say, is trend stationary one has to check that the unit vector (1,0,...,0) is contained in the β space. If the hypothesis is rejected a better description of the variable is as an I(1) variable plus a trend. A different way of describing the same model is by the equations

$$A(L)(X_t - m_0 - m_1 t) = \epsilon_t. \qquad (2.9)$$

If the model is written in the form (2.1) with condition (2.2) we find that

$$\mu_0 = \alpha\beta' m_0 + (\alpha\beta' + \Gamma)m_1,$$
$$\mu_1 = \alpha\beta' m_1.$$

Thus model (2.9) has μ_1 restricted by $\alpha_\perp'\mu_1 = 0$ and is in fact model $H_0^*(r)$ in a different parameterization. The parameterization given by (2.9) is sometimes preferred, but one should note that the parameter m_0 is not identified, since only the r combinations $\beta' m_0$ enter the model, whereas m_1 is identified.

In model $H_1(r)$ where $\mu_1 = 0$ we find $\tau_2 = 0$ and in this case the process X_t still has a linear trend, determined by $\tau_1 = C\mu_0 = \beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp'\alpha_\perp\gamma_0$. This trend is eliminated by the cointegrating relations β, and the process contains no trend stationary components.

In model $H_1^*(r)$ there are no trends, since $\alpha_\perp'\mu_0 = 0$, but a constant term is allowed in the cointegrating relations. Finally in model $H_2(r)$ all stationary components have mean zero.

The conclusion of this discussion is that the role of the deterministic term of (2.1) is seen to depend on the relation between α and the coefficients μ_0 and μ_1. This changing role of the deterministic term for the interpretation of the process defined by the equations has implications for the statistical analysis.

3. THE STATISTICAL ANALYSIS AND THE LIKELIHOOD RATIO TESTS

This section contains a brief description of reduced rank regression and demonstrates how this procedure is applied to derive estimators and test statistics for models with various restrictions on the deterministic terms.

3.1 Reduced rank regression

The statistical analysis of all the models is given by the same procedure which is called a reduced rank regression, and which was introduced by Anderson (1951) in the context of independent variables, and has been applied by Ahn and Reinsel (1988) for stationary processes, and Johansen (1988) for nonstationary processes, that is, the model $H_2(r)$. See Johansen and Juselius (1990) and Johansen (1992a) for a discussion of the models $H_1^*(r)$ and $H_1(r)$.

DEFINITION *By a reduced rank regression of U_t on V_t corrected for Z_t we understand the following statistical calculations: Regress U_t and V_t on Z_t to form residuals R_{ut}, and R_{vt}, and solve the reduced rank regression*

$$R_{ut} = \alpha\beta' R_{vt} + \epsilon_t,$$

by defining product moment matrices

$$S_{ij} = T^{-1}\Sigma_1^T R_{it}R'_{jt},\ i,j = u,v,$$

and solving the eigenvalue problem

$$|\lambda S_{vv} - S_{vu}S_{uu}^{-1}S_{uv}| = 0, \tag{3.1}$$

for eigenvalues $1 > \lambda_1 > \ldots > \lambda_p > 0$, *and eigenvectors* $w = (w_1,\ldots,w_p)$. *That is, the vectors* w_i *satisfy*

$$\lambda_i S_{vv}w_i = S_{vu}S_{uu}^{-1}S_{uv}w_i,\ i = 1,\ldots,p,$$

and are normalized such that $w'S_{vv}w = I$, *so that* $w'S_{vu}S_{uu}^{-1}S_{uv}w = \mathrm{diag}(\lambda_1,\ldots,\lambda_p)$. *The reduced rank estimators are given by* $\hat{\beta} = (w_1,\ldots,w_r)$ *and* $\hat{\alpha} = S_{uv}\hat{\beta}$.

We show in the next section that the models (2.4) – (2.8) can all be estimated by reduced rank regression.

3.2 Derivation of test statistics for cointegration rank

The problem treated in this section is to test for cointegration rank under various assumptions on the deterministic part of the process. The main result is that the statistical calculations are all performed by reduced rank regression, but as shown in the next section, the asymptotic distribution of the test statistics differs for the various models.

We first consider the test of $H_0(r)$ in $H_0(p)$ that is the test for cointegration rank, or $\Pi = \alpha\beta'$, when there is an unrestricted linear term in model (2.1). The Gaussian errors in equation (2.1) give rise to a likelihood analysis which leads to a regression, and for the analysis of $H_0(r)$, where $\Pi = \alpha\beta'$, this is seen to be a reduced rank regression of ΔX_t on X_{t-1} corrected for lagged differences, constant and linear term. The estimator of the error variance is given by $\hat{\Omega} = S_{uu} - \hat{\alpha}\hat{\alpha}'$, and the maximized likelihood is given, apart from a constant, by

$$L_{\max}^{-2/T}(r) = |S_{uu}| \prod_{i=1}^{r} (1-\lambda_i). \tag{3.2}$$

By dividing two expressions like (3.2) for r and r = p, we find that the likelihood ratio test $Q\{H_0(r)|H_0(p)\}$ of the model $H_0(r)$ versus the unrestricted autoregressive model, $H_0(p)$, is given by

$$-2\ln Q\{H_0(r)|H_0(p)\} = -T\Sigma_{r+1}^{p}\log(1-\lambda_i). \qquad (3.3)$$

The same analysis holds for $H_1(r)$ in $H_1(p)$, only the reduced rank regression is of ΔX_t on X_{t-1} corrected for lagged differences and constant term, and in the analysis of $H_2(r)$ we only correct for the lagged differences. Thus in all three cases we get the test statistic (3.3) only with differently calculated eigenvalues.

Next consider $H_0^*(r)$, where μ_1 is restricted by $\alpha_\perp'\mu_1 = 0$, or $\gamma_1 = 0$. In order to show that the analysis of $H_0^*(r)$ is also given by reduced rank regression we use the restriction $\gamma_1 = 0$ to write

$$\alpha\beta'X_t + \alpha\beta_1 t = \alpha(\beta'X_t + \beta_1 t) = \alpha(\beta',\beta_1)(X_t',t)' = \alpha\beta^{*\prime}X_t^*,$$

where $\beta^* = (\beta',\beta_1)'$ and $X_t^* = (X_t',t)'$. With this notation the model is

$$\Delta X_t = \alpha\beta^{*\prime}X_{t-1}^* + \Sigma_1^{k-1}\Gamma_i\,\Delta X_{t-i} + \mu_0 + \epsilon_t,$$

such that the statistical analysis consists of a reduced rank regression of ΔX_t on X_{t-1}^* corrected for lagged differences and the constant. Hence the test statistic for the hypothesis $H_0^*(r)$ in $H_0^*(p)$ is again of the form (3.3). Finally in model $H_1^*(r)$, where $\mu_0 = \alpha\beta_0$, we write the equation as

$$\alpha\beta'X_t + \alpha\beta_0 = \alpha\beta^{**\prime}X_t^{**},$$

where $\beta^{**} = (\beta',\beta_0)'$ and $X_t^{**} = (X_t',1)'$. With this notation it is seen that the model is analyzed by a reduced rank regression of ΔX_t on X_{t-1}^{**} corrected for the lagged differences, and again we get a test statistic of the form (3.3).

3.3 Test for reduction in the degree of the trend

In this section we want to compare the model with quadratic trend

$H_0(r)$ with the model without a quadratic trend, $\overset{*}{H}_0(r)$, by testing $\alpha'_\perp\mu_1$ = 0. The likelihood ratio test statistic is given by

$$Q\{\overset{*}{H}_0(r)|H_0(r)\} = \frac{\overset{*}{L}_{max}(r)}{L_{max}(r)} = \frac{\overset{*}{L}_{max}(r)/\overset{*}{L}_{max}(p)}{L_{max}(r)/L_{max}(p)}\frac{\overset{*}{L}_{max}(p)}{L_{max}(p)}.$$

The last factor is 1, since for any set of values $\Pi = \alpha\beta'$ and μ_1 with $\alpha'_\perp\mu_1 = 0$, we can find a $(\tilde{\Pi},\tilde{\mu}_1)$ close to (Π,μ_1) such that the two conditions are not satisfied. Thus the maximization gives the same whether μ_1 is restricted or not. Hence the test of the hypothesis $\overset{*}{H}_0(r)$ in $H_0(r)$, or the test that $\alpha'_\perp\mu_1 = 0$, gives a test statistic equal to

$$-2\ln Q\{\overset{*}{H}_0(r)|H_0(r)\} = T\Sigma_{r+1}^{p}\ln\{(1-\overset{*}{\lambda}_i)/(1-\lambda_i)\}. \qquad (3.4)$$

Since also $\overset{*}{H}_1(p)$ and $H_1(p)$ give the same maximum, we find similarly that the test of $\overset{*}{H}_1(r)$ in $H_1(r)$, or that $\alpha'_\perp\mu_0 = 0$, is given by (3.4) with eigenvalues calculated differently.

3.4 Test of the absence of the trend in the trend stationary components

Model $H_0(r)$ allows for a linear trend in the cointegrating relations and we want to test that this linear trend vanishes, this is a comparison of model $H_1(r)$ with $\overset{*}{H}_0(r)$ or a test of the hypothesis that $\beta_1 = 0$, see (2.5) and (2.6). We find the likelihood ratio test

$$Q\{H_1(r)|\overset{*}{H}_0(r)\} = \frac{L^1_{max}(r)}{L^{0*}_{max}(r)} = \frac{L^1_{max}(r)/L^1_{max}(0)}{L^{0*}_{max}(r)/L^{0*}_{max}(0)}\frac{L^1_{max}(0)}{L^{0*}_{max}(0)}.$$

The last factor is 1 since if r = 0 then $\alpha = 0$, and $\mu_t = \mu_0$ for both models, see (2.5) and (2.6). Thus we find the test statistic

$$-2\ln Q\{H_1(r)|\overset{*}{H}_0(r)\} = T\Sigma_1^r\ln\{(1-\lambda_i^1)/(1-\lambda_i^{0*})\}.$$

A similar analysis of $H_2(r)$ in $\overset{*}{H}_1(r)$ or $\beta_0 = 0$, will test that the stationary components have no intercept.

3.5 Test of trend stationarity

In all the models above it is possible to test hypotheses on β and α, in analogy with the treatment given in Johansen and Juselius (1992). The idea there is to investigate the class of hypotheses which give rise to models that can be solved by reduced rank regressions. That is, the point of view is not to solve all possible hypotheses, but to find the hypotheses that can be solved by the same procedure. Surprisingly often in applications the interesting hypotheses are of such a form. Thus for instance the hypothesis $\beta = H\varphi$, for some known H is seen immediately to lead to a reduced rank regression of ΔX_t on $H'X_{t-1}$ corrected for lagged differences and a linear trend. We here treat one such hypothesis in detail, namely the hypothesis of trend stationarity.

The hypothesis of trend stationarity of a given set of r_1 variables $b'X_t$ can be tested as a restriction on the cointegrating vectors in the model $H_0^*(r)$ where $\mu_t = \mu_0 + \alpha\beta_1 t$, which allows for a linear trend in all components but no quadratic trend. In this case the hypothesis that $sp(b) \subset sp(\beta)$, or that

$$\beta = (b, \varphi_1) \tag{3.5}$$

where b is $(p \times r_1)$ and φ_1 is $(p \times r_2)$ $(r_1 + r_2 = r)$, can be formulated in terms of $\beta^* = (\beta', \beta_1)'$ as the hypothesis

$$\beta^* = (H\psi, \varphi) \tag{3.6}$$

where ψ is $(r_1+1) \times r_1$, φ is $(p+1) \times r_2$ and H is a $(p+1) \times (r_1+1)$ matrix defined by

$$H = \begin{bmatrix} b & 0 \\ 0 & 1 \end{bmatrix}.$$

To see this let $\alpha = (\alpha_1, \alpha_2)$ and $\psi' = (\psi_1', \psi_2')$, $\varphi' = (\varphi_1', \varphi_2')$ such that (3.6) is equivalent to

$$\begin{aligned} \alpha\beta^{*\prime} X_t^* &= \alpha_1 \psi' H' X_t^* + \alpha_2 \varphi' X_t^* \\ &= \alpha_1 \psi_1' b' X_t + \alpha_1 \psi_2' t + \alpha_2 \varphi_1' X_t + \alpha_2 \varphi_2' t \\ &= \tilde{\alpha}_1 b' X_t + \alpha_2 \varphi_1' X_t + \alpha \gamma_1 t. \end{aligned}$$

Here the square matrix ψ_1 has been absorbed into $\tilde{\alpha}_1 = \alpha_1\psi_1'$, and $\alpha\gamma_1 = \alpha_1\psi_2' + \alpha_2\varphi_2'$. Thus the restriction (3.6) is seen to be a restriction on some β–vectors, that is of the form (3.5). As shown in Johansen and Juselius (1992) the parameters under this hypothesis can be estimated by a simple switching algorithm, since the equation (3.5) is a regression involving two reduced rank matrices $\alpha_1\psi'$ and $\alpha_2\varphi'$. The algorithm consists of fixing either ψ of φ and solve for the other, and then iterate.

4. ASYMPTOTIC ANALYSIS OF I(1) MODELS

This section contains the asymptotic analysis of the various test statistics and estimators considered in section 3. The proofs are sketchy, since they mimic already existing proofs, and only a few intermediate results are given which show how the various statistics should be normalized when trends are present, see Johansen (1988,1992a), and Ahn and Reinsel (1990) and Reinsel and Ahn (1990). The main conclusion is that the asymptotic distribution of the test statistics depends on which assumptions we make on the deterministic components of the process, and a few tables are provided. Most of the tests and tables are described in Johansen and Juselius (1990) so we focus on the model H_0, but give the other results for completeness.

4.1 The asymptotic results for estimators and test statistics

The results involve a standard Brownian motion B(t) in p–r dimensions, and we introduce the notation $\bar{B}_i = \int_0^1 B_i(t)dt$, and $A_i = \int_0^1 (t-\frac{1}{2})B_i(t)dt / \int_0^1 (t-\frac{1}{2})^2 dt$, such that the Brownian motion corrected for trend is

$$B_i(t) - \bar{B}_i - A_i(t-\tfrac{1}{2}).$$

The function t^2 corrected for trend is given by

$$t^2 - t - {}^1/_6.$$

THEOREM 1 *On the assumption that the cointegrating rank is r, the asymptotic distribution of the likelihood ratio test statistic for cointegration rank is given by*

$$tr\{\int_0^1 (dB)F'[\int_0^1 FF'du]^{-1}\int_0^1 F(dB)'\}. \tag{4.1}$$

Here B is a standard p-r dimensional Brownian motion on the unit interval and F depends on B and on which of the hypotheses is being tested.

If $\mu_t = \mu_0 + \mu_1 t$, $\alpha_\perp'\mu_1 \neq 0$ and if $\Pi = \alpha\beta'$ is tested in the general VAR model, that is, $H_0(r)$ in $H_0(p)$ then

$$F_i(t) = B_i(t) - \bar{B}_i - A_i(t-\tfrac{1}{2}),\ i = 1,\dots,p-r-1, \tag{4.2}$$
$$F_{p-r}(t) = t^2 - t - {}^1/_6. \tag{4.3}$$

The distribution of the test statistic is tabulated by simulation in TABLE IV.

If $\mu_t = \mu_0 + \alpha\beta_1 t$, and if $\Pi = \alpha\beta'$ is tested in $H_0^*(p)$ then

$$F_i(t) = B_i(t) - \bar{B}_i,\ i = 1,\dots,p-r, \tag{4.4}$$
$$F_{p-r+1}(t) = t - \tfrac{1}{2}. \tag{4.5}$$

The distribution of the test statistic is tabulated by simulation in TABLE V.

If $\mu_t = \mu_0$, $\alpha_\perp'\mu_0 \neq 0$, and if $\Pi = \alpha\beta'$ is tested in $H_1(p)$ then

$$F_i(t) = B_i(t) - \bar{B}_i,\ i = 1,\dots,p-r-1, \tag{4.6}$$
$$F_{p-r}(t) = t - \tfrac{1}{2}. \tag{4.7}$$

If $\mu_t = \alpha\beta_0$, and if $\Pi = \alpha\beta'$ is tested in $H_1^*(p)$ then

$$F_i(t) = B_i(t),\ i = 1,\dots,p-r, \tag{4.8}$$
$$F_{p-r+1}(t) = 1. \tag{4.9}$$

Finally if $\mu_t = 0$, and $\Pi = \alpha\beta'$ is tested , that is, if $H_2(r)$ is tested in $H_2(p)$ then

$$F_i(t) = B_i(t),\ i = 1,\dots,p-r. \tag{4.10}$$

It is seen that the asymptotic distribution reflects the statistical calculations, and the model that is being tested. Thus in the unrestricted model X_t has a quadratic trend, but is corrected for a constant and a linear trend in the calculations, such that F in this case is given by a Brownian motion with one direction replaced by a quadratic trend, and the whole process corrected for a constant and a linear trend.

THEOREM 2 *The asymptotic distribution of $\hat{\beta}$ from $H_i(r)$ $i = 0,1,2$ and $\hat{\beta}^*$ from $H_i^*(r)$, $i = 0,1$ is mixed Gaussian. If $K'\beta = 0$ for K $(p\times m)$ it holds that*

$$(K'S_{11}^{-1}K)^{-\frac{1}{2}}T^{\frac{1}{2}}(K'\hat{\beta})(\hat{\Lambda}^{-1}-I)^{\frac{1}{2}} \xrightarrow{w} N_{m\times r}(0,I).$$

Here $\hat{\Lambda} = diag(\lambda_1,\ldots,\lambda_r)$, and S_{11} can be replaced by $\hat{v}\hat{v}'$, where the eigenvectors of (3.1) are decomposed as $w = (\hat{\beta},\hat{v})$. A similar result holds for $\hat{\beta}^*$.

This result can be directly applied to test hypotheses of the form $K'\beta = 0$, using the Gaussian distribution. One can show that $\hat{\alpha}'\hat{\Omega}^{-1}\hat{\alpha} = \hat{\Lambda}^{-1} - I$. Note that the factor $T^{\frac{1}{2}}$ comes from the normalization of S_{11}, so that a different way of stating the result is as

$$(K'\sum_{t=1}^{T} X_tX_t'K)^{-\frac{1}{2}}(K'\hat{\beta})(\hat{\Lambda}^{-1}-I)^{\frac{1}{2}} \xrightarrow{w} N_{m\times r}(0,I),$$

which shows that by normalizing the estimated contrasts $K'\hat{\beta}$ using the square root of the observed information, which can be calculated as

$$(\sum_{t=1}^{T} X_tX_t)^{-1} \otimes \alpha'\Omega^{-1}\alpha,$$

one can act as if the asymptotic distribution of $\hat{\beta}$ were Gaussian. We end this secion with some results for the test statistics:

THEOREM 3 *The asymptotic distribution of the likelihood ratio tests for the hypotheses $\gamma_i = 0$, that is, $H_i^*(r)$ versus $H_i(r)$ is χ^2 with p-r degrees of freedom, $i = 0,1,2$. The asymptotic distribution of the likelihood ratio tests for the hypotheses $\beta_i = 0$, that is, $H_i(r)$ versus $H_{i-1}^*(r)$ is χ^2 with r degrees of freedom, $i = 1,2$.*

THEOREM 4 *The asymptotic distribution of the test for trend stationarity of the vectors* b $(p{\times}r_1)$ *in model* $H_0^*(r)$ *is* χ^2 *with* $(p-r)r_1$ *degrees of freedom.*

4.2. Proofs of the asymptotic results

We start with the proof of Theorem 1. The proof will be outlined for the model $H_0(r)$ and $H_0^*(r)$, since the models $H_1(r)$, $H_1(r)$ and $H_2(r)$ are treated in detail in Johansen (1989, 1992a). The proof follows closely the proof given there. Let $U_t = \Delta X_t$, $V_t = X_{t-1}$, and $Z_t' = (\Delta X_{t-1}',\ldots,\Delta X_{t-k+1}',1,t)$. Consider first the test statistic (3.3) as given by the eigenvalues λ_i, derived from the equation

$$|\lambda S_{vv} - S_{vu}S_{uu}^{-1}S_{uv}| = 0. \tag{4.11}$$

The process R_{vt}, which is X_{t-1} corrected for lagged differences, constant and linear term, and which enters the matrices S_{ij}, has a quadratic trend in the direction $\tau = \tau_2 = C\mu_2$, see (2.3), but orthogonal to this there is no quadratic trend left. There is also no linear trend in the residuals, since they are corrected for a linear trend. We define the directions $\gamma(p{\times}(p-r-1))$ such that (β,γ,τ) span all of R^p. Then $\tau' R_{vt}$ has a quadratic trend, whereas $\gamma' R_{vt}$ has no deterministic trend but a random trend, and $\beta' R_{vt}$ is stationary. This is spelled out in Lemma 1, where the results areexpressed in terms of the Brownian motion W_t defined by the ϵ_t's, i.e.

$$T^{-\frac{1}{2}}\Sigma_1^{[Tt]}\epsilon_i \overset{W}{\to} W_t.$$

LEMMA 1 *The residuals* R_{vt} *satisfy*

$$T^{-2}\bar{\tau}' R_{v[Tt]} \overset{w}{\to} \tfrac{1}{2}(t^2 - t - {}^1/_6) = G_2(t),$$

say, and

$$T^{-\frac{1}{2}}\bar{\gamma}' R_{v[Tt]} \overset{w}{\to} \bar{\gamma}' C(W - \bar{W} - A_w(t-\tfrac{1}{2})) = G_1(t),$$

say, where $\bar{W}$ *and* A_w *are determined by regressing* W *on 1 and t.*

Finally $\beta' R_{vt}$ *behaves asymptotically as a stationary process. Hence for* $G' = (G_1', G_2)$, *and* $C_T = (\bar{\gamma}, T^{-3/2}\bar{\tau})$ *it holds that*

$$T^{-1} C_T' S_{vv} C_T \overset{w}{\to} \int_0^1 GG' du \text{ and } C_T'(S_{vu} - S_{vv}\alpha\beta') \overset{w}{\to} \int_0^1 G(dW)'.$$

By expanding the likelihood function and using the result that we can make inference about β *for fixed values of the other parameters, we find that*

$$T\gamma'(\hat{\beta}-\beta) \overset{w}{\to} [\int_0^1 G_{1.2}G_{1.2}' du]^{-1} \int_0^1 G_{1.2}(dV_\alpha)'$$

$$T^{5/2}\tau'(\hat{\beta}-\beta) \overset{w}{\to} [\int_0^1 G_{2.1}G_{2.1}' du]^{-1} \int_0^1 G_{2.1}(dV_\alpha)',$$

where $V_\alpha = (\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}W$ *is independent of G. We use the notation that* $G_{1.2}$ *is* G_1 *corrected for* G_2, *and* $G_{2.1}$ *is* G_2 *corrected for* G_1.

Using these results it is not difficult to show that the p–r smallest roots, normalized by T, of the equation (4.11) will converge to the roots of the equation

$$|\rho\int_0^1 GG'du - \int_0^1 G(dW)'\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\alpha_\perp'\int_0^1(dW)G'| = 0$$

Now define $B = (\alpha_\perp'\Omega\alpha_\perp)^{-\frac{1}{2}}\alpha_\perp' W$, which has variance matrix I, then it follows that the limit distribution of the test statistic (3.3), which is asymptotically equivalent to the sum of the p–r smallest eigenvalues normalized by T, converges as stated in the theorem with F defined by (4.2) and (4.3). Note that if in fact the parameter γ_1 is zero, then of course the process has no quadratic trend, and the above argument has to be modified, by leaving out the last component, such that the limit distribution of the test statistic is given by

$$\text{tr}\{\int_0^1(dB)(B-a-bt)'[\int_0^1(B-a-bt)(B-a-bt)']^{-1}\int_0^1(B-a-bt)d(B)'\} \quad (4.12)$$

where a and b are determined by correcting B for a linear trend and constant term. If instead we consider the hypothesis $H_0^*(r)$ and want to test it in $H_0^*(p)$, we define the residuals R_{1t}^* as the residuals from a regression of $X_t^* = (X_t', t)'$ on the lagged differences and the constant. In this case the residuals contain a linear trend, but they are still corrected for their mean. Since the last component of X_t^* is a linear trend it is convenient to transform the residuals before describing the limit distributions. We define the norming matrix $C_T^* = \begin{bmatrix} \beta_\perp & 0 \\ -\tau_1'\beta_\perp & T^{-1} \end{bmatrix}$.

LEMMA 2 *The residuals R_{vt}^* satisfy*

$$T^{-\frac{1}{2}}(\beta_\perp', -\beta_\perp'\tau_1)R_{vt}^* \overset{w}{\to} \beta_\perp' C(W_t - \bar{W}) = G_1^*(t),$$

such that

$$T^{-1} C_T^{*\prime} S_{vv} C_T^* \overset{w}{\to} \int_0^1 G^* G^{*\prime} du \text{ and } C_T^{*\prime}(S_{vu}^* - S_{vv}^* \alpha\beta^{*\prime}) \overset{w}{\to} \int_0^1 G^*(dW)'$$

where the last component of G^* is $t - \frac{1}{2}$. Furthermore

$$C_T^{*\prime}(\hat{\beta}^* - \beta^*) \overset{w}{\to} [\int_0^1 G^* G^{*\prime} du]^{-1} \int_0^1 G^*(dV_a)'.$$

With these fragments of a proof one can reproduce the proof in Johansen (1992a), and prove that the limit distribution is given by (4.1) with F given by (4.4) and (4.5). This completes the proof of Theorem 1.

The results of Theorem 2 follow from Lemma 1 and 2. Theorem 3 is proved as follows: In order to find the distribution of the test for the hypothesis $H_0^*(r)$ versus $H_0(r)$, one can compare the hypotheses with $H_0^*(p)$ and $H_0(p)$ respectively, as is done in the derivation of the test (3.4). The test statistic has to be evaluated under the assumption that $\alpha_\perp'\mu_1 = 0$, so that no quadratic trend is present.

From the relation

$$\begin{bmatrix} B_t - a \\ t - \frac{1}{2} \end{bmatrix}' [\int_0^1 \begin{bmatrix} B_t - a \\ t - \frac{1}{2} \end{bmatrix} \begin{bmatrix} B_t - a \\ t - \frac{1}{2} \end{bmatrix}' dt]^{-1} \begin{bmatrix} B_s - a \\ s - \frac{1}{2} \end{bmatrix} = 12(t-\tfrac{1}{2})(s-\tfrac{1}{2}) \quad (4.13)$$
$$+ (B_t - a - bt)' [\int_0^1 (B_t - a - bt)(B_t - a - bt)' dt]^{-1} (B_s - a - bs)$$

the test of $H_0^*(r)$ within $H_0(r)$ is equal to

$$-2\ln Q\{H_0^*(r) | H_0^*(p)\} - (-2\ln Q\{H_0(r) | H_0(p)\},$$

which by (4.12) and (4.13) converges towards

$$\mathrm{tr}\{12\int_0^1(dB)(t-\tfrac{1}{2})\int_0^1(s-\tfrac{1}{2})(dB)'\} = \chi^2(p-r).$$

The remaining part of Theorem 3 as well as Theorem 4 concern hypotheses on β, and the results follow since the limit distribution of the estimator of β is mixed Gaussian, such that usual χ^2 inference can be conducted.

5. THE I(2) MODELS, THEIR INTERPRETATION AND STATISTICAL ANALYSIS

This section contains a very brief description of some autoregressive models for I(2) variables. The basic theory is presented in Johansen (1991c,d). The theory is quite involved, but has been illustrated in an analysis of the purchasing power parity between Australia and USA, Johansen (1992b). The basic reason for introducing these models is that in analyzing price series, it turned out that the inflation rate was best described by nonstationary processes such that the prices series were nonstationary but their differences were also nonstationary. In analyzing such processes we also need to take account of the deterministic trend, and we therefore in the next section use a representation theorem for I(2) processes to discuss the influence on the process of the deterministic terms in the model. From the application point of view, we certainly want to have the possibility to model a linear trend in the variables, but a quadratic trend in the variables means a linear trend in the differences, which at least for price series seems unreasonable. This application will determine the models we want to consider.

5.1. The definition and basic properties of I(2) models

In order to describe the I(2) models we consider model (2.1) written as a reduced form error correction model

$$\Delta^2 X_t = \Gamma\Delta X_{t-1} + \Pi X_{t-2} + \Sigma_1^{k-2}\Phi_i\Delta^2 X_{t-i} + \mu_0 + \epsilon_t,\ t = 1,2,\ldots,T. \quad (5.1)$$

We define $\Phi = I - \Sigma_1^{k-2}\Phi_i$. For the model (5.1) to allow I(2) variables we need two reduced rank conditions, and it can be shown, see Johansen

(1992c), that if

$$\Pi = \alpha\beta' \text{ has rank r}, \tag{5.2}$$

and

$$\alpha_\perp'\Gamma\beta_\perp = \varphi\eta' \text{ has rank s}, \tag{5.3}$$

and if finally $\varphi_\perp'\alpha_\perp'(\Phi + \Gamma\bar{\beta}\bar{\alpha}\Gamma)\beta_\perp\eta_\perp$ has full rank (p–r–s), then the process has the representation

$$X_t = C_2\sum_{j=1}^{t}\sum_{i=1}^{j}(\epsilon_i + \mu) + C_1\sum_{i=1}^{t}(\epsilon_i + \mu) + C_2(L)(\epsilon_t + \mu) + A + Bt,$$

for suitable matrices C_1 and C_2. The coefficients A and B are determined by the initial conditions. The reduced rank conditions (5.2) and (5.3) allow us to define vectors $(\alpha,\alpha_1,\alpha_2)$ and (β,β_1,β_2) as follows

$$\beta_1 = \bar{\beta}_\perp\eta,\ \beta_2 = \beta_\perp\eta_\perp,\ \alpha_1 = \bar{\alpha}_\perp\varphi \text{ and } \alpha_2 = \alpha_\perp\varphi_\perp.$$

Then $(\alpha,\alpha_1,\alpha_2)$ are mutually orthogonal and span R^p, and the same is true for (β,β_1,β_2). The matrix C_2 is expressed as

$$C_2 = \beta_2(\alpha_2'(\Phi + \Gamma\bar{\beta}\bar{\alpha}\Gamma)\beta_2)^{-1}\alpha_2',$$

such that the quadratic trend is given by $\tau_2 = C_2\sum_{j=1}^{t}\sum_{i=1}^{j}\mu = \frac{1}{2}C_2t(t+1)$.

Ignoring the deterministic terms in the model for a while, the main conclusion of this analysis is that under suitable restrictions, see (5.2) and (5.3), on the parameters of the autoregressive model we can generate an I(2) process by equation (5.1). The process has different properties in the directions given by (β,β_1,β_2). This can be expressed by saying that $\beta_2'X_t$ is an I(2) process that does not cointegrate. The process $(\beta,\beta_1)'X_t$ is an I(1) process, thus the order of the process is reduced from 2 to 1 by the linear combination in (β,β_1). The linear combinations $\beta_1'X_t$ do not cointegrate, but $\beta'X_t$ cointegrate with the differences in such a way that $\beta'X_t + \bar{\alpha}'\Gamma\bar{\beta}_2\beta_2'\Delta X_t$ is stationary. Thus we do not in general get linear combinations that are stationary, but if we can find a vector ξ such that $\xi'\bar{\alpha}'\Gamma\bar{\beta}_2 = 0$, then of course $\xi'\beta'X_t$ is stationary.

5.2. *The statistical analysis of I(2) models*

Before we discuss the restrictions on the constant term and the role it plays for the process X_t, we give a brief account of the statistical analysis as described in Johansen (1992d). The analysis is based on the likelihood function, but instead of maximizing the likelihood function directly, which is difficult due to the two reduced rank problems we suggest a two stage procedure, whereby the analysis can be performed by a repeated application of the analysis of the I(1) model, that is, by reduced rank regression. To see this, consider again the equation (5.1) but now multiplied by $\alpha_\perp'$:

$$\alpha_\perp'\Delta^2X_t = \alpha_\perp'\Gamma\Delta X_{t-1} + \Sigma_1^{k-2}\alpha_\perp'\Phi_i\Delta^2X_{t-i} + \alpha_\perp'\mu_0 + \alpha_\perp'\epsilon_t, \qquad (5.4)$$

These equations constitute p—r equations involving the differences of the process. Now decompose the first term on the right hand side as

$$\alpha_\perp'\Gamma\Delta X_{t-1} = \alpha_\perp'\Gamma(\beta\bar{\beta}' + \beta_\perp\bar{\beta}_\perp')\Delta X_{t-1} = (\alpha_\perp'\Gamma\beta)\bar{\beta}'\Delta X_{t-1} + \varphi\eta'\bar{\beta}'\Delta X_{t-1}$$

such that

$$\alpha_\perp'\Delta^2X_t = (\alpha_\perp'\Gamma\beta)\bar{\beta}'\Delta X_{t-1} + \varphi\eta'\bar{\beta}_\perp'\Delta X_{t-1} + \Sigma_1^{k-2}\alpha_\perp'\Phi_i\Delta^2X_{t-i} + \alpha_\perp'\mu_0 + \alpha_\perp'\epsilon_t, \qquad (5.5)$$

From (5.5) it is seen, that if only r, α and β, and hence $\alpha_\perp$ and $\beta_\perp$, were known, it would be easy to analyze equation (5.5) by reduced rank regression of $\alpha_\perp'\Delta^2X_t$ on $\bar{\beta}_\perp'\Delta X_{t-1}$ corrected for $\bar{\beta}'\Delta X_{t-1}$ as well as lagged second differences and the constant. It turns out that one can determine estimates of r, α and β from an initial I(1) analysis and then apply the estimated values in the analysis at the second step. Thus in summary the proposed I(2) analysis is the following:

A. Perform a reduced rank regression of ΔX_t on ΔX_{t-1} corrected for $\Delta X_{t-1},\dots,\Delta X_{t-k+1}$, and 1. This determines $\hat{r}$, $\hat{\alpha}$, and $\hat{\beta}$.

B. Perform a reduced rank regression of $\hat{\alpha}_\perp'\Delta^2X_t$ on $\hat{\beta}_\perp'\Delta X_{t-1}$ corrected for $\hat{\beta}'\Delta X_{t-1}$, $\Delta^2X_{t-1},\dots,\Delta^2X_{t-k+2}$, and 1. This determines $\hat{s}$, $\hat{\varphi}$ and $\hat{\eta}$.

The analysis has the obvious advantage that the statistical calculations for the analysis of the I(2) model are the same as those for the I(1) analysis.

The second step of the analysis is very natural, since if the process is I(2) it seems reasonable to take differences and then analyze the differences by an I(1) model. This analysis would appear as a result of a likelihood analysis if we impose $\Pi = 0$ in (5.1), so that we can see that the preliminary I(1) analysis allows us to exploit the information in the levels of the process, as expressed in the reduced rank of Π. Having found the r relations $\beta' X_t$ we then focus on the p–r relation $\alpha_\perp' X_t$, which are the combinations of the variables that evolve without taking into account the disequilibrium error $\beta' X_t$, that is, the equation (5.4) and (5.5) only involves differences. The relations $\beta' X_t$ have the property that $\Delta\beta' X_t$ is stationary, and hence the analysis in the second step is a reduced rank regression where one corrects for the stationary relations $\Delta\bar{\beta}' X_t$ as well as the other stationary terms consisting of the second differences.

5.3 *Statistical properties of the I(2) analysis*

We now discuss briefly the properties of the proposed method for the analysis of the I(2) model. Let Q_r denote the test statistic $-2\ln Q\{H_1(r)|H_1(p)\}$ calculated from (3.3) for testing $H_1(r)$ in $H_1(p)$ from the I(1) analysis given by (5.1). Let $Q_{r,s}$ denote the corresponding test statistic for testing $H_{r,s}$ defined by (5.2) and (5.3) in H_r determined by an I(1) analysis of equation (5.5) for fixed r, α and β. It follows from the limit theory of section 4 that the asymptotic distribution of $Q_{r,s}$ is given by (4.1) with p–r–s degrees of freedom, and with F given by (4.6) and (4.7) since the constant is included in the analysis and the differences have a linear trend. The correction for the stationary term $\beta' \Delta X_{t-1}$ does not change the limit theory which is dominated by the nonstationary contributions.

According to Theorem 7 in Johansen (1992d) it holds that the limit distribution of $Q_{r,s}$ is the same if we replace r, α and β by their estimates from the initial I(1) analysis. It is of course the superconsistency of the estimates that allows this phenomenon to take place.

The test statistics are arranged in a convenient way in Table 2, where the formal analysis is performed as follows: First $\hat{r}$ is determined by comparing Q_0 with its quantile c_0, then Q_1 with its quantile c_1 etc. stopping the first time the test statistic is less than its quantile. Thus for instance $\hat{r} = 2$ on the set

$$\{\hat{r} = 2\} = \{Q_0 > c_0,\ Q_1 > c_1,\ Q_2 < c_2\}.$$

The quantile c_r is determined from the distribution (4.1) with F given by (4.6) and (4.7) and p − r degrees of freedom. Having determined $\hat{r} = 2$ the value of s is determined by reading the row with r = 2 from left to right and comparing the test statistics $Q_{2,s}$, s = 0,1 with its quantile given in the second last row. Thus $\hat{s} = 1$ on the set

$$\{Q_{2,0} > c_2, Q_{2,1} < c_1\}.$$

TABLE II

Test statistics for the I(2) analysis, p = 4 with the constant unrestricted

r						quantile	p—r
0	$Q_{0,0}$	$Q_{0,1}$	$Q_{0,2}$	$Q_{0,3}$	Q_0	c_0	4
1		$Q_{1,0}$	$Q_{1,1}$	$Q_{1,2}$	Q_1	c_1	3
2			$Q_{2,0}$	$Q_{2,1}$	Q_2	c_2	2
3				$Q_{3,0}$	Q_3	c_3	1
quantile	c_0	c_1	c_2	c_3			
p—r—s	4	3	2	1			

Note that the quantiles are the same as in the determination of r, since the second step of the analysis also allows a constant in the model, and hence a linear trend in the process $\alpha_\perp'\Delta X_t$. If we find that also the last statistic $Q_{2,1}$ is rejected, then $\hat{s} = p-r = 4-2 = 2$, and there are no I(2) components in the system.

5.4 The role of the constant term in I(2) models

Now let us return to the discussion of the constant term in model (5.1). It is seen that a constant term in the model implies a quadratic trend in the process. The slope is given by $\tau_2 = \frac{1}{2}C_2\mu$, which vanishes if $\alpha_2'\mu = 0$. Let us therefore decompose μ in the direction $(\alpha, \alpha_1, \alpha_2)$:

$$\mu = \alpha\beta_0 + \alpha_1\gamma_0 + \alpha_2\delta_0$$

The quadratic trend is determined by δ_0, such that $\tau_2 = 0$ if $\delta_0 = 0$. If this condition holds, there is a linear trend in the process determined by γ_0. If $\gamma_0 = 0$ then there is only a constant level left in the process. The model that we shall discuss here is defined by $\delta_0 = 0$, such that no quadratic trend is in the process. The reason for this is that I(2)–ness is needed for processes whose differences are nonstationary, but we have yet to find an example where the differences also have a linear trend. If such an example appears it is of course possible to extend the analysis given below.

The two step procedure outlined in section 5.3 can be briefly summarized by saying that first we analyze the data without the restriction on the matrix Γ and calculate estimates of r, α and β, and then we transform the equations by $\hat{\alpha}_\perp$ and impose all the restrictions on Γ but now for known values of r, α and β. This last analysis is then the analysis of model $H_1^*(s)$, see (3.4). This formulation immediately suggests how the analysis should be made in the model where Γ as well as μ are restricted.

A. Perform a reduced rank regression of ΔX_t on ΔX_{t-1} corrected for $\Delta X_{t-1}, \ldots, \Delta X_{t-k+1}$, and 1. This determines $\hat{r}$, $\hat{\alpha}$, and $\hat{\beta}$.

B. Perform a reduced rank regression of $\hat{\alpha}_{\perp}'\Delta^2 X_t$ *on* $\hat{\beta}_{\perp}'\Delta X_{t-1}$ *with 1 appended, corrected for* $\hat{\beta}'\Delta X_{t-1}$ *and* $\Delta^2 X_{t-1},\ldots,\Delta^2 X_{t-k+2}$. *This determines* $\hat{s}$, $\hat{\varphi}$ *and* $\hat{\eta}$.

The consequences of this is that in TABLE II the quantiles used in determining the rank r are the ones we get from (4.1) with F defined by (4.6) and (4.7), that is, c_r. The quantiles we need to determine s are given by the distribution (4.1) but now with F defined by (4.8) and (4.9). They will be called c_r^*. Thus for an I(2) analysis which allows a linear trend, but no quadratic trend, the rank determination is performed as in TABLE III.

TABLE III

Test statistics for the I(2) analysis, p = 4
with the constant restricted in the second step

r						quantile	p−r
0	$Q_{0,0}$	$Q_{0,1}$	$Q_{0,2}$	$Q_{0,3}$	Q_0	c_0	4
1		$Q_{1,0}$	$Q_{1,1}$	$Q_{1,2}$	Q_1	c_1	3
2			$Q_{2,0}$	$Q_{2,1}$	Q_2	c_2	2
3				$Q_{3,0}$	Q_3	c_3	1
quantile	c_0^*	c_1^*	c_2^*	c_3^*			
p−r−s	4	3	2	1			

ACKNOWLEDGEMENTS

This paper was written while the author was visiting Department of Statistics, University of Helsinki, and completed during a visit to Department of Statistics, The Australian National University, Canberra. The visit to Helsinki was supported by the Danish Social Sciences Council acc. nr. 14–5793.

REFERENCES

Ahn, S. K., & Reinsel, G. C., (1988). Nested reduced–rank autoregressive models for multiple time series. American Statistical Association, 83, 849–856.

Ahn, S. K., & Reinsel, G. C.,(1990). Estimation for Partially Non–stationary Multivariate Autoregressive Models. Journal of the American Statistical Association, 85, 813–823.

Anderson, T. W., (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. Annals of Mathematical Statistics, 22, 327–351.

Engle, R. F. & Granger C. W. J., (1987). Co–integration and Error Correction:Representation, Estimation and Testing. Econometrica, 55, 251–276.

Johansen, S., (1988). Statistical Analysis of Cointegration Vectors. Journal of Economic Dynamics and Control, 12, 231–254.

Johansen, S., (1989). Likelihood based inference on cointegration. Lecture notes, University of Copenhagen.

Johansen, S., (1992a). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. To appear in Econometrica.

Johansen, S., (1992b). An I(2) cointegration analysis of the purchasing power parity between Australia and USA. Preprint, University of Copenhagen.

Johansen, S., (1992c). A representation of vector autoregressive processes integrated of order 2. To appear in Econometric Theory.

Johansen, S., (1992d). A statistical analysis of cointegration for I(2) variables. To appear in Econometric Theory.

Johansen, S. & Juselius, K., (1990). Maximum Likelihood estimation and Inference on Cointegration – with Applications to the demand for money. Oxford Bulletin of Economics and Statistics, 52, 169–210.

Johansen, S. & Juselius, K., (1992). Structural tests in a multivariate cointegration analysis of the PPP and UIP for UK. To appear in Journal of Econometrics.

TABLE IV

The quantiles of the asymptotic distribution of the test statistic (4.1) with F equal to B , but last component replaced by t^2, and corrected for trend, see (4.2) and (4.3). Number of terms in the series is T = 400, and the number of simulations is 6000.

Lambda–max

dim	50%	80%	90%	95%	97.5%	99%	mean	Var
1	.44	1.67	2.70	3.84	5.25	6.98	1.01	2.12
2	8.73	12.37	14.64	16.69	18.84	20.88	9.38	15.37
3	14.82	18.92	21.44	23.75	25.68	28.31	15.33	21.05
4	20.35	24.87	27.39	29.93	32.22	35.57	20.86	25.89
5	25.73	30.47	33.45	36.46	39.00	41.87	26.23	30.68

Trace

1	.44	1.67	2.70	3.84	5.25	6.98	1.01	2.12
2	9.61	13.43	15.74	18.08	20.26	22.40	10.21	17.38
3	22.73	28.31	31.67	34.27	36.98	40.10	23.38	37.37
4	39.65	46.66	50.62	54.02	57.01	61.03	40.27	61.03
5	60.33	68.80	73.73	77.61	81.29	85.56	60.95	91.55

TABLE V

The quantiles of the asymptotic distribution of the test statistic (4.1) with F equal to B extended by t, and corrected for the mean, see (4.4) and (4.5). Number of terms in the series is T = 400, and the number of simulations is 6000.

Lambda max

1	5.73	8.64	10.59	12.49	14.06	16.42	6.33	10.52
2	11.01	14.65	16.93	19.16	20.87	23.66	11.58	16.56
3	16.35	20.52	23.11	25.44	27.67	30.38	16.79	22.04
4	21.51	26.13	29.04	31.53	34.24	37.15	22.14	27.10
5	26.73	31.71	34.82	37.75	40.05	42.78	27.35	31.70

Trace

1	5.73	8.64	10.59	12.49	14.06	16.42	6.33	10.52
2	15.75	20.33	22.95	25.43	27.82	30.55	16.33	25.47
3	29.57	35.69	39.01	42.35	45.23	48.99	30.24	45.91
4	47.33	54.51	58.98	62.71	66.36	70.63	47.85	71.22
5	68.62	77.40	82.29	86.71	90.70	95.19	69.26	100.37

PREPRINTS 1991

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, DK-2100 COPENHAGEN Ø, DENMARK, TELEPHONE + 45 31 35 31 33.

No. 1 Johansen, Søren: Determination of Cointegration Rank in the Presence of a Linear Trend.

No. 2 Johansen, Søren: A Statistical Analysis of Cointegration for I(2) Variables.

No. 3 Jacobsen, Martin: Homogeneous Gaussian Diffusions in Finite Dimensions.

No. 4 Johansen, Søren: Testing Weak Exogeneity and the Order of Cointegration in UK Money Demand Data.

No. 5 Johansen, Søren: An I(2) Cointegration Analysis of the Purchasing Power Parity between Australia and USA.

No. 6 Johansen, Søren: Estimating Systems of Trending Variables.

No. 7 Jacobsen, Martin and Keiding, Niels: Random Censoring and Coarsening at Random.

PREPRINTS 1992

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, DK-2100 COPENHAGEN Ø, DENMARK. TELEPHONE + 45 31 35 31 33.

No. 1 Johansen, Søren: The Role of the Constant Term in Cointegration Analysis of Nonstationary Variables.