DISCUSSION PAPERS
Department of Economics
University of Copenhagen

02-20

Analyzing I(2) Systems by Transformed Vector Autoregressions

Hans Christian Kongsted
Heino Bohn Nielsen

Studiestræde 6, DK-1455 Copenhagen K., Denmark
Tel. +45 35 32 30 82 - Fax +45 35 32 30 00
http://www.econ.ku.dk

Analyzing I(2) Systems by Transformed Vector Autoregressions

Hans Christian Kongsted*
Institute of Economics, University of Copenhagen

Heino Bohn Nielsen
Institute of Economics, University of Copenhagen

November 29, 2002

Abstract

We characterize the restrictions imposed by the minimal I(2)-to-I(1) transformation that underlies much applied work, e.g. on money demand relationships or open-economy pricing relationships. The relationship between the parameters of the original I(2) vector autoregression, including the coefficients of polynomially cointegrating relationships, and the transformed I(1) model is characterized. We discuss estimation of the transformed model subject to restrictions as well as the more commonly used approach of unrestricted reduced rank regression. Only a minor loss of efficiency is incurred by ignoring the restrictions in the empirical example and a simulation study. A properly transformed vector autoregression thus provides a practical and effective means for inference on the parameters of the I(2) model.

Key words: Cointegration, stochastic trend, price homogeneity, nominal, real, Monte Carlo experiment.

JEL classification: C32, C51, C52, F41

*Studiestræde 6, DK-1455 Copenhagen K, Denmark. Phone +45 3532 3076, fax +45 3532 3064. E-mail Hans.Christian.Kongsted@econ.ku.dk. We thank seminar participants at the ESEM 2002 meeting in Venice and the European Central Bank for comments. Kongsted gratefully acknowledges financial support from the Danish Social Sciences Research Council under the project "Macroeconomic transmission mechanisms in Europe: Empirical applications and econometric methods".

1 Introduction

This paper is motivated by a rich empirical literature applying cointegration analysis in examining the levels and the growth rates of macroeconomic variables and their relationships. Main examples are relationships that involve the growth rates of nominal variables, e.g., the rate of inflation, wage growth, or the money growth rate, and real or relative magnitudes of such variables, e.g. real wages, real money, or the markup, see Coenen and Vega (2001), Crowder, Hoffman, and Rasche (1999), or Doornik, Hendry and Nielsen (1998) for examples. Other studies consider so-called stock-flow relationships, e.g. between income, consumption, and wealth as in Hendry and von Ungern-Sternberg (1981), or between sales, production, and inventories at the industry level as in Granger and Lee (1989).

The time series of the rate of inflation—at least over the post-WWII period—is often treated as being integrated of order one, denoted I(1),[1] in the literature on relationships between nominal variables. This carries an immediate implication that price levels are I(2) and has prompted a flurry of statistical research into models of I(2) variables, see Haldrup (1998) for a recent survey. Recent research has also established that the study of stock-flow relationships ought to be conducted within an I(2) framework, see Engsted and Johansen (1999). Still, with few exceptions the full-blown I(2) analysis tends to be avoided by the applied literature. Instead most studies rely on transformations that partly difference the data vector. The widespread use of transformations in dealing with I(2) variables seems related to the fact that inference in I(2) models is difficult in the sense that few hypotheses allow the usual asymptotic χ^2 inference. In particular, this holds for hypotheses on the so-called polynomially cointegrating relationships that relate levels and growth rates of the process.

The present paper offers a general and formal characterization of the partly differencing approach. Specifically, we derive the properties of a transformed vector process obtained by partly differencing an I(2) process. The transformation eliminates the I(2) trends while retaining possible cointegrating relationships among the variables. A case is examined in which the original process is generated by a vector autoregression (VAR). The transformation examined is minimal in terms of the amount of *a priori* information on the parameters required to achieve a valid reduction from I(2) to I(1). Throughout, the validity of *a priori* parameter restrictions will be taken as given. Clearly, in empirical applications their validity should be tested.[2] A properly transformed process will satisfy a VAR and inference on the full set of cointegrating parameters can be achieved by standard I(1) methods.

The parameters of the transformed VAR are subject to certain restrictions.

[1] A process is integrated of order d, denoted I(d), if it becomes stationary only after first-differencing d times, see Johansen (1996) for the formal definition.

[2] Kongsted (2002) examines the properties of $\tilde{X}_t$ under invalid *a priori* parameter restrictions. Tests of the validity of the transformation are derived by Kongsted (1998) based on the two-step I(2) algorithm of Johansen (1995a) and by Johansen (2002) based on the maximum likelihood estimator of vector autoregressions with I(2) restrictions.

One set of restrictions is shown to fit directly into the standard I(1) reduced rank regression analysis, see Johansen (1996). Another set requires different and more elaborate estimation techniques and is commonly ignored in applied work. Moreover, we find that standard methods for inference on the cointegration rank employ an alternative hypothesis which is irrelevant under the assumptions maintained in transforming the data. The likely consequences of ignoring these considerations in terms of the resulting efficiency loss are explored in an empirical example and a small-scale simulation experiment. The main objective lies in analyzing if the transformed model can provide a practical and efficient means for inference on the parameters of the original I(2) model.

The paper is outlined as follows: Section 2 defines the I(2) model assumed to generate the original data and the class of transformations analyzed. The structure of cointegrating relationships and common stochastic trends in the transformed model is also derived in this section and the parameters of the I(2) model are recovered. Section 3 collects the results on parameter restrictions implied by the transformation and outline some estimation algorithms that impose those restrictions. Section 4 provides an empirical illustration based on Banerjee, Cockerell and Russell (2001) and uses it to set up a small-scale simulation experiment.

Some notation and definitions are used throughout: For a $p \times r$ matrix α of rank r, let $\alpha_\perp$ denote a basis of the $p \times (p-r)$ orthogonal complement and define $\bar{\alpha} = \alpha(\alpha'\alpha)^{-1}$. For β $(p \times r)$ and η $(p-r \times s)$, $s < p-r$, define $\beta_1 = \bar{\beta}_\perp \eta$ and $\beta_2 = \beta_\perp \eta_\perp$. The matrices β, β_1, and β_2 are thus mutually orthogonal. Also note the relationship $I = V(B'V)^{-1}B' + b(v'b)^{-1}v'$ where $V = v_\perp$ is $p \times (r+s)$, $B = b_\perp$, and $|v'b| \neq 0$, see Hansen and Johansen (1998).

2 A Minimal Transformation from I(2) to I(1)

This section derives the process obtained by a minimal transformation from I(2) to I(1), relating its cointegration properties and common stochastic trend structure to the original I(2) process. Because the precise I(2) conditions play a major role in deriving the implied restrictions, the section will briefly set up the VAR of the original data based on Johansen (1992) and Rahbek, Kongsted and Jørgensen (1999). Then, a minimal I(2)-to-I(1) transformation is defined and the transformed VAR is derived.

2.1 The Original I(2) Process

The starting point for the analysis is a p-dimensional I(2) vector time series, X_t. The original process satisfies the kth order vector autoregression written in a parameterization suitable for I(2) processes,

$$\Delta^2 X_t = \Pi X_{t-1} - \Gamma \Delta X_{t-1} + \sum_{i=1}^{k-2} \Psi_i \Delta^2 X_{t-i} + \mu_0 + \mu_1 t + \varepsilon_t, \tag{1}$$

for $t = 1, \ldots, T$. For the statistical analysis ε_t is assumed to be identically and independently distributed $N(0, \Omega)$ terms and the initial observations, $X_{-k+1}, \ldots, X_0$, are taken to be fixed.

Assuming that the roots of the characteristic polynomial of (1) are either at one or outside the unit circle and maintaining a further rank condition, see Johansen (1992), the parameters of (1) should satisfy the following reduced rank conditions for X_t to be an I(2) process,

$$\Pi = \alpha\beta' \quad \text{and} \quad \alpha'_\perp \Gamma \beta_\perp = \xi\eta'. \tag{2}$$

Here, α and β are $p \times r$ matrices of full rank, whereas ξ and η are $(p-r) \times r$ and also of full rank.

The cointegration structure of X_t and the structure of its common stochastic trends are determined by (2) as derived in Johansen (1996, section 4.3). There are $q = p - r - s$ common I(2) trends embodied in X_t, represented as $\alpha'_2 \sum\sum \varepsilon_i$ with $\alpha_2 = \alpha_\perp \xi_\perp$. They are loaded into the X_t process by a matrix which is proportional to $\beta_2 = \beta_\perp \eta_\perp$. The common I(2) trends are eliminated by the full set of $r + s$ cointegrating vectors (β, β_1) where $\beta_1 = \bar{\beta}_\perp \eta$. Both sets of linear combinations, $\beta' X_t$ and $\beta'_1 X_t$ are I(1) in general and include the first-differenced I(2) component, $\alpha'_2 \sum \varepsilon_i$. The s linear combinations, $\beta'_1 X_t$, in addition includes the genuine I(1) trend of the system, $\alpha'_1 \sum \varepsilon_i$, where $\alpha_1 = \bar{\alpha}_\perp \xi$, and do therefore not cointegrate any further. The r linear combinations, $\beta' X_t$, on the contrary, cointegrate to stationarity with ΔX_t, producing a second layer of cointegration reflected in the r polynomially cointegrating relations,

$$S_t = \beta' X_t - \delta\beta'_2 \Delta X_t, \tag{3}$$

which are I(0). This relationship includes the polynomially cointegrating parameter, $\delta = \bar{\alpha}' \Gamma \bar{\beta}_2$, of dimensions $r \times q$. Note that if $r > q$, there are fewer I(2) trends than polynomially cointegrating relationships and $r - q$ directly cointegrating relationships can be defined among the X_t variables as $\delta'_\perp \beta' X_t$.

In terms of the deterministic terms in (1) a specification will be considered that allows the process X_t to be linearly trending in all directions whereas, by assumption, no quadratic trends can be present. This is the model of Rahbek et al. (1999), which requires the parameters of the constant and linear drift terms in (1) to be restricted as

$$\alpha'_\perp \mu_0 = -\xi\eta'_0 - \alpha'_\perp \Gamma \bar{\beta} \beta'_0 \tag{4}$$

and

$$\mu_1 = \alpha\beta'_0, \tag{5}$$

where η'_0 and β'_0 are vectors of dimensions s and r. Transforming the process X_t will in most cases also impose restrictions on the deterministic part.

2.2 The Transformation

The transformed process is defined and analyzed under a specific set of assumptions on the parameters of the original process, X_t. Those assumptions reflect a situation in which there are strong *a priori* expectations of some number of common I(2) trends being shared in certain known proportions by a (sub)set of variables in X_t.[3]

A common example would be that one nominal trend is reflected by several I(2) variables in equal proportions, e.g. by the price level and the money stock, or by several price measures, as in the empirical application below. The transformed process then includes variables that reduce to I(1) either by linear transformation, e.g. the real money stock and relative prices (along with any real variables in X_t), or by first-differencing as for instance the rate of inflation.

In general terms, the transformation starts from a known matrix b of dimension $p \times q$. The transformed vector process, $\tilde{X}_t$, is defined by

$$\tilde{X}_t = \begin{pmatrix} B'X_t \\ v'\Delta X_t \end{pmatrix} \equiv \begin{pmatrix} Z_t \\ U_t \end{pmatrix} \tag{6}$$

where $B = b_\perp$ is $p \times (r+s)$. The $p \times q$ matrix v that defines the first-differenced term should satisfy $|v'b| \neq 0$. Throughout b is assumed to satisfy orthogonality in terms of the full set of cointegrating vectors,

$$b'(\beta, \beta_1) = 0. \tag{7}$$

2.3 The Structure of the Transformed Process

Under the condition (7) the process $\tilde{X}_t$ will be I(1) with cointegrating rank r. This is shown by Kongsted (2002) who also examines the general case of a potentially invalid transformation that does not achieve the reduction from I(2) to I(1). The present paper examines the case that (7) is indeed satisfied. Next, we characterize the parameter restrictions implied by the fact that $\tilde{X}_t$ derives from the I(2) process X_t by this particular transformation. Moreover, it will be shown explicitly how to recover the parameters of the original I(2) model.

The matrix of loadings of the I(2) common trends, β_2, is known (up to a normalization) under the condition (7) and b is a valid basis for β_2. Similarly, the full set of cointegrating vectors can be given the representation

$$(\beta, \beta_1) = B(\varphi, (B'B)^{-1}\varphi_\perp) \tag{8}$$

for some $(r+s) \times r$ matrix φ of full rank. The condition imposed on the matrix v ensures that linear combinations of first-differenced process that are needed in order to recover the polynomially cointegrating relationships (3) are in fact included in U_t.

The transformation requires b and thus the number of common I(2) trends in the system, $q = p - r - s$, to be known. The number of polynomially cointegrating relationships, r, and the number of genuine I(1) common trends amongst the variables, s, are only restricted by their sum, $r + s$. This reflects a view that often less *a priori* information is available on r and s. Note that

[3] The transformation is denoted a nominal-to-real transformation in Kongsted (2002) since it typically involves going from a system of nominal variables to a real system.

the transformation leaves unrestricted the relationship between B and each of the sets of cointegrating vectors, β and β_1. In that sense, (6) is the minimal transformation that achieves the reduction from I(2) to I(1).

Next, the parameters of the transformed process $\tilde{X}_t$ will be derived. To this end, the conditions $\Pi = \alpha\beta'$ and $\mu_1 = \alpha\beta_0'$ are imposed in (1) which is then premultiplied by the non-singular matrix $M = (B, v)'$ to obtain

$$\begin{aligned}\begin{pmatrix} B'\Delta^2 X_t \\ v'\Delta^2 X_t \end{pmatrix} = \begin{pmatrix} B'\alpha\beta' X_{t-1} \\ v'\alpha\beta' X_{t-1} \end{pmatrix} &- \begin{pmatrix} B'\Gamma\Delta X_{t-1} \\ v'\Gamma\Delta X_{t-1} \end{pmatrix} + \sum_{i=1}^{k-2} \begin{pmatrix} B'\Psi_i\Delta^2 X_{t-i} \\ v'\Psi_i\Delta^2 X_{t-i} \end{pmatrix} \\ &+ \begin{pmatrix} B'\alpha\beta_0' t \\ v'\alpha\beta_0' t \end{pmatrix} + \begin{pmatrix} B'(\mu_0 + \varepsilon_t) \\ v'(\mu_0 + \varepsilon_t) \end{pmatrix}. \end{aligned} \tag{9}$$

Then the definitions of Z_t and U_t are applied and $A = V(B'V)^{-1}$ and $a = b(v'b)^{-1}$ are defined. Substituting $\Delta X_t = A\Delta Z_t + aU_t$ in (9), and finally collecting terms, a set of equations for $\tilde{X}_t$ is obtained,

$$\Delta\tilde{X}_t = \tilde{\Pi}\tilde{X}_{t-1} + \sum_{i=1}^{k-1} \tilde{\Gamma}_i \Delta\tilde{X}_{t-i} + \tilde{\mu}_1 t + \tilde{\mu}_0 + \tilde{\varepsilon}_t, \tag{10}$$

where $\tilde{\varepsilon}_t = M\varepsilon_t$, $\tilde{\mu}_0 = M\mu_0$ and $\tilde{\mu}_1 = M\alpha\beta_0'$. This is a VAR($k$) for the transformed variables in error correction format, the standard representation for cointegration analysis of I(1) processes, see Johansen (1996). Comparing (9) and (10) the parameters of the transformed VAR are given by:

$$\tilde{\Pi} = \begin{pmatrix} B'\alpha\varphi' & -B'\Gamma a \\ v'\alpha\varphi' & -v'\Gamma a \end{pmatrix},$$

$$\tilde{\Gamma}_1 = \begin{pmatrix} I + B'(\Psi_1 - \Gamma)A & B'\Psi_1 a \\ v'(\Psi_1 - \Gamma)A & v'\Psi_1 a \end{pmatrix},$$

$$\tilde{\Gamma}_i = \begin{pmatrix} B'(\Psi_i - \Psi_{i-1})A & B'\Psi_i a \\ v'(\Psi_i - \Psi_{i-1})A & v'\Psi_i a \end{pmatrix}, \quad i = 2, \ldots, k-2,$$

$$\tilde{\Gamma}_{k-1} = \begin{pmatrix} -B'\Psi_{k-2}A & 0 \\ -v'\Psi_{k-2}A & 0 \end{pmatrix}.$$

The structure of the cointegrating relationships in the transformed model can be analyzed by applying a result from Johansen (1992), that for the I(2) process it holds that

$$\Gamma = \Gamma\bar{\beta}\beta' + (\alpha\bar{\alpha}'\Gamma\bar{\beta}_1 + \alpha_1)\beta_1' + \alpha\bar{\alpha}'\Gamma\bar{\beta}_2\beta_2'. \tag{11}$$

Post-multiplying by $a = b(v'b)^{-1}$ will eliminate the first two terms because b is a valid basis for $sp(\beta_2)$. Thus, $\Gamma a = \alpha\tilde{\delta}$ with $\tilde{\delta} = \delta b'b(v'b)^{-1}$ where δ reflects the rule adopted for normalizing β_2 in the I(2) model. Substituting for Γa in the above expression for $\tilde{\Pi}$ it is seen to be the product of $p \times r$ matrices

$$\tilde{\beta} = \begin{pmatrix} \varphi \\ -\tilde{\delta}' \end{pmatrix} \text{ and } \tilde{\alpha} = M\alpha = \begin{pmatrix} B'\alpha \\ v'\alpha \end{pmatrix}. \tag{12}$$

The full set of I(2) cointegrating parameters is seen to be recovered from the cointegrating parameters of the transformed process, $\tilde{\beta}$. Specifically, $\beta = B\varphi$ and $\beta_1 = \bar{B}\varphi_\perp$. A particularly useful result concerns the parameter δ of the polynomially cointegrating relationship (3) which can be recovered as $\delta = \tilde{\delta}v'b(b'b)^{-1}$. It enters the transformed model as a standard I(1) cointegrating parameter for which there is a well-developed theory of inference, see Johansen (1996).

It can also be noted that $sp(\tilde{\delta})$ equals $sp(\delta)$ which means that if $r > q$ we can define $\tilde{\delta}_\perp = \delta_\perp$ and produce $r-q$ linear combinations of $\tilde{\beta}$ with a zero coefficient for U_t. These reflect the directly cointegrating relationships that may exist in the I(2) model. In the general case, restrictions on elements of $\tilde{\delta}$ can be imposed as restrictions on $\tilde{\beta}$.[4] Note that a zero restriction on any linear combination of the last $p-r-s$ rows of $\tilde{\beta}$ implies that certain linear combination of the differenced common I(2) trend does not enter the polynomially cointegrating relationship of the I(2) model.

The matrix of equilibrium-correction loadings of the levels term in (1), α, is recovered as

$$\alpha = M^{-1}\tilde{\alpha}$$

where $M^{-1} = (V(B'V)^{-1}, b(v'b)^{-1})$. Because $\tilde{X}_t$ is I(1) and satisfies the VAR (10) we can apply a result on weak exogeneity from Johansen (1992b): If $\tilde{c}'\tilde{\alpha} = 0$ for some $p \times (p-m)$ matrix $\tilde{c}$ with $m \geq r$, then the process $\tilde{c}'\tilde{X}_t$ is weakly exogenous for $\tilde{\beta}$. The equivalent condition in terms of α is $c'\alpha = 0$ with $c = M'\tilde{c}$. This shows that if β_2 is known then imposing a condition on α—or, equivalently, $\alpha_\perp$—is sufficient for efficient inference on the cointegrating parameters. This simplifies matters considerably as compared to the unrestricted I(2) case for which the general result of Paruolo and Rahbek (1999) holds that $c'X_t$ is weakly exogenous for the parameters (β, β_1, δ) if the condition $c'(\alpha, \alpha_1, \Gamma\bar{\beta}) = 0$ is satisfied.

Finally, in order to recover α_1 and α_2 it is useful to specify the common trends structure of the transformed process. The $p-r$ common I(1) trends of the transformed process are given by $\tilde{\alpha}'_\perp \sum \tilde{\varepsilon}_i$ with $\tilde{\alpha}_\perp = (M^{-1})'\alpha_\perp$ and therefore

$$\tilde{\alpha}'_\perp \sum \tilde{\varepsilon}_i = \alpha'_\perp M^{-1} \sum M\varepsilon_i = \alpha'_\perp \sum \varepsilon_i. \tag{13}$$

As $sp(\alpha_\perp) = sp(\alpha_1, \alpha_2)$ this shows that the common trends of the original I(2) process are fully recovered by the transformed I(1) process. Both sets of common trends now enter as I(1) stochastic trends. To separately identify α_1 and α_2 based on the transformed process we examine the matrix of loadings of the common trends. It is proportional to the orthogonal complement, $\tilde{\beta}_\perp$, and conveniently represented as

$$\tilde{\beta}_\perp = \begin{pmatrix} \varphi_\perp & \bar{\varphi}\delta \\ 0 & v'b(b'b)^{-1} \end{pmatrix}. \tag{14}$$

[4] This includes identifying restrictions which can be imposed on $\tilde{\delta}$ with the usual caveat that the structure is in fact identifying according the conditions laid out by Johansen (1995b). In particular, one should not be able to impose a full row of zeros in $\tilde{\delta}'$. This is testable by standard I(1) tools as a row restriction on $\tilde{\beta}$.

Two observations on the common trends structure emerge directly from (14). First, the fact that $|v'b| \neq 0$ implies that $U_t = v'\Delta X_t$ in itself is non-cointegrating, that is, any linear combination of the components the q-dimenstional process U_t would remain I(1). Essentially, the full set of common I(2) trends of the original process carry over in first differences via U_t. Secondly, the particular representation in (14) separates the last q columns related to the differenced I(2) trend, $\alpha_2' \sum \varepsilon_i$, from the first s columns related to the genuine I(1) trend, $\alpha_1' \sum \varepsilon_i$. To see this, note that $\alpha_2' \sum \varepsilon_i$ is the only common trend left in U_t due to first-differencing. Thus, $\alpha_1' \sum \varepsilon_i$ produce zero loadings in U_t which can be used as part of the identification scheme of a structural common trends model along the lines of King, Stock, Plosser and Watson (1991), see also Warne (1993). Economic theory may suggest alternative identifying assumptions but the different roles assigned to $\alpha_1'\varepsilon_t$ and $\alpha_2'\varepsilon_t$ due to their very different effects in the I(2) model would seem suggestive for their economic interpretation as well.

3 Restrictions and Estimation

The parameters of the transformed VAR in (10) are subject to certain restrictions that derive from the structure of the original process, X_t, and the transformation itself. The restrictions can be categorized in two groups in practical terms.

A first group of restrictions relate to the reduced rank of the coefficient matrix of the levels term, $\tilde{\Pi}$. This leads to the parameterization $\tilde{\Pi} = \tilde{\alpha}\tilde{\beta}'$. Moreover, the coefficient of the linear trend can be seen to be restricted accordingly, that is,

$$\tilde{\mu}_1 = \tilde{\alpha}\beta_0'. \tag{15}$$

The reduced rank of $\tilde{\Pi}$ and (15) are straightforward to implement in the reduced rank regression algorithm for I(1) models with a restricted linear term, see Johansen (1996).

A second group of parameter restrictions on the transformed VAR do not fall naturally into the reduced rank regression framework. Evidently, there are zero restrictions on the coefficients of ΔU_{t-k+1}, restricting the last q columns of the last lag coefficient, $\tilde{\Gamma}_{k-1}$. Moreover, if the I(2) process has restricted deterministic terms, conditions such as (4) carry over to the transformed model. Specifically, as (4) serves to exclude the possibility of quadratic trends in X_t, the first-differenced component in $\tilde{X}_t$, $U_t = v'\Delta X_t$, can have no linear trend.

The latter group of restrictions are commonly ignored in applied work. The aim of the empirical example and the simulation experiment below is to assess the importance of the resulting efficiency loss. Before turning to this part we outline the estimation algorithms adopted.

3.1 Reduced Rank Regression

For completeness we first outline the standard case of reduced rank regression. This is based on a VAR(k) of the transformed variables, $\tilde{X}_t$, with a restriction on the linear trend term similar to (5). Maximum likelihood estimation of this model amounts to solving the eigenvalue problem

$$\left|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}\right| = 0,$$

where $S_{ij} = T^{-1} \sum_{t=1}^{T} R_{it} R'_{jt}$ are sample moment matrices, and R_{0t} and R_{1t} are least squares residuals of regressing $\Delta \tilde{X}_t$ and $(\tilde{X}'_{t-1}, t)'$ respectively on $W_t = (\Delta \tilde{X}'_{t-1}, \Delta \tilde{X}'_{t-2}, \ldots, \Delta \tilde{X}'_{t-k+1}, 1)'$, see Johansen (1996, chapter 6). This yields $p + 1$ ordered eigenvalues $1 > \widehat{\lambda}_1 > \widehat{\lambda}_2 > \ldots > \widehat{\lambda}_p > \widehat{\lambda}_{p+1} = 0$. The MLE of $(\tilde{\beta}', \beta_0)'$ is given by the eigenvectors corresponding to the r largest eigenvalues. Furthermore the likelihood ratio test for a reduced rank of r compared to the full rank alternative can be written as a function of the eigenvalues as the Trace test statistic

$$Q_r = -2 \log Q \left(\operatorname{rank}(\tilde{\Pi}) \leq r \mid \operatorname{rank}(\tilde{\Pi}) \leq p\right) = -T \sum_{i=r+1}^{p} \log\left(1 - \widehat{\lambda}_i\right). \quad (16)$$

Note that by not imposing the restriction on U_{t-k+1} one more initial observation is necessary for the unrestricted model.

3.2 Restriction on the Lagged First-Differences

The redundancy of U_{t-k+1} is implied by the hypothesis

$$H_0 : \widetilde{\Gamma}_{k-1} = \left(\gamma_{p \times (r+s)}, 0_{p \times q}\right),$$

where γ contains the free parameters. To impose the restriction we modify the reduced rank regression described above. In particular we modify the vector of unrestricted variables to obtain

$$W_t^* = \left(\Delta \tilde{X}'_{t-1}, \Delta \tilde{X}'_{t-2}, \ldots, \Delta \tilde{X}'_{t-k+1} \begin{pmatrix} I_{r+s} \\ 0_{q \times (r+s)} \end{pmatrix}, 1\right)'$$

whereby U_{t-k+1} is excluded. Note that each equation of the VAR still contains the same set of variables and the effect can still be partialled out using least squares.

3.3 Restriction on the Trend Term

In order to impose that a linear trend is absent in U_t we first rewrite (10) as

$$\Delta \widetilde{Y}_t = \widetilde{\alpha} \widetilde{\beta}' \widetilde{Y}_{t-1} + \sum_{i=1}^{k-1} \widetilde{\Gamma}_i \Delta \widetilde{Y}_{t-i} + \widetilde{\alpha} \psi + \widetilde{\epsilon}_t \quad (17)$$

$$\widetilde{X}_t = \widetilde{Y}_t + \theta t. \quad (18)$$

Non-zero means in all directions are allowed for by the constant term restricted to the cointegrating relations in (17) and the linear trend is added in the factor representation (18). The restriction is that the last q elements in θ are zero, i.e.

$$H_1 : \theta = N\vartheta = \begin{pmatrix} I_{r+s} \\ 0_{q\times(r+s)} \end{pmatrix} \vartheta = \begin{pmatrix} \vartheta \\ 0_{q\times 1} \end{pmatrix}$$

where ϑ contains the free trend parameters.

Maximum likelihood estimation of the model under H_1 can be performed by applying the switching algorithm of Nielsen (2002). The idea is that conditional on an estimate, $\widehat{\theta}$, of the parameters to the linear trend, θ, the parameters of (17) can be estimated using a usual reduced rank regression of the corrected data $\widetilde{Y}_t = \widetilde{X}_t - \widehat{\theta}t$. With these estimates we can construct the estimated characteristic polynomial, $\widehat{A}(L)$, and the estimated residual, $\widehat{e}_t = \widehat{A}(L)\widetilde{X}_t - \widehat{\widetilde{\alpha}}\widehat{\psi}$, which under the model can be written as

$$\widehat{e}_t = \widehat{A}(L)\,N\vartheta t + \widetilde{\epsilon}_t. \tag{19}$$

Since t is a scalar variable we can rewrite (19) as

$$\widehat{e}_t = \widehat{H}_t\vartheta + \widetilde{\epsilon}_t, \tag{20}$$

where $\widehat{H}_t \equiv \widehat{A}(L)\,Nt$. The likelihood function conditional on $\widehat{\widetilde{\alpha}}\widehat{\psi}$ and the parameters in $\widehat{A}(L)$ is maximized over ϑ by a GLS estimation in (20), i.e.

$$\widehat{\vartheta} = \left(\sum_{i=1}^{T}\left(\widehat{H}_t'\widehat{\Omega}^{-1}\widehat{H}_t\right)\right)^{-1}\left(\sum_{i=1}^{T}\left(\widehat{H}_t'\widehat{\Omega}^{-1}\widehat{e}_t\right)\right), \tag{21}$$

see Tsay, Peña and Pankratz (2000) and Saikkonen and Lütkepohl (2000) for a similar GLS step used in a two step estimator. Here we follow Nielsen (2002) and iterate between the two conditional steps until convergence.

4 Empirical Application and Simulations

This section provides an empirical assessment of likely efficiency losses associated with the usual practice of applying standard I(1) methods to the unrestricted VAR of $\tilde{X}_t$, ignoring the fact that it was derived by transforming an I(2) process. First, the empirical analysis of Banerjee, Cockerell, and Russell (2001), BCR in the following, is reexamined in view of the above results. Secondly, a small scale Monte Carlo experiment is set up to provide simulation evidence on the issue. It uses the estimated BCR model as a realistic data generating process (DGP).[5]

[5] The computations have been conducted in Ox 3.0, see Doornik (2001).

4.1 Empirical Illustration

BCR used Australian data to analyze the relation between inflation and the markup of prices on costs. The analysis included a set of three core variables: The log of consumer prices, p_t, the log of unit labor costs, u_t, and the log of import prices, m_t.[6] The data has 94 effective quarterly observations for the period $t = 1972:1 - 1995:2$. BCR analyzed a VAR(2) with the deterministic specification of Paruolo (1996), applying the two-step estimator of Johansen (1995a).

For the purpose of illustration we make a few modifications to the specification. In particular, four conditioning variables are excluded so that $X_t = (p_t, u_t, m_t)'$, and the deterministic specification of Rahbek et al. (1999) is applied. Moreover, we apply the maximum likelihood (ML) estimator of the I(2) model, see Johansen (1997), to ensure that differences between the I(2) model and the transformed I(1) model do not reflect inefficiencies in the estimation of the I(2) model.

BCR impose the rank indices $r = 1$ and $s = 1$, and linear homogeneity between the variables, i.e. $b = (1, 1, 1)'$. The chosen rank indices are also consistent with the simplified specification.[7] The six eigenvalues of the characteristic polynomial of the restricted model have moduli given by

$$1.000;\ 1.000;\ 1.000;\ .540;\ .277;\ .060.$$

Two of the unit roots are associated with the common I(2) trend and one with the genuine I(1) trend of the system. The largest unrestricted eigenvalue is far from unit circle, reflecting a fast dynamic adjustment.

The estimates of the polynomially cointegration parameters are reported in Table 1 in terms of β' and $-\tilde{\delta}$. The estimates in row (iii) are obtained by applying the ML estimator to the simplified model. The original estimates from BCR are reported in rows (i) and (ii). The results are similar although the import share of the present analysis is larger. The differences reflect the different deterministic specifications, the exclusion of conditioning variables, and the use of the MLE rather than the two-step estimator.

A nominal-to-real transformation is performed by BCR using the matrices

$$B = \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} \text{ and } v = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tag{22}$$

which produces a transformed data set given by $\widetilde{X}_t = (p_t - u_t, p_t - m_t, \Delta p_t)'$. It includes the markup on unit labor costs, the inverse of the real import prices and the rate of change in the consumer prices. The transformation satisfies the requirement that $|v'b| \neq 0$.

[6] In addition, four variables assumed to be stationary and weakly exogenous, were included.

[7] The test of homogeneity is one of the few tests that can be conducted by standard methods in the I(2) model, see Kongsted (1998) and Johansen (2002). Using the ML estimator, homogeneity is accepted with a test statistic of 1.43 corresponding to a $p-$value of .49 according to a $\chi^2(2)$ distribution.

Applying this transformation to the simplified version of the BCR model and estimating the unrestricted I(1) model yields the estimates presented in row (*iv*). The log-likelihood of the unrestricted model is only marginally higher than the likelihood of the homogeneous I(2) model. The formal test statistic for the test of 4 restrictions—three restricted lag coefficients and one restricted trend coefficient—is around two, which is far from significant in a $\chi^2(4)$ distribution.

Row (*v*) reports the results of imposing zero restrictions on the last lag on $v'\Delta X_t$, i.e.

$$H_0 : \widetilde{\Gamma}_1 = \left(\gamma_{3\times 2}, 0_{3\times 1}\right),$$

while row (*vi*) reports the results obtained by imposing the restriction that the transformed variables $v'\Delta X_t$ have no linear trend,

$$H_1 : \theta = \left(\vartheta'_{1\times 2}, 0\right)'.$$

Finally, row (*vii*) reports the results of imposing both restrictions on the model. The fully restricted I(1) model for the transformed data is simply a reparametrization of the homogeneous I(2) model and the results in rows (*iii*) and (*vii*) are seen to be identical. The partly restricted models show that in the present case the importance of the trend restriction, H_1, is negligible whereas the lag restriction, H_0, is somewhat more important. This could simply be a result of the fact that H_0 impose three restrictions on the model whereas H_1 impose only one. In total, the results indicate that inference on β' and $\tilde{\delta}'$ can be effectively performed in an unrestricted I(1) analysis of the transformed data set.

Subject to the limitation that $|v'b| \neq 0$, the choice of v only matters for the interpretation of the model. This is illustrated in the lower part of Table 1, which reports estimation results for an alternative choice of v. The average inflation rate, $v = \frac{1}{3}(1,1,1)'$, now represents the first difference of the I(2) trend. If the full set of restrictions is imposed as in row (*xi*), the results are indeed identical to row (*iii*) and (*vii*) and a different choice of v amounts to a reparametrization. However, in unrestricted or partly restricted models some differences in the results may appear. The effects of including a redundant lag of $v'\Delta X_t$ depend on the sample correlation between that particular variable and the other terms in the model, and therefore also on the specific choice of v. Similarly, the effects of the redundant linear trend allowed for in $v'\Delta X_t$ depend on the particular sample values.[8]

4.2 Simulations

The loss of efficiency from ignoring the derived nature of the transformed process will now be characterized along two dimensions: Cointegration rank determination and estimation of the polynomially cointegrating parameters. A small

[8] A grid search over possible vectors $v = (1, v_2, v_3)'$, where v_2 and v_3 take values between -100 and 100, resulted in log-likelihood values between 1195.13 and 1196.24 for the unrestricted RRR estimator, the extremes being obtained for the vectors $(1, 0.2, -0.4)'$ and $(1, -1, -0.2)'$ respectively.

scale Monte Carlo simulation is set up for this. It employs the homogeneous I(2) model reported in row (*iii*) of Table 1 as its DGP.[9]

Samples of $T + 100$ observations, $t = -101, -100, ..., 0, 1, 2, ..., T$, are generated by replacing ε_t with pseudo-random independent $N\left(0, \widehat{\Omega}\right)$ drawings. Sample sizes of $T = \{50, 75, 100, 200, 400\}$ effective observations are considered, 100 presample observations are discarded to eliminate the importance of the initial values $Y_{-101}, Y_{-100} = 0$, and 10.000 Monte Carlo replications are used for each case.

The determination of cointegration rank based on the transformed process is the starting point in most applications. Kongsted (2002) showed that subject to (7), the rank index r will indeed be correctly identified as rank$(\tilde{\Pi})$. However, simply applying the trace test (16) would employ rank$(\tilde{\Pi}) \leq p$ as the alternative hypothesis when, under the set of assumptions maintained in transforming the process, in fact it holds that rank$(\tilde{\Pi}) \leq p - q$. A potential efficiency loss is evident already from the fact that applying the standard trace test could well result in an estimated cointegration rank in the infeasible interval $p - q < r \leq p$.

Alternatively, one could take rank$(\tilde{\Pi}) \leq p - q$ as the alternative hypothesis and consider a modified trace statistic,

$$\tilde{Q}_r = -2 \log Q \left(\text{rank}(\tilde{\Pi}) \leq r \mid \text{rank}(\tilde{\Pi}) \leq p - q\right) = -T \sum_{i=r+1}^{p-q} \log\left(1 - \widehat{\lambda}_i\right). \tag{23}$$

For the special case of testing rank$(\tilde{\Pi}) \leq p - q - 1$ against rank$(\tilde{\Pi}) \leq p - q$ this is the so-called λ_{max} statistic and the asymptotic distribution is tabulated e.g. in MacKinnon, Haug and Michelis (1999) and Doornik (1998). For more general cases the asymptotic distribution is not readily available but can be easily simulated.[10]

The BCR-based DGP that underlies the simulation experiment has $p = 3$ and $q = 1$. The modified trace statistic $\tilde{Q}_r$ employs rank$(\tilde{\Pi}) \leq 2$ as the alternative hypothesis. The relevant asymptotic distribution is the λ_{max} distribution with two degrees of freedom. Figure 1 compares the empirical rejection frequencies of two tests of the (true) null that rank$(\tilde{\Pi}) \leq 1$, Q_1 and $\tilde{Q}_1$, at different sample sizes. The tests can hardly be distinguished. The rejection frequencies of Q_0 and $\tilde{Q}_0$ for the (false) null that rank$(\tilde{\Pi}) = 0$ are also depicted in Figure 1. For small and moderate sample sizes there is an efficiency gain since the power of $\tilde{Q}_0$ is marginally higher. Still, the differences are minor and, overall,

[9] This is chosen as a simple yet realistic DGP. Due to the fast dynamic adjustment in the model, it is well-behaved and shows reasonable size and power properties even in fairly small samples. Still, the emphasis is on the comparison between unrestricted and restricted estimators rather than their performance in absolute terms.

[10] We obtained the asymptotic distributions employed in the simulations below by simulations with 100.000 replications, using a response surface for the mean and the variance based on 26 different values of T between 50 and 5.000. The critical values are then calculated from a Gamma approximation to the asymptotic distribution, see Doornik (1998). The 95 per cent quantiles thus obtained for the case of $p = 3$ and $q = 1$ are: $\tilde{Q}_1$: 19.25 ; Q_1: 25.74 ; $\tilde{Q}_0$: 37.47; Q_0: 42.77.

the loss of efficiency from using the standard trace test seems very limited. We will therefore make use of the standard test in the following.

The second issue concerns efficiency in estimating the polynomially cointegrating parameters from the transformed model by unrestricted reduced rank regression. In each Monte Carlo replication we applied the estimators outlined in Section 3 to the simulated data. To evaluate the properties of the estimators the average angle between the estimated $\widetilde{\beta}$ and the true vector of the DGP, $\widetilde{\beta} = (0.795, 0.205, 7.887)'$, is reported together with the rejection frequency of the LR test for the (true) restrictions on the last lag and the trend coefficient. Furthermore, we report the actual size and power at a nominal 5 per cent level of the standard trace test for rank determination based on the different estimators.

Table 2 reports the estimation results for a transformation given by B and v as defined in (22). In the unrestricted RRR estimation, reported in Panel A, the average angle between the true and the estimated $\widetilde{\beta}$ is 6 degrees for $T = 50$ and converging to zero relatively fast. The results for the rank determination reflect fast dynamic adjustment in the DGP. With $T = 100$ observations the power of the Trace test for the hypothesis $r = 0$ is 95 per cent and the size is close to 5 per cent for all sample lengths.[11]

Imposing the lag restriction, H_0, improves the average precision of the estimates in small samples, cf. the results reported of Panel B in Table 2. For 50 observations the restriction improves the average angle from 6 to 4.7 degrees. For 100 observations the difference is down to .1 degrees and for $T = 200$ the results are almost identical. In the restricted model the distribution of the Trace test is apparently moved a little to the right, implying a higher power and higher size of the test. The LR test for the lag restriction, H_0, which is $\chi^2(3)$ distributed, is somewhat oversized in small samples with a rejection frequency of 10 to 15 per cent against the nominal size of 5 per cent.

Imposing the restriction, H_1, that the variable $v'\Delta X_t$ has no linear trend, yields little or no improvement. The results reported in Panel C are almost identical to the results of the unrestricted RRR estimation. Due to a faster rate of convergence of the trend coefficient it matters very little if the restriction is imposed or not. Similarly, the results in Panel D of imposing the full set of restrictions are almost identical to the results obtained under the lag restriction alone reflecting the minor importance of the trend restriction.

5 Conclusions

This paper has derived the restrictions that apply to a transformed vector autoregression obtained by a minimal I(2)-to-I(1) transformation. The relationship between the parameters of the I(2) vector autoregression and the transformed model is characterized, including the coefficients of polynomially cointegrating relationships.

[11] Note that a small negative approximation error in terms of the 95 per cent quantile seems to be implied by the Gamma approximation used here.

In applied work it is common to use unrestricted reduced rank regression and to apply standard tools for inference on the cointegrating rank in the transformed model. We find that there is only a small gain from excluding the alternative hypotheses which are irrelevant under the assumptions maintained in transforming the data. Moreover, unrestricted reduced rank regression is shown to yield only a minor loss of efficiency compared to imposing the restrictions in the simulation experiment. Most efficiency is gained by imposing the absence of redundant lags in the differenced I(2) process, which is a fairly simple restriction in terms of the restricted estimation procedure. Imposing the restriction on the trend coefficients, which requires a more involved iterative estimation algorithm, leads to little or no efficiency gain. It appears fairly safe to ignore this restriction in applied work.

In conclusion, a properly transformed vector autoregression provides a practical and effective means for inference on the parameters of the I(2) model.

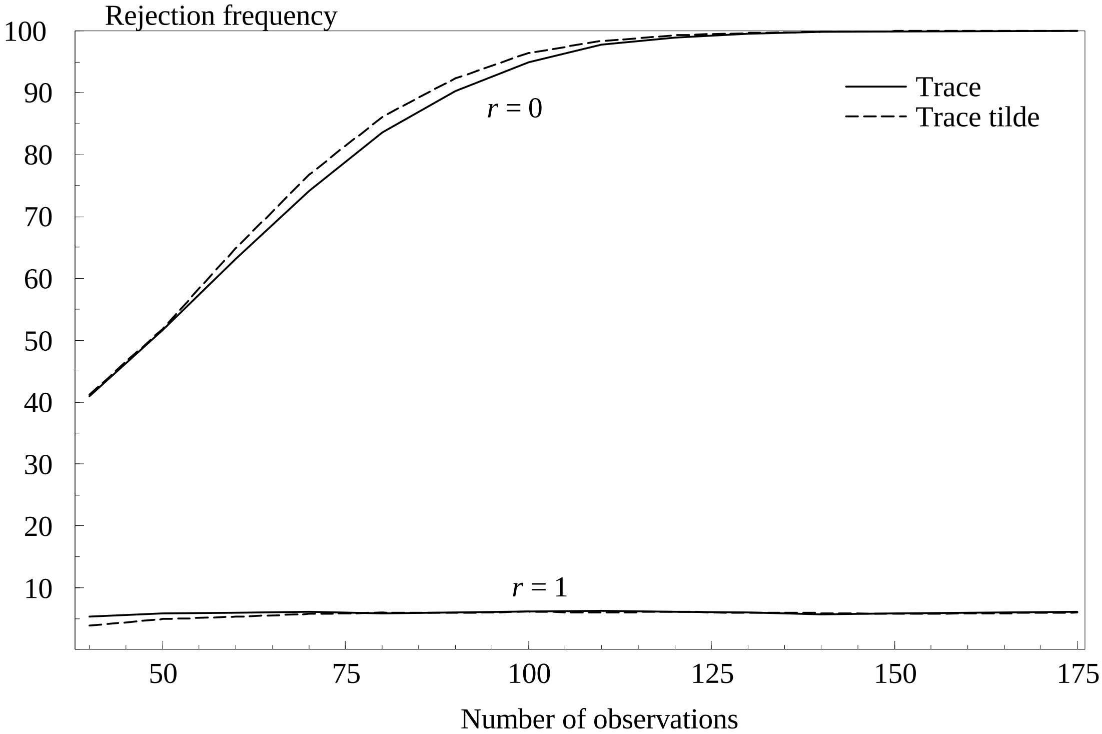

Figure 1: Rejection frequencies for tests of the true null of $r = 1$ and the false null of $r = 0$ using the conventional trace test statistic, Q_r, and the modified statistic, $\widetilde{Q}_r$.

References

[1] Banerjee, A., Cockerell, L., and Russell, B. (2001), An I(2) analysis of inflation and the markup, Journal of Applied Econometrics, 16, 221-240.

[2] Coenen, G., and Vega, J.-L., (2001), The Demand for M3 in the Euro Area, Journal of Applied Econometrics, 16, 727-748.

[3] Crowder, W., Hoffman, D., Rasche, R., (1999), Identification, Long-Run Relations, and Fundamental Innovations in a Simple Cointegrated System, Review of Economics and Statistics, 81(1), 109-21.

[4] Doornik, J.A. (1998), Approximations to the Asymptotic Distributions of Cointegration Tests, Journal of Economic Surveys 12(5), 573-93.

[5] Doornik, J.A. (2001), Object-Oriented Matrix Programming Using Ox, 4rd ed. London: Timberlake Consultants Press and Oxford: www.nuff.ox.ac.uk/Users/Doornik.

Table 1: Estimation on the Banerjee et al. (2001) data.

		β'			$-\tilde{\delta}$	Log-	Difference to
		p_t	u_t	m_t	Δp_t	likelihood	I(2) model
	Banerjee et al. (2001)						
(i)	Homogeneous I(2)	1.000	$-.868$	$-.132$	7.026		
(ii)	I(1)	1.000	$-.901$	$-.099$	7.201		
(iii)	Homogeneous I(2)	1.000	$-.795$	$-.205$	7.887	1194.7875	...
	$v = (1,0,0)'$						
(iv)	Unrestricted	1.000	$-.770$	$-.230$	8.252	1195.8085	1.0210
(v)	Lag restriction	1.000	$-.788$	$-.212$	8.180	1194.9731	.1856
(vi)	Trend restriction	1.000	$-.782$	$-.218$	7.742	1195.6406	.8532
(vii)	Both restrictions	1.000	$-.795$	$-.205$	7.887	1194.7875	.0000
	$v = \frac{1}{3}(1,1,1)'$						
$(viii)$	Unrestricted	1.000	$-.768$	$-.232$	8.706	1195.9540	1.1665
(ix)	Lag restriction	1.000	$-.788$	$-.212$	8.180	1194.9731	.1856
(x)	Trend restriction	1.000	$-.775$	$-.225$	8.319	1195.7618	.9743
(xi)	Both restrictions	1.000	$-.795$	$-.205$	7.887	1194.7875	.0000

Table 2: Simulations results, 10.000 replications

T	Average angle to true $\tilde{\beta}$	Difference in angle to Panel A	Rejection frequency (per cent): LR test of VAR restriction	Trace test $r = 0$	Trace test $r \leq 1$
	A. Unrestricted RRR estimation				
50	6.194	...	...	50.7	5.2
75	2.801	...	...	78.7	5.9
100	1.743	...	...	95.3	6.1
200	0.744	...	...	100.0	5.8
400	0.342	...	...	100.0	6.0
	B. Lag restriction imposed				
50	4.701	-1.493	14.5	75.8	7.0
75	2.379	-0.423	10.2	97.3	6.6
100	1.617	-0.125	8.6	99.9	6.2
200	0.725	-0.020	7.0	100.0	5.8
400	0.338	-0.004	6.2	100.0	6.1
	C. Trend restriction imposed				
50	6.269	0.075	4.5	47.3	4.8
75	2.838	0.037	4.7	76.2	5.5
100	1.739	-0.003	5.1	93.9	5.7
200	0.741	-0.003	4.9	100.0	5.3
400	0.343	-0.000	5.0	100.0	5.7
	D. Both restrictions imposed (Homogeneous I(2) model)				
50	4.835	-1.359	12.6	72.6	6.3
75	2.408	-0.393	9.4	96.5	6.1
100	1.619	-0.123	8.1	99.9	5.9
200	0.722	-0.022	6.5	100.0	5.4
400	0.338	-0.004	6.0	100.0	5.7

[6] Doornik, J., Hendry, D. and Nielsen, B. (1998), Inference in Cointegrating Models: UK M1 Revisited, Journal of Economic Surveys, 12, 533-572.

[7] Engsted, T., and Johansen, S. (1999): Granger's Representation Theorem and Multicointegration, in Engle, R., and White, H.: Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W.J. Granger, Oxford University Press.

[8] Granger, C., Lee,T., Investigation of Production, Sales and Inventory Relationships Using Multicointegration and Non-symmetric Error Correction Models, Journal of Applied Econometrics, 4, S145-59.

[9] Haldrup, N. (1998), An Econometric Analysis of I(2) Variables, Journal of Economic Surveys, 12, 595-650.

[10] Hansen, P. R., and Johansen, S (1998): Workbook on cointegration, Oxford University Press.

[11] Johansen, S. (1992), A Representation of Vector Autoregressive Processes Integrated of Order 2, Econometric Theory, 8, 188-202.

[12] Johansen, S. (1995a), A Statistical Analysis of Cointegration for I(2) Variables, Econometric Theory, 11, 25-59.

[13] Johansen, S. (1995b), Identifying restrictions of linear equations: with applications to simultaneous equations and cointegration, Journal of Econometrics, 11, 111-132.

[14] Johansen, S. (1996), Likelihood Based Inference in Cointegrated Vector Autoregressive Models, 2nd edition, Oxford University Press, Oxford.

[15] Johansen, S. (1997), Likelihood Analysis of the I(2) Model, Scandinavian Journal of Statistics, 433-462.

[16] Johansen, S. (2002), The statistical analysis of hypotheses on the cointegrating relations in the I(2) model, mimeo, Department of Statistics and Operations Research, University of Copenhagen.

[17] King, R., Plosser, C., Stock, J., and Watson, M. (1988), Stochastic Trends and Economic Fluctuations, American Economic Review, 81, 4, 819-840.

[18] Kongsted, H. C. (1998), An I(2) Cointegration Analysis of Small-Country Import Price Determination, Discussion paper 98-22, December 1998, Institute of Economics, University of Copenhagen.

[19] Kongsted, H. C. (2002), Testing the nominal-to-real transformation, Discussion Paper 02-06, Institute of Economics, University of Copenhagen.

[20] MacKinnon, J., Haug, A., Michelis,L. (1999), Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration, Journal of Applied Econometrics, 14(5), 563-77.

[21] Nielsen, Bent and Anders Rahbek (2000), Similarity Issues in Cointegration Analysis, Oxford Bulletin of Economics and Statistics, 62:1, 5-22.

[22] Nielsen, Heino (2002), Cointegration Analysis in the Presence of Outliers, Mimeo, Institute of Economics, University of Copenhagen.

[23] Paruolo, P. (1996), On the Determination of Integration Indices in I(2) Systems, Journal of Econometrics 72(1-2), 313-56.

[24] Paruolo, P., Rahbek, A. (1999), Weak Exogeneity in I(2) VAR Systems, Journal of Econometrics, 93, 281-308.

[25] Rahbek, A., Kongsted, H. C., and Jørgensen, C. (1999), Trend Stationarity in the I(2) Cointegration Model, Journal of Econometrics, 90, 265-289.

[26] Saikkonen, Pentti and Helmut Lütkepohl (2000), Trend Adjustment Prior to Testing for the Cointegration Rank of a VAR Process, Journal of Time Series Analysis, 21:4, 435-456.

[27] Tsay, Ruey S. , Daniel Peña and Alan E. Pankratz (2000), Outliers in multivariate time series, Biometrika, 87:4, 789-804.

[28] Warne, A. (1993), A common trends model: Identification, Estimation and Inference, Seminar paper 555, Institute of International Studies, Stockholm University.