Edinburgh School of Economics
Discussion Paper Series
Number 38

*Structural analysis of vector error correction models
with exogenous $I(1)$ variables*

Mohammad Hashem Pesaran (Trinity College, Cambridge)

Richard J Smith (University of Bristol)

Yongcheol Shin (University of Edinburgh)

Date
October 1999

Published by

School of Economics
University of Edinburgh
30 -31 Buccleuch Place
Edinburgh EH8 9JT
+44 (0)131 650 8361

<http://www.ed.ac.uk/schools-departments/economics>

THE UNIVERSITY *of* EDINBURGH

Structural Analysis of Vector Error Correction Models with Exogenous $I(1)$ Variables*

M. Hashem Pesaran
Trinity College, Cambridge

Yongcheol Shin
Department of Economics, University of Edinburgh

Richard J. Smith
Department of Economics, University of Bristol

October 1999

Abstract

This paper generalizes the existing cointegration analysis literature in two respects. Firstly, the problem of efficient estimation of vector error correction models containing exogenous $I(1)$ variables is examined. The asymptotic distributions of the (log-)likelihood ratio statistics for testing cointegrating rank are derived under different intercept and trend specifications and their respective critical values are tabulated. Tests for the presence of an intercept or linear trend in the cointegrating relations are also developed together with model misspecification tests. Secondly, efficient estimation of vector error correction models when the short-run dynamics may differ within and between equations is considered. A re-examination of the purchasing power parity and the uncovered interest rate parity hypotheses is conducted using U.K. data under the maintained assumption of exogenously given foreign and oil prices.

JEL Classification: C12, C13, C32.

Key Words: Structural Vector Error Correction Model, Cointegration, Unit Roots, Likelihood Ratio Statistics, Critical Values, Seemingly Unrelated Regression, Monte Carlo Simulations, Purchasing Power Parity, Uncovered Interest Rate Parity.

*Partial financial support from the ESRC (Grant Nos. R000233608 and R000237334) and the Isaac Newton Trust of Trinity College, Cambridge, for this research is gratefully acknowledged. We are also grateful to an Editor, two anonymous referees, Peter Boswijk, Cheng Hsiao, Bent Nielsen, and Michael Binder for helpful comments on the previous version of this paper [Pesaran, Shin and Smith (1997)].

1 Introduction

This paper generalizes the analysis of cointegrated systems advanced by Johansen (1991, 1995) in two important respects. Firstly, we consider a sub-system approach in which we regard a subset of random variables which are integrated of order one ($I(1)$) as structurally *exogenous*; that is, any cointegrating vectors present do not appear in the sub-system *vector error correction* model (VECM) for these exogenous variables and the error terms in this sub-system are uncorrelated with those in the rest of the system.¹ This generalization is particularly relevant in the macroeconometric analysis of “small open” economies where it is plausible to assume that some of the $I(1)$ forcing variables, for example, foreign income and prices, are exogenous. Similar considerations arise in the empirical analysis of sectoral and regional models where some of the economy-wide $I(1)$ forcing variables may also be viewed as exogenous. This extension paves the way for a more efficient multivariate analysis of economic time series for which data are typically only available over relatively short periods. Secondly, we allow constraints on the short-run dynamics in the VECM. This extension is also important in applied contexts where, due to data limitations, researchers may wish to use *a priori* restrictions or model selection criteria to choose the lag orders of the stationary variables in the model. As Abadir, Hadri and Tzavalis (1999) demonstrate, the inclusion of irrelevant stationary terms in a VECM may result in substantial small sample estimator bias.

The plan of the paper is as follows. The basic vector autoregressive (VAR) model and other notation are set out in section 2. The importance of an appropriate specification of deterministic terms in the VAR model is also highlighted here. In particular, unless the coefficients associated with the intercept or the linear deterministic trend are restricted to lie in the column space of the long-run multiplier matrix, the VAR model has the unsatisfactory feature that quite different deterministic behavior should be observed in the levels of the variables for differing values of the cointegrating rank. The problem of efficient conditional estimation of a VECM containing $I(1)$ exogenous variables is addressed in section 3. This section distinguishes between five different cases which are classified by the deterministic behaviour in the levels of the underlying variables; that is, Case I: zero intercept and linear trend coefficients, Case II: restricted intercept and zero linear trend coefficients, Case III: unrestricted intercept and zero linear trend coefficients, Case IV: unrestricted intercept and restricted linear trend coefficients, Case V: unrestricted intercept and unrestricted linear trend coefficients. These cases have been considered by Johansen (1995) when all the $I(1)$ variables in the VAR are treated as endogenous. Section 4 weakens the distributional assumptions of earlier sections and develops tests of cointegration rank for all five cases in sub-sections 4.1 and 4.2; relevant asymptotic critical values are provided in Tables T.1-T.5. Modifications to the tests of sub-sections 4.1 and 4.2 necessitated in the presence of exogenous variables which are integrated of order zero are briefly addressed in sub-section 4.3. Tests for the absence of an intercept or a trend in the cointegrating relations are discussed in sub-section 4.4, and particular misspecification tests concerning the intercept and lin-

¹A similar approach is taken in Harbo *et al.* (1998), an earlier version of which we became aware of after the first version of this paper [Pesaran, Shin and Smith (1997)] was completed. This revision identifies the areas of overlap between the two papers.

ear trend coefficients are developed in sub-section 4.5 together with misspecification tests for the weak exogeneity assumption. The problem of efficient conditional estimation of a VECM subject to restrictions on the short-run dynamics is considered in section 5. The subsequent two sections consider the empirical relevance of the proposed tests. Section 6 addresses the issue of the small sample performance of the proposed trace and maximum eigenvalue tests using a limited set of Monte Carlo experiments. It compares the size and power performance of the standard Johansen procedure with the new tests that take account of exogenous $I(1)$ variables as well as possible restrictions on the short-run coefficients. Section 7 presents an empirical re-examination of the validity of the Purchasing Power Parity (PPP) and the Uncovered Interest Parity (UIP) hypotheses using UK quarterly data over the period 1972(1)-1987(2) which was previously analyzed by Johansen and Juselius (1992) and Pesaran and Shin (1996). In contrast to this earlier work, foreign prices are assumed to be exogenously determined and a more satisfactory treatment of oil price changes is provided in the analysis. Section 8 concludes the paper. Proofs of results are collected in Appendix A and Appendix B describes the simulation method for the computation of the asymptotic critical values provided in Tables T.1-T.5.

2 The Treatment of Trends in VAR Models

Let $\{\mathbf{z}_t\}_{t=1}^{\infty}$ denote an m -vector random process. The data generating process (DGP) for $\{\mathbf{z}_t\}_{t=1}^{\infty}$ is the *vector autoregressive* model of order p (VAR(p)) described by

$$\Phi(L)(\mathbf{z}_t - \boldsymbol{\mu} - \boldsymbol{\gamma}t) = \mathbf{e}_t, \quad t = 1, 2, \dots, \quad (2.1)$$

where L is the lag operator, $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ are m -vectors of unknown coefficients, and the (m, m) matrix lag polynomial of order p , $\Phi(L) \equiv \mathbf{I}_m - \sum_{i=1}^p \Phi_i L^i$, comprises the unknown (m, m) coefficient matrices $\{\Phi_i\}_{i=1}^p$. For the purposes of exposition in this and the following section, the error process $\{\mathbf{e}_t\}_{t=-\infty}^{\infty}$ is assumed to be $IN(\mathbf{0}, \Omega)$, Ω positive definite. The analysis that follows is conducted given the initial values $\mathbf{Z}_0 \equiv (\mathbf{z}_{-p+1}, \dots, \mathbf{z}_0)$.

It is convenient to re-express the lag polynomial $\Phi(L)$ in a form which arises in the vector error correction model discussed in section 3; *viz.*

$$\Phi(L) \equiv -\Pi L + \Gamma(L)(1 - L). \quad (2.2)$$

In (2.2), we have defined the long-run multiplier matrix

$$\Pi \equiv -\left(\mathbf{I}_m - \sum_{i=1}^p \Phi_i\right) \quad (2.3)$$

and the short-run response matrix lag polynomial $\Gamma(L) \equiv \mathbf{I}_m - \sum_{i=1}^{p-1} \Gamma_i L^i$, $\Gamma_i = -\sum_{j=i+1}^p \Phi_j$, $i = 1, \dots, p-1$. Hence, the VAR(p) model (2.1) may be rewritten in the following form

$$\Phi(L)\mathbf{z}_t = \mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{e}_t, \quad t = 1, 2, \dots, \quad (2.4)$$

where

$$\mathbf{a}_0 \equiv -\Pi\boldsymbol{\mu} + (\Gamma + \Pi)\boldsymbol{\gamma}, \quad \mathbf{a}_1 \equiv -\Pi\boldsymbol{\gamma}, \quad (2.5)$$

and the sum of the short-run coefficient matrices Γ is given by

$$\Gamma \equiv \mathbf{I}_m - \sum_{i=1}^{p-1} \Gamma_i = -\Pi + \sum_{i=1}^p i\Phi_i. \quad (2.6)$$

The cointegration rank hypothesis is defined by

$$H_r : \text{Rank}[\Pi] = r, \quad r = 0, \dots, m, \quad (2.7)$$

where $\text{Rank}[\cdot]$ denotes the rank of $[\cdot]$. Under H_r of (2.7), we may express

$$\Pi = \alpha\beta', \quad (2.8)$$

where α and β are (m, r) matrices of full column rank. Correspondingly we may define $(m, m-r)$ matrices of full column rank $\alpha_\perp$ and $\beta_\perp$ whose columns form bases for the null spaces (kernels) of α and β respectively; in particular, $\alpha'\alpha_\perp = \mathbf{0}$ and $\beta'\beta_\perp = \mathbf{0}$.

We now adopt the following assumptions.

Assumption 2.1 *The (m, m) matrix polynomial $\Phi(z) = \mathbf{I}_m - \sum_{i=1}^p \Phi_i z^i$ is such that the roots of the determinantal equation $|\Phi(z)| = 0$ satisfy $|z| > 1$ or $z = 1$.*

Assumption 2.1 rules out the possibility that the random process $\{(\mathbf{z}_t - \boldsymbol{\mu} - \gamma t)\}_{t=1}^\infty$ admits explosive roots or seasonal unit roots except at the zero frequency.

Assumption 2.2 *The $(m-r, m-r)$ matrix $\alpha'_\perp \Gamma \beta_\perp$ is full rank.*

Under Assumption 2.1, Assumption 2.2 is a necessary and sufficient condition for the processes $\{\beta'_\perp(\mathbf{z}_t - \boldsymbol{\mu} - \gamma t)\}_{t=1}^\infty$ and $\{\beta'(\mathbf{z}_t - \boldsymbol{\mu} - \gamma t)\}_{t=1}^\infty$ to be integrated of orders one and zero respectively.² Moreover, Assumption 2.2 specifically excludes the process $\{(\mathbf{z}_t - \boldsymbol{\mu} - \gamma t)\}_{t=1}^\infty$ being integrated of order two. Together these assumptions permit the infinite order moving average representations described below. See Johansen (1991, Theorem 4.1, p.1559) and Johansen (1995, Theorem 4.2, p.49).

The differenced process $\{\Delta \mathbf{z}_t\}_{t=1}^\infty$ may be expressed under Assumptions 2.1 and 2.2 from (2.4) as the infinite vector moving average process

$$\Delta \mathbf{z}_t = \mathbf{C}(L)(\mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{e}_t) = \mathbf{b}_0 + \mathbf{b}_1 t + \mathbf{C}(L)\mathbf{e}_t, \quad t = 1, 2, \dots, \quad (2.9)$$

where $\mathbf{b}_0 \equiv \mathbf{C}\mathbf{a}_0 + \mathbf{C}^*\mathbf{a}_1$, $\mathbf{b}_1 \equiv \mathbf{C}\mathbf{a}_1$. The matrix lag polynomial $\mathbf{C}(L)$ is given by³

$$\mathbf{C}(L) \equiv \mathbf{I}_m + \sum_{j=1}^{\infty} \mathbf{C}_j L^j = \mathbf{C} + (1-L)\mathbf{C}^*(L), \quad \mathbf{C}^*(L) \equiv \sum_{j=0}^{\infty} \mathbf{C}_j^* L^j,$$

²See Johansen (1995, Definitions 3.2 and 3.3, p.35). That is, defining the difference operator $\Delta \equiv (1-L)$, the processes $\{\beta'_\perp[\Delta(\mathbf{z}_t - \boldsymbol{\mu} - \gamma t)]\}_{t=1}^\infty$, and $\{\beta'(\mathbf{z}_t - \boldsymbol{\mu} - \gamma t)\}_{t=1}^\infty$ admit stationary and invertible ARMA representations; see also Engle and Granger (1987, Definition, p.252).

³The matrices $\{\mathbf{C}_i\}$ can be obtained from the recursions $\mathbf{C}_i = \sum_{j=1}^p \mathbf{C}_{i-j}\Phi_j$, $i > 1$, $\mathbf{C}_0 = \mathbf{I}_m$, $\mathbf{C}_1 = -(\mathbf{I}_m - \Phi_1)$, defining $\mathbf{C}_i = \mathbf{0}$ for $i < 0$. Similarly, for the matrices $\{\mathbf{C}_j^*\}$, $\mathbf{C}_j^* = \mathbf{C}_j + \mathbf{C}_{j-1}^*$, $j > 0$, $\mathbf{C}_0^* = \mathbf{I}_m - \mathbf{C}$.

$$\mathbf{C} \equiv \sum_{j=0}^{\infty} \mathbf{C}_j, \mathbf{C}^* \equiv \sum_{j=0}^{\infty} \mathbf{C}_j^*. \quad (2.10)$$

Now, as $\mathbf{C}(L)\Phi(L) = \Phi(L)\mathbf{C}(L) = (1-L)\mathbf{I}_m$, $\Pi\mathbf{C} = \mathbf{0}$ and $\mathbf{C}\Pi = \mathbf{0}$; in particular, $\mathbf{C} = \boldsymbol{\beta}_{\perp}(\boldsymbol{\alpha}'_{\perp}\Gamma\boldsymbol{\beta}_{\perp})^{-1}\boldsymbol{\alpha}'_{\perp}$. Re-expressing (2.9) in levels,

$$\mathbf{z}_t = \mathbf{z}_0 + \mathbf{b}_0 t + \mathbf{b}_1 \frac{t(t+1)}{2} + \mathbf{C}\mathbf{s}_t + \mathbf{C}^*(L)(\mathbf{e}_t - \mathbf{e}_0), \quad (2.11)$$

where the partial sum $\mathbf{s}_t \equiv \sum_{s=1}^t \mathbf{e}_s$, $t = 1, 2, \dots$

Adopting the VAR(p) formulation (2.1) rather than the more usual (2.4), in which $\mathbf{a}_0$ and $\mathbf{a}_1$ are unrestricted, reveals immediately from (2.11) that the restrictions (2.5) on $\mathbf{a}_1$ induce $\mathbf{b}_1 = \mathbf{0}$ and ensure that the nature of the deterministic trending behavior of the level process $\{\mathbf{z}_t\}_{t=1}^{\infty}$ remains invariant to the rank r of the long-run multiplier matrix Π ; that is, it is *linear*. Hence, the infinite moving average representation for the level process $\{\mathbf{z}_t\}_{t=1}^{\infty}$ is⁴

$$\mathbf{z}_t = \boldsymbol{\mu} + \boldsymbol{\gamma}t + \mathbf{C}\mathbf{s}_t + \mathbf{C}^*(L)\mathbf{e}_t, \quad (2.12)$$

where we have used the initialization $\mathbf{z}_0 \equiv \boldsymbol{\mu} + \mathbf{C}^*(L)\mathbf{e}_0$.⁵ See also Johansen (1994) and Johansen (1995, Section 5.7, pp.80-84).⁶ If, however, $\mathbf{a}_1$ were not subject to the restrictions (2.5), the quadratic trend term would be present in the level equation (2.11) apart from in the full rank stationary case $H_m : \text{Rank}[\Pi] = m$ or $\mathbf{C} = \mathbf{0}$. However, $\mathbf{b}_1$ would be unconstrained under the null hypothesis of *no* cointegration; that is, $H_0 : \text{Rank}[\Pi] = 0$, and $\mathbf{C}$ full rank. In the general case $H_r : \text{Rank}[\Pi] = r$ of (2.7), this would imply the unsatisfactory conclusion that quite different deterministic trending behavior should be observed in the levels process $\{\mathbf{z}_t\}_{t=1}^{\infty}$ for differing values of the cointegrating rank r , the number of independent quadratic deterministic trends, $m - r$, decreasing as r increases.

The above analysis further reveals that because cointegration is only concerned with the elimination of stochastic trends it does not therefore rule out the possibility of deterministic trends in the cointegrating relations. Pre-multiplying both sides of (2.12) by the cointegrating matrix $\boldsymbol{\beta}'$, we obtain the cointegrating relations

$$\boldsymbol{\beta}'\mathbf{z}_t = \boldsymbol{\beta}'\boldsymbol{\mu} + (\boldsymbol{\beta}'\boldsymbol{\gamma})t + \boldsymbol{\beta}'\mathbf{C}^*(L)\mathbf{e}_t, \quad t = 1, 2, \dots, \quad (2.13)$$

which are trend-stationary. In general, the *co-trending* restriction [Park (1992)] $\boldsymbol{\beta}'\boldsymbol{\gamma} = \mathbf{0}$ if and only if $\mathbf{a}_1 = \mathbf{0}$. In this case, the representation for the VAR(p) model, (2.4), and the cointegrating regression, (2.13), will contain no deterministic trends. However, the restriction $\boldsymbol{\beta}'\boldsymbol{\gamma} = \mathbf{0}$ may not prove to be satisfactory in practice. Therefore, it is important that the composite $\boldsymbol{\beta}'\boldsymbol{\gamma}$ in (2.13) or, equivalently, $\mathbf{a}_1 = -\Pi\boldsymbol{\gamma}$ in (2.4) is estimated along with the other parameters of the model and the co-trending restriction tested; see sub-section 4.4.

⁴From (2.2), as $\mathbf{C}(L)\Phi(L) = (1-L)\mathbf{I}_m$ and, in particular, $\mathbf{C}\Pi = \mathbf{0}$, $\mathbf{C}\Gamma - \mathbf{C}^*\Pi = \mathbf{I}_m$.

⁵Of course, the levels equation (2.11) could also have been obtained directly from (2.1) by noting $\Delta(\mathbf{z}_t - \boldsymbol{\mu} - \boldsymbol{\gamma}t) = \mathbf{C}(L)\mathbf{e}_t$, $t = 1, 2, \dots$

⁶As the cointegration rank hypothesis (2.7) may be alternatively and equivalently expressed as $H'_r : \text{Rank}[\mathbf{C}] = m - r$, $r = 0, \dots, m$, it is interesting to note that, from (2.4) and (2.5), there are r linearly independent deterministic trends and, from (2.11), $m - r$ independent stochastic trends $\mathbf{C}\mathbf{s}_t$, the combined total of which is m .

3 Efficient Estimation of a Structural Error Correction Model

We now partition the m -vector of random variables $\mathbf{z}_t$ into the n -vector $\mathbf{y}_t$ and the k -vector $\mathbf{x}_t$, where $k \equiv m - n$; that is, $\mathbf{z}_t = (\mathbf{y}'_t, \mathbf{x}'_t)'$, $t = 1, 2, \dots$. The primary concern of this paper is the structural modelling of the vector $\mathbf{y}_t$ conditional on its past, $\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots$, and current and past values of the vector of random variables $\mathbf{x}_t, \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \dots$, $t = 1, 2, \dots$. The assumption of section 2 concerning the error process $\{\mathbf{e}_t\}_{t=-\infty}^{\infty}$, $\mathbf{e}_t \sim IN(\mathbf{0}, \Omega)$, $t = 0, \pm 1, \dots$, where Ω is positive definite, permits a likelihood analysis and the conditional model interpretation given below; see Harbo *et al.* (1998) for a similar development.

It is convenient to re-express the VAR(p) of (2.4) as the *vector error correction* model (VECM):

$$\Delta \mathbf{z}_t = \mathbf{a}_0 + \mathbf{a}_1 t + \sum_{i=1}^{p-1} \Gamma_i \Delta \mathbf{z}_{t-i} + \Pi \mathbf{z}_{t-1} + \mathbf{e}_t, t = 1, 2, \dots, \quad (3.1)$$

where the short-run response matrices $\{\Gamma_i\}_{i=1}^{p-1}$ and the long-run multiplier matrix Π are defined below (2.2).

By partitioning the error term $\mathbf{e}_t$ conformably with $\mathbf{z}_t = (\mathbf{y}'_t, \mathbf{x}'_t)'$ as $\mathbf{e}_t = (\mathbf{e}'_{yt}, \mathbf{e}'_{xt})'$ and its variance matrix as

$$\Omega = \begin{pmatrix} \Omega_{yy} & \Omega_{yx} \\ \Omega_{xy} & \Omega_{xx} \end{pmatrix}$$

we are able to express $\mathbf{e}_{yt}$ conditionally in terms of $\mathbf{e}_{xt}$ as

$$\mathbf{e}_{yt} = \Omega_{yx} \Omega_{xx}^{-1} \mathbf{e}_{xt} + \mathbf{u}_t, \quad (3.2)$$

where $\mathbf{u}_t \sim IN(\mathbf{0}, \Omega_{uu})$, $\Omega_{uu} \equiv \Omega_{yy} - \Omega_{yx} \Omega_{xx}^{-1} \Omega_{xy}$ and $\mathbf{u}_t$ is independent of $\mathbf{e}_{xt}$. Substitution of (3.2) into (3.1) together with a similar partitioning of the parameter vectors and matrices $\mathbf{a}_0 = (\mathbf{a}'_{y0}, \mathbf{a}'_{x0})'$, $\mathbf{a}_1 = (\mathbf{a}'_{y1}, \mathbf{a}'_{x1})'$, $\Pi = (\Pi'_y, \Pi'_x)'$, $\Gamma = (\Gamma'_y, \Gamma'_x)'$, $\Gamma_i = (\Gamma'_{yi}, \Gamma'_{xi})'$, $i = 1, \dots, p-1$, provides a conditional model for $\Delta \mathbf{y}_t$ in terms of $\mathbf{z}_{t-1}, \Delta \mathbf{x}_t, \Delta \mathbf{z}_{t-1}, \Delta \mathbf{z}_{t-2}, \dots$; *viz.*

$$\Delta \mathbf{y}_t = \mathbf{c}_0 + \mathbf{c}_1 t + \Lambda \Delta \mathbf{x}_t + \sum_{i=1}^{p-1} \Psi_i \Delta \mathbf{z}_{t-i} + \Pi_{yy.x} \mathbf{z}_{t-1} + \mathbf{u}_t, \quad (3.3)$$

$t = 1, 2, \dots$, where $\mathbf{c}_0 \equiv \mathbf{a}_{y0} - \Omega_{yx} \Omega_{xx}^{-1} \mathbf{a}_{x0}$, $\mathbf{c}_1 \equiv \mathbf{a}_{y1} - \Omega_{yx} \Omega_{xx}^{-1} \mathbf{a}_{x1}$, $\Lambda = \Omega_{yx} \Omega_{xx}^{-1}$, $\Psi_i \equiv \Gamma_{yi} - \Omega_{yx} \Omega_{xx}^{-1} \Gamma_{xi}$, $i = 1, \dots, p-1$, and $\Pi_{yy.x} \equiv \Pi_y - \Omega_{yx} \Omega_{xx}^{-1} \Pi_x$.

Following Johansen (1992) and Boswijk (1992, Ch. 3), we assume that the process $\{\mathbf{x}_t\}_{t=1}^{\infty}$ is weakly exogenous with respect to the matrix of long-run multiplier parameters Π ; *viz.*⁷

Assumption 3.1 $\Pi_x = \mathbf{0}$.

Therefore,

$$\Pi_{yy.x} = \Pi_y. \quad (3.4)$$

⁷Specification tests for Assumption 3.1 are presented in sub-section 4.5 below.

Consequently, under Assumption 3.1, from (3.1) and (3.3), the system of equations are rendered as

$$\Delta \mathbf{y}_t = \mathbf{c}_0 + \mathbf{c}_1 t + \Lambda \Delta \mathbf{x}_t + \sum_{i=1}^{p-1} \Psi_i \Delta \mathbf{z}_{t-i} + \Pi_y \mathbf{z}_{t-1} + \mathbf{u}_t, \quad (3.5)$$

$$\Delta \mathbf{x}_t = \mathbf{a}_{x0} + \sum_{i=1}^{p-1} \Gamma_{xi} \Delta \mathbf{z}_{t-i} + \mathbf{e}_{xt}, \quad (3.6)$$

$t = 1, 2, \dots$, where now, because $\mathbf{a}_{x1} = \mathbf{0}$, $\mathbf{c}_1 \equiv \mathbf{a}_{y1}$ and the relations (2.5) are modified to

$$\mathbf{c}_0 = -\Pi_y \boldsymbol{\mu} + (\Gamma_y - \Omega_{yx} \Omega_{xx}^{-1} \Gamma_x + \Pi_y) \boldsymbol{\gamma}, \quad \mathbf{c}_1 = -\Pi_y \boldsymbol{\gamma}. \quad (3.7)$$

The restriction $\Pi_x = \mathbf{0}$ of Assumption 3.1 implies that the elements of the vector process $\{\mathbf{x}_t\}_{t=1}^{\infty}$ are not cointegrated among themselves as is evident from (3.6). Moreover, the information available from the differenced VAR($p-1$) model (3.6) for $\{\mathbf{x}_t\}_{t=1}^{\infty}$ is redundant for efficient conditional estimation and inference concerning the long-run parameters Π_y as well as the deterministic and short-run parameters $\mathbf{c}_0$, $\mathbf{c}_1$, Λ and Ψ_i , $i = 1, \dots, p-1$, of (3.5). Furthermore, under Assumption 3.1, we may regard $\{\mathbf{x}_t\}_{t=1}^{\infty}$ as *long run forcing* for $\{\mathbf{y}_t\}_{t=1}^{\infty}$; see Granger and Lin (1995).⁸

The cointegration rank hypothesis (2.7) is therefore restated in the context of (3.5) as

$$H_r : \text{Rank}[\Pi_y] = r, \quad r = 0, \dots, n. \quad (3.8)$$

We differentiate between and delineate five cases of interest; *viz.*

- **Case I:** (No Intercepts; No Trends.) $\mathbf{c}_0 = \mathbf{0}$ and $\mathbf{c}_1 = \mathbf{0}$. That is, $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\gamma} = \mathbf{0}$. Hence, the structural VECM (3.5) becomes

$$\Delta \mathbf{y}_t = \Lambda \Delta \mathbf{x}_t + \sum_{i=1}^{p-1} \Psi_i \Delta \mathbf{z}_{t-i} + \Pi_y \mathbf{z}_{t-1} + \mathbf{u}_t. \quad (3.9)$$

- **Case II:** (Restricted Intercepts; No Trends.) $\mathbf{c}_0 = -\Pi_y \boldsymbol{\mu}$ and $\mathbf{c}_1 = \mathbf{0}$. Here, $\boldsymbol{\gamma} = \mathbf{0}$. The structural VECM (3.5) is

$$\Delta \mathbf{y}_t = (-\Pi_y \boldsymbol{\mu}) + \Lambda \Delta \mathbf{x}_t + \sum_{i=1}^{p-1} \Psi_i \Delta \mathbf{z}_{t-i} + \Pi_y \mathbf{z}_{t-1} + \mathbf{u}_t. \quad (3.10)$$

- **Case III:** (Unrestricted Intercepts; No Trends.) $\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 = \mathbf{0}$. Again, $\boldsymbol{\gamma} = \mathbf{0}$. In this case, the intercept restriction $\mathbf{c}_0 = -\Pi_y \boldsymbol{\mu}$ is ignored and the structural VECM estimated is

$$\Delta \mathbf{y}_t = \mathbf{c}_0 + \Lambda \Delta \mathbf{x}_t + \sum_{i=1}^{p-1} \Psi_i \Delta \mathbf{z}_{t-i} + \Pi_y \mathbf{z}_{t-1} + \mathbf{u}_t. \quad (3.11)$$

⁸Note that this restriction does not preclude $\{\mathbf{y}_t\}_{t=1}^{\infty}$ being *Granger-causal* for $\{\mathbf{x}_t\}_{t=1}^{\infty}$ in the *short run*.

- **Case IV:** (Unrestricted Intercepts; Restricted Trends.) $\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 = -\Pi_y \boldsymbol{\gamma}$. Thus

$$\Delta \mathbf{y}_t = \mathbf{c}_0 + (-\Pi_y \boldsymbol{\gamma})t + \Lambda \Delta \mathbf{x}_t + \sum_{i=1}^{p-1} \Psi_i \Delta \mathbf{z}_{t-i} + \Pi_y \mathbf{z}_{t-1} + \mathbf{u}_t. \quad (3.12)$$

- **Case V:** (Unrestricted Intercepts; Unrestricted Trends.) $\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 \neq \mathbf{0}$. Here, the deterministic trend restriction $\mathbf{c}_1 = -\Pi_y \boldsymbol{\gamma}$ is ignored and the structural VECM estimated is

$$\Delta \mathbf{y}_t = \mathbf{c}_0 + \mathbf{c}_1 t + \Lambda \Delta \mathbf{x}_t + \sum_{i=1}^{p-1} \Psi_i \Delta \mathbf{z}_{t-i} + \Pi_y \mathbf{z}_{t-1} + \mathbf{u}_t. \quad (3.13)$$

It should be emphasised that the DGPs for Cases II and III are identical as are those for Cases IV and V. However, as in the test for a unit root proposed by Dickey and Fuller (1979) compared with that of Dickey and Fuller (1981) for univariate models, estimation and hypothesis testing in Cases III and V proceed ignoring the constraints linking respectively the intercept and trend coefficient vectors, $\mathbf{c}_0$ and $\mathbf{c}_1$, to the parameter matrix Π_y whereas Cases II and IV fully incorporate the restrictions in (3.7).

We concentrate on Case IV, that is, (3.12), which may be simply revised to yield the remainder. Firstly, note that under (3.8) we may express

$$\Pi_y = \boldsymbol{\alpha}_y \boldsymbol{\beta}', \quad (3.14)$$

where the (n, r) loadings matrix $\boldsymbol{\alpha}_y$ and the (m, r) matrix of cointegrating vectors $\boldsymbol{\beta}$ are each full column rank and identified up to an arbitrary (r, r) non-singular matrix.⁹

The estimation of the cointegrating matrix $\boldsymbol{\beta}$ has been the subject of much intensive research for the case in which $n = m$ or $k = 0$, that is, *no* exogenous variables. See, for example, Engle and Granger (1987), Johansen (1988, 1991, 1995), Phillips (1991), Ahn and Reinsel (1990), Phillips and Hansen (1990), Park (1992) and Pesaran and Shin (1999). More recently, Harbo *et al.* (1998) have also considered the cointegration rank hypothesis (3.8) in the context of a conditional model; that is, when $k > 0$.

The procedure elucidated below is an adaptation for the case in which $k > 0$ which gives the results for $n = m$ as a special case. Rewrite (3.12) as

$$\Delta \mathbf{y}_t = \mathbf{c}_0 + \Lambda \Delta \mathbf{x}_t + \sum_{i=1}^{p-1} \Psi_i \Delta \mathbf{z}_{t-i} + \Pi_{y*} \mathbf{z}_{t-1}^* + \mathbf{u}_t, \quad t = 1, 2, \dots, \quad (3.15)$$

where $\mathbf{z}_{t-1}^* = (t, \mathbf{z}_{t-1}')'$, and $\Pi_{y*} = \Pi_y(-\boldsymbol{\gamma}, \mathbf{I}_m)$. Note that $\text{Rank}[\Pi_{y*}] = \text{Rank}[\Pi_y]$ and, thus, from (3.14)¹⁰

$$\Pi_{y*} = \boldsymbol{\alpha}_y \boldsymbol{\beta}_*', \quad (3.16)$$

⁹That is, $(\boldsymbol{\alpha}_y \mathbf{K}^{-1})(\mathbf{K} \boldsymbol{\beta}') = \boldsymbol{\alpha}_y \boldsymbol{\beta}'$ for any (r, r) non-singular matrix $\mathbf{K}$.

¹⁰The trend parameter vector $\boldsymbol{\gamma}$ is no longer identified in system (3.15) and (3.6). Only $\boldsymbol{\beta}' \boldsymbol{\gamma}$ and $\mathbf{a}_{x0} = \Gamma_x \boldsymbol{\gamma}$ may be identified.

where

$$\beta_* = \begin{pmatrix} -\gamma' \\ \mathbf{I}_m \end{pmatrix} \beta.$$

Consequently, we may therefore restate the cointegration rank hypothesis (3.14) as

$$H_r : \text{Rank}[\Pi_{y*}] = r, \quad r = 0, \dots, n. \quad (3.17)$$

Hence, we may adapt the reduced rank techniques of Johansen (1995) to estimate the revised system (3.15); see also Boswijk (1995) and Harbo *et al.* (1998).

If T observations are available, stacking the structural VECM (3.15) results in

$$\Delta \mathbf{Y} = \mathbf{c}_0 \boldsymbol{\iota}'_T + \Psi \Delta \mathbf{Z}_- + \Pi_{y*} \mathbf{Z}_{-1}^* + \mathbf{U}, \quad (3.18)$$

where $\Delta \mathbf{Y} \equiv (\Delta \mathbf{y}_1, \dots, \Delta \mathbf{y}_T)$, $\boldsymbol{\iota}_T$ is a T -vector of ones, $\Delta \mathbf{X} \equiv (\Delta \mathbf{x}_1, \dots, \Delta \mathbf{x}_T)$, $\Delta \mathbf{Z}_{-i} \equiv (\Delta \mathbf{z}_{1-i}, \dots, \Delta \mathbf{z}_{T-i})$, $i = 1, \dots, p-1$, $\Psi \equiv (\Lambda, \Psi_1, \dots, \Psi_{p-1})$, $\Delta \mathbf{Z}_- \equiv (\Delta \mathbf{X}', \Delta \mathbf{Z}'_{-1}, \dots, \Delta \mathbf{Z}'_{1-p})'$, $\mathbf{Z}_{-1}^* \equiv (\boldsymbol{\tau}_T, \mathbf{Z}'_{-1})'$, $\boldsymbol{\tau}_T \equiv (1, \dots, T)'$, $\mathbf{Z}_{-1} \equiv (\mathbf{z}_0, \dots, \mathbf{z}_{T-1})$ and $\mathbf{U} \equiv (\mathbf{u}_1, \dots, \mathbf{u}_T)$.

The log-likelihood function of the structural VECM model (3.18) is given by

$$\ell_T(\boldsymbol{\psi}; r) = -\frac{nT}{2} \ln 2\pi - \frac{T}{2} \ln |\Omega_{uu}| - \frac{1}{2} \text{Trace}(\Omega_{uu}^{-1} \mathbf{U} \mathbf{U}'), \quad (3.19)$$

where the parameter vector $\boldsymbol{\psi}$ collects together the unknown parameters in Ω_{uu} , $\mathbf{c}_0$, Ψ and Π_{y*} . Successively concentrating out Ω_{uu} , $\mathbf{c}_0$ and Ψ , and $\boldsymbol{\alpha}_y$ in (3.19) results in the concentrated log-likelihood function

$$\ell_T^c(\boldsymbol{\beta}_*; r) = -\frac{nT}{2} (1 + \ln 2\pi) - \frac{T}{2} \ln \left| T^{-1} \Delta \hat{\mathbf{Y}} \left(\mathbf{I}_T - \hat{\mathbf{Z}}_{-1}^* \boldsymbol{\beta}_* \left(\boldsymbol{\beta}_*' \hat{\mathbf{Z}}_{-1}^* \hat{\mathbf{Z}}_{-1}^{*'} \boldsymbol{\beta}_* \right)^{-1} \boldsymbol{\beta}_*' \hat{\mathbf{Z}}_{-1}^* \right) \Delta \hat{\mathbf{Y}}' \right|, \quad (3.20)$$

where $\Delta \hat{\mathbf{Y}}$ and $\hat{\mathbf{Z}}_{-1}^*$ are respectively the OLS residuals from regressions of $\Delta \mathbf{Y}$ and $\mathbf{Z}_{-1}^*$ on $(\boldsymbol{\iota}_T, \Delta \mathbf{Z}_-)$.¹¹ Defining the sample moment matrices

$$\mathbf{S}_{YY} \equiv T^{-1} \Delta \hat{\mathbf{Y}} \Delta \hat{\mathbf{Y}}', \quad \mathbf{S}_{YZ} \equiv T^{-1} \Delta \hat{\mathbf{Y}} \hat{\mathbf{Z}}_{-1}^{*'}, \quad \mathbf{S}_{ZZ} \equiv T^{-1} \hat{\mathbf{Z}}_{-1}^* \hat{\mathbf{Z}}_{-1}^{*'}, \quad (3.21)$$

the maximization of the concentrated log-likelihood function $\ell_T^c(\boldsymbol{\beta}_*; r)$ of (3.20) reduces to the minimization of

$$\left| \mathbf{S}_{YY} - \mathbf{S}_{YZ} \boldsymbol{\beta}_* (\boldsymbol{\beta}_*' \mathbf{S}_{ZZ} \boldsymbol{\beta}_*)^{-1} \boldsymbol{\beta}_*' \mathbf{S}_{YZ} \right| = \frac{|\mathbf{S}_{YY}| |\boldsymbol{\beta}_*' (\mathbf{S}_{ZZ} - \mathbf{S}_{YZ} \mathbf{S}_{YY}^{-1} \mathbf{S}_{YZ}) \boldsymbol{\beta}_*|}{|\boldsymbol{\beta}_*' \mathbf{S}_{ZZ} \boldsymbol{\beta}_*|}$$

with respect to $\boldsymbol{\beta}_*$. The solution $\hat{\boldsymbol{\beta}}_*$ to this minimization problem, that is, the maximum likelihood (ML) estimator for $\boldsymbol{\beta}_*$, is given by the eigenvectors corresponding to the r largest eigenvalues $\hat{\lambda}_1 > \dots > \hat{\lambda}_r > 0$ of

$$\left| \hat{\lambda} \mathbf{S}_{ZZ} - \mathbf{S}_{YZ} \mathbf{S}_{YY}^{-1} \mathbf{S}_{YZ} \right| = 0; \quad (3.22)$$

¹¹The corresponding estimators are given by $\hat{\Omega}_{uu} = T^{-1} \hat{\mathbf{U}} \hat{\mathbf{U}}'$, where $\hat{\mathbf{U}}$ is defined via (3.18) and is a function of the unknown parameter matrices $\mathbf{c}_0$, Ψ and Π_{y*} , $\hat{\mathbf{c}}_0(\Pi_{y*}) = (\Delta \mathbf{Y} - \hat{\Psi}(\Pi_{y*}) \Delta \mathbf{Z}_- - \Pi_{y*} \mathbf{Z}_{-1}^*) \boldsymbol{\iota}_T (\boldsymbol{\iota}_T' \boldsymbol{\iota}_T)^{-1}$, $\hat{\Psi}(\Pi_{y*}) = (\Delta \mathbf{Y} - \Pi_{y*} \mathbf{Z}_{-1}^*) \bar{\mathbf{P}}_\iota \Delta \mathbf{Z}_- (\Delta \mathbf{Z}_- \bar{\mathbf{P}}_\iota \Delta \mathbf{Z}_-)^{-1}$, where $\bar{\mathbf{P}}_\iota \equiv \mathbf{I}_T - \boldsymbol{\iota}_T (\boldsymbol{\iota}_T' \boldsymbol{\iota}_T)^{-1} \boldsymbol{\iota}_T'$, and $\hat{\boldsymbol{\alpha}}_y(\boldsymbol{\beta}_*) = \Delta \hat{\mathbf{Y}} \hat{\mathbf{Z}}_{-1}^{*'} \boldsymbol{\beta}_* (\boldsymbol{\beta}_*' \hat{\mathbf{Z}}_{-1}^* \hat{\mathbf{Z}}_{-1}^{*'} \boldsymbol{\beta}_*)^{-1}$.

cf. Johansen (1991, pp.1553-1554). The ML estimator $\hat{\beta}_*$ is identified up to post-multiplication by an (r, r) non-singular matrix; that is, r^2 just-identifying restrictions on β_* are required for exact identification.¹² The resultant maximized concentrated log-likelihood function $\ell_T^c(\beta_*; r)$ at $\hat{\beta}_*$ of (3.20) is

$$\ell_T^c(r) = -\frac{nT}{2}(1 + \ln 2\pi) - \frac{T}{2} \ln |\mathbf{S}_{YY}| - \frac{T}{2} \sum_{i=1}^r \ln(1 - \hat{\lambda}_i). \quad (3.23)$$

Note that the maximized value of the log-likelihood $\ell_T^c(r)$ is only a function of the cointegration rank r (and n and k) through the eigenvalues $\{\hat{\lambda}_i\}_{i=1}^r$ defined by (3.22). See also Harbo *et al.* (1998, Section 2).

For the other four cases of interest, we need to modify our definitions of $\Delta\hat{\mathbf{Y}}$ and $\hat{\mathbf{Z}}_{-1}^*$ and, consequently, the sample moment matrices $\mathbf{S}_{YY}$, $\mathbf{S}_{YZ}$ and $\mathbf{S}_{ZZ}$ given by (3.21). We state the required definitions below.

- **Case I:** $\mathbf{c}_0 = \mathbf{0}$ and $\mathbf{c}_1 = \mathbf{0}$. $\Delta\hat{\mathbf{Y}}$ and $\hat{\mathbf{Z}}_{-1}^*$ are the OLS residuals from the regression of $\Delta\mathbf{Y}$ and $\mathbf{Z}_{-1}$ on $\Delta\mathbf{Z}_{-}$.
- **Case II:** $\mathbf{c}_0 = -\Pi_y \boldsymbol{\mu}$ and $\mathbf{c}_1 = \mathbf{0}$. $\Delta\hat{\mathbf{Y}}$ and $\hat{\mathbf{Z}}_{-1}^*$ are the OLS residuals from the regression of $\Delta\mathbf{Y}$ and $\mathbf{Z}_{-1}^*$ on $\Delta\mathbf{Z}_{-}$, where $\mathbf{Z}_{-1}^* = (\boldsymbol{\nu}_T, \mathbf{Z}'_{-1})'$.
- **Case III:** $\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 = \mathbf{0}$. $\Delta\hat{\mathbf{Y}}$ and $\hat{\mathbf{Z}}_{-1}^*$ are the OLS residuals from the regression of $\Delta\mathbf{Y}$ and $\mathbf{Z}_{-1}$ on $(\boldsymbol{\nu}_T, \Delta\mathbf{Z}'_{-})'$.
- **Case IV:** $\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 = -\Pi_y \boldsymbol{\gamma}$. $\Delta\hat{\mathbf{Y}}$ and $\hat{\mathbf{Z}}_{-1}^*$ are the OLS residuals from the regression of $\Delta\mathbf{Y}$ and $\mathbf{Z}_{-1}^*$ on $(\boldsymbol{\nu}_T, \Delta\mathbf{Z}'_{-})'$, where $\mathbf{Z}_{-1}^* = (\boldsymbol{\tau}_T, \mathbf{Z}'_{-1})'$.
- **Case V:** $\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 \neq \mathbf{0}$. $\Delta\hat{\mathbf{Y}}$ and $\hat{\mathbf{Z}}_{-1}^*$ are the OLS residuals from the regression of $\Delta\mathbf{Y}$ and $\mathbf{Z}_{-1}^*$ on $(\boldsymbol{\nu}_T, \boldsymbol{\tau}_T, \Delta\mathbf{Z}'_{-})'$.

4 Structural Tests for Cointegration and Tests of Specification

Our interest in this section is five-fold. Firstly, sub-section 4.1 addresses testing the null hypothesis of cointegration rank r , H_r of (3.8), against the alternative hypothesis

$$H_{r+1} : \text{Rank}[\Pi_y] = r + 1, \quad r = 0, \dots, n - 1,$$

in the structural VECM (3.5). Secondly, sub-section 4.2 presents a test of the null hypothesis of cointegration rank r , H_r above, $r = 0, \dots, n - 1$, against the alternative hypothesis of stationarity; that is

$$H_n : \text{Rank}[\Pi_y] = n.$$

¹²Pesaran and Shin (1999) provide a comprehensive treatment of the imposition of exactly- and over-identifying (non-linear) restrictions on β when $n = m$ and, thus, $k = 0$. In principle, their approach may be adapted for our problem; see section 7.

Tables T.1-T.5 of Appendix B provide the relevant asymptotic critical values. Thirdly, sub-section 4.3 discusses testing H_r of (3.8) in the presence of weakly exogenous explanatory variables which are integrated of order *zero*. Fourthly, testing whether an intercept should be present in Case II, that is, $\mathbf{c}_0 = \mathbf{0}$ or $\beta'\mu = \mathbf{0}$, or whether a trend should be present in Case IV, that is, the co-trending restriction $\mathbf{c}_1 = \mathbf{0}$ or $\beta'\gamma = \mathbf{0}$, is considered in sub-section 4.4. Finally, sub-section 4.5 is concerned with providing specification tests for the various assumptions embodied in our approach. Proofs of the results in this section may be found in Appendix A.¹³

We now weaken the independent normal distributional assumption of sections 2 and 3 on the error process $\{\mathbf{e}_t\}_{t=-\infty}^{\infty}$ in (2.1) and, hence, on the structural error process $\{\mathbf{u}_t\}_{t=-\infty}^{\infty}$ in (3.2).

Assumption 4.1 *The error process $\{\mathbf{e}_t\}_{t=-\infty}^{\infty}$ is such that:*

- (a) (i) $E\{\mathbf{e}_t | \{\mathbf{z}_{t-i}\}_{i=1}^{t-1}, \mathbf{Z}_0\} = \mathbf{0}$, (ii) $\text{var}\{\mathbf{e}_t | \{\mathbf{z}_{t-i}\}_{i=1}^{t-1}, \mathbf{Z}_0\} = \Omega$, Ω positive definite;
- (b) (i) $E\{\mathbf{u}_t | \mathbf{x}_t, \{\mathbf{z}_{t-i}\}_{i=1}^{t-1}, \mathbf{Z}_0\} = \mathbf{0}$, (ii) $\text{var}\{\mathbf{u}_t | \mathbf{x}_t, \{\mathbf{z}_{t-i}\}_{i=1}^{t-1}, \mathbf{Z}_0\} = \Omega_{uu}$, where $\mathbf{u}_t \equiv \mathbf{e}_{yt} - \Omega_{yx}\Omega_{xx}^{-1}\mathbf{e}_{xt}$ and $\Omega_{uu} \equiv \Omega_{yy} - \Omega_{yx}\Omega_{xx}^{-1}\Omega_{xy}$;
- (c) $\sup_t E\{\|\mathbf{e}_t\|^s\} < \infty$ for some $s > 2$.

Assumption 4.1(a) states that the error process $\{\mathbf{e}_t\}_{t=-\infty}^{\infty}$ is a martingale difference sequence with constant conditional variance; hence, $\{\mathbf{e}_t\}_{t=-\infty}^{\infty}$ is an uncorrelated process. Therefore, the VECM (3.1) represents a conditional model for $\Delta\mathbf{z}_t$ given $\{\Delta\mathbf{z}_{t-i}\}_{i=1}^{p-1}$ and $\mathbf{z}_{t-1}$, $t = 1, 2, \dots$. Assumption 4.1(b) is a linear conditional mean condition; that is, under Assumption 4.1(b)(i), $E\{\mathbf{e}_{yt} | \mathbf{x}_t, \{\mathbf{z}_{t-i}\}_{i=1}^{t-1}, \mathbf{Z}_0\} = \Omega_{yx}\Omega_{xx}^{-1}\mathbf{e}_{xt}$ which, together with Assumption 4.1(b)(ii), also ensures that $\text{var}\{\mathbf{e}_{yt} | \mathbf{x}_t, \{\mathbf{z}_{t-i}\}_{i=1}^{t-1}, \mathbf{Z}_0\} = \Omega_{uu}$. Therefore, under this assumption, (3.15) can still be interpreted as a conditional model for $\Delta\mathbf{y}_t$ given $\Delta\mathbf{x}_t$, $\{\Delta\mathbf{z}_{t-i}\}_{i=1}^{p-1}$ and $\mathbf{z}_{t-1}$, $t = 1, 2, \dots$. Hence, (3.15) remains appropriate for conditional inference. Moreover, the error process $\{\mathbf{u}_t\}_{t=-\infty}^{\infty}$ is also a martingale difference process with constant conditional variance and is uncorrelated with the $\{\mathbf{e}_{xt}\}_{t=-\infty}^{\infty}$ process. Thus, Assumptions 4.1(a)(ii) and 4.1(b)(ii) rule out any conditional heteroskedasticity. Assumption 4.1(c) is quite standard and, together with Assumption 4.1(a), is required for the multivariate invariance principle stated in (4.1) below; see Phillips and Solo (1992, Theorem 3.15(a), p.983). Assumption 4.1(b) together with Assumption 4.1(c) implies the multivariate invariance principle (4.2) below. Assumption 4.1(c) embodies a slight strengthening of that in Phillips and Durlauf (1986, Theorem 2.1(d), p.475) which together with Assumption 4.1(a) firstly allows an invariance principle to be stated in terms of the $\{\mathbf{z}_t\}_{t=1}^{\infty}$ process itself as the $VAR(p)$ form (2.1) together with Assumptions 2.1 and 2.2 yields $\sum_{j=0}^{\infty} j \|\mathbf{C}_j\| < \infty$, where $\|\mathbf{A}\| \equiv [\text{tr}(\mathbf{A}'\mathbf{A})]^{1/2}$; see Phillips and Solo (1992, Theorem 3.15(b), p.983).¹⁴ Secondly, terms involving stationary components are asymptotically negligible relative to those involving components integrated of order one. See Phillips and Solo (1992, Theorem 3.16 and Remark 3.17(iii), p.983). Note

¹³See Harbo *et al.* (1998, Theorem 1 and Appendix) for a statement and proof of Theorem 4.2 below for Cases I, II and IV under the distributional assumption of sections 2 and 3. This analysis may be straightforwardly adapted for Theorem 4.1 below under the same assumption. Case III which corresponds to their $c = 0$ is stated in Harbo *et al.* (1998, Theorem 2).

¹⁴We are grateful to Peter Boswijk for helpful discussions on this result; a proof of this assertion is available from the authors on request.

also that $\sum_{j=0}^{\infty} \|\mathbf{C}_j^*\| < \infty$ and $\|\mathbf{C}\| < \infty$, Phillips and Solo (1992, Lemma 2.1, p.972), which excludes the cointegrating relations (2.13) being fractionally integrated of positive order.

We define the partial sum process

$$\mathcal{S}_T^e(a) \equiv T^{-1/2} \sum_{s=1}^{[Ta]} \mathbf{e}_t,$$

where $[Ta]$ denotes the integer part of Ta , $a \in [0, 1]$. Under Assumption 4.1, $\mathcal{S}_T^e(a)$ satisfies the multivariate invariance principle [Phillips and Durlauf (1986, Theorem 2.1, p.475)]

$$\mathcal{S}_T^e(a) \Rightarrow \mathbf{B}_m(a), a \in [0, 1], \quad (4.1)$$

where $\mathbf{B}_m(\cdot)$ denotes an m -dimensional Brownian motion with variance matrix Ω . We partition $\mathcal{S}_T^e(a) = (\mathcal{S}_T^y(a)', \mathcal{S}_T^x(a)')'$ conformably with $\mathbf{z}_t = (\mathbf{y}_t', \mathbf{x}_t')'$ and the Brownian motion $\mathbf{B}_m(a) = (\mathbf{B}_n(a)', \mathbf{B}_k(a)')'$ likewise, $a \in [0, 1]$. Define $\mathcal{S}_T^u(a) \equiv T^{-1/2} \sum_{s=1}^{[Ta]} \mathbf{u}_t$, $a \in [0, 1]$. Hence, as $\mathbf{u}_t \equiv \mathbf{e}_{yt} - \Omega_{yx}\Omega_{xx}^{-1}\mathbf{e}_{xt}$,

$$\mathcal{S}_T^u(a) \Rightarrow \mathbf{B}_n^*(a), \quad (4.2)$$

where $\mathbf{B}_n^*(a) \equiv \mathbf{B}_n(a) - \Omega_{yx}\Omega_{xx}^{-1}\mathbf{B}_k(a)$ is a Brownian motion with variance matrix Ω_{uu} which is independent of $\mathbf{B}_k(a)$, $a \in [0, 1]$. Consequently, the results described in Harbo *et al.* (1998) remain valid under Assumption 4.1.

Under Assumption 3.1, the $(m, m-r)$ matrix $\boldsymbol{\alpha}_{\perp} \equiv \text{diag}(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}_x^{\perp})$, where $\boldsymbol{\alpha}_x^{\perp}$ is a (k, k) non-singular matrix, is a basis for the orthogonal complement of the (m, r) loadings matrix $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_y', \mathbf{0}')'$. Hence, we define the $(m-r)$ -dimensional standard Brownian motion $\mathbf{W}_{m-r}(a) \equiv (\mathbf{W}_{n-r}(a)', \mathbf{W}_k(a)')'$ partitioned into the $(n-r)$ - and k -dimensional sub-vector independent standard Brownian motions $\mathbf{W}_{n-r}(a) \equiv (\boldsymbol{\alpha}_y^{\perp'}\Omega_{uu}\boldsymbol{\alpha}_y^{\perp})^{-1/2}\boldsymbol{\alpha}_y^{\perp'}\mathbf{B}_n^*(a)$ and $\mathbf{W}_k(a) \equiv (\boldsymbol{\alpha}_x^{\perp'}\Omega_{xx}\boldsymbol{\alpha}_x^{\perp})^{-1/2}\boldsymbol{\alpha}_x^{\perp'}\mathbf{B}_k(a)$, $a \in [0, 1]$. See Pesaran, Shin and Smith (1997, Appendix A) for further details. We will also require the corresponding de-meaned $(m-r)$ -vector standard Brownian motion

$$\tilde{\mathbf{W}}_{m-r}(a) \equiv \mathbf{W}_{m-r}(a) - \int_0^1 \mathbf{W}_{m-r}(a) da, \quad (4.3)$$

and de-meaned and de-trended $(m-r)$ -vector standard Brownian motion

$$\hat{\mathbf{W}}_{m-r}(a) \equiv \tilde{\mathbf{W}}_{m-r}(a) - 12(a - \frac{1}{2}) \int_0^1 (a - \frac{1}{2}) \tilde{\mathbf{W}}_{m-r}(a) da, \quad (4.4)$$

and their respective partitioned counterparts $\tilde{\mathbf{W}}_{m-r}(a) = (\tilde{\mathbf{W}}_{n-r}(a)', \tilde{\mathbf{W}}_k(a)')'$, and $\hat{\mathbf{W}}_{m-r}(a) = (\hat{\mathbf{W}}_{n-r}(a)', \hat{\mathbf{W}}_k(a)')'$, $a \in [0, 1]$.

4.1 Testing H_r against H_{r+1}

The (log-) likelihood ratio statistic for testing $H_r : \text{Rank}[\Pi_y] = r$ against $H_{r+1} : \text{Rank}[\Pi_y] = r+1$ is given by

$$\mathcal{LR}(H_r|H_{r+1}) = -T \ln(1 - \hat{\lambda}_{r+1}), \quad (4.5)$$

where $\hat{\lambda}_r$ is the r th largest eigenvalue from the determinantal equation (3.22), $r = 0, \dots, n-1$, with the appropriate definitions of $\Delta \hat{\mathbf{Y}}$ and $\hat{\mathbf{Z}}_{-1}^*$ and, thus, the sample moment matrices $\mathbf{S}_{YY}$, $\mathbf{S}_{YZ}$ and $\mathbf{S}_{ZZ}$ given by (3.21) to cover Cases I-V.

Theorem 4.1 (*Limit Distribution of $\mathcal{LR}(H_r|H_{r+1})$.*) Under H_r defined by (3.8) and Assumptions 2.1, 2.2, 3.1 and 4.1, the limit distribution of $\mathcal{LR}(H_r|H_{r+1})$ of (4.5) for testing H_r against H_{r+1} is given by the distribution of the maximum eigenvalue of

$$\int_0^1 d\mathbf{W}_{n-r}(a) \mathbf{F}_{m-r}(a)' \left(\int_0^1 \mathbf{F}_{m-r}(a) \mathbf{F}_{m-r}(a)' da \right)^{-1} \int_0^1 \mathbf{F}_{m-r}(a) d\mathbf{W}_{n-r}(a)', \quad (4.6)$$

where

$$\mathbf{F}_{m-r}(a) = \left\{ \begin{array}{ll} \mathbf{W}_{m-r}(a) & \text{Case I} \\ (\mathbf{W}_{m-r}(a)', 1)' & \text{Case II} \\ \tilde{\mathbf{W}}_{m-r}(a) & \text{Case III} \\ (\tilde{\mathbf{W}}_{m-r}(a)', a - \frac{1}{2})' & \text{Case IV} \\ \hat{\mathbf{W}}_{m-r}(a) & \text{Case V} \end{array} \right\}, \quad a \in [0, 1],$$

$r = 0, \dots, n-1$, where Cases I-V are defined following (3.8).

4.2 Testing H_r against H_n

The (log-) likelihood ratio statistic for testing $H_r : \text{Rank}[\Pi_y] = r$ against $H_n : \text{Rank}[\Pi_y] = n$ is given by

$$\mathcal{LR}(H_r|H_n) = -T \sum_{i=r+1}^n \ln(1 - \hat{\lambda}_i), \quad (4.7)$$

where $\hat{\lambda}_i$ is the i th largest eigenvalue from the determinantal equation (3.22).

Theorem 4.2 (*Limit Distribution of $\mathcal{LR}(H_r|H_n)$.*) Under H_r defined by (3.8) and Assumptions 2.1, 2.2, 3.1 and 4.1, the limit distribution of $\mathcal{LR}(H_r|H_n)$ of (4.7) for testing H_r against H_n is given by the distribution of

$$\text{Trace} \left\{ \int_0^1 d\mathbf{W}_{n-r}(a) \mathbf{F}_{m-r}(a)' \left(\int_0^1 \mathbf{F}_{m-r}(a) \mathbf{F}_{m-r}(a)' da \right)^{-1} \int_0^1 \mathbf{F}_{m-r}(a) d\mathbf{W}_{n-r}(r)' \right\},$$

where $\mathbf{F}_{m-r}(a)$, $a \in [0, 1]$, is defined in Theorem 4.1 for Cases I-V, $r = 0, \dots, n-1$.

4.3 Testing H_r in the Presence of $I(0)$ Weakly Exogenous Regressors

The (log-) likelihood ratio tests described above require that the process $\{\mathbf{x}_t\}_{t=1}^\infty$ is integrated of order one as noted below (3.7) and implied by the weak exogeneity Assumption 3.1; see (3.6). However, many applications will include current and lagged values of weakly exogenous

regressors which are integrated of order zero as explanatory variables in (3.5). In such circumstances, Theorems 4.1 and 4.2 no longer apply with the limiting distributions of the (log-) likelihood ratio tests for cointegration now being dependent on nuisance parameters. However, the above analysis may easily be adapted with this difficulty. Let $\{\mathbf{w}_t\}_{t=1}^{\infty}$ denote a k_w -vector process of weakly exogenous explanatory variables which is integrated of order zero. Therefore, the partial sum vector process $\{\sum_{s=1}^t \mathbf{w}_s\}_{t=1}^{\infty}$ is integrated of order one. Defining $\sum_{s=1}^t \mathbf{w}_s$ as a sub-vector of $\mathbf{x}_t$ with the corresponding subvector of $\Delta \mathbf{x}_t$ as $\mathbf{w}_t$, $t = 1, 2, \dots$, allows the above analysis to proceed unaltered. With these re-definitions of $\mathbf{x}_t$ and $\Delta \mathbf{x}_t$ to include $\sum_{s=1}^t \mathbf{w}_s$ and $\mathbf{w}_t$, $t = 1, 2, \dots$, respectively, the partial sum $\sum_{s=1}^t \mathbf{w}_s$ will now appear in the cointegrating relations (2.13) and the lagged level term $\mathbf{z}_{t-1}$ in (3.5) although economic theory may indicate its absence; that is, the corresponding (k_w, r) block of the cointegrating matrix β is null. This constraint on the cointegrating matrix β is straightforwardly tested using a likelihood ratio statistic which will possess a limiting chi-squared distribution with rk_w degrees of freedom under H_r . See Rahbek and Mosconi (1999) for further discussion.

4.4 Testing the Co-Trending Hypothesis

The cointegration rank hypothesis does not preclude the presence of deterministic trends in the cointegrating relations; see (2.13). In particular, section 2 demonstrates that in order to preserve similar deterministic trending behavior in the level process $\{\mathbf{z}_t\}_{t=1}^{\infty}$ for differing values of the cointegration rank r it is necessary to restrict the trend (intercept) coefficients. Accordingly, in the context of the VECM formulation (3.5), the trend parameter is rendered as $\mathbf{c}_1 = -\Pi_y \gamma$ (see (3.7)). In the absence of a deterministic trend in (2.1), $\gamma = \mathbf{0}$, then the intercept parameter in the VECM formulation (3.5) is similarly restricted as $\mathbf{c}_0 = -\Pi_y \mu$, and the cointegrating relationship (2.13) has an intercept given by $\beta' \mu$. Therefore, when the cointegrating rank is r , out of the n trend (intercept) terms, r of them can take any values and the remaining $n - r$ terms must satisfy prior restrictions.

This sub-section is concerned with the (log-) likelihood ratio test for the *co-trending* restriction, that is, the absence of a trend in the cointegrating relationship (2.13); *viz.* $\beta' \gamma = \mathbf{0}$ or equivalently $\mathbf{c}_1 = \mathbf{0}$ in the VECM formulation (3.5) for Case IV which yields Case III. In the absence of a trend, $\gamma = \mathbf{0}$ (Case II), we also consider the likelihood ratio test for the absence of an intercept in the cointegrating relationship (2.13); *viz.* $\beta' \mu = \mathbf{0}$ or equivalently $\mathbf{c}_0 = \mathbf{0}$ in the VECM formulation (3.5) which yields Case I. In both cases, the requisite ML estimators are obtained under the cointegrating hypothesis H_r of (3.8). See Johansen (1995, Theorem 11.3, p.162) for similar results when $n = m$ and Harbo *et al.* (1998, Theorem 3) which states Theorem 4.3 below under the assumptions of sections 2 and 3.

The following theorems detail the limiting behavior of both likelihood ratio tests.

Theorem 4.3 (*Limit Distribution of the Likelihood Ratio Statistic for $\mathbf{c}_1 = \mathbf{0}$ in Case IV*). *In Case IV, under H_r defined by (3.8), Assumptions 2.1, 2.2, 3.1 and 4.1, and $\mathbf{c}_1 = \mathbf{0}$, the likelihood ratio statistic for $\mathbf{c}_1 = \mathbf{0}$ has a limiting chi-squared distribution with r degrees of freedom, $r = 1, \dots, n$.*

Similarly,

Theorem 4.4 (*Limit Distribution of the Likelihood Ratio Statistic for $\mathbf{c}_0 = \mathbf{0}$ in Case II*). In Case II, under H_r defined by (3.8), Assumptions 2.1, 2.2, 3.1 and 4.1, and $\mathbf{c}_0 = \mathbf{0}$, the likelihood ratio statistic for $\mathbf{c}_0 = \mathbf{0}$ has a limiting chi-squared distribution with r degrees of freedom, $r = 1, \dots, n$.

We also note that under H_r of (3.8) the tests described in this sub-section are asymptotically independent of the corresponding cointegration rank tests for H_r of sub-sections 4.1 and 4.2. Consequently, the overall asymptotic size of the resultant induced test for H_r and the co-trending or intercept hypothesis of this sub-section based on the above likelihood ratio statistics is simply calculated from the asymptotic sizes of the individual tests.

4.5 Specification Tests

The framework adopted in this paper has imposed certain assumptions which may be overly restrictive. For completeness, this sub-section describes specification tests for these assumptions.

Firstly, as discussed in section 2, we require that the trend parameters obey certain parametric restrictions in order that the deterministic trending behavior of the level process $\{\mathbf{z}_t\}_{t=1}^\infty$ is invariant to the cointegrating rank r . In the absence of a deterministic trend, the intercept parameter vector is similarly restricted. In order to test these restrictions, we may use likelihood ratio tests calculated under the cointegration rank hypothesis H_r defined by (3.8): Case IV against Case V in the first instance and Case III against Case II in the second. See Johansen (1995, Corollary 11.2, p.163) for the full system case $n = m$.

Theorem 4.5 (*Limit Distribution of the Likelihood Ratio Statistic for $\mathbf{c}_1 = -\Pi_y \gamma$*). In Case V (defined by (3.13)), under H_r defined by (3.8), Assumptions 2.1, 2.2, 3.1 and 4.1, and $\mathbf{c}_1 = -\Pi_y \gamma$, the likelihood ratio statistic for $\mathbf{c}_1 = -\Pi_y \gamma$ has a limiting chi-squared distribution with $n - r$ degrees of freedom, $r = 0, \dots, n - 1$.

Similarly,

Theorem 4.6 (*Limit Distribution of the Likelihood Ratio Statistic for $\mathbf{c}_0 = -\Pi_y \mu$*). In Case III (defined by (3.11)), under H_r defined by (3.8), Assumptions 2.1, 2.2, 3.1 and 4.1, and $\mathbf{c}_0 = -\Pi_y \mu$, the likelihood ratio statistic for $\mathbf{c}_0 = -\Pi_y \mu$ has a limiting chi-squared distribution with $n - r$ degrees of freedom, $r = 0, \dots, n - 1$.

Secondly, in Assumption 3.1, we imposed the weak exogeneity restriction $\Pi_x = \mathbf{0}$; that is, the level process $\{\mathbf{x}_t\}_{t=1}^\infty$ is integrated of order one and long-run forcing for $\{\mathbf{y}_t\}_{t=1}^\infty$. This restriction takes two forms. Firstly, as noted in section 3, the level process $\{\mathbf{x}_t\}_{t=1}^\infty$ is not mutually cointegrated. Secondly, the lagged cointegrating relationship does not enter the evolution of the differenced process $\{\Delta \mathbf{x}_t\}_{t=1}^\infty$ in (3.6). Our discussion below is concerned with Case IV defined analogously to that of earlier sections, the remaining cases are defined likewise and their associated results may be obtained by a similar analysis to that for Case IV.

Firstly, consider the following sub-system regression model

$$\Delta \mathbf{x}_t = \mathbf{a}_{x0} + \sum_{i=1}^{p-1} \Gamma_{xi} \Delta \mathbf{z}_{t-i} + \Pi_{xx*} \mathbf{x}_{t-1}^* + \mathbf{e}_{xt}, \quad (4.8)$$

where $\Pi_{xx*} \equiv \Pi_{xx}(-\gamma_x, \mathbf{I}_k)$, $\Pi_x \equiv (\Pi_{xy}, \Pi_{xx})$, $\gamma \equiv (\gamma'_y, \gamma'_x)'$ and $\mathbf{x}_{t-1}^* \equiv (t, \mathbf{x}_{t-1})'$, $t = 1, 2, \dots$; cf. (3.6). Equation (4.8) embodies the possibility that the level process $\{\mathbf{x}_t\}_{t=1}^\infty$ is mutually cointegrated. We denote the cointegration rank hypothesis in (4.8) by $H_r^x : \text{Rank}[\Pi_x] = r$, $r = 0, \dots, k$. Therefore, similarly to Theorems 4.1 and 4.2 [cf. Johansen (1991)],

Theorem 4.7 (*Limit Distribution of $\mathcal{LR}(H_0^x | H_1^x)$.*) Under H_r defined in (3.8) and Assumptions 2.1, 2.2, 3.1 and 4.1, the limit distribution of the likelihood ratio statistic $\mathcal{LR}(H_0^x | H_1^x)$ for testing H_0^x against H_1^x in (4.8) is the distribution of the maximum eigenvalue of

$$\int_0^1 d\mathbf{W}_k(a) \mathbf{F}_k(a)' \left(\int_0^1 \mathbf{F}_k(a) \mathbf{F}_k(a)' da \right)^{-1} \int_0^1 \mathbf{F}_k(a) d\mathbf{W}_k(a)',$$

where

$$\mathbf{F}_k(a) = \left\{ \begin{array}{ll} \mathbf{W}_k(a) & \text{Case I} \\ (\mathbf{W}_k(a)', 1)' & \text{Case II} \\ \tilde{\mathbf{W}}_k(a) & \text{Case III} \\ (\tilde{\mathbf{W}}_k(a)', a - \frac{1}{2})' & \text{Case IV} \\ \hat{\mathbf{W}}_k(a) & \text{Case V} \end{array} \right\}, a \in [0, 1],$$

and Cases I-V are defined following (3.8).

Theorem 4.8 (*Limit Distribution of $\mathcal{LR}(H_0^x | H_k^x)$.*) Under H_r defined in (3.8) and Assumptions 2.1, 2.2, 3.1 and 4.1, the limit distribution of the likelihood ratio statistic $\mathcal{LR}(H_0^x | H_k^x)$ for testing H_0^x against H_k^x in (4.8) is the distribution of

$$\text{Trace} \left\{ \int_0^1 d\mathbf{W}_k(a) \mathbf{F}_k(a)' \left(\int_0^1 \mathbf{F}_k(a) \mathbf{F}_k(a)' da \right)^{-1} \int_0^1 \mathbf{F}_k(a) d\mathbf{W}_k(r)' \right\},$$

where $\mathbf{F}_k(a)$, $a \in [0, 1]$, is defined in Theorem 4.3 for Cases I-V.

Secondly, consider the sub-system regression model

$$\Delta \mathbf{x}_t = \mathbf{a}_{x0} + \sum_{i=1}^{p-1} \Gamma_{xi} \Delta \mathbf{z}_{t-i} + \alpha_{xy} \hat{\beta}'_* \mathbf{z}_{t-1}^* + \hat{\mathbf{e}}_{xt}, \quad (4.9)$$

$t = 1, \dots, T$, where $\hat{\beta}_*$ denotes the ML estimator for the cointegrating vector β_* under H_r in Cases I-V of the earlier sections. Equation (4.9) allows for the alternative possibility that the lagged cointegrating relationships $\{\beta'_* \mathbf{z}_t^*\}_{t=1}^\infty$ might enter (3.6).¹⁵

Theorem 4.9 (*Limit Distribution of the Likelihood Ratio Statistic for $\alpha_{xy} = \mathbf{0}$.*) Under H_r defined in (3.8) and Assumptions 2.1, 2.2, 3.1 and 4.1, the likelihood ratio statistic for $\alpha_{xy} = \mathbf{0}$ in (4.9) has a limiting chi-squared distribution with kr degrees of freedom for Cases I-V, $r = 1, \dots, n$.

¹⁵Boswijk (1995) provides a Lagrange multiplier test for $\alpha_{xy} = \mathbf{0}$ in (4.9). See also Johansen (1992) and Harbo *et al.* (1998, sub-section 4.1).

5 Efficient Estimation of a Structural Error Correction Model Subject to Constraints on the Short-Run Dynamics

The structural VECM model described in section 3 does not permit any restrictions to be imposed on the short run dynamics of the model. In many applications, due to data limitations or *a priori* restrictions, researchers would potentially wish to impose such restrictions on the VECM form (3.5). Furthermore, Abadir, Hadri and Tzavalis (1999) suggest that the inclusion of irrelevant terms in systems such as the VECM (3.5) may result in substantial estimator bias. As the restrictions only concern parameter vectors and matrices associated with stationary variables, the results of the previous section continue to hold under H_r of (3.8) as will become evident below.

As in earlier sections, our exposition is in terms of Case IV. Consider the following linear constraints on the parameter matrix Ψ associated with the short-run dynamics in the VECM (3.18):

$$vec(\Psi') = \mathbf{S}\boldsymbol{\varphi}, \quad (5.1)$$

where $\mathbf{S}$ is a $(nmp - n^2, s)$ matrix of full column rank s with known elements and $\boldsymbol{\varphi}$ a s -vector of unknown parameters, $s \leq nmp - n^2$. For example, consider the j th equation of (3.18), revised to incorporate exclusion restrictions on the short-run dynamics,

$$\Delta \mathbf{y}'_j = c_{0j} \boldsymbol{\iota}'_T + \boldsymbol{\varphi}^{j'} \mathbf{S}^{j'} \Delta \mathbf{Z}_- + \boldsymbol{\pi}'_{y*j} \mathbf{Z}_{-1}^* + \mathbf{u}'_j \quad (5.2)$$

where $\Delta \mathbf{y}_j \equiv (\Delta y_{1j}, \dots, \Delta y_{Tj})'$ and $\mathbf{u}_j \equiv (u_{1j}, \dots, u_{Tj})'$, $j = 1, \dots, n$, and $\Pi_{y*} = (\boldsymbol{\pi}_{y*1}, \dots, \boldsymbol{\pi}_{y*n})'$. The formulation (5.2) allows certain elements of the stationary component $\Delta \mathbf{Z}_-$ to be excluded from the j th equation and, thus, also the lag lengths to differ across equations. These constraints are summarised by $\boldsymbol{\psi}^j = \mathbf{S}^j \boldsymbol{\varphi}^j$, with $\mathbf{S}^j$ an appropriately defined selection matrix, $j = 1, \dots, n$, and $\Psi = (\boldsymbol{\psi}^1, \dots, \boldsymbol{\psi}^n)'$. Stacking these n equations yields:

$$\Delta \mathbf{Y} = \mathbf{c}_0 \boldsymbol{\iota}'_T + \Psi \Delta \mathbf{Z}_- + \Pi_{y*} \mathbf{Z}_{-1}^* + \mathbf{U}, \quad (5.3)$$

where $\Delta \mathbf{Y} \equiv (\Delta \mathbf{y}_1, \dots, \Delta \mathbf{y}_n)'$, subject to the constraints $vec(\Psi') = \mathbf{S}\boldsymbol{\varphi}$ as in (5.1) with $\mathbf{S} \equiv diag(\mathbf{S}^1, \dots, \mathbf{S}^n)$ and $\boldsymbol{\varphi} = (\boldsymbol{\varphi}^1, \dots, \boldsymbol{\varphi}^n)'$; cf. (3.18). Note that the constraints (5.1) also permit the possibility of cross-equation restrictions.

Let $\tilde{\boldsymbol{\varphi}}$ denote a $T^{1/2}$ -consistent estimator for $\boldsymbol{\varphi}$ in the *seemingly unrelated regressions* (SUR) VECM (5.3) subject to the constraints $vec(\Psi') = \mathbf{S}\boldsymbol{\varphi}$ of (5.1); that is, $T^{1/2}(\tilde{\boldsymbol{\varphi}} - \boldsymbol{\varphi) = O_P(1)$ and, thus, $T^{1/2}(\tilde{\Psi} - \Psi) = O_P(1)$ where $vec(\tilde{\Psi}') = \mathbf{S}\tilde{\boldsymbol{\varphi}}$. For example, under Assumptions 2.1, 2.2, 3.1 and 4.1, a $T^{1/2}$ -consistent estimator $\tilde{\boldsymbol{\varphi}}$ may be obtained using SUR estimation on (5.3) subject to $vec(\Psi') = \mathbf{S}\boldsymbol{\varphi}$, with Π_{y*} unrestricted and Ω_{uu} replaced by an initial $T^{1/2}$ -consistent estimator based on residuals computed from ordinary least squares equation-by-equation estimation, once again treating Π_{y*} as unrestricted.¹⁶ The next step

¹⁶ A proof of $T^{1/2}$ -consistency of the SUR estimator of $\boldsymbol{\varphi}$ can be established along the following lines. Using results in Park and Phillips (1989), in particular, Theorems 3.1 and 3.2, p.102, and their Sections 5.2

subtracts the *estimated* short-run dynamics from the left hand side of (5.3), and applies the reduced rank technique to the following synthetic multivariate regression

$$(\Delta \mathbf{Y} - \tilde{\Psi} \Delta \mathbf{Z}_-) = \mathbf{c}_0 \boldsymbol{\iota}'_T + \Pi_{y*} \mathbf{Z}_{-1}^* + \tilde{\mathbf{U}}, \quad (5.4)$$

where $vec(\tilde{\Psi}') = \mathbf{S}\tilde{\varphi}$ and $\tilde{\mathbf{U}} \equiv \mathbf{U} - (\tilde{\Psi} - \Psi)\Delta \mathbf{Z}_-$.

For this purpose consider the minimisation with respect to β_* under H_r of (3.17) of the determinantal criterion function

$$\left| T^{-1} \Delta \tilde{\mathbf{Y}} \left(\mathbf{I}_T - \tilde{\mathbf{Z}}_{-1}' \beta_* \left(\beta_*' \tilde{\mathbf{Z}}_{-1}^* \tilde{\mathbf{Z}}_{-1}' \beta_* \right)^{-1} \beta_*' \tilde{\mathbf{Z}}_{-1}^* \right) \Delta \tilde{\mathbf{Y}}' \right|,$$

cf. (3.20), where $\Delta \tilde{\mathbf{Y}} \equiv (\Delta \mathbf{Y} - \tilde{\Psi} \Delta \mathbf{Z}_-) \bar{\mathbf{P}}_\iota$ and $\tilde{\mathbf{Z}}_{-1}^* \equiv \mathbf{Z}_{-1}^* \bar{\mathbf{P}}_\iota$ are respectively the de-meanded versions of $\Delta \mathbf{Y} - \tilde{\Psi} \Delta \mathbf{Z}_-$ and $\mathbf{Z}_{-1}^*$, $\bar{\mathbf{P}}_\iota \equiv \mathbf{I}_T - \boldsymbol{\iota}_T (\boldsymbol{\iota}'_T \boldsymbol{\iota}_T)^{-1} \boldsymbol{\iota}'_T$, which results from (5.4) by concentrating out α_y after substitution of $\Pi_{y*} = \alpha_y \beta_*'$. Defining the moment matrices

$$\tilde{\mathbf{S}}_{YY} \equiv T^{-1} \Delta \tilde{\mathbf{Y}} \Delta \tilde{\mathbf{Y}}', \quad \tilde{\mathbf{S}}_{YZ} \equiv T^{-1} \Delta \tilde{\mathbf{Y}} \tilde{\mathbf{Z}}_{-1}', \quad \tilde{\mathbf{S}}_{ZZ} \equiv T^{-1} \tilde{\mathbf{Z}}_{-1}^* \tilde{\mathbf{Z}}_{-1}', \quad (5.5)$$

cf. (3.21), the solution $\tilde{\beta}_*$ to the above minimization problem is given by the eigenvectors corresponding to the r largest eigenvalues $\tilde{\lambda}_1 > \dots > \tilde{\lambda}_r > 0$ of

$$\left| \tilde{\lambda} \tilde{\mathbf{S}}_{ZZ} - \tilde{\mathbf{S}}_{ZY} \tilde{\mathbf{S}}_{YY}^{-1} \tilde{\mathbf{S}}_{YZ} \right| = 0; \quad (5.6)$$

cf. (3.22).

The modifications of the definitions of $\Delta \tilde{\mathbf{Y}}$ and $\tilde{\mathbf{Z}}_{-1}^*$ and, thus, the sample moment matrices $\tilde{\mathbf{S}}_{YY}$, $\tilde{\mathbf{S}}_{YZ}$ and $\tilde{\mathbf{S}}_{ZZ}$ of (5.5) necessary in the determinantal equation (5.6) for the five cases of interest are given as follows.

- **Case I:** $\mathbf{c}_0 = \mathbf{0}$ and $\mathbf{c}_1 = \mathbf{0}$. $\Delta \tilde{\mathbf{Y}} \equiv (\Delta \mathbf{Y} - \tilde{\Psi} \Delta \mathbf{Z}_-)$ and $\tilde{\mathbf{Z}}_{-1}^* \equiv \mathbf{Z}_{-1}$ where $vec(\tilde{\Psi}') = \mathbf{S}\tilde{\varphi}$ and $\tilde{\varphi}$ is a $T^{1/2}$ -consistent estimator for φ of $vec(\Psi') = \mathbf{S}\varphi$ obtained from $\Delta \mathbf{Y} = \Psi \Delta \mathbf{Z}_- + \Pi_y \mathbf{Z}_{-1} + \mathbf{U}$.
- **Case II:** $\mathbf{c}_0 = -\Pi_y \boldsymbol{\mu}$ and $\mathbf{c}_1 = \mathbf{0}$. $\Delta \tilde{\mathbf{Y}} \equiv (\Delta \mathbf{Y} - \tilde{\Psi} \Delta \mathbf{Z}_-)$ and $\tilde{\mathbf{Z}}_{-1}^* \equiv \mathbf{Z}_{-1}^*$ where $vec(\tilde{\Psi}') = \mathbf{S}\tilde{\varphi}$, $\tilde{\varphi}$ is a $T^{1/2}$ -consistent estimator for φ of $vec(\Psi') = \mathbf{S}\varphi$ obtained from $\Delta \mathbf{Y} = \Psi \Delta \mathbf{Z}_- + \Pi_{y*} \mathbf{Z}_{-1}^* + \mathbf{U}$ and $\mathbf{Z}_{-1}^* = (\boldsymbol{\iota}_T, \mathbf{Z}'_{-1})'$.
- **Case III:** $\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 = \mathbf{0}$. $\Delta \tilde{\mathbf{Y}} \equiv (\Delta \mathbf{Y} - \tilde{\Psi} \Delta \mathbf{Z}_-) \bar{\mathbf{P}}_\iota$ and $\tilde{\mathbf{Z}}_{-1}^* \equiv \mathbf{Z}_{-1} \bar{\mathbf{P}}_\iota$ are respectively $\Delta \mathbf{Y} - \tilde{\Psi} \Delta \mathbf{Z}_-$ and $\mathbf{Z}_{-1}$ de-meanded, $\bar{\mathbf{P}}_\iota \equiv \mathbf{I}_T - \boldsymbol{\iota}_T (\boldsymbol{\iota}'_T \boldsymbol{\iota}_T)^{-1} \boldsymbol{\iota}'_T$, where $vec(\tilde{\Psi}') = \mathbf{S}\tilde{\varphi}$, $\tilde{\varphi}$ is a $T^{1/2}$ -consistent estimator for φ of $vec(\Psi') = \mathbf{S}\varphi$ obtained from $\Delta \mathbf{Y} = \mathbf{c}_0 \boldsymbol{\iota}'_T + \Psi \Delta \mathbf{Z}_- + \Pi_y \mathbf{Z}_{-1} + \mathbf{U}$.

and 5.4, an estimator for Ω_{uu} based on residuals from equation-by-equation ordinary least squares treating Π_{y*} as unrestricted is $T^{1/2}$ -consistent; see Park and Phillips (1989, Theorem 3.4, p.107). Consequently, an adaptation of Park and Phillips (1989, Proof of Theorem 3.1, p.126) may be used to show that the resultant SUR estimator for φ is $T^{1/2}$ -consistent.

- **Case IV:** $\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 = -\Pi_y \gamma$. $\Delta \tilde{\mathbf{Y}} \equiv (\Delta \mathbf{Y} - \tilde{\Psi} \Delta \mathbf{Z}_-) \bar{\mathbf{P}}_\iota$ and $\tilde{\mathbf{Z}}_{-1}^* \equiv \mathbf{Z}_{-1}^* \bar{\mathbf{P}}_\iota$ are respectively $\Delta \mathbf{Y} - \tilde{\Psi} \Delta \mathbf{Z}_-$ and $\mathbf{Z}_{-1}^*$ de-meaned, where $\text{vec}(\tilde{\Psi}') = \mathbf{S} \tilde{\varphi}$, $\tilde{\varphi}$ is a $T^{1/2}$ -consistent estimator for φ of $\text{vec}(\Psi') = \mathbf{S} \varphi$ obtained from $\Delta \mathbf{Y} = \mathbf{c}_0 \iota'_T + \Psi \Delta \mathbf{Z}_- + \Pi_{y*} \mathbf{Z}_{-1}^* + \mathbf{U}$ and $\mathbf{Z}_{-1}^* = (\tau_T, \mathbf{Z}'_{-1})'$.
- **Case V:** $\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 \neq \mathbf{0}$. $\Delta \tilde{\mathbf{Y}} \equiv (\Delta \mathbf{Y} - \tilde{\Psi} \Delta \mathbf{Z}_-) \bar{\mathbf{P}}_{\iota, \tau}$ and $\tilde{\mathbf{Z}}_{-1}^* \equiv \mathbf{Z}_{-1} \bar{\mathbf{P}}_{\iota, \tau}$ are respectively $\Delta \mathbf{Y} - \tilde{\Psi} \Delta \mathbf{Z}_-$ and $\mathbf{Z}_{-1}$ de-meaned and de-trended, $\bar{\mathbf{P}}_\iota \equiv \bar{\mathbf{P}}_\iota - \bar{\mathbf{P}}_{\iota, \tau} \tau_T' (\tau_T' \bar{\mathbf{P}}_\iota \tau_T)^{-1} \tau_T' \bar{\mathbf{P}}_\iota$, where $\text{vec}(\tilde{\Psi}') = \mathbf{S} \tilde{\varphi}$, $\tilde{\varphi}$ is a $T^{1/2}$ -consistent estimator for φ of $\text{vec}(\Psi') = \mathbf{S} \varphi$ obtained from $\Delta \mathbf{Y} = \mathbf{c}_0 \iota'_T + \mathbf{c}_1 \tau_T' + \Psi \Delta \mathbf{Z}_- + \Pi_y \mathbf{Z}_{-1} + \mathbf{U}$.

The estimated eigenvalues $\{\tilde{\lambda}_i\}_{i=1}^n$ from (5.6) may be then used to test H_r against H_{r+1} , cf. Theorem 4.1, and H_r against H_n , cf. Theorem 4.2.

Theorem 5.1 (*Limit Distributions of $\widetilde{\mathcal{LR}}(H_r|H_{r+1})$ and $\widetilde{\mathcal{LR}}(H_r|H_n)$.*) If $\tilde{\varphi}$ is a $T^{1/2}$ -consistent estimator for φ , then, under H_r defined by (3.8) and Assumptions 2.1, 2.2, 3.1 and 4.1, the limit distributions of the pseudo-(log)-likelihood ratio statistics $\widetilde{\mathcal{LR}}(H_r|H_{r+1}) \equiv -T \ln(1 - \tilde{\lambda}_{r+1})$ for testing H_r against H_{r+1} and $\widetilde{\mathcal{LR}}(H_r|H_n) \equiv -T \sum_{i=r+1}^n \ln(1 - \tilde{\lambda}_i)$ for testing H_r against H_n are identical to those of $\mathcal{LR}(H_r|H_{r+1})$ of (4.5) and $\mathcal{LR}(H_r|H_n)$ of (4.7) stated in Theorems 4.1 and 4.2 respectively for Cases I-V, $r = 0, \dots, n-1$.

For Case IV, let $\tilde{\beta}_*$ denote the $(m+1, r)$ matrix of *just-identified* eigenvectors under H_r of (3.8) corresponding to the eigenvalues $\tilde{\lambda}_i$, $i = 1, \dots, r$, of (5.6). Consider the multivariate regression

$$\Delta \mathbf{Y} = \mathbf{c}_0 \iota'_T + \Psi \Delta \mathbf{Z}_- + \alpha_y \tilde{\beta}_*' \mathbf{Z}_{-1}^* + \tilde{\mathbf{U}}, \quad (5.7)$$

subject to the constraints $\text{vec}(\Psi') = \mathbf{S} \varphi$ of (5.1); cf. (5.3). Asymptotically, the inclusion of the lagged estimated cointegrating relationships $\tilde{\beta}_*' \mathbf{z}_{t-1}^*$ is the same as the inclusion of the actual cointegrating relationships $\beta_*' \mathbf{z}_{t-1}^*$, $t = 1, \dots, T$. Therefore, as the regression contains only variables which are asymptotically stationary, SUR estimation of (5.7) will be asymptotically efficient for the short-run parameters φ and the (n, r) loadings matrix α_y . The modifications of (5.7) required for Cases I-III and V are obvious and are therefore omitted.

6 Finite Sample Properties

This section provides some preliminary evidence on the finite sample properties of the Case IV cointegration rank statistics developed in the previous sections using Monte Carlo techniques. The experimental design is an extension of that used by Yamada and Toda (1998), who consider a VAR(2) model with 2 endogenous $I(1)$ variables, that is, $\mathbf{y}_t = (y_{1t}, y_{2t})'$. Their VAR(2) model is augmented by a scalar exogenous $I(1)$ variable x_t , which results in the following DGP:

$$(\mathbf{I}_3 - \mathbf{A}L)(\mathbf{I}_3 - \mathbf{B}L)\mathbf{z}_t = \mathbf{e}_t, \quad (6.1)$$

where $\mathbf{z}_t = (\mathbf{y}'_t, x_t)'$,

$$\mathbf{A} = \begin{pmatrix} a_{11} & 0 & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & b_{12} & 0 \\ b_{21} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (6.2)$$

and $\mathbf{e}_t \sim IN(\mathbf{0}, \Omega)$. An advantage of this formulation is that the integration and cointegration properties of the system (6.1) are controlled *via* a_{11} and a_{22} if we assume $|b_{21}b_{12}| < 1$. To ensure that the cointegration rank is unity, that is, $Rank[\Pi] = 1$, we must have either $a_{11} = 1$ or $a_{22} = 1$ (but not both) and, for reasons which will become clear below, we additionally require $a_{13} = -a(1 - a_{11})$ and $a_{23} = aa_{21}$. As all of the Case IV cointegration rank statistics considered below are invariant to the intercept and trend parameter vectors $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$, we have set $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\gamma} = \mathbf{0}$ in (6.1).

Consequently, the structural VECM (3.5) is given by

$$\Delta \mathbf{y}_t = \boldsymbol{\lambda} \Delta x_t + \Psi_1 \Delta \mathbf{z}_{t-1} + \Pi_y \mathbf{z}_{t-1} + \mathbf{u}_t, \quad (6.3)$$

where $\mathbf{u}_t \sim IN(\mathbf{0}, \Omega_{uu})$,

$$\boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}, \quad \Psi_1 = \begin{pmatrix} 0 & a_{11}b_{12} & 0 \\ a_{22}b_{21} & a_{21}b_{12} & 0 \end{pmatrix},$$

and, for a cointegration rank of unity, $Rank[\Pi_y] = 1$, either, if $a_{11} = 1$ and $|a_{22}| < 1$:

$$\Pi_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} a_{21} + b_{21}[1 - a_{22}], & -[1 - a_{22}] - a_{21}b_{12}, & aa_{21} \end{pmatrix},$$

or, if $a_{22} = 1$ and $|a_{11}| < 1$:

$$\Pi_y = - \begin{pmatrix} 1 - a_{11} \\ -a_{21} \end{pmatrix} \begin{pmatrix} 1, & -b_{12}, & a \end{pmatrix}.$$

Because the cointegration rank statistics of earlier sections are invariant to scale, we may set

$$\Omega_{uu} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

and $var(e_{xt}) = 1$ yielding λ_1 and λ_2 as the covariances of y_{1t} and y_{2t} with x_t respectively.

The above DGP, characterized *via* $\boldsymbol{\lambda}$, Ψ_1 , Π_y and Ω_{uu} , contains 9 free parameters. It is clearly beyond the scope of the present paper to examine all the possible parameter configurations that this design entails. Hence, for illustrative purposes, in what follows we fix $a_{21} = 0.1$, $b_{12} = 0.2$, $b_{21} = -0.2$, $\lambda_1 = 0.3$ and $\lambda_2 = -0.2$. The remaining parameter values are chosen to ensure that the cointegrating rank is unity, $Rank[\Pi_y] = 1$, and that the population R^2 for each of the structural VECM equations falls in the range 0.10 – 0.30,

which reflects the sort of values generally encountered in practice.^{17,18} A summary of the parameter values considered and the implied population R^2 is given below:

	Experiments: Set I				Experiments: Set II			
Parameters	E_{1a}	E_{1b}	E_{1c}	E_{1d}	E_{2a}	E_{2b}	E_{2c}	E_{2d}
a_{11}	1.0	1.0	1.0	1.0	0.8	0.8	0.8	0.8
a_{22}	0.8	0.8	0.8	0.8	1.0	1.0	1.0	1.0
a	0	0	1	1	0	0	1	1
ρ	0	0.5	0	0.5	0	0.5	0	0.5
Implied R^2								
$R^2_{\Delta y_{1t}}$	0.121	0.121	0.123	0.122	0.198	0.172	0.290	0.270
$R^2_{\Delta y_{2t}}$	0.163	0.152	0.198	0.188	0.087	0.083	0.138	0.134

For each of these eight experiments we computed three sets of Case IV (log-) likelihood ratio statistics for testing H_r against H_{r+1} and H_n :

- (i) *LR1*: the statistics due to Johansen (1991) which treat all three variables in $\mathbf{z}_t$ as endogenous;
- (ii) *LR2*: the statistics developed in sub-sections 4.1 and 4.2 which correctly treat x_t as an exogenous $I(1)$ variable;
- (iii) *LR3*: the statistics of section 5 which also incorporate the zero restrictions on the short run dynamics in the matrix Ψ_1 which is estimated by equation-by-equation ordinary least squares treating Π_y as unrestricted.

Sample sizes $T = 30, 50, 75, 100, 150$ and 300 are considered. The nominal size of each of the tests is set at 0.05 and the number of replications at $20,000$.¹⁹

Tables 1(a) – 1(d) about here

The rejection frequencies for testing $H_r : \text{Rank}(\Pi_y) = r$ against H_{r+1} and H_n , $r = 0, 1$, are summarized in Tables 1(a)-1(d) for the experiments in Set I. The results in Table 1(a)

¹⁷Population R^2 values of the structural VECM equations are obtained as

$$R^2_{\Delta y_{it}} = 1 - \frac{\text{Var}(\Delta y_{it} | \Delta x_t, \Delta \mathbf{z}_{t-1}, \mathbf{z}_{t-1})}{\text{Var}(\Delta y_{it})} = 1 - \frac{1}{\text{Var}(\Delta y_{it})}, \quad i = 1, 2.$$

The unconditional variances $\text{Var}(\Delta y_{it})$, $i = 1, 2$, are given by the first two diagonal elements of the following matrix

$$\text{Cov}(\Delta \mathbf{z}_t) = \sum_{j=0}^{\infty} \mathbf{C}_j \Omega \mathbf{C}'_j;$$

see below (2.9) and fn.3.

¹⁸Other experiments with slightly larger values for these population R^2 were also conducted and yielded similar results to those described below.

¹⁹Similar results for Cases I, II, III and V to those for Case IV given below were also found and are available from the authors on request.

for Experiment E_{1a} show that all tests tend to be under-sized when $T \leq 150$. Because the size distortion is fairly uniform across the tests, it is possible to directly compare their power properties. The power of the $LR3$ tests of section 5 is significantly higher than their $LR1$ and $LR2$ counterparts with the power of the $LR2$ tests of sub-sections 4.1 and 4.2 being slightly better than the $LR1$ tests of Johansen (1991) in most cases. There seems little to choose between the $LR3$ tests for H_r against H_{r+1} and H_n although the former test is slightly more powerful for $T > 50$. Similar conclusions may be drawn from Experiment E_{1b} summarized in Table 1(b).

In Experiment E_{1c} , see Table 1(c), all tests are slightly over-sized for $T \geq 150$. The $LR3$ tests still perform best in terms of power with the $LR2$ test also slightly more powerful than the $LR1$ tests. In relative terms, the results are similar to those in Tables 1(a) but the powers of the tests are now significantly greater than those obtained for Experiment E_{1a} . Consequently, a non-zero value for a (which results in a higher value of the population $R_{\Delta y_{it}}^2$, $i = 1, 2$) may improve the finite sample power performance of the tests. As above, there is little difference between the results for Experiments E_{1c} and E_{1d} in Tables 1(c) and 1(d) respectively indicating that the structural error correlation parameter ρ has little effect on the finite sample performance of these tests for the experiments in Set I.

Tables 2(a) – 2(d) about here

Tables 2(a)-2(d) summarize the experiments in Set II. The results in Table 2(a) for Experiment E_{2a} are very similar to those in Tables 1(a) and 1(b). However, Table 2(b) for Experiment E_{2b} indicates that as T increases all tests, from being under-sized, become slightly over-sized and then approach nominal size. Unlike the experiments in Set I, the non-zero value of ρ , although it does not result in a higher value for the population $R_{\Delta y_{it}}^2$, $i = 1, 2$, improves the power of all tests significantly.

Table 2(c) for Experiment E_{2c} demonstrates that the under-size problem of previous experiments is not a general feature of these tests in small samples. Both $LR3$ tests are now over-sized with the $LR1$ and $LR2$ tests oversized for $T \geq 100$, for example, the $LR3$ tests have empirical size 0.096 for $T = 75$. Consequently, power comparisons of the different tests are problematic and the significantly large power of the $LR3$ tests over the others' should be discounted. Subject to this caveat, the comparisons between the tests are quite similar to those reported earlier. Table 2(d) for Experiment E_{2d} reveals similar conclusions. However, the results for the experiments in Set II tend to be more sensitive to non-zero-values of a and ρ than those for those in Set I.

In summary, the above results suggest that all tests perform reasonably satisfactorily in most of the cases considered. Albeit quite slowly, the empirical sizes of all tests tend to the nominal level as the sample size increases. Furthermore, the $LR3$ tests of section 5 are more powerful than the other tests although there are situations where these tests tend to over-reject in small samples.

7 An Empirical Application: A Re-examination of the Long-Run Validity of the PPP and UIP Hypotheses

In this section we re-examine the empirical evidence on the long-run validity of the Purchasing Power Parity (PPP) and Uncovered Interest Parity (UIP) hypotheses presented in Johansen and Juselius (1992) and Pesaran and Shin (1996). In these applications domestic and foreign prices and interest rates are assumed to be endogenously determined. Such a symmetric treatment of domestic and foreign variables does not, however, seem to be necessary in the case of small open economies where it is unlikely that changes in domestic variables have a significant impact on the long-run evolution of foreign (world) prices or interest rates. This earlier analysis also included changes in the logarithm of oil prices and its lagged values as exogenous $I(0)$ variables in the underlying cointegrating VAR model. However, as discussed in sub-section 4.3, the appropriate method of allowing for such effects is to include an integrated version of the $I(0)$ variables in the model which, in the context of the present application, implies adding the logarithm of oil prices to the list of $I(1)$ variables, and then testing the validity of excluding the level of oil prices from the cointegrating relations.

For comparability purposes, we use the same quarterly observations employed by Johansen and Juselius. The data set relates to the UK economy and covers the period 1972(1) - 1987(2). The included $I(1)$ variables are: the logarithm of the effective exchange rate, e_t , the logarithm of the UK wholesale price index, p_t , the logarithm of the trade-weighted foreign wholesale price index, p_t^* , the logarithm of oil prices, po_t , and the domestic and foreign interest rate variables, $r_t = \ln(1 + R_t/100)$ and $r_t^* = \ln(1 + R_t^*/100)$, where R_t is the UK three-month Treasury Bill rate, and R_t^* is the three-month Eurodollar rate. Here we carry out a cointegration analysis on the six $I(1)$ variables, e_t , p_t , r_t , r_t^* , p_t^* , and po_t , treating the last two variables as weakly exogenous and, thus, long-run forcing, in the sense described in section 3. We also considered treating the foreign interest rate variable r_t^* as exogenous but, given the importance of the UK in world financial markets, it was felt that this was unlikely to be justified.

Consequently, the appropriate framework for this application is the model defined by (3.5) and (3.6). Following Johansen and Juselius (1992) we set $p = 2$ but did not include seasonal dummies in the model. The seasonal effects were only marginally significant and given the limited sample size available it was thought best to leave them out. With the foreign price and oil price variables, p_t^* and po_t , treated as $I(1)$ and exogenous, we estimated the following system of equations over the period 1972(3) -1987(2):

$$\Delta \mathbf{y}_t = \mathbf{c}_0 + (-\Pi_y \boldsymbol{\gamma})t + \Lambda \Delta \mathbf{x}_t + \Psi_1 \Delta \mathbf{z}_{t-1} + \Pi_y \mathbf{z}_{t-1} + \mathbf{u}_t, \quad (7.1)$$

$$\Delta \mathbf{x}_t = \mathbf{a}_{x0} + \Gamma_{x1} \Delta \mathbf{z}_{t-1} + \mathbf{e}_{xt},$$

where $\mathbf{y}_t = (p_t, e_t, r_t, r_t^*)'$, $\mathbf{x}_t = (p_t^*, po_t)'$, $\mathbf{z}_t = (\mathbf{y}_t', \mathbf{x}_t')'$, Π_y and Ψ_1 are each $(4, 6)$ coefficient matrices, and Λ a $(4, 2)$ coefficient matrix. Note that in the above specification the trend coefficients $(-\Pi_y \boldsymbol{\gamma})$ are restricted as in Case IV defined by (3.12) and, hence, the level of $\mathbf{y}_t$ will exhibit linear deterministic trends for all values of the cointegration rank $r = \text{Rank}(\Pi_y)$. These restrictions on the trend coefficients can be tested using standard chi-squared tests set

out in Theorem 4.5. The (log-) likelihood ratio statistics for testing the trend restrictions (Case IV) against the unrestricted trend specification (Case V) for $r = 0, 1, 2, 3$ are 7.61[9.49], 7.26[7.81], 4.73[5.99], and 2.24[3.84], respectively. The figures in [.] are the 0.05 critical values of the chi-squared distribution with $4 - r$ degrees of freedom. Therefore, irrespective of the value of r the trend restrictions cannot be rejected.

Table 3 reports the cointegration rank test statistics defined by $\mathcal{LR}(H_r|H_n)$ and $\mathcal{LR}(H_r|H_{r+1})$ of (4.7) and (4.5) respectively, together with the corresponding asymptotic critical values at the 0.05 and 0.10 significance levels reproduced from Table T.4 with $k = 2$ and $n - r = 4, 3, 2, 1$.

Table 3 about here

Neither statistic rejects the hypothesis of no cointegration at the 0.05 significance level but $\mathcal{LR}(H_r|H_n)$ rejects the null of no cointegration at the 0.10 level.²⁰ If we confined our analysis solely to these statistical tests, we could only conclude that there is *at most* one cointegrating relation between the six $I(1)$ variables under investigation.

However, economic theory suggests the existence of two long-run (cointegrating) relations in the above system, namely the PPP and UIP relations defined as $p_t - e_t - p_t^*$ and $r_t - r_t^*$ respectively. Since in this application the sample size is relatively small ($T = 60$) and the dimension of the system is relatively large ($m = 6$), the evidence against the hypothesis of two cointegrating relations does not seem to be particularly strong. Moreover, the Monte Carlo simulation results of section 6 indicate that these cointegrating rank test statistics generally tend to under-reject in small samples. We therefore examine the case in which restrictions on short-run dynamics are available and are imposed in the construction of the (log-) likelihood ratio tests of the cointegrating rank restrictions; see section 5. A preliminary analysis of the regressions for Δp_t , Δe_t , Δr_t and Δr_t^* suggest the following 24 zero restrictions on the short-run coefficients:

$$\Lambda = \begin{bmatrix} 0 & * \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}; \Psi_1 = \begin{bmatrix} * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & * & 0 & 0 & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & * & 0 \end{bmatrix}, \quad (7.2)$$

where unrestricted coefficients are denoted by *. The maximized log-likelihood values of the restricted and the unrestricted model together with the associated Akaike Information Criterion (AIC) and the Schwarz Bayesian Criterion (SBC) are given by

	Unrestricted	Restricted
LL	725.01	716.49
AIC	661.01	676.49
SBC	593.99	634.61

²⁰These results are based on the assumption that p^* and po are $I(1)$ and non-cointegrated. However, it is possible to test this assumption (see Theorems 4.7 and 4.8). Using an unrestricted trend VAR(3) model in these exogenous variables, augmented with two lagged changes in the endogenous variables we could not reject the hypothesis that p^* and po are non-cointegrated. The choice of the order of the VAR in this application was based on the Schwarz Bayesian Criterion and the choice of trend specification was based on the specification test of Theorem 4.5.

It is clear that the 24 restrictions are not rejected and that both AIC and SBC also select the restricted model.

Table 4 reports the cointegration rank statistics $\widetilde{\mathcal{LR}}(H_r|H_n)$ and $\widetilde{\mathcal{LR}}(H_r|H_{r+1})$ of section 5 based on the above restricted model.

Table 4 about here

These statistics do not reject the null hypothesis of $r = 2$ which is in line with the prediction of economic theory. Consequently, we proceed as if there are two cointegrating relations. We return subsequently to re-consider the case of cointegrating rank $r = 1$ to see the extent to which our conclusions based on $r = 2$ regarding the long-run validity of the PPP hypothesis are affected by this choice of r .

We now examine the validity of the PPP and UIP hypotheses using the long-run structural modelling techniques advanced in Pesaran and Shin (1999). This approach enables us to test the validity of these hypotheses and to identify factors that might be responsible for their possible breakdown. Denote the two cointegrating vectors comprising $\beta^* = (-\gamma, \mathbf{I}_m)' \beta$ associated with $\mathbf{z}_t^* = (t, p_t, e_t, r_t, r_t^*, p_t^*, po_t)'$ by $\beta_1^* = (\beta_{01}^*, \beta_{11}, \beta_{21}, \beta_{31}, \beta_{41}, \beta_{51}, \beta_{61})'$ and $\beta_2^* = (\beta_{02}^*, \beta_{12}, \beta_{22}, \beta_{32}, \beta_{42}, \beta_{52}, \beta_{62})'$ respectively viewing β_1^* as explaining domestic prices and β_2^* as explaining the domestic interest rate. Exact identification of these vectors requires the imposition of two restrictions per vector. We chose the following exactly identifying constraints:

$$H_E : \beta_*' = \begin{pmatrix} * & 1 & * & * & 0 & * & * \\ * & 0 & * & 1 & * & * & * \end{pmatrix},$$

which yielded the estimate:

$$\hat{\beta}_{*E}' = \begin{pmatrix} .0195 & 1 & -.701 & -6.005 & 0 & -2.564 & .206 \\ (.0229) & & (.345) & (2.088) & & (1.943) & (.274) \\ -0.0062 & 0 & -.091 & 1 & -.791 & 0.493 & -.054 \\ (.0038) & & (.095) & & (.326) & (0.334) & (.046) \end{pmatrix}, LL_E = 714.50,$$

where LL_E is the maximized value of the log-likelihood function for the just-identified case. Asymptotic standard errors are given in parentheses.²¹

A number of hypotheses of interest may now be tested using the above exactly identified model. Consider firstly the co-trending hypothesis H_{co} , namely that the trend coefficients are zero in the two cointegrating relations:

$$H_{co} : \beta_*' = \begin{pmatrix} 0 & 1 & * & * & 0 & * & * \\ 0 & 0 & * & 1 & * & * & * \end{pmatrix},$$

Under H_{co} , we obtained the estimate:

$$\hat{\beta}_{*co}' = \begin{pmatrix} 0 & 1 & -0.881 & -5.511 & 0 & -0.854 & -0.025 \\ & & (0.333) & (1.748) & & (0.280) & (0.048) \\ 0 & 0 & -0.037 & 1 & -0.929 & -0.067 & 0.023 \\ & & (0.101) & & (0.414) & (0.053) & (0.015) \end{pmatrix}, LL_{co} = 713.30.$$

²¹All the computations reported in this section are carried out using *Microfit* 4.0. See Pesaran and Pesaran (1997).

Therefore, the log-likelihood ratio statistic for testing the co-trending hypothesis is equal to $2(714.50 - 713.30) = 2.40$, which is below the 0.05 critical value of the chi-squared distribution with 2 degrees of freedom. Hence, the co-trending hypothesis is not rejected.

We now consider the hypothesis that level of oil prices do not enter the cointegrating relations which we denote by H_{po} .²² To save space, we only report the test of this hypothesis jointly with the co-trending hypothesis, $H_{co} \cap H_{po}$, which yields:

$$\hat{\beta}'_{*co,po} = \begin{pmatrix} 0 & 1 & -0.834 & -5.646 & 0 & -0.925 & 0 \\ & & (0.327) & (1.850) & & (0.227) & \\ 0 & 0 & -0.062 & 1 & -0.803 & -0.005 & 0 \\ & & (0.129) & & (0.432) & (0.063) & \end{pmatrix}, LL_{co,po} = 711.49.$$

The (log-) likelihood ratio statistic associated with these 4 over-identifying restrictions is 6.01 which, thus, does not reject $H_{co} \cap H_{po}$; similar results obtain for the individual test of H_{po} undertaken separately from H_{co} . Accordingly, in what follows we carry out tests of the PPP and the UIP hypotheses given $H_{co} \cap H_{po}$.

Under the UIP hypothesis and $H_{co} \cap H_{po}$:

$$H_{UIP} : \beta'_* = \begin{pmatrix} 0 & 1 & * & * & 0 & * & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix},$$

the estimate of which is:

$$\hat{\beta}'_{*UIP} = \begin{pmatrix} 0 & 1 & -0.965 & -5.132 & 0 & -0.943 & 0 \\ & & (0.281) & (1.257) & & (0.167) & \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix}, LL_{UIP} = 709.61.$$

Therefore, the (log-) likelihood ratio statistic for testing the UIP hypotheses is 9.78 which is well below the 0.05 critical value of the chi-squared distribution with 7 degrees of freedom.²³ Thus, the hypothesis of UIP (jointly with $H_{co} \cap H_{po}$) is not rejected even at the 0.01 level.

Next, testing the PPP hypothesis given $H_{co} \cap H_{po}$:

$$H_{PPP} : \beta'_* = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & * & 1 & * & * & 0 \end{pmatrix},$$

we obtained the value of 33.37 for the (log-) likelihood ratio statistic which is well above the 0.05 critical value of the chi-squared distribution with 7 degrees of freedom. Therefore, the PPP hypothesis considered jointly with $H_{co} \cap H_{po}$ is strongly rejected.

However, as can be easily deduced from the above exactly-identified estimation results, the rejection of the PPP hypothesis is mainly due to the statistically significant effect of domestic interest rates on the real exchange rate (defined by $e_t + p_t^* - p_t$). To see this,

²²Recall that the primary motivation behind the inclusion of oil prices in this model is to capture the short-run effects of changes in oil prices on the PPP and the UIP relations.

²³Notice that the number of over-identifying restrictions in this case is equal to 7 ($= 11 - 4$).

consider the following modified version of the PPP hypothesis (denoted by H_{PPP^*}) that allows the effect of domestic interest rate on the real exchange rate to be unrestricted:

$$H_{PPP^*} : \beta'_* = \begin{pmatrix} 0 & 1 & -1 & * & 0 & -1 & 0 \\ 0 & 0 & * & 1 & * & * & 0 \end{pmatrix}.$$

Under these restrictions, we obtained the following estimate:

$$\hat{\beta}'_{*PPP^*} = \begin{pmatrix} 0 & 1 & -1 & -5.363 & 0 & -1 \\ & & & (0.876) & & \\ 0 & 0 & -0.022 & 1 & -0.778 & -0.002 \\ & & (0.097) & & (0.325) & (0.045) \end{pmatrix}, \quad LL_{PPP^*} = 709.81.$$

The (log-) likelihood ratio statistic for testing H_{PPP^*} is 9.37 which is below the 0.05 critical value of the chi-squared distribution with 6 degrees of freedom. Thus, the main cause of the breakdown of the PPP hypothesis in the present application seems to be the existence of a statistically significant negative relation between the real exchange rate and the domestic interest rate.²⁴

Finally, we estimated the cointegrating relations under the modified PPP and the UIP restrictions (jointly with $H_{co} \cap H_{po}$) yielding the following estimate:

$$\hat{\beta}'_{*PPP^*,UIP} = \begin{pmatrix} 0 & 1 & -1 & -4.775 & 0 & -1 & 0 \\ & & & (0.627) & & & \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix}, \quad LL_{PPP^*,UIP} = 709.03. \quad (7.3)$$

The log-likelihood ratio statistic is now 10.93 which is well below the 0.05 critical value of the chi-squared distribution with 9 degrees of freedom. Hence, we are unable to reject the modified PPP and the UIP hypotheses.

Conditional on the above long-run estimates we have the following expressions for the error correction terms

$$\hat{\beta}'_1 \mathbf{z}_{t-1} = p_{t-1} - e_{t-1} - p_{t-1}^* - 4.775r_{t-1},$$

$$\hat{\beta}'_2 \mathbf{z}_{t-1} = r_{t-1} - r_{t-1}^*.$$

To check the resultant model's statistical adequacy, we also estimated the following VECM

$$\begin{aligned} \Delta \mathbf{y}_t &= \mathbf{c}_0 - \alpha_y \beta' \gamma t + \Lambda \Delta \mathbf{x}_t + \Psi_1 \Delta \mathbf{z}_{t-1} + \alpha_y \beta' \mathbf{z}_{t-1} + \mathbf{u}_t \\ &= \mathbf{c}_0 - \alpha_y \beta' \gamma t + \Lambda \Delta \mathbf{x}_t + \Psi_1 \Delta \mathbf{z}_{t-1} + \alpha_{1y} \beta'_1 \mathbf{z}_{t-1} + \alpha_{2y} \beta'_2 \mathbf{z}_{t-1} + \mathbf{u}_t, \end{aligned}$$

where α_{1y} and α_{2y} are 4-dimensional vectors of adjustment (error correction) coefficients. The estimates of these adjustment coefficients together with a number of diagnostic test

²⁴Similar results also follow if instead of the domestic interest rate the coefficient of the foreign interest rate variable is left unrestricted.

statistics are presented in Table 5 below.²⁵

Table 5 about here

The figures in $(\cdot)$ are estimated asymptotic standard errors whereas those in $[\cdot]$ are the corresponding p -values. The bold-faced estimates denote statistical significance at the 0.05 level. The error correction equations pass almost all of the diagnostic tests with the exception of the equation for Δr_t^* , which conclusively fails the normality test and marginally the residual serial correlation test. The domestic price equation performs best, explaining 0.78 of the price variation over the sample period. The error correction term associated with the first cointegrating relation explaining long-run price movements also has a significant but small negative impact on current price changes, suggesting an equilibrating but slow adjustment process for UK prices in response to changes in the domestic interest rate, the exchange rate and foreign prices.

We now return to consider the estimation results under cointegration rank $r = 1$ as suggested by the test results in Table 3. For this case, we write $\beta_* = (\beta_0^*, \beta_1, \dots, \beta_5, \beta_6)'$. Under the just-identifying normalization restriction $\beta_1 = 1$, we obtained the estimates

$$\hat{\beta}'_* = \begin{pmatrix} 0.004 & 1 & -0.923 & -3.567 & -1.929 & -1.362 & 0.074 \\ (0.018) & & (0.323) & (1.891) & (1.434) & (1.553) & (0.207) \end{pmatrix}, \quad LL = 701.70.$$

Only the coefficients of e_t and r_t are statistically significant. Imposing the restrictions implied by $H_{co} \cap H_{po}$ and setting the coefficient of r_t^* to zero yields

$$\hat{\beta}'_* = \begin{pmatrix} 0 & 1 & -0.782 & -5.198 & 0 & -0.998 & 0 \\ & & (0.273) & (1.519) & & (0.182) & \end{pmatrix}, \quad LL = 700.43.$$

The log-likelihood ratio statistic for testing these three restrictions is 2.52. Further, imposing the restrictions that the coefficients of e_t and p_t^* are both equal to -1 , as indicated by the PPP hypothesis, we finally obtain

$$\hat{\beta}'_{*PPP} = \begin{pmatrix} 0 & 1 & -1 & -5.401 & 0 & -1 & 0 \\ & & & (0.861) & & & \end{pmatrix}, \quad LL = 699.26.$$

This result is remarkably close to that obtained assuming $r = 2$ for the modified PPP relation given above by (7.3). Therefore, as far as the test of the PPP hypothesis is concerned our main conclusion would seem to be unaffected by assuming the cointegration rank to be $r = 1$ or $r = 2$. The above result continues to hold if instead of setting $\beta_1 = 1$, we normalize on the coefficients of e_t , r_t or p_t^* . Clearly, we cannot normalize on the coefficient of r_t^* if we wish to test its possible significance in the PPP relationship.

However, normalizing on the coefficient of r_t^* is not inappropriate if we believe that the single cointegrating relationship in fact represents the *UIP* hypothesis. Using this normalization and estimating the cointegrating relation subject to the UIP restrictions (jointly with

²⁵Full estimation results are available on request. $\chi^2_{SC}(4)$ is the Lagrange multiplier statistic for testing the null of no serial correlation, $\chi^2_{FF}(1)$ is Ramsey's RESET test statistic, $\chi^2_N(2)$ is the Jarque-Bera statistic for testing the null of Gaussian errors, and $\chi^2_H(1)$ is the statistic for testing the null of no heteroskedasticity. The number in $(\cdot)$ indicates the degrees of freedom.

$H_{co} \cap H_{po}$), namely, $\beta_3 = 1$ and $\beta_0^* = \beta_1 = \beta_2 = \beta_5 = \beta_6 = 0$, we obtain:

$$\hat{\beta}'_{*UIP} = (0 \quad 0 \quad 0 \quad -1 \quad 1 \quad 0 \quad 0), \quad LL = 692.93.$$

The log-likelihood ratio statistic for testing these 6 restrictions is equal to $2(701.70 - 692.93) = 17.54$ and strongly rejects the UIP hypothesis if $r = 1$. We also tried a number of modifications of the UIP hypothesis but they were all rejected. Thus, we conclude that with $r = 1$ only a modified version of the PPP hypothesis would appear to be compatible with the data.

8 Concluding Remarks

The results derived in this paper are intended to complement the pioneering contributions of Johansen reviewed recently in Johansen (1995). We extend his work in two directions: we relax the assumption that all the $I(1)$ variables in the system are endogenously determined and allow for parametric restrictions on the coefficients of the stationary variables in the vector error correction form of the cointegrating VAR model. We also highlight the importance of appropriate specifications of intercepts and/or trend coefficients in these models. Our re-examination of the validity of the PPP and UIP relations provides an illustration of the utility of these extensions in practice where the available time series observations are often limited relative to the size of the VAR models being estimated. As with many other contributions in this area our results are asymptotic. In this paper we also provide some limited Monte Carlo evidence on the small sample performance of alternative tests of the cointegrating rank restrictions. These results are encouraging and suggest that tests that allow for the presence of exogenous $I(1)$ variables and/or restrictions on the short-run coefficients are likely to perform better in small samples. However, more work in this area is clearly necessary before more definite conclusions can be reached.

Appendix A: Proofs of Results

In order to prove Theorems 4.1-4.9 and 5.1, we need a number of additional subsidiary results. We consider Case IV and briefly detail the alterations necessary for the other cases below Lemma A.1.

Define the $(m+1, m-r)$ matrix δ as

$$\delta \equiv \begin{pmatrix} -\gamma' \\ \mathbf{I}_m \end{pmatrix} \beta_{\perp},$$

where $\beta_{\perp}$ is a $(m, m-r)$ matrix whose columns are a basis for the orthogonal complement of β . Hence, $(\beta, \beta_{\perp})$ is a basis for $\mathcal{R}^m$. Let ξ be the $(m+1)$ -unit vector $(1, \mathbf{0}')'$. Then, (β_*, ξ, δ) is a basis for $\mathcal{R}^{m+1}$. It therefore follows from (2.11) that for $\frac{t-1}{T} \leq a < \frac{t}{T}$, $a \in [0, 1]$

$$T^{-1/2} \delta' \mathbf{z}_t^* = T^{-1/2} \beta'_{\perp} \mu + \beta'_{\perp} \mathbf{C} \mathbf{S}_t^e + \beta'_{\perp} T^{-1/2} \mathbf{C}^*(L) \mathbf{e}_t \Rightarrow \beta'_{\perp} \mathbf{C} \mathbf{B}_m(a), \quad (\text{A.1})$$

see Phillips and Solo (1992, Theorem 3.15, p.983). Also

$$T^{-1} \xi' \mathbf{z}_t^* = T^{-1} t \Rightarrow a. \quad (\text{A.2})$$

Similarly, we note that

$$\beta'_* \mathbf{z}_t^* = \beta' \mu + \beta' \mathbf{C}^*(L) \mathbf{e}_t = O_P(1). \quad (\text{A.3})$$

Hence, from (A.3) and Phillips and Solo (1992, Theorem 3.16, p.983), it follows that

$$\beta'_* \mathbf{S}_{ZZ} \beta_* \xrightarrow{P} \Sigma_{\beta\beta},$$

where $\Sigma_{\beta\beta}$ is a (r, r) positive definite matrix. Similarly, as $\beta'_* \mathbf{S}_{ZY} = \beta'_* \mathbf{S}_{ZZ} \Pi'_{y*} + T^{-1} \beta'_* \hat{\mathbf{Z}}_{-1}^* \mathbf{U}'$,

$$\beta'_* \mathbf{S}_{ZY} \xrightarrow{P} \Sigma_{\beta y} = \Sigma_{\beta\beta} \alpha'_y,$$

and, as $\mathbf{S}_{YY} = T^{-1} (\Pi_{y*} \hat{\mathbf{Z}}_{-1}^* + \mathbf{U}) (\bar{\mathbf{P}}_t - \bar{\mathbf{P}}_t \Delta \mathbf{Z}'_{-} (\Delta \mathbf{Z}_{-} \bar{\mathbf{P}}_t \Delta \mathbf{Z}'_{-})^{-1} \Delta \mathbf{Z}_{-} \bar{\mathbf{P}}_t) (\Pi_{y*} \hat{\mathbf{Z}}_{-1}^* + \mathbf{U})'$,

$$\mathbf{S}_{YY} \xrightarrow{P} \Sigma_{yy} = \Omega_{uu} + \alpha_y \Sigma_{\beta\beta} \alpha'_y.$$

Cf. Johansen (1991, Lemma A.3, p.1569) and Johansen (1995, Lemma 10.3, p.146).

Lemma A.1 *Let $\mathbf{B}_T \equiv (\delta, T^{-1/2} \xi)$ and define $\mathbf{G}(a) = (\mathbf{G}_1(a)', \mathbf{G}_2(a)')$, where $\mathbf{G}_1(a) \equiv \beta'_{\perp} \mathbf{C} \tilde{\mathbf{B}}_m(a)$, $\tilde{\mathbf{B}}_m(a) = (\tilde{\mathbf{B}}_n(a)', \tilde{\mathbf{B}}_k(a)')' = \mathbf{B}_m(a) - \int_0^1 \mathbf{B}_m(a) da$, and $\mathbf{G}_2(a) \equiv a - \frac{1}{2}$, $a \in [0, 1]$. Then*

$$T^{-1} \mathbf{B}'_T \mathbf{S}_{ZZ} \mathbf{B}_T \Rightarrow \int_0^1 \mathbf{G}(a) \mathbf{G}(a)' da, \quad (\text{A.4})$$

$$\mathbf{B}'_T (\mathbf{S}_{ZY} - \mathbf{S}_{ZZ} \Pi'_{y*}) \Rightarrow \int_0^1 \mathbf{G}(a) d\tilde{\mathbf{B}}_n^*(a)', \quad (\text{A.5})$$

where $\tilde{\mathbf{B}}_n^*(a) \equiv [\tilde{\mathbf{B}}_n(a) - \Omega_{yx} \Omega_{xx}^{-1} \tilde{\mathbf{B}}_k(a)]$, $a \in [0, 1]$, and $\mathbf{S}_{ZZ}$, $\mathbf{S}_{ZY}$ and $\mathbf{S}_{YY}$ are defined in (3.21).

Proof. These statements follow from (A.1) and (A.2) by applying the results of Phillips and Solo (1992, Theorem 3.15, p.983) and using (4.1) and (4.2). Cf. Johansen (1991, Lemma A.3, p.1569) and Johansen (1995, Lemma 10.3, p.146) and Phillips and Durlauf (1986). ■

For the remaining cases, we need only make minor modifications to Case IV. In Case I, $\delta = \beta_{\perp}$ with $(\beta, \beta_{\perp})$ a basis for $\mathcal{R}^m$ and $\mathbf{B}_T = \delta$. For Case II, where $\mathbf{Z}_{-1}^* = (\iota_T, \mathbf{Z}'_{-1})'$, we have $\beta_* = \begin{pmatrix} -\mu' \\ \mathbf{I}_m \end{pmatrix} \beta$ and,

consequently, we define ξ as in Case IV, $\delta = \begin{pmatrix} -\mu' \\ \mathbf{I}_m \end{pmatrix} \beta_\perp$ and $\mathbf{B}_T = (\delta, \xi)$. Case III is similar to Case I as is Case V.

Proof of Theorems 4.1 and 4.2: See Pesaran, Shin and Smith (1997, Appendix A). ■

Lemmas A.2-A.4 below follow the analysis of Johansen (1995, Chapter 13, pp. 177-200). See also Johansen (1994). We concentrate on the representation for the limiting distribution of the ML estimator $\hat{\beta}_*$ in Case IV under $H_r : \text{rank}[\Pi_y] = r$ of (3.8).

Firstly, we note that we may decompose the ML estimator $\hat{\beta}_*$ via orthogonal projection as

$$\hat{\beta}_* = \beta_* \bar{\beta}'_* \hat{\beta}_* + \delta \bar{\delta}' \hat{\beta}_* + \xi \bar{\xi}' \hat{\beta}_*,$$

where $\bar{\beta}_* \equiv \mathbf{P}_{(\delta, \xi)}^\perp \beta_* \left(\beta'_* \mathbf{P}_{(\delta, \xi)}^\perp \beta_* \right)^{-1}$ with $\mathbf{P}_{(\delta, \xi)}^\perp$ the orthogonal projector onto the orthogonal complement of (δ, ξ) and $\bar{\delta}$ and $\bar{\xi}$ defined similarly. We then define

$$\begin{aligned} \tilde{\beta}_* &\equiv \hat{\beta}_* \left(\bar{\beta}'_* \hat{\beta}_* \right)^{-1} \\ &= \beta_* + \delta \bar{\delta}' \tilde{\beta}_* + \xi \bar{\xi}' \tilde{\beta}_* = \beta_* + \mathbf{B}_T \mathbf{U}_T, \end{aligned} \quad (\text{A.6})$$

where $\mathbf{B}_T$ is defined in Lemma A.1 and

$$\mathbf{U}_T \equiv \left(\bar{\delta}, T^{1/2} \bar{\xi} \right)' \tilde{\beta}_*;$$

cf. Johansen (1995, (13.1), (13.2), p.179). The normalization $\bar{\beta}'_* \tilde{\beta}_* = \mathbf{I}_r$ provides the r^2 restrictions required to just-identify $\tilde{\beta}_*$. The ML estimator for Ω_{uu} is given from (3.20) by

$$\hat{\Omega}_{uu} = T^{-1} \Delta \hat{\mathbf{Y}} \left(\mathbf{I}_T - \hat{\mathbf{Z}}_{-1}^* \hat{\beta}_* \left(\hat{\beta}'_* \hat{\mathbf{Z}}_{-1}^* \hat{\mathbf{Z}}_{-1}^* \hat{\beta}_* \right)^{-1} \hat{\beta}'_* \hat{\mathbf{Z}}_{-1}^* \right) \Delta \hat{\mathbf{Y}}'$$

Lemma A.2 Under H_r defined in (3.8) and Assumptions 2.1, 2.2, 3.1 and 4.1,

$$\tilde{\beta}_* - \beta_* = o_P(T^{-1/2}), \tilde{\beta}'_* \mathbf{S}_{ZZ} \tilde{\beta}_* = \Sigma_{\beta\beta} + o_P(1), \tilde{\beta}'_* \mathbf{S}_{ZY} = \Sigma_{\beta y} + o_P(1), \hat{\Omega}_{uu} = \Omega_{uu} + o_P(1).$$

Proof. Defining $\mathbf{A}_T \equiv (\beta_*, T^{-1/2} \mathbf{B}_T)$ and noting that $\mathbf{A}_T^{-1} = (\bar{\beta}_*, T^{1/2} \bar{\delta}, T \bar{\xi})'$, these results are obtained by the application of a similar argument to that used in the proof of Lemma 13.1 of Johansen (1995, pp.180-1). ■

Lemma A.3 Under H_r defined in (3.8) and Assumptions 2.1, 2.2, 3.1 and 4.1,

$$T \mathbf{U}_T = \left(T \bar{\delta}, T^{3/2} \bar{\xi} \right)' \left(\tilde{\beta}_* - \beta_* \right) \Rightarrow \left(\int_0^1 \mathbf{G}(a) \mathbf{G}(a)' da \right)^{-1} \int_0^1 \mathbf{G}(a) d\mathbf{B}_n^*(a)' \Omega_{uu}^{-1} \alpha_y (\alpha_y' \Omega_{uu}^{-1} \alpha_y)^{-1}, \quad (\text{A.7})$$

which is mixed normal with conditional variance given by

$$(\alpha_y' \Omega_{uu}^{-1} \alpha_y)^{-1} \otimes \left(\int_0^1 \mathbf{G}(a) \mathbf{G}(a)' da \right)^{-1}, \quad (\text{A.8})$$

where $\mathbf{B}_n^*(a) \equiv [\mathbf{B}_n(a) - \Omega_{yx} \Omega_{xx}^{-1} \mathbf{B}_k(a)]$ and $\mathbf{G}(a)$ is defined in Lemma A.1, $a \in [0, 1]$. Hence, $(\tilde{\beta}_* - \beta_*) = O_P(T^{-1})$.

Proof. It is straightforward to show that $\tilde{\beta}_*$ satisfies the likelihood equations for $\hat{\beta}_*$ derived from (3.20); viz.

$$\tilde{\alpha}'_y \hat{\Omega}_{uu}^{-1} [\mathbf{S}_{YZ} - \tilde{\alpha}_y \tilde{\beta}'_* \mathbf{S}_{ZZ}] = \mathbf{0}, \quad (\text{A.9})$$

where $\tilde{\alpha}_y \equiv \hat{\alpha}_y(\hat{\beta}_*) \hat{\beta}'_* \bar{\beta}_*$ and, from Lemma A.2, $\tilde{\alpha}_y - \alpha_y = o_P(1)$. Equation (A.9) after post-multiplication by $\mathbf{B}_T$ may be rearranged as

$$\tilde{\alpha}'_y \hat{\Omega}_{uu}^{-1} [(\mathbf{S}_{YZ} - \alpha_y \beta'_* \mathbf{S}_{ZZ}) - \tilde{\alpha}_y (\tilde{\beta}_* - \beta'_*)' \mathbf{S}_{ZZ} - (\tilde{\alpha}_y - \alpha_y) \beta'_* \mathbf{S}_{ZZ}] \mathbf{B}_T = \mathbf{0}.$$

Now, $\beta'_* \mathbf{S}_{ZZ} \mathbf{B}_T = O_P(1)$; cf. Johansen (1995, Lemma 10.3, (10.18), p.146). Substituting for $\tilde{\beta}_* - \beta_*$ from (A.6) yields

$$\begin{aligned} T\mathbf{U}_T &= (T^{-1} \mathbf{B}'_T \mathbf{S}_{ZZ} \mathbf{B}_T)^{-1} \mathbf{B}'_T (\mathbf{S}_{YZ} - \mathbf{S}_{ZZ} \beta'_* \alpha'_y) \Omega_{uu}^{-1} \alpha_y (\alpha'_y \Omega_{uu}^{-1} \alpha_y)^{-1} + o_P(1) \\ &\Rightarrow \left(\int_0^1 \mathbf{G}(a) \mathbf{G}(a)' da \right)^{-1} \int_0^1 \mathbf{G}(a) d\tilde{\mathbf{B}}_n^*(a) \Omega_{uu}^{-1} \alpha_y (\alpha'_y \Omega_{uu}^{-1} \alpha_y)^{-1}, \end{aligned}$$

by the consistency of $\hat{\Omega}_{uu}$, Lemma A.1 and $\tilde{\alpha}_y - \alpha_y = o_P(1)$. The first equality in (A.7) results from the definition of $\mathbf{U}_T$ and the representation (A.7) follows by noting that $\tilde{\mathbf{B}}_n^*(.)$ may be replaced by $\mathbf{B}_n^*(.)$ in (A.5). As $\mathbf{G}(.)$ is independent of $(\alpha'_y \Omega_{uu}^{-1} \alpha_y)^{-1} \alpha'_y \Omega_{uu}^{-1} \mathbf{B}_n^*(.)$, $T\mathbf{U}_T$ is asymptotically mixed normal with conditional variance (A.8). Cf. Johansen (1995, Lemma 13.2, pp.181-3). ■

Lemma A.4 Under H_r defined in (3.8) and Assumptions 2.1, 2.2, 3.1 and 4.1, the likelihood ratio statistic

$$\begin{aligned} &-2[\ell_T^c(\beta_*; r) - \ell_T^c(\hat{\beta}_*; r)] \\ &\Rightarrow \text{Trace} \left\{ (\alpha'_y \Omega_{uu}^{-1} \alpha_y)^{-1} \alpha'_y \Omega_{uu}^{-1} \int_0^1 d\mathbf{B}_n^*(a) \mathbf{G}(a)' \left(\int_0^1 \mathbf{G}(a) \mathbf{G}(a)' da \right)^{-1} \int_0^1 \mathbf{G}(a) d\mathbf{B}_n^*(a) \Omega_{uu}^{-1} \alpha_y \right\} \quad (\text{A.10}) \end{aligned}$$

which is distributed as a chi-squared random variable with $r(m - r + 1)$ degrees of freedom.

Proof. The proof mirrors that of Lemma 13.8 of Johansen (1995, pp.192-4). Applications of Lemmas A.2 and A.3 allow us to show that $-2[\ell_T^c(\beta_*; r) - \ell_T^c(\hat{\beta}_*; r)]$ equals

$$\begin{aligned} &= T \times \text{Trace} \left\{ \left[(\Sigma_{\beta\beta} - \Sigma_{\beta y} \Sigma_{yy}^{-1} \Sigma_{y\beta})^{-1} - \Sigma_{\beta\beta}^{-1} \right] (\tilde{\beta}_* - \beta'_*)' \mathbf{S}_{ZZ} (\tilde{\beta}_* - \beta'_*) \right\} + O_P(T^{-1}) \\ &= \text{Trace} \left\{ \alpha'_y \Omega_{uu}^{-1} \alpha_y T\mathbf{U}'_T (T^{-1} \mathbf{B}'_T \mathbf{S}_{ZZ} \mathbf{B}_T) T\mathbf{U}_T \right\} + O_P(T^{-1}); \end{aligned}$$

cf. Johansen (1995, (10.9), p.142). The representation (A.10) follows from the representations (A.5) of Lemma A.1 and (A.7) of Lemma A.3. Now, conditional on $\mathbf{G}(.)$, the random $(m - r + 1)r$ -vector $\text{vec} \left(\alpha'_y \Omega_{uu}^{-1} \int_0^1 d\mathbf{B}_n^*(a) \mathbf{G}(a)' \right)$ is normally distributed with mean zero and conditional variance matrix

$$\int_0^1 \mathbf{G}(a) \mathbf{G}(a)' da \otimes \alpha'_y \Omega_{uu}^{-1} \alpha_y.$$

Hence, (A.10) is conditionally $\chi^2[r(m - r + 1)]$ which is independent of $\mathbf{G}(.)$ which implies the result holds unconditionally. ■

Proof of Theorems 4.3 and 4.4: We discuss the proof of Theorem 4.4; Theorem 4.3 follows similarly. In Case III and under H_r , analogously to (A.10), the likelihood ratio statistic $-2[\ell_T^c(\beta; r) - \ell_T^c(\hat{\beta}; r)]$, where $\hat{\beta}$ denotes the ML estimator for β in Case III, has the limiting representation

$$\text{Trace} \left\{ (\alpha'_y \Omega_{uu}^{-1} \alpha_y)^{-1} \alpha'_y \Omega_{uu}^{-1} \int_0^1 d\mathbf{B}_n^*(a) \mathbf{G}_1(a)' \left(\int_0^1 \mathbf{G}_1(a) \mathbf{G}_1(a)' da \right)^{-1} \int_0^1 \mathbf{G}_1(a) d\mathbf{B}_n^*(a) \Omega_{uu}^{-1} \alpha_y \right\},$$

[A.3]

which is $\chi^2[r(m-r)]$ distributed. The likelihood ratio statistic for $\mathbf{c}_1 = \mathbf{0}$ is given by $-2[\ell_T^c(\hat{\beta}; r) - \ell_T^c(\hat{\beta}_*; r)]$ which by partitioned inversion of $\left(\int_0^1 \mathbf{G}(a)\mathbf{G}(a)'da\right)^{-1}$ has the limiting representation under H_r and $\mathbf{c}_1 = \mathbf{0}$ given by

$$\text{Trace} \left\{ (\alpha_y' \Omega_{uu}^{-1} \alpha_y)^{-1} \alpha_y' \Omega_{uu}^{-1} \int_0^1 d\mathbf{B}_n^*(a) \mathbf{G}_{2.1}(a) \left(\int_0^1 \mathbf{G}_{2.1}(a)^2 ds \right)^{-1} \int_0^1 \mathbf{G}_{2.1}(a) d\mathbf{B}_n^*(a)' \Omega_{uu}^{-1} \alpha_y \right\},$$

where $\mathbf{G}_{2.1}(a) \equiv \mathbf{G}_2(a) - \int_0^1 \mathbf{G}_2(a) \mathbf{G}_1(a)' da \left(\int_0^1 \mathbf{G}_1(a) \mathbf{G}_1(a)' da \right)^{-1} \mathbf{G}_1(a)$, $a \in [0, 1]$. Now, conditional on $\mathbf{G}(\cdot)$, the random r -vector $\int_0^1 \mathbf{G}_{2.1}(a) d\mathbf{B}_n^*(a)$ is normally distributed with mean zero and conditional variance matrix $\alpha_y' \Omega_{uu}^{-1} \alpha_y \left(\int_0^1 \mathbf{G}_{2.1}(a)^2 da \right)$. Hence, in Case IV, under H_r and $\mathbf{c}_1 = \mathbf{0}$, the likelihood ratio statistic $-2[\ell_T^c(\hat{\beta}; r) - \ell_T^c(\hat{\beta}_*; r)]$ has a limiting $\chi^2(r)$ distribution. ■

Proofs of Theorems 4.5 and 4.6: We discuss the proof of Theorem 4.6; Theorem 4.5 follows by a similar argument. The likelihood ratio statistic is the difference of likelihood ratio statistics for H_r against H_n in Cases IV and V as the maximized likelihoods for H_n are identical in both Cases. Hence, from Theorem 4.2, the limiting representation for the likelihood ratio statistic for H_r against H_n in Case IV is given by partitioned inversion as

$$12 \int_0^1 \left(a - \frac{1}{2} \right) d\mathbf{W}_{n-r}(a)' \int_0^1 \left(a - \frac{1}{2} \right) d\mathbf{W}_{n-r}(a) \\ + \text{Trace} \left\{ \int_0^1 d\mathbf{W}_{n-r}(a) \hat{\mathbf{W}}_{m-r}(a)' \left[\int_0^1 \hat{\mathbf{W}}_{m-r}(a) \hat{\mathbf{W}}_{m-r}(a)' da \right]^{-1} \int_0^1 \hat{\mathbf{W}}_{m-r}(a) d\mathbf{W}_{n-r}(a)' \right\},$$

the latter being the limiting representation for the likelihood ratio statistic for H_r against H_n in Case V. As $\int_0^1 \left(a - \frac{1}{2} \right) d\mathbf{W}_{n-r}(a) \sim N(\mathbf{0}, \frac{1}{12} \mathbf{I}_{n-r})$, the result follows. Cf. the proof of Theorem 2.3 of Johansen (1991, p.1571). ■

Proof of Theorems 4.7 and 4.8: These results follow by similar arguments to those in the Proofs of Theorems 4.1 and 4.2 given in Pesaran, Shin and Smith (1997, Appendix A). ■

Proof of Theorem 4.9: By Lemma A.3, $\tilde{\beta}_* - \beta_* = O_P(T^{-1})$. Hence, asymptotically, the inclusion of the lagged estimated cointegrating relationships $\tilde{\beta}_* \mathbf{z}_{t-1}^*$ is the same as the inclusion of the actual cointegrating relationships $\beta_* \mathbf{z}_{t-1}^*$, $t = 1, \dots, T$. Therefore, as the regression contains only variables which are stationary, the result follows; see Phillips and Durlauf (1986) and Phillips and Solo (1992). ■

Proof of Theorem 5.1: By a similar argument to that above Lemma A.1, $\beta_*' \tilde{\mathbf{S}}_{ZZ} \beta_* \xrightarrow{p} \tilde{\Sigma}_{\beta\beta}$, where $\tilde{\Sigma}_{\beta\beta}$ is a (r, r) positive definite matrix. As $\beta_*' \tilde{\mathbf{S}}_{ZY} = \beta_*' \tilde{\mathbf{S}}_{ZZ} \Pi_{y*}' + T^{-1} \beta_*' \tilde{\mathbf{Z}}_{-1}^* \tilde{\mathbf{U}}'$, $T^{-1} \beta_*' \tilde{\mathbf{Z}}_{-1}^* \tilde{\mathbf{U}}' = T^{-1} \beta_*' \tilde{\mathbf{Z}}_{-1}^* \mathbf{U}' - T^{-1} \beta_*' \tilde{\mathbf{Z}}_{-1}^* \Delta \mathbf{Z}'_-(\tilde{\Psi} - \Psi)'$ and $T^{1/2}(\tilde{\Psi} - \Psi) = O_P(1)$, $\beta_*' \tilde{\mathbf{S}}_{ZY} \xrightarrow{p} \tilde{\Sigma}_{\beta y} = \tilde{\Sigma}_{\beta\beta} \alpha_y'$. Similarly, as $\tilde{\mathbf{S}}_{Y Y} = T^{-1}(\Pi_{y*} \tilde{\mathbf{Z}}_{-1}^* + \tilde{\mathbf{U}}) \bar{\mathbf{P}}_t (\Pi_{y*} \tilde{\mathbf{Z}}_{-1}^* + \tilde{\mathbf{U}})'$, $\tilde{\mathbf{S}}_{Y Y} \rightarrow^P \tilde{\Sigma}_{yy} = \Omega_{uu} + \alpha_y \tilde{\Sigma}_{\beta\beta} \alpha_y'$. It also straightforwardly seen that $T^{-1} \mathbf{B}_T' \tilde{\mathbf{S}}_{ZZ} \mathbf{B}_T = T^{-2} \mathbf{B}_T' \mathbf{Z}_{-1}^* \bar{\mathbf{P}}_t \mathbf{Z}_{-1}^{*'} \mathbf{B}_T \Rightarrow \int_0^1 \mathbf{G}(a) \mathbf{G}(a)' da$ and $\mathbf{B}_T' (\tilde{\mathbf{S}}_{ZY} - \tilde{\mathbf{S}}_{ZZ} \Pi_{y*}') = T^{-1} \mathbf{B}_T' \mathbf{Z}_{-1}^* \bar{\mathbf{P}}_t \tilde{\mathbf{U}}' \Rightarrow \int_0^1 \mathbf{G}(a) d\tilde{\mathbf{B}}_n^*(a)'$, where $\mathbf{B}_T$ is defined in Lemma A.1; cf. (A.4) and (A.5). Therefore, the result obtains by employing similar arguments to those in the Proof of Theorems 4.1 and 4.2 given in Pesaran, Shin and Smith (1997, Appendix A).

Appendix B. Computation of Critical Values

The calculations of asymptotic critical values for the (log-) likelihood ratio cointegration rank statistics $\mathcal{LR}(H_r|H_{r+1})$ and $\mathcal{LR}(H_r|H_n)$, defined by (4.5) and (4.7) respectively, are carried out as follows.

For a given cointegrating rank hypothesis H_r defined by (3.8), we generate $m - r$ independent random walk processes

$$\mathbf{z}_t = \mathbf{z}_{t-1} + \mathbf{e}_t, \quad t = 1, 2, \dots, T,$$

with $\mathbf{z}_0 = \mathbf{0}$ and $\mathbf{e}_t \sim IN(\mathbf{0}, \mathbf{I}_m)$. We then partition $\mathbf{z}_t = (\mathbf{y}'_t, \mathbf{x}'_t)'$ where $\mathbf{y}_t$ and $\mathbf{x}_t$ are respectively $(n - r)$ - and k -vectors of random walk variables and $\mathbf{e}_t = (\mathbf{w}'_t, \mathbf{v}'_t)'$ is partitioned conformably with $\mathbf{z}_t = (\mathbf{y}'_t, \mathbf{x}'_t)'$, $t = 1, 2, \dots, T$. For Cases I-V, we compute the following statistic

$$\sum_{t=1}^T \mathbf{w}_t \mathbf{f}'_t \left(\sum_{t=1}^T \mathbf{f}_t \mathbf{f}'_t \right)^{-1} \sum_{t=1}^T \mathbf{f}_t \mathbf{w}'_t, \quad (\text{B.1})$$

where

$$\mathbf{f}_t \equiv \left\{ \begin{array}{ll} \mathbf{z}_{t-1} & \text{Case I} \\ (\mathbf{z}'_{t-1}, 1)' & \text{Case II} \\ \tilde{\mathbf{z}}_{t-1} & \text{Case III} \\ (\tilde{\mathbf{z}}'_{t-1}, t - \frac{1}{2})' & \text{Case IV} \\ \hat{\mathbf{z}}_{t-1} & \text{Case V} \end{array} \right\},$$

$\tilde{\mathbf{z}}_{t-1}$ is de-meaned $\mathbf{z}_{t-1}$ (or the OLS residual from the regression of $\mathbf{z}_{t-1}$ on 1) and $\hat{\mathbf{z}}_{t-1}$ is de-meaned and de-trended $\mathbf{z}_{t-1}$ (or the OLS residual from the regression of $\mathbf{z}_{t-1}$ on $(1, t)$), $t = 1, 2, \dots, T$. We compute the maximum eigenvalue and the trace of the matrix (B.1). Using $T = 500$, the above steps are repeated 10,000 times. Tables T.1- T.5 provide the 0.05 and 0.10 upper critical values from these empirical distributions, denoted $\lambda_{\max}$ and λ_{trace} respectively, as estimates of the corresponding critical values of the limiting distributions of the cointegration rank statistics $\mathcal{LR}(H_r|H_{r+1})$ and $\mathcal{LR}(H_r|H_n)$ for Cases I-V with $n - r = 1, 2, \dots, 12$ and $k = 0, \dots, 5$.

For Cases I, II and IV, the asymptotic critical values reported in Tables T.1, T.2 and T.4 are similar to those in Tables 3-6 of Harbo *et al.* (1998). However, for Cases III and V, Tables T.3 and T.5, the critical values when $k = 0$ differ from those tabulated in Johansen (1995, Chapter 15) due to differences in the specification of the deterministic components which are summarized below:

	Intercept		Linear Trend		Quadratic Trend	
	PSS	JOH	PSS	JOH	PSS	JOH
Case I	N	N	N	N	N	N
Case II	restricted	restricted	N	N	N	N
Case III	unrestricted	unrestricted	N	Y	N	N
Case IV	unrestricted	unrestricted	restricted	restricted	N	N
Case V	unrestricted	unrestricted	unrestricted	unrestricted	N	Y

In Case III, Johansen computes critical values including a linear trend in the VAR, but we do not. Similarly, in Case V, the critical values reported by Johansen (and Osterwald-Lenum (1992)) are based on a VAR model containing both linear and quadratic trends, but in our computations we only include linear trends with unrestricted coefficients. See also section 2 and the discussion in Perron and Campbell (1993). We note in passing that the critical values reported for Cases I, III and V may be regarded as multivariate generalizations of the t -statistics $\hat{\tau}$, $\hat{\tau}_\mu$ and $\hat{\tau}_\tau$ in Dickey and Fuller (1979) while those reported for Cases II and IV are multivariate equivalents of the F -statistics Φ_1 and Φ_3 in Dickey and Fuller (1981).

TABLE T.1*

Case I: No Intercepts and No Trends**0.05**

$n - r \backslash k$	λ_{trace}						λ_{max}					
	0	1	2	3	4	5	0	1	2	3	4	5
12	304.4	329.4	353.5	378.0	402.0	426.2	71.25	74.99	78.30	81.38	84.68	87.66
11	258.7	280.5	303.1	326.0	347.9	370.3	65.86	68.98	72.21	75.45	78.45	81.60
10	215.4	235.8	256.7	277.4	298.7	318.6	59.62	63.21	66.35	69.65	72.58	75.43
9	175.9	195.9	214.9	233.5	252.1	271.0	53.91	57.54	60.62	63.82	67.16	70.20
8	141.2	158.2	175.3	192.1	209.3	225.7	47.94	51.30	54.53	57.81	60.80	63.78
7	110.1	125.9	141.4	156.7	171.4	185.6	42.30	45.45	48.81	51.99	55.38	58.28
6	83.18	96.79	110.2	123.3	136.0	148.7	36.27	39.76	42.86	46.26	49.13	52.07
5	59.33	69.87	81.45	92.92	104.2	115.3	29.95	33.14	36.59	39.92	42.92	46.19
4	39.81	49.56	59.19	68.70	77.60	86.38	23.92	27.27	30.60	33.98	37.24	40.14
3	24.05	32.16	39.40	46.79	53.72	60.47	17.68	21.16	24.50	27.93	30.65	33.42
2	12.36	17.79	23.37	28.64	33.68	38.10	11.03	17.79	18.08	21.25	24.04	27.03
1	4.16	8.13	11.41	14.19	17.24	19.76	4.16	8.13	11.41	14.19	17.24	19.76

0.10

$n - r \backslash k$	λ_{trace}						λ_{max}					
	0	1	2	3	4	5	0	1	2	3	4	5
12	295.7	320.6	344.8	368.9	392.1	415.8	67.95	71.51	74.74	77.88	80.98	83.91
11	251.2	273.0	295.4	317.2	338.9	360.6	62.62	65.76	68.99	72.03	74.99	78.07
10	208.0	228.8	249.3	269.6	289.9	310.1	56.60	59.83	63.13	66.21	69.29	72.04
9	169.8	189.1	207.7	226.2	244.7	262.7	50.60	54.30	57.41	60.60	63.55	66.56
8	135.6	152.3	168.9	185.6	202.3	218.5	45.00	48.23	51.43	54.44	57.57	60.60
7	105.4	120.1	135.5	150.4	164.7	179.0	39.39	42.72	45.88	48.86	52.11	55.05
6	78.47	91.90	104.9	117.5	130.5	143.0	33.48	36.93	40.02	43.16	46.03	48.96
5	55.42	66.14	77.11	87.89	99.28	109.9	27.57	30.70	33.93	37.15	40.07	43.14
4	36.69	45.95	55.07	64.13	72.93	81.39	21.58	24.84	28.16	31.26	34.37	37.25
3	21.46	29.08	36.20	43.24	49.87	56.55	15.57	18.96	22.18	25.23	28.20	31.05
2	10.25	15.83	20.93	25.81	30.68	35.26	9.28	12.83	16.01	18.95	21.72	24.48
1	3.04	6.49	9.52	12.21	14.91	17.40	3.04	6.49	9.52	12.21	14.91	17.40

*The critical values λ_{max} and λ_{trace} statistics are computed as the maximum eigenvalue and the trace of the matrix defined by (B.1) for sample size $T = 500$ and 10,000 replications. **Case I** is defined by (3.9). n and k refer to the number of endogenous and exogenous variables, respectively, and r is the number of cointegrating relations defined by (3.8). The critical values at other quantiles are available from the authors upon request.

TABLE T.2*

Case II: Restricted Intercepts and No Trends**0.05**

$n - r \backslash k$	λ_{trace}						λ_{max}					
	0	1	2	3	4	5	0	1	2	3	4	5
12	341.2	365.1	389.4	413.6	437.1	461.5	75.51	78.57	81.74	84.82	87.90	91.11
11	291.7	314.2	336.4	358.7	381.0	403.4	69.76	73.11	76.02	79.23	82.36	85.20
10	245.7	265.9	287.0	307.2	327.8	348.5	64.11	66.85	70.17	73.47	76.20	79.28
9	204.0	223.8	242.3	260.8	279.5	297.9	57.95	61.34	64.47	67.77	70.86	73.65
8	166.1	182.7	199.7	216.6	232.8	249.3	52.06	55.26	58.21	61.59	64.40	67.45
7	132.5	148.0	163.1	178.0	192.5	206.5	46.47	49.57	52.87	55.81	58.85	61.74
6	102.6	116.3	128.9	141.7	154.4	167.0	40.53	43.76	46.90	49.97	52.88	55.65
5	75.98	86.58	97.57	108.6	119.8	130.8	34.40	37.48	40.57	43.62	46.77	49.80
4	53.48	62.75	72.15	81.25	90.60	99.37	28.27	31.48	34.69	37.83	40.19	43.58
3	34.87	42.40	49.43	56.28	63.10	69.72	22.04	25.54	28.49	31.56	34.15	37.19
2	20.18	25.23	30.46	35.46	39.94	44.56	15.87	18.88	21.92	24.97	27.82	30.50
1	9.16	12.45	15.27	17.80	20.63	23.02	9.16	12.45	15.27	17.80	20.63	23.02

0.10

$n - r \backslash k$	λ_{trace}						λ_{max}					
	0	1	2	3	4	5	0	1	2	3	4	5
12	332.0	356.4	380.3	403.8	427.3	450.6	72.09	75.35	78.55	81.40	84.36	87.24
11	284.2	305.4	328.1	350.1	371.6	393.4	66.47	69.48	72.75	75.71	78.68	81.72
10	238.3	259.0	279.0	299.3	319.2	338.9	60.66	63.77	66.78	69.98	72.86	75.88
9	197.7	216.4	234.7	253.0	271.4	289.3	54.91	58.17	61.21	64.27	67.35	70.32
8	160.3	176.4	192.9	209.2	226.1	241.8	49.04	52.23	55.38	58.31	61.28	64.24
7	127.2	142.2	157.0	171.5	185.7	199.5	43.44	46.59	49.67	52.67	55.61	58.57
6	97.87	110.5	123.5	136.0	148.3	160.6	37.65	40.93	44.05	47.09	49.85	52.67
5	71.81	82.17	92.93	103.6	114.4	125.2	31.73	34.99	37.81	40.86	43.80	46.87
4	49.95	59.07	67.83	76.69	85.34	93.92	25.80	29.01	32.00	35.08	38.03	40.84
3	31.93	39.12	45.89	52.71	59.23	65.69	19.86	22.98	26.08	28.83	31.73	34.56
2	17.88	22.76	27.58	32.28	36.84	41.33	13.81	16.74	19.67	22.54	25.27	27.87
1	7.53	10.50	13.21	15.68	18.24	20.57	7.53	10.50	13.21	15.68	18.24	20.57

***Case II** is defined by (3.10). Also see the footnote to Table T.1.

TABLE T.3*

Case III: Unrestricted Intercepts and No Trends**0.05**

$n - r \backslash k$	λ_{trace}						λ_{max}					
	0	1	2	3	4	5	0	1	2	3	4	5
12	328.5	352.4	376.5	400.8	424.3	448.8	74.62	77.72	80.87	83.93	87.13	90.23
11	279.8	302.2	324.4	347.7	369.5	391.5	68.91	72.31	75.22	78.42	81.56	84.41
10	235.0	255.5	276.5	297.2	317.4	337.7	63.32	65.98	69.40	72.56	75.48	78.58
9	194.4	214.4	232.7	251.6	270.4	288.6	57.20	60.66	63.68	67.04	70.20	72.98
8	157.8	174.2	191.6	208.3	224.8	241.2	51.15	54.44	57.54	60.80	63.54	66.61
7	124.6	140.5	155.9	170.7	185.6	199.8	45.53	48.91	52.22	55.04	58.05	61.03
6	95.87	109.6	122.8	135.4	148.4	160.9	39.83	42.95	46.09	49.10	52.06	54.84
5	70.49	81.45	92.42	103.7	114.6	125.9	33.64	36.80	39.85	42.92	45.98	49.02
4	48.88	58.63	68.06	77.21	86.42	95.14	27.42	30.71	33.87	37.08	39.99	42.68
3	31.54	38.93	46.44	53.41	60.19	66.94	21.12	24.59	27.75	30.74	33.63	36.38
2	17.86	23.32	28.42	33.35	38.15	42.73	14.88	18.06	21.07	24.22	26.94	29.79
1	8.07	11.47	14.35	16.90	19.73	22.19	8.07	11.47	14.35	16.90	19.73	22.19

0.10

$n - r \backslash k$	λ_{trace}						λ_{max}					
	0	1	2	3	4	5	0	1	2	3	4	5
12	319.6	343.5	367.5	391.1	415.0	438.0	71.36	74.61	77.65	80.60	83.55	86.49
11	272.7	294.1	316.9	338.8	360.1	381.8	65.63	68.80	71.95	75.02	77.86	80.86
10	227.7	248.3	268.5	288.8	309.1	328.9	59.85	63.00	65.97	69.13	72.08	74.91
9	187.9	207.1	225.3	243.7	262.1	280.2	54.10	57.56	60.49	63.59	66.59	69.53
8	152.0	168.4	184.8	201.3	217.7	233.6	48.23	51.47	54.54	57.55	60.46	63.40
7	119.7	134.9	149.7	164.2	178.3	192.3	42.70	45.74	48.96	51.93	54.68	57.85
6	91.40	104.4	117.3	130.0	142.1	154.6	36.84	40.21	43.25	46.30	49.06	51.90
5	66.23	76.95	87.93	98.45	109.3	120.4	31.02	34.10	37.15	40.02	42.95	45.97
4	45.70	54.84	63.57	72.69	81.40	89.90	24.99	28.27	31.30	34.30	37.21	39.99
3	28.78	35.88	42.67	49.56	56.08	62.65	19.02	22.15	25.21	28.11	30.85	33.67
2	15.75	20.75	25.63	30.37	35.06	39.59	12.98	15.98	18.78	21.67	24.37	27.05
1	6.50	9.53	12.27	14.76	17.39	19.63	6.50	9.53	12.27	14.76	17.39	19.63

***Case III** is defined by (3.11). Also see the footnote to Table T.1.

TABLE T.4*

Case IV: Unrestricted Intercepts and Restricted Trends**0.05**

$n - r \backslash k$	λ_{trace}						λ_{max}					
	0	1	2	3	4	5	0	1	2	3	4	5
12	364.8	389.2	413.5	437.3	460.7	484.3	78.42	81.85	84.73	87.82	90.74	93.75
11	314.1	335.5	358.0	380.5	402.3	423.7	72.50	75.77	78.97	81.99	85.08	87.98
10	265.8	286.8	307.0	327.7	347.9	367.8	67.05	70.20	73.82	76.15	78.87	81.89
9	222.6	242.4	260.4	279.3	297.3	315.6	61.27	64.53	67.67	70.71	73.62	76.51
8	183.0	199.1	215.8	232.9	249.2	265.3	55.14	58.08	61.22	64.42	67.45	70.48
7	147.3	163.0	177.8	192.8	207.1	221.4	49.32	52.62	55.83	59.01	62.00	64.61
6	115.9	128.8	141.7	154.3	166.9	179.6	43.61	46.97	50.10	53.08	55.83	58.78
5	87.17	97.83	108.9	120.0	130.6	141.2	37.86	40.89	43.72	46.84	49.76	52.63
4	63.00	72.10	81.20	90.02	99.11	107.6	31.79	34.70	37.85	40.98	43.75	46.66
3	42.34	49.36	56.43	63.54	69.84	76.82	25.42	28.72	31.68	34.65	37.44	40.12
2	25.77	30.77	35.37	40.37	45.10	49.52	19.22	22.16	24.88	27.80	30.55	33.26
1	12.39	15.44	18.08	20.47	23.17	25.70	12.39	15.44	18.08	20.47	23.17	25.70

0.10

$n - r \backslash k$	λ_{trace}						λ_{max}					
	0	1	2	3	4	5	0	1	2	3	4	5
12	355.9	379.8	403.3	426.8	450.3	473.6	75.02	78.09	81.31	84.34	87.34	90.08
11	305.8	327.1	349.2	371.0	392.4	414.1	69.45	72.72	75.59	78.45	81.52	84.23
10	258.0	278.9	298.9	319.2	338.7	358.9	63.60	66.78	69.90	72.94	75.64	78.53
9	215.9	235.1	253.2	271.2	289.5	306.9	58.09	61.42	64.28	67.43	70.30	73.07
8	176.9	192.8	209.1	225.5	241.7	257.5	52.08	55.25	58.18	61.19	63.99	67.02
7	141.8	157.0	171.6	185.9	199.9	214.0	46.54	49.70	52.69	55.51	58.58	61.37
6	110.6	123.3	136.2	148.4	160.9	173.0	40.76	44.01	47.08	49.78	52.73	55.62
5	82.88	93.13	103.7	114.7	125.1	135.8	35.04	37.92	40.94	43.92	46.74	49.59
4	59.16	68.04	76.68	85.59	93.98	102.5	29.13	32.12	35.04	38.04	41.01	43.66
3	39.34	46.00	52.71	59.39	65.90	72.33	27.10	26.10	29.00	31.89	34.66	37.28
2	23.08	27.96	32.51	37.07	41.57	46.10	17.18	19.79	22.53	25.28	27.86	30.54
1	10.55	13.31	15.82	18.19	20.73	23.11	10.55	13.31	15.82	18.19	20.73	23.11

***Case IV** is defined by (3.12). Also see the footnote to Table T.1.

TABLE T.5*

Case V: Unrestricted Intercepts and Unrestricted Trends**0.05**

$n - r \backslash k$	λ_{trace}						λ_{max}					
	0	1	2	3	4	5	0	1	2	3	4	5
12	352.1	376.8	401.2	425.3	449.1	472.3	77.73	81.00	83.98	87.05	90.01	92.88
11	302.4	323.9	346.6	368.8	390.8	412.4	71.84	75.01	77.97	81.18	84.29	87.25
10	255.1	276.4	296.3	317.5	337.4	357.6	66.17	69.38	72.46	75.49	78.09	81.04
9	213.4	233.3	251.2	270.1	288.1	306.1	60.48	63.80	66.96	69.81	72.80	75.70
8	174.9	190.9	207.9	224.8	241.2	257.2	54.17	57.38	60.37	63.62	66.67	69.60
7	140.2	156.0	170.6	185.3	199.9	214.3	48.57	51.97	54.94	58.20	61.30	63.86
6	109.2	122.3	135.5	148.3	160.8	173.3	42.67	46.11	49.29	52.24	54.97	58.00
5	82.23	92.96	104.0	115.1	125.5	136.5	37.07	40.00	42.94	46.03	49.03	51.83
4	58.93	68.10	77.14	86.30	94.88	103.8	31.00	33.97	37.08	40.06	43.03	45.77
3	39.33	46.58	53.48	60.54	67.22	74.01	24.35	27.96	30.92	33.81	36.59	39.28
2	23.83	28.76	33.39	38.50	43.45	47.84	18.33	21.21	24.18	26.95	29.74	32.53
1	11.54	14.53	17.14	19.62	22.35	24.82	11.54	14.53	17.14	19.62	22.35	24.82

0.10

$n - r \backslash k$	λ_{trace}						λ_{max}					
	0	1	2	3	4	5	0	1	2	3	4	5
12	343.4	367.4	391.0	415.1	438.3	461.4	74.25	77.27	80.43	83.50	86.53	89.28
11	294.1	316.2	337.9	359.9	381.6	402.9	68.56	71.83	74.79	77.61	80.64	83.42
10	247.8	268.7	289.0	308.8	328.4	348.5	62.80	66.07	69.17	72.11	74.91	77.79
9	206.4	226.0	244.0	261.9	279.9	297.6	57.37	60.54	63.50	66.62	69.43	72.30
8	168.2	184.6	201.0	217.4	234.1	249.4	51.26	54.46	57.43	60.32	63.16	66.24
7	134.5	149.9	164.5	178.6	193.0	206.8	45.75	48.97	52.03	54.75	57.71	60.60
6	104.3	117.2	130.2	142.3	155.0	167.0	39.90	43.22	46.30	49.07	51.94	54.75
5	77.55	88.25	98.82	109.7	120.5	130.8	34.16	37.13	40.16	43.12	46.01	48.89
4	55.01	63.98	72.71	81.72	90.31	98.57	28.32	31.31	34.25	37.33	40.13	42.86
3	36.28	43.01	49.86	56.52	63.09	69.44	22.26	25.28	28.28	31.11	33.85	36.41
2	21.23	26.02	30.54	35.06	39.63	44.31	16.28	18.98	21.69	24.47	27.06	29.60
1	9.75	12.43	15.04	17.31	19.79	22.31	9.75	12.43	15.04	17.31	19.79	22.31

***Case V** is defined by (3.13). Also see the footnote to Table T.1.

Table 1(a): Rejection Frequencies for Experiment E_{1a}

		<i>LR1</i>		<i>LR2</i>		<i>LR3</i>	
<i>T</i>	<i>H_r</i>	<i>H_n</i>	<i>H_{r+1}</i>	<i>H_n</i>	<i>H_{r+1}</i>	<i>H_n</i>	<i>H_{r+1}</i>
30	$r = 0$	.204	.166	.196	.172	.265	.263
	$r = 1$	.024	.015	.020	.020	.024	.024
50	$r = 0$	.151	.123	.164	.140	.254	.254
	$r = 1$	.017	.009	.015	.015	.019	.019
75	$r = 0$	.199	.159	.226	.206	.345	.352
	$r = 1$	.020	.013	.022	.022	.025	.025
100	$r = 0$	.280	.248	.342	.326	.475	.492
	$r = 1$	.029	.019	.031	.031	.034	.034
150	$r = 0$	.418	.430	.524	.549	.723	.764
	$r = 1$	.033	.031	.040	.040	.043	.043
300	$r = 0$	.957	.988	.988	.999	.997	.999
	$r = 1$	.050	.051	.053	.053	.053	.053

Table 1(b): Rejection Frequencies for Experiment E_{1b}

		<i>LR1</i>		<i>LR2</i>		<i>LR3</i>	
<i>T</i>	<i>H_r</i>	<i>H_n</i>	<i>H_{r+1}</i>	<i>H_n</i>	<i>H_{r+1}</i>	<i>H_n</i>	<i>H_{r+1}</i>
30	$r = 0$	.193	.164	.184	.165	.304	.309
	$r = 1$	.021	.014	.018	.018	.025	.025
50	$r = 0$	.135	.112	.144	.125	.280	.284
	$r = 1$	.013	.008	.012	.012	.020	.020
75	$r = 0$	.162	.133	.187	.169	.350	.356
	$r = 1$	.017	.010	.017	.017	.025	.025
100	$r = 0$	.227	.196	.278	.257	.461	.478
	$r = 1$	.022	.016	.025	.025	.032	.032
150	$r = 0$	.419	.429	.519	.547	.724	.768
	$r = 1$	.036	.035	.036	.036	.042	.042
300	$r = 0$	.958	.991	.989	.998	.997	1.00
	$r = 1$	.048	.047	.049	.049	.050	.050

Table 1(c): Rejection Frequencies for Experiment E_{1c}

		<i>LR1</i>		<i>LR2</i>		<i>LR3</i>	
T	H_r	H_n	H_{r+1}	H_n	H_{r+1}	H_n	H_{r+1}
30	$r = 0$	.231	.182	.221	.192	.291	.279
	$r = 1$	.028	.019	.024	.024	.029	.029
50	$r = 0$	.203	.159	.221	.188	.323	.308
	$r = 1$	.024	.015	.022	.022	.030	.030
75	$r = 0$	.299	.257	.351	.329	.486	.490
	$r = 1$	.033	.024	.034	.034	.044	.044
100	$r = 0$	.443	.431	.535	.541	.680	.707
	$r = 1$	.044	.036	.047	.047	.055	.055
150	$r = 0$	.763	.824	.858	.896	.934	.959
	$r = 1$	.054	.053	.058	.058	.060	.060
300	$r = 0$	1.00	1.00	1.00	1.00	1.00	1.00
	$r = 1$	.056	.055	.057	.057	.060	.060

Table 1(d): Rejection Frequencies for Experiment E_{1d}

		<i>LR1</i>		<i>LR2</i>		<i>LR3</i>	
T	H_r	H_n	H_{r+1}	H_n	H_{r+1}	H_n	H_{r+1}
30	$r = 0$	.226	.182	.219	.188	.304	.309
	$r = 1$	.027	.017	.022	.022	.025	.025
50	$r = 0$	.198	.157	.220	.188	.351	.339
	$r = 1$	.023	.015	.022	.022	.033	.033
75	$r = 0$	.289	.252	.345	.322	.515	.521
	$r = 1$	.033	.023	.036	.036	.047	.047
100	$r = 0$	.433	.422	.526	.529	.707	.734
	$r = 1$	.041	.033	.046	.046	.060	.060
150	$r = 0$	.751	.811	.848	.888	.946	.968
	$r = 1$	.054	.052	.057	.057	.063	.063
300	$r = 0$	1.00	1.00	1.00	1.00	1.00	1.00
	$r = 1$	.054	.056	.057	.057	.060	.060

Table 2(a): Rejection Frequencies for Experiment E_{2a}

		<i>LR1</i>		<i>LR2</i>		<i>LR3</i>	
T	H_r	H_n	H_{r+1}	H_n	H_{r+1}	H_n	H_{r+1}
30	$r = 0$	.202	.163	.191	.169	.227	.216
	$r = 1$	.022	.015	.018	.018	.021	.021
50	$r = 0$	.156	.128	.166	.146	.217	.201
	$r = 1$	.015	.009	.014	.014	.019	.019
75	$r = 0$	.197	.164	.233	.210	.295	.278
	$r = 1$	.021	.012	.022	.022	.028	.029
100	$r = 0$	.276	.251	.340	.332	.424	.425
	$r = 1$	.027	.022	.029	.029	.038	.038
150	$r = 0$	.521	.554	.642	.681	.747	.796
	$r = 1$	.041	.036	.043	.043	.050	.050
300	$r = 0$	.986	.998	.998	1.00	1.00	1.00
	$r = 1$	.054	.054	.053	.053	.056	.056

Table 2(b): Rejection Frequencies for Experiment E_{2b}

		<i>LR1</i>		<i>LR2</i>		<i>LR3</i>	
T	H_r	H_n	H_{r+1}	H_n	H_{r+1}	H_n	H_{r+1}
30	$r = 0$	.258	.197	.247	.211	.276	.251
	$r = 1$	.035	.023	.028	.028	.030	.030
50	$r = 0$	.262	.210	.291	.254	.349	.323
	$r = 1$	.032	.021	.030	.030	.038	.038
75	$r = 0$	.408	.377	.474	.469	.559	.559
	$r = 1$	.047	.034	.046	.046	.054	.054
100	$r = 0$	.597	.616	.691	.718	.773	.804
	$r = 1$	.055	.049	.055	.055	.066	.066
150	$r = 0$	.897	.942	.951	.972	.977	.989
	$r = 1$	.060	.057	.057	.057	.060	.060
300	$r = 0$	1.00	1.00	1.00	1.00	1.00	1.00
	$r = 1$	.060	.060	.057	.057	.059	.059

Table 2(c): Rejection Frequencies for Experiment E_{2c}

		<i>LR1</i>		<i>LR2</i>		<i>LR3</i>	
T	H_r	H_n	H_{r+1}	H_n	H_{r+1}	H_n	H_{r+1}
30	$r = 0$	.285	.209	.274	.229	.348	.298
	$r = 1$	.044	.026	.033	.033	.057	.057
50	$r = 0$	.324	.261	.364	.325	.502	.465
	$r = 1$	.044	.030	.042	.042	.077	.077
75	$r = 0$	.524	.501	.600	.600	.768	.781
	$r = 1$	.058	.047	.061	.061	.096	.096
100	$r = 0$	.735	.769	.814	.842	.932	.948
	$r = 1$	.066	.062	.065	.065	.094	.094
150	$r = 0$	.965	.986	.984	.993	.998	.999
	$r = 1$	.064	.060	.060	.060	.079	.079
300	$r = 0$	1.00	1.00	1.00	1.00	1.00	1.00
	$r = 1$	.061	.060	.058	.058	.066	.066

Table 2(d): Rejection Frequencies for Experiment E_{2d}

		<i>LR1</i>		<i>LR2</i>		<i>LR3</i>	
T	H_r	H_n	H_{r+1}	H_n	H_{r+1}	H_n	H_{r+1}
30	$r = 0$	.392	.288	.388	.328	.496	.435
	$r = 1$	.071	.044	.056	.056	.093	.093
50	$r = 0$	.539	.490	.595	.579	.770	.763
	$r = 1$	.074	.057	.072	.072	.119	.119
75	$r = 0$	.814	.837	.875	.893	.964	.972
	$r = 1$	.085	.073	.080	.080	.113	.113
100	$r = 0$	.958	.978	.978	.987	.997	.998
	$r = 1$	.082	.078	.078	.078	.103	.103
150	$r = 0$	1.00	1.00	1.00	1.00	1.00	1.00
	$r = 1$	.068	.066	.066	.066	.079	.079
300	$r = 0$	1.00	1.00	1.00	1.00	1.00	1.00
	$r = 1$	.063	.061	.061	.061	.068	.068

Table 3: Cointegration Rank Statistics

	$\widehat{\mathcal{LR}}(H_r H_n)$			$\widehat{\mathcal{LR}}(H_r H_{r+1})$		
H_r	Statistic	0.05 CV	0.10 CV	Statistic	0.05 CV	0.10 CV
$r = 0$	80.71	81.20	76.68	34.09	37.85	35.04
$r = 1$	46.62	56.43	52.71	25.59	31.68	29.00
$r = 2$	21.03	35.37	32.51	14.10	24.88	22.53
$r = 3$	6.93	18.08	15.82	6.93	18.08	15.82

Table 4: Cointegration Rank Statistics for Short-run Restricted Model

	$\widehat{\mathcal{LR}}(H_r H_n)$			$\widehat{\mathcal{LR}}(H_r H_{r+1})$		
H_0	Statistic	0.05 CV	0.10 CV	Statistic	0.05 CV	0.10 CV
$r = 0$	139.62	81.20	76.68	75.73	37.85	35.04
$r = 1$	63.89	56.43	52.71	37.08	31.68	29.00
$r = 2$	26.80	35.37	32.51	19.60	24.88	22.53
$r = 3$	7.21	18.08	15.82	7.21	18.08	15.82

Table 5: Estimates of the Error Correction Coefficients and Diagnostic Statistics

Equation	α_{1y}	α_{2y}	R^2	$\chi^2_{SC}(4)$	$\chi^2_{FF}(1)$	$\chi^2_N(2)$	$\chi^2_{HE}(1)$
Δp_t	-0.062 (0.016)	-0.063 (0.066)	0.784	3.35 [0.50]	3.31 [0.07]	3.15 [0.21]	0.02 [0.88]
Δe_t	0.124 (0.070)	0.049 (0.285)	0.154	0.76 [0.94]	0.72 [0.40]	0.13 [0.94]	0.06 [0.81]
Δr_t	0.044 (0.025)	-.036 (0.103)	0.070	1.87 [0.76]	1.35 [0.25]	5.02 [0.08]	0.13 [0.72]
Δr_t^*	0.075 (0.028)	0.421 (0.114)	0.172	9.75 [0.045]	1.35 [0.25]	29.32 [0.00]	0.023 [0.88]

References

- Abadir, K.M., Hadri, K., Tzavalis, E., 1999. The influence of VAR dimensions on estimator biases. *Econometrica* 67, 163-181.
- Ahn, S.K., Reinsel, G.C., 1990. Estimation of partially nonstationary multivariate autoregressive models. *Journal of the American Statistical Association* 85, 813-823.
- Boswijk, H.P., 1992. Cointegration, Identification and Exogeneity: Inference in Structural Error Correction Models. Tinbergen Institute Research Series.
- Boswijk, H.P., 1995. Efficient inference on cointegration parameters in structural error correction models. *Journal of Econometrics* 69, 133-158.
- Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427-431.
- Dickey, D.A., Fuller, W.A., 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* 49, 1057-1072.
- Engle, R.F., Granger, C.W.J., 1987. Co-integration and error correction: representation, estimation and testing. *Econometrica* 55, 251-276.
- Granger, C.W.J., Lin, J.L., 1995. Causality in the long run. *Econometric Theory* 11, 530-536.
- Harbo, I., Johansen, S., Nielsen, B., Rahbek, A., 1998. Asymptotic inference on cointegrating rank in partial systems. *Journal of Business Economics and Statistics* 16, 388-399.
- Johansen, S., 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12, 231-254.
- Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* 59, 1551-1580.
- Johansen, S., 1992. Cointegration in partial systems and the efficiency of single-equation analysis," *Journal of Econometrics* 52: 389-402.
- Johansen, S., 1994. The role of the constant and linear terms in cointegration analysis of nonstationary variables. *Econometric Reviews* 13, 205-229.
- Johansen, S., 1995. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press, Oxford.
- Johansen, S., Juselius, K., 1992. Testing structural hypotheses in a multivariate cointegration analysis of the PPP and UIP for UK. *Journal of Econometrics* 53: 211-244.
- Osterwald-Lenum, M., 1992. A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics. *Oxford Bulletin of Economics and Statistics* 54, 461-471.
- Park, J.Y., 1992. Canonical cointegrating regressions. *Econometrica* 60, 119-143.
- Park, J.Y., Phillips, P.C.B., 1989. Statistical inference in regressions with integrated processes: Part 2. *Econometric Theory* 5, 95-131.
- Perron, P., Campbell, J., 1993. A note on Johansen's cointegration procedure when trends are present. *Empirical Economics* 18, 777-789.
- Pesaran, M.H., Pesaran, B., 1997. Working with Microfit 4.0: Interactive Econometric Analysis. Oxford University Press, Oxford.
- Pesaran, M.H., Shin, Y., 1996. Cointegration and speed of convergence to equilibrium. *Journal of Econometrics* 71: 117-143.
- Pesaran, M.H., Shin, Y., 1999. Long-run structural modelling. Manuscript. University of Cambridge. <http://www.econ.cam.ac.uk/faculty/pesaran/>
- Pesaran, M.H., Shin, Y., Smith, R.J., 1997. Structural analysis of vector error correction models with exogenous I(1) variables. D.A.E. Working Paper No. 9706. University of Cambridge.
- Phillips, P.C.B., 1991. Optimal inference in cointegrated systems. *Econometrica* 59, 283-306.
- Phillips, P.C.B., Durlauf, S.N., 1986. Multiple time series regression with integrated processes. *Review of Economic Studies* 53, 473-495.
- Phillips, P.C.B., Hansen, B.E., 1990. Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57, 99-125.
- Phillips, P.C.B., Solo, V., 1992. Asymptotics for linear processes. *Annals of Statistics* 20, 971-1001.

- Rahbek, A., Mosconi, R., 1999. Cointegration rank inference with stationary regressors in VAR models. *The Econometrics Journal* 2.
- Yamada, H., Toda, H.Y., 1998. Inference in possibly integrated vector autoregressive models: some finite sample evidence. *Journal of Econometrics* 86, 55-95.