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

M. Hashem Pesaran$^a$,*, Yongcheol Shin$^b$, Richard J. Smith$^c$

$^a$Faculty of Economics and Politics, Austin Robinson Building, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DD, UK
$^b$Department of Economics, University of Edinburgh, UK
$^c$Department of Economics, University of Bristol, UK

Received 1 April 1997; received in revised form 1 October 1999; accepted 1 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. © 2000 Elsevier Science S.A. All rights reserved.

JEL classification: C12; C13; C32

Keywords: 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

*Corresponding author. Tel.: +44 1223 335216; fax: +44 1223 335471.
E-mail address: hashem.pesaran@econ.cam.ac.uk (M.H. Pesaran).

0304-4076/00/$-see front matter © 2000 Elsevier Science S.A. All rights reserved.
PII: S 0 3 0 4 - 4 0 7 6 ( 9 9 ) 0 0 0 7 3 - 1
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. 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 et al. (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 behavior in the levels of the underlying variables; that is, Case I: zero intercepts and linear trend coefficients, Case II: restricted intercepts and zero linear trend coefficients, Case III: unrestricted intercepts and zero linear trend coefficients, Case IV: unrestricted intercepts and restricted linear trend coefficients, Case V: unrestricted intercepts and unrestricted linear trend coefficients. These cases have been considered by

---

1 A 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.
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 Sections 4.1 and 4.2; relevant asymptotic critical values are provided in Tables 6(a)–6(e). Modifications to the tests of Sections 4.1 and 4.2 necessitated in the presence of exogenous variables which are integrated of order zero are briefly addressed in Section 4.3. Tests for the absence of an intercept or a trend in the cointegrating relations are discussed in Section 4.4, and particular misspecification tests concerning the intercept and linear trend coefficients are developed in 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 U.K. 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 6(a)–(e).

2. The treatment of trends in VAR models

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

$$\Phi(L)(z_t - \mu - \gamma t) = \epsilon_t, \quad t = 1, 2, \ldots, \tag{2.1}$$

where $L$ is the lag operator, $\mu$ and $\gamma$ are $m$-vectors of unknown coefficients, and the $(m,m)$ matrix lag polynomial of order $p$, $\Phi(L) \equiv I_m - \sum_{i=1}^p \Phi_i L^i$, 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 $\{\epsilon_t\}_{t=-\infty}^\infty$ is assumed to be $\text{IN}(0, \Omega)$, $\Omega$ positive definite. The analysis that follows is conducted given the initial values $Z_0 \equiv (z_{-p+1}, \ldots, z_0)$. 


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.

$$\Phi(L) \equiv - P L + \Gamma(L)(1 - L).$$

(2.2)

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

$$P \equiv - \left( I_m - \sum_{i=1}^{p} \Phi_i \right)$$

(2.3)

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

$$\Phi(L)z_t = a_0 + a_1 t + e_t, \quad t = 1, 2, \ldots,$$

(2.4)

where

$$a_0 \equiv - \Pi \mu + (\Gamma + \Pi)\gamma, \quad a_1 \equiv - \Pi \gamma,$$

(2.5)

and the sum of the short-run coefficient matrices $\Gamma$ is given by

$$\Gamma \equiv I_m - \sum_{i=1}^{p-1} \Gamma_i = - \Pi + \sum_{i=1}^{p} i \Phi_i.$$ 

(2.6)

The cointegration rank hypothesis is defined by

$$H_r: \quad \text{Rank}[\Pi] = r, \quad r = 0, \ldots, m,$$

(2.7)

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

$$\Pi = \alpha \beta',$$

(2.8)

where $\alpha$ and $\beta$ 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 $\alpha$ and $\beta$ respectively; in particular, $\alpha' \alpha_\perp = 0$ and $\beta' \beta_\perp = 0$.

We now adopt the following assumptions.

**Assumption 2.1.** The $(m, m)$ matrix polynomial $\Phi(z) = I_m - \sum_{i=1}^{p} \Phi_i z^i$ 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 \{$(z_t - \mu - \gamma t)$\}$_{t=1}^{\infty}$ 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 \( \{\beta_t(z_t - \mu - \gamma t)\}_{t=1}^{\infty} \) and \( \{\beta'(z_t - \mu - \gamma t)\}_{t=1}^{\infty} \) to be integrated of orders one and zero respectively.\(^2\) Moreover, Assumption 2.2 specifically excludes the process \( \{(z_t - \mu - \gamma t)\}_{t=1}^{\infty} \) 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 \( \{\Delta z_t\}_{t=1}^{\infty} \) may be expressed under Assumptions 2.1 and 2.2 from (2.4) as the infinite vector moving average process

\[
\Delta z_t = C(L)(a_0 + a_1 t + e_t) = b_0 + b_1 t + C(L)e_t, \quad t = 1, 2, \ldots ,
\tag{2.9}
\]

where \( b_0 \equiv Ca_0 + C^*a_1, b_1 \equiv Ca_1 \). The matrix lag polynomial \( C(L) \) is given by\(^3\)

\[
C(L) \equiv I_m + \sum_{j=1}^{\infty} C_j L^j = C + (1 - L)C^*(L), \quad C^*(L) \equiv \sum_{j=0}^{\infty} C^*_j L^j,
\tag{2.10}
\]

Now, as \( C(L) \Phi(L) = \Phi(L)C(L) = (1 - L)I_m, \quad \Pi C = 0 \) and \( \Pi H = 0; \quad \Pi H = 0; \quad \Pi = 0; \) in particular, \( C = \beta_1(\alpha_1 \Gamma \beta_1)^{-1}\alpha_1 \). Re-expressing (2.9) in levels,

\[
z_t = z_0 + b_0 t + b_1 \frac{t(t + 1)}{2} + Cs_t + C^*(L)(e_t - e_0),
\tag{2.11}
\]

where the partial sum \( s_t \equiv \sum_{i=1}^{t} e_i, \quad t = 1, 2, \ldots . \)

Adopting the VAR(\( p \)) formulation (2.1) rather than the more usual (2.4), in which \( a_0 \) and \( a_1 \) are unrestricted, reveals immediately from (2.11) that the restrictions (2.5) on \( a_1 \) induce \( b_1 = 0 \) and ensure that the nature of the deterministic trend behavior of the level process \( \{z_t\}_{t=1}^{\infty} \) remains invariant to the rank \( r \) of the long-run multiplier matrix \( \Pi \); that is, it is linear. Hence, the infinite moving average representation for the level process \( \{z_t\}_{t=1}^{\infty} \) is\(^4\)

\[
z_t = \mu + \gamma t + Cs_t + C^*(L)e_t,
\tag{2.12}
\]

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

\(^3\) The matrices \( \{C_i\} \) can be obtained from the recursions \( C_i = \sum_{j=1}^{r} C_{i-j} \Phi_j, i > 1, \quad C_0 = I_m, \quad C_1 = -(I_m - \Phi_1), \) defining \( C_i = 0 \) for \( i < 0 \). Similarly, for the matrices \( \{C_j\} \), \( C_j = C_j + C_{j-1}, j > 0, \quad C_0 = I_m - C. \)

\(^4\) From (2.2), as \( C(L) \Phi(L) = (1 - L)I_m \) and, in particular, \( \Pi = 0, \quad C \Gamma = C^* \Pi = I_m. \)
where we have used the initialization $z_0 \equiv \mu + C^*(L)e_0$.\(^5\) See also Johansen (1994) and Johansen (1995, Section 5.7, pp. 80–84).\(^6\) If, however, $a_1$ 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 $C = 0$. However, $b_1$ would be unconstrained under the null hypothesis of no cointegration; that is, $H_0$: $\text{Rank}[\Pi] = 0$, and $C$ 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 $\{z_t\}_{t=1}^\infty$ 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 $\beta'$, we obtain the cointegrating relations

$$
\beta'z_t = \beta'\mu + (\beta'\gamma)t + \beta' C^*(L)e_t, \quad t = 1, 2, \ldots,
$$

(2.13)

which are trend stationary. In general, the co-trending restriction (Park, 1992) $\beta'\gamma = 0$ if and only if $a_1 = 0$. 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 $\beta'\gamma = 0$ may not prove to be satisfactory in practice. Therefore, it is important that the composite $\beta'\gamma$ in (2.13) or, equivalently, $a_1 = -\Pi\gamma$ in (2.4) is estimated along with the other parameters of the model and the co-trending restriction tested; see Section 4.4.

3. Efficient estimation of a structural error correction model

We now partition the $m$-vector of random variables $z_t$ into the $n$-vector $y_t$ and the $k$-vector $x_t$, where $k \equiv m - n$; that is, $z_t = (y_t', x_t')'$, $t = 1, 2, \ldots$. The primary concern of this paper is the structural modelling of the vector $y_t$ conditional on its past, $y_{t-1}, y_{t-2}, \ldots$, and current and past values of the vector of random variables $x_t, x_{t-1}, x_{t-2}, \ldots$, $t = 1, 2, \ldots$. The assumption of Section 2 concerning the error process $\{e_t\}_{t=-\infty}^\infty, e_t \sim \text{IN}(0, \Omega), t = 0, \pm 1, \ldots$, where $\Omega$ is positive

---

\(^5\) Of course, the levels equation (2.12) could also have been obtained directly from (2.1) by noting $A(z_t - \mu - \gamma t) = C(L)e_t, t = 1, 2, \ldots$.

\(^6\) As the cointegration rank hypothesis (2.7) may be alternatively and equivalently expressed as $H_r$: $\text{Rank}[C] = m - r$, $r = 0, \ldots, m$, 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 $C$, the combined total of which is $m$. 

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):

\[ \Delta z_t = a_0 + a_1 t + \sum_{i=1}^{p-1} \Gamma_i \Delta z_{t-i} + \Pi z_{t-1} + e_t, \quad t = 1, 2, \ldots, \quad (3.1) \]

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

By partitioning the error term \( e_t \) conformably with \( z_t = (y_t', x_t')' \) as \( e_t = (e_{yt}', e_{xt}')' \) and its variance matrix as

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

we are able to express \( e_{yt} \) conditionally in terms of \( e_{xt} \) as

\[ e_{yt} = \Omega_{yx} \Omega_{xx}^{-1} e_{xt} + u_t, \quad (3.2) \]

where \( u_t \sim \text{IN}(0, \Omega_{uu}) \), \( \Omega_{uu} = \Omega_{yy} - \Omega_{yx} \Omega_{xx}^{-1} \Omega_{xy} \) and \( u_t \) is independent of \( e_{xt} \).

Substitution of (3.2) into (3.1) together with a similar partitioning of the parameter vectors and matrices \( a_0 = (a_{0y}', a_{0x}')', \quad a_1 = (a_{1y}', a_{1x}')', \quad \Pi = (\Pi_y', \Pi_x')', \quad \Gamma = (\Gamma_y', \Gamma_x')', \quad \Gamma_i = (\Gamma_{yi}', \Gamma_{xi}')', \quad i = 1, \ldots, p - 1 \), provides a conditional model for \( \Delta y_t \) in terms of \( z_{t-1}, \Delta x_t, \Delta z_{t-1}, \Delta z_{t-2}, \ldots \); viz.

\[ \Delta y_t = c_0 + c_1 t + \Lambda \Delta x_t + \sum_{i=1}^{p-1} \Psi_i \Delta z_{t-i} + \Pi_{yy.x} z_{t-1} + u_t, \quad (3.3) \]

where \( c_0 \equiv a_{0y} - \Omega_{yx} \Omega_{xx}^{-1} a_{0x}, \quad c_1 \equiv a_{y1} - \Omega_{yx} \Omega_{xx}^{-1} a_{x1}, \quad \Lambda = \Omega_{yy} \Omega_{xx}^{-1}, \quad \Psi_i \equiv \Gamma_{yi} - \Omega_{yy} \Omega_{xx}^{-1} \Gamma_{xi}, \quad i = 1, \ldots, p - 1 \), and \( \Pi_{yy.x} \equiv \Pi_y - \Omega_{yy} \Omega_{xx}^{-1} \Pi_x \).

Following Johansen (1992) and Boswijk (1992, Chapter 3), we assume that the process \( \{x_t\}_{i=1}^{\infty} \) is weakly exogenous with respect to the matrix of long-run multiplier parameters \( \Pi \); viz. \(^7\)

**Assumption 3.1.** \( \Pi_x = 0 \).

Therefore,

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

\(^7\)Specification tests for Assumption 3.1 are presented in Section 4.5 below.
Consequently, under Assumption 3.1, from (3.1) and (3.3), the system of equations are rendered as

\[ \Delta y_t = c_0 + c_1 t + \Lambda \Delta x_t + \sum_{i=1}^{p-1} \Psi_i \Delta z_{t-i} + \Pi_y z_{t-1} + u_t, \]  

(3.5)

\[ \Delta x_t = a_{x0} + \sum_{i=1}^{p-1} \Gamma_{xi} \Delta z_{t-i} + e_{xt}, \quad t = 1, 2, \ldots, \]  

(3.6)

where now, because \( a_{x1} = 0 \), \( c_1 \equiv a_{y1} \) and the relations (2.5) are modified to

\[ c_0 = - \Pi_y \mu + (\Gamma_y - \Omega_{yx} \Omega_{xx}^{-1} \Gamma_x + \Pi_y) \gamma, \quad c_1 = - \Pi_y \gamma. \]  

(3.7)

The restriction \( \Pi_x = 0 \) of Assumption 3.1 implies that the elements of the vector process \( \{x_t\}_{i=1}^{\infty} \) 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 \( \{x_t\}_{i=1}^{\infty} \) is redundant for efficient conditional estimation and inference concerning the long-run parameters \( \Pi_y \) as well as the deterministic and short-run parameters \( c_0, c_1, \Lambda \) and \( \Psi_i, i = 1, \ldots, p - 1 \), of (3.5). Furthermore, under Assumption 3.1, we may regard \( \{x_t\}_{i=1}^{\infty} \) as long run forcing for \( \{y_t\}_{i=1}^{\infty} \); see Granger and Lin (1995).\(^8\)

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

\[ H_r: \text{Rank}[\Pi_y] = r, \quad r = 0, \ldots, n. \]  

(3.8)

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

Case I: (No intercepts; no trends.) \( c_0 = 0 \) and \( c_1 = 0 \). That is, \( \mu = 0 \) and \( \gamma = 0 \).

Hence, the structural VECM (3.5) becomes

\[ \Delta y_t = \Lambda \Delta x_t + \sum_{i=1}^{p-1} \Psi_i \Delta z_{t-i} + \Pi_y z_{t-1} + u_t. \]  

(3.9)

Case II: (Restricted intercepts; no trends.) \( c_0 = -\Pi_y \mu \) and \( c_1 = 0 \). Here, \( \gamma = 0 \). The structural VECM (3.5) is

\[ \Delta y_t = ( - \Pi_y \mu) + \Lambda \Delta x_t + \sum_{i=1}^{p-1} \Psi_i \Delta z_{t-i} + \Pi_y z_{t-1} + u_t. \]  

(3.10)

Case III: (Unrestricted intercepts; no trends.) \( c_0 \neq 0 \) and \( c_1 = 0 \). Again, \( \gamma = 0 \).

In this case, the intercept restriction \( c_0 = -\Pi_y \mu \) is ignored and the structural

---

\(^8\) Note that this restriction does not preclude \( \{y_t\}_{i=1}^{\infty} \) being Granger-causal for \( \{x_t\}_{i=1}^{\infty} \) in the short run.
VECM estimated is
\[ \Delta y_t = c_0 + \Lambda \Delta x_t + \sum_{i=1}^{p-1} \Psi_i \Delta z_{t-i} + \Pi_y z_{t-1} + u_t. \] (3.11)

**Case IV:** (Unrestricted intercepts; restricted trends.) \( c_0 \neq 0 \) and \( c_1 = -\Pi_y \gamma \). Thus
\[ \Delta y_t = c_0 + ( -\Pi_y \gamma) t + \Lambda \Delta x_t + \sum_{i=1}^{p-1} \Psi_i \Delta z_{t-i} + \Pi_y z_{t-1} + u_t. \] (3.12)

**Case V:** (Unrestricted intercepts; unrestricted trends.) \( c_0 \neq 0 \) and \( c_1 \neq 0 \). Here, the deterministic trend restriction \( c_1 = -\Pi_y \gamma \) is ignored and the structural VECM estimated is
\[ \Delta y_t = c_0 + c_1 t + \Lambda \Delta x_t + \sum_{i=1}^{p-1} \Psi_i \Delta z_{t-i} + \Pi_y z_{t-1} + u_t. \] (3.13)

It should be emphasized 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, \( c_0 \) and \( c_1 \), to the parameter matrix \( \Pi_y \) 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
\[ \Pi_y = \mathbf{x}_y \mathbf{\beta}^\prime, \] (3.14)
where the \((n, r)\) loading matrix \( \mathbf{x}_y \) and the \((m, r)\) matrix of cointegrating vectors \( \mathbf{\beta} \) are each full column rank and identified up to an arbitrary \((r, r)\) non-singular matrix. 9

The estimation of the cointegrating matrix \( \mathbf{\beta} \) 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
\[ \Delta y_t = c_0 + \Lambda \Delta x_t + \sum_{i=1}^{p-1} \Psi_i \Delta z_{t-i} + \Pi_y \gamma z_{t-1}^* + u_t, \quad t = 1, 2, \ldots, \] (3.15)

---

9 That is, \( (\mathbf{x}, \mathbf{K}^{-1}) (\mathbf{K} \mathbf{\beta}) = \mathbf{x}_y \mathbf{\beta}^\prime \) for any \((r, r)\) non-singular matrix \( \mathbf{K} \).
where \( z_{t-1}^* = (t, z'_{t-1})' \), and \( \Pi_{y*} = \Pi_y(-\gamma, I_m) \). Note that \( \text{Rank}[\Pi_{y*}] = \text{Rank}[\Pi_y] \) and, thus, from (3.14):\(^{10}\)

\[
\Pi_{y*} = \alpha_y\beta_y^*,
\]

where

\[
\beta_y^* = \left(-\gamma I_m\right)^{-1}\beta.
\]

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

\[
H_r: \text{Rank}[\Pi_{y*}] = r, \quad r = 0, \ldots, n.
\]

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

\[
\Delta Y = \alpha_0 1_T + \Psi \Delta Z - \Pi_{y*}Z_{t-1}^* + U,
\]

where \( \Delta Y \equiv (\Delta y_1, \ldots, \Delta y_T), 1_T \) is a \( T \)-vector of ones, \( \Delta X \equiv (\Delta x_1, \ldots, \Delta x_T), \Delta Z_{t-i} \equiv (\Delta z_{t-i}, \ldots, \Delta z_{t-1}) \), \( i = 1, \ldots, p - 1 \), \( \Psi \equiv (\Psi_1, \ldots, \Psi_{p-1}) \), \( \Delta Z_{t-i} \equiv (\Delta x, \Delta z_{t-i}, \ldots, \Delta z_{t-p})' \), \( Z_{t-1}^* \equiv (\tau_T, Z_{t-1})' \), \( \tau_T \equiv (1, \ldots, T) \), and \( U \equiv (u_1, \ldots, u_T) \).

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

\[
\ell_T(\psi; r) = -\frac{nT}{2} \ln 2\pi - \frac{T}{2} \ln|\Omega_{yy}| - \frac{1}{2} Tr(\Omega_{yy}^{-1}UU'),
\]

where the parameter vector \( \psi \) collects together the unknown parameters in \( \Omega_{yy}, \alpha_0, \Psi \) and \( \Pi_{y*} \). Successively concentrating out \( \Omega_{yy}, \alpha_0 \) and \( \Psi \), and \( \alpha_y \) in (3.19) results in the concentrated log-likelihood function

\[
\ell_T^c(\beta_y^*; r) = -\frac{nT}{2} (1 + \ln 2\pi) - \frac{T}{2} \ln|T^{-1} \hat{\Delta} \hat{Y}(1_T - \hat{Z}_{t-1}^* \beta_y^*) \times (\beta_y^* \hat{Z}_{t-1}^* \beta_y^*)^{-1} \hat{Z}_{t-1}^* \Delta \hat{Y}|,
\]

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

\[
S_{YY} = T^{-1} \Delta \hat{Y} \Delta \hat{Y}' \quad S_{YZ} = T^{-1} \Delta \hat{Y} \hat{Z}_{t-1}' \quad S_{ZZ} = T^{-1} \hat{Z}_{t-1}' \hat{Z}_{t-1},
\]

\(^{10}\) The trend parameter vector \( \gamma \) is no longer identified in system (3.15) and (3.6). Only \( \beta_y^* \) and \( a_{y0} = \Gamma_y \gamma \) may be identified.

\(^{11}\) The corresponding estimators are given by \( \hat{\Omega}_{yy} = T^{-1} \hat{U}_T \hat{U}'_T \), where \( \hat{U} \) is defined via (3.18) and is a function of the unknown parameter matrices \( \alpha_0, \Psi \) and \( \Pi_{y*} \). In the case \( \alpha_0, \Pi_{y*} \)

\[
\hat{\Psi} = (\Delta Y - \hat{\Psi}(\Pi_{y*}) \Delta Z - \Pi_{y*} \hat{Z}_{t-1}) (1_{T \times T})^{-1},
\]

\[
\hat{\Pi}_{y*} = (\Delta Y - \Pi_{y*} \hat{Z}_{t-1}) P - \Delta Z (\Delta Z' P' \Delta Z')^{-1},
\]

where \( \hat{\Sigma} = \hat{\Psi}(\Pi_{y*}) \Delta Z_{t-1}^* \), \( \hat{\gamma} = \Delta \hat{Y} \hat{Z}_{t-1}' \beta_y^* \hat{Z}_{t-1} \hat{Z}_{t-1}' \beta_y^* \).
the maximization of the concentrated log-likelihood function \( \ell_T^* (\beta_*; r) \) of (3.20) reduces to the minimization of

\[
|S_{YY} - S_{YZ} \beta_* (b_* S_{ZZ} \beta_*)^{-1} \beta_* S_{SY}| = |S_{YY}| \frac{| \beta_* (S_{ZZ} - S_{ZY} S_{YY}^{-1} S_{YZ}) \beta_* |}{| \beta_* S_{ZZ} \beta_* |}
\]

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

\[
|\hat{\beta}_* S_{ZZ} - S_{ZY} S_{YY}^{-1} S_{YZ}| = 0; \quad (3.22)
\]

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 \( \beta_* \) are required for exact identification.\(^{12}\) The resultant maximized concentrated log-likelihood function \( \ell_T^* (\beta_*; r) \) at \( \hat{\beta}_* \) of (3.20) is

\[
\ell_T^* (r) = - \frac{n T}{2} (1 + \ln 2\pi) - \frac{T}{2} \ln |S_{YY}| - \frac{T}{2} \sum_{i=1}^r \ln (1 - \hat{\lambda}_i). \quad (3.23)
\]

Note that the maximized value of the log-likelihood \( \ell_T^* (r) \) 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 \( \Delta \hat{Y} \) and \( \hat{Z}_{-1} \) and, consequently, the sample moment matrices \( S_{YY}, S_{YZ} \) and \( S_{ZZ} \) given by (3.21). We state the required definitions below.

**Case I:** \( c_0 = 0 \) and \( c_1 = 0 \). \( \Delta \hat{Y} \) and \( \hat{Z}_{-1} \) are the OLS residuals from the regression of \( \Delta Y \) and \( Z_{-1} \) on \( \Delta Z_{-} \).

**Case II:** \( c_0 = - \Pi_s \mu \) and \( c_1 = 0 \). \( \Delta \hat{Y} \) and \( \hat{Z}_{-1} \) are the OLS residuals from the regression of \( \Delta Y \) and \( Z_{-1} \) on \( \Delta Z_{-} \), where \( Z_{-1} = (t_T, Z_{-1}') \).

**Case III:** \( c_0 \neq 0 \) and \( c_1 = 0 \). \( \Delta \hat{Y} \) and \( \hat{Z}_{-1} \) are the OLS residuals from the regression of \( \Delta Y \) and \( Z_{-1} \) on \( (t_T, \Delta Z_{-}') \).

**Case IV:** \( c_0 \neq 0 \) and \( c_1 = - \Pi_r \gamma \). \( \Delta \hat{Y} \) and \( \hat{Z}_{-1} \) are the OLS residuals from the regression of \( \Delta Y \) and \( Z_{-1} \) on \( (t_T, \Delta Z_{-}') \), where \( Z_{-1} = (\tau_T, Z_{-1}') \).

**Case V:** \( c_0 \neq 0 \) and \( c_1 \neq 0 \). \( \Delta \hat{Y} \) and \( \hat{Z}_{-1} \) are the OLS residuals from the regression of \( \Delta Y \) and \( Z_{-1} \) on \( (t_T, \tau_T, \Delta Z_{-}') \).

### 4. Structural tests for cointegration and tests of specification

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

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

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

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

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

Tables 6(a)–(e) of Appendix B provide the relevant asymptotic critical values. Thirdly, 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, \( c_0 = 0 \) or \( \beta'\mu = 0, \) or whether a trend should be present in Case IV, that is, the co-trending restriction \( c_1 = 0 \) or \( \beta'\gamma = 0, \) is considered in Section 4.4. Finally, 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 \{\( e_t \)\}\( _{t=-\infty}^{\infty} \) in (2.1) and, hence, on the structural error process \{\( u_t \)\}\( _{t=-\infty}^{\infty} \) in (3.2).

**Assumption 4.1.** The error process \{\( e_t \)\}\( _{t=-\infty}^{\infty} \) is such that

(a) (i) \( \mathbb{E}\{e_t|\{z_{t-i}\}_{i=1}^{t-1}, Z_0\} = 0, \) (ii) \( \text{var}\{e_t|\{z_{t-i}\}_{i=1}^{t-1}, Z_0\} = \Omega, \) \( \Omega \) positive definite;

(b) (i) \( \mathbb{E}\{u_t|x_t, \{z_{t-i}\}_{i=1}^{t-1}, Z_0\} = 0, \) (ii) \( \text{var}\{u_t|x_t, \{z_{t-i}\}_{i=1}^{t-1}, Z_0\} = \Omega_{uu}, \) where \( u_t \equiv e_{yt} - \Omega_{yx}\Omega_{xx}^{-1}e_{xt} \) and \( \Omega_{uu} \equiv \Omega_{yy} - \Omega_{yx}\Omega_{xx}^{-1}\Omega_{xy}; \)

(c) \( \text{sup}\, \mathbb{E}\{|\|e_t\|\|^s\} < \infty \) for some \( s > 2. \)

Assumption 4.1(a) states that the error process \{\( e_t \)\}\( _{t=-\infty}^{\infty} \) is a martingale difference sequence with constant conditional variance; hence, \{\( e_t \)\}\( _{t=-\infty}^{\infty} \) is an uncorrelated process. Therefore, the VECM (3.1) represents a conditional model for \( \Delta z_t \) given \{\( \Delta z_{t-i} \)\}\( _{i=1}^{t-1} \) and \( z_{t-1}, t = 1, 2, \ldots. \) Assumption 4.1(b) is a linear conditional mean condition; that is, under Assumption 4.1(b)(i),

\[ \mathbb{E}\{e_{yt}|x_t, \{z_{t-i}\}_{i=1}^{t-1}, Z_0\} = \Omega_{yx}\Omega_{xx}^{-1}e_{xt} \]

which, together with Assumption 4.1(b)(ii), also ensures that \( \text{var}\{e_{yt}|x_t, \{z_{t-i}\}_{i=1}^{t-1}, Z_0\} = \Omega_{uu}. \) Therefore, under this assumption, (3.15) can still be interpreted as a conditional model for \( \Delta y_t \) given \( \Delta x_t, \{\Delta z_{t-i}\}_{i=1}^{t-1} \) and \( z_{t-1}, t = 1, 2, \ldots. \) Hence, (3.15) remains appropriate for conditional inference. Moreover, the error process \{\( u_t \)\}\( _{t=-\infty}^{\infty} \) is also a martingale

---

\(^{13}\) See Harbo et al. (1998, Theorem 1, 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).
difference process with constant conditional variance and is uncorrelated with the \( \{e_{it}\}_{t=-\infty}^{\infty} \) 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 \( \{z_t\}_{t=1}^{\infty} \) process itself as the VAR(1) form (2.1) together with Assumptions 2.1 and 2.2 yields \( \sum_{j=0}^{\infty} ||C_j|| < \infty \), where \( ||A|| = [\text{tr}(A^tA)]^{1/2} \); see Phillips and Solo (1992, Theorem 3.15(b), p. 983). 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, Remark 3.17(iii), p. 983). Note also that \( \sum_{j=0}^{\infty} ||C_j^u|| < \infty \), and \( ||C|| < \infty \), 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

\[
\mathcal{S}_y^n(a) \equiv T^{-1/2} \sum_{s=1}^{[Ta]} e_s,
\]

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

\[
\mathcal{S}_y^n(a) \Rightarrow B_m(a), \quad a \in [0, 1],
\]

(4.1)

where \( B_m(.) \) denotes an \( m \)-dimensional Brownian motion with variance matrix \( \Omega \). We partition \( \mathcal{S}_y^n(a) = (\mathcal{S}_y^n(a'), \mathcal{S}_y^n(a'))' \) conformably with \( z_t = (y'_t, x'_t)' \) and the Brownian motion \( B_m(a) = (B_n(a'), B_k(a))' \) likewise, \( a \in [0, 1] \). Define \( \mathcal{S}_u^n(a) \equiv T^{-1/2} \sum_{s=1}^{[Ta]} u_s, \ a \in [0,1] \). Hence, as \( u_t \equiv e_{xt} - \Omega_{yx} \Omega_{xx}^{-1} e_{xt}, \)

\[
\mathcal{S}_u^n(a) \Rightarrow B_u^n(a),
\]

(4.2)

where \( B_u^n(a) \equiv B_n(a) - \Omega_{yx} \Omega_{xx}^{-1} B_k(a) \) is a Brownian motion with variance matrix \( \Omega_{uu} \) which is independent of \( B_k(a), \ a \in [0, 1] \). Consequently, the results described in Harbo et al. (1998) remain valid under Assumption 4.1.

Under Assumption 3.1, the \((m, m-r)\) matrix \( \mathbf{a} = \text{diag}(\mathbf{a}_y, \mathbf{a}_x^\perp) \), where \( \mathbf{a}_x^\perp \) is a \((k, k)\) non-singular matrix, is a basis for the orthogonal complement of the \((m, r)\) loadings matrix \( \mathbf{a} = (\mathbf{a}_y', \mathbf{0})' \). Hence, we define the \((m - r)\)-dimensional standard
Brownian motion $W_{m-r}(a) \equiv (W_{n-r}(a)'$, $W_k(a)')'$ partitioned into the $(n-r)$- and $k$-dimensional sub-vector independent standard Brownian motions $W_{n-r}(a) \equiv (x_{1r}^\top\Omega_{xU}x_{1r})^{-1/2}x_{1r}^\top B_{n}^X(a)$ and $W_k(a) \equiv (x_{2r}^\top\Omega_{xU}x_{2r})^{-1/2}x_{2r}^\top B_{k}^X(a)$, $a \in [0, 1]$. See Pesaran et al. (1997, Appendix A) for further details. We will also require the corresponding de-meaned (m - r)-vector standard Brownian motion

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

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

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

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

4.1. Testing $H_r$ against $H_{r+1}$

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

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

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

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

$$\int_0^1 dW_{n-r}(a) F_{m-r}(a) \left( \int_0^1 F_{m-r}(a) F_{m-r}(a)' \, da \right)^{-1} \int_0^1 F_{m-r}(a) \, dW_{n-r}(a)' , \quad (4.6)$$

where

$$F_{m-r}(a) = \begin{cases} W_{m-r}(a) & \text{Case I} \\ (W_{m-r}(a)', 1)' & \text{Case II} \\ \tilde{W}_{m-r}(a) & \text{Case III}, \quad a \in [0, 1], \\ (\hat{W}_{m-r}(a)', a - \frac{1}{2})' & \text{Case IV} \\ \hat{W}_{m-r}(a) & \text{Case V} \end{cases}$$

$r = 0, \ldots, n - 1$, 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

$$LR(H_rD_H_n) = -T \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_i),$$

(4.7)

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

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

$$\text{Trace} \left\{ \int_0^1 dW_{n-r}(a)F_{m-r}(a) \left( \int_0^1 F_{m-r}(a)dW_{n-r}(a)' \right)^{-1} \times \int_0^1 F_{m-r}(a)dW_{n-r}(r) \right\},$$

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

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 $\{x_t\}_{t=1}^{\infty}$ 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 to deal with this difficulty. Let $\{w_t\}_{t=1}^{\infty}$ denote a $k_w$-vector process of weakly exogenous explanatory variables which is integrated of order zero. Therefore, the partial sum vector process $\sum_{s=1}^{t} w_s$ is integrated of order one. Defining $\sum_{s=1}^{t} w_s$ as a sub-vector of $x_t$ with the corresponding subvector of $\Delta x_t$ as $w_t, t = 1, 2, \ldots$, allows the above analysis to proceed unaltered. With these re-definitions of $x_t$ and $\Delta x_t$ to include $\sum_{s=1}^{t} w_s$ and $w_t, t = 1, 2, \ldots$, respectively, the partial sum $\sum_{s=1}^{t} w_s$ will now appear in the cointegrating relations (2.13) and the lagged level term $z_{t-1}$ in (3.5) although economic theory may indicate its absence; that is, the corresponding $(k_w, r)$ block of the cointegrating matrix $\beta$ is null. This constraint on the cointegrating matrix $\beta$ 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 \( \{ z_t \}_{t=1}^{\infty} \) 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 \( c_1 = - \Pi y \) (see (3.7)). In the absence of a deterministic trend in (2.1), \( y = 0 \), then the intercept parameter in the VECM formulation (3.5) is similarly restricted as \( c_0 = - \Pi y \), 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. \( \beta' y = 0 \) or equivalently \( c_1 = 0 \) in the VECM formulation (3.5) for Case IV which yields Case III. In the absence of a trend, \( y = 0 \) (Case II), we also consider the likelihood ratio test for the absence of an intercept in the cointegrating relationship (2.13); viz. \( \beta' \mu = 0 \) or equivalently \( c_0 = 0 \) 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 \( c_1 = 0 \) in Case IV). In Case IV, under \( H_r \) defined by (3.8), Assumptions 2.1, 2.2, 3.1 and 4.1, and \( c_1 = 0 \), the likelihood ratio statistic for \( c_1 = 0 \) has a limiting chi-squared distribution with \( r \) degrees of freedom, \( r = 1, \ldots, n \).

Similarly,

**Theorem 4.4** (Limit distribution of the likelihood ratio statistic for \( c_0 = 0 \) in Case II). In Case II, under \( H_r \) defined by (3.8), Assumptions 2.1, 2.2, 3.1 and 4.1, and \( c_0 = 0 \), the likelihood ratio statistic for \( c_0 = 0 \) has a limiting chi-squared distribution with \( r \) degrees of freedom, \( r = 1, \ldots, n \).

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 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 \( \{z_t\}_{t=1}^{\infty} \) 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 \( c_1 = -\Pi_x^t \gamma \)). 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 \( c_1 = -\Pi_x^t \gamma \), the likelihood ratio statistic for \( c_1 = -\Pi_x^t \gamma \) has a limiting chi-squared distribution with \( n - r \) degrees of freedom, \( r = 0, \ldots, n - 1 \).

Similarly,

**Theorem 4.6** (Limit distribution of the likelihood ratio statistic for \( c_0 = -\Pi_x^t \mu \)). 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 \( c_0 = -\Pi_x^t \mu \), the likelihood ratio statistic for \( c_0 = -\Pi_x^t \mu \) has a limiting chi-squared distribution with \( n - r \) degrees of freedom, \( r = 0, \ldots, n - 1 \).

Secondly, in Assumption 3.1, we imposed the weak exogeneity restriction \( \Pi_x = 0 \); that is, the level process \( \{x_t\}_{t=1}^{\infty} \) is integrated of order one and long-run forcing for \( \{y_t\}_{t=1}^{\infty} \). This restriction takes two forms. Firstly, as noted in Section 3, the level process \( \{x_t\}_{t=1}^{\infty} \) is not mutually cointegrated. Secondly, the lagged cointegrating relationship does not enter the evolution of the differenced process \( \{\Delta x_t\}_{t=1}^{\infty} \) 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:

\[ \Delta x_t = a_{x0} + \sum_{i=1}^{p-1} \Gamma_{xi} \Delta z_{t-i} + \Pi_{xx*} x_{t-1}^* + e_{xt}, \]  

(4.8)

where \( \Pi_{xx*} \equiv \Pi_{xx}(y_x, I_k), \Pi_x \equiv (\Pi_{sx}, \Pi_{xx}), \gamma \equiv (y_{x}' y_x)' \) and \( x_{t-1}^* \equiv (t, x_{t-1})' \), \( t = 1, 2, \ldots \); cf. (3.6). Eq. (4.8) embodies the possibility that the level process \( \{x_t\}_{t=1}^\infty \) is mutually cointegrated. We denote the cointegration rank hypothesis in (4.8) by \( H_{xr} \): \( \text{Rank}[\Pi_x] = r, r = 0, \ldots, k \). Therefore, similarly to Theorems 4.1 and 4.2 (cf. Johansen, 1991).

**Theorem 4.7** (Limit distribution of \( \mathcal{L}(H_0^*|H_1^*|) \)). 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 \( \mathcal{L}(H_0^*|H_1^*) \) for testing \( H_0^* \) against \( H_1^* \) in (4.8) is the distribution of the maximum eigenvalue of

\[ \int_0^1 \mathbf{dW}_k(a) \mathbf{F}_k(a)' \left( \int_0^1 \mathbf{F}_k(a) \mathbf{F}_k(a)' \mathbf{d}a \right)^{-1} \int_0^1 \mathbf{F}_k(a) \mathbf{dW}_k(a)', \]

where

\[ \mathbf{F}_k(a) = \begin{cases} \mathbf{W}_k(a) & \text{Case I} \\ (\mathbf{W}_k(a)', 1)' & \text{Case II} \\ (\mathbf{W}_k(a)', a - \frac{1}{2})' & \text{Case IV} \\ \hat{\mathbf{W}}_k(a) & \text{Case V} \end{cases} \]

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

**Theorem 4.8** (Limit distribution of \( \mathcal{L}(H_0^*|H_1^*|) \)). 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 \( \mathcal{L}(H_0^*|H_1^*) \) for testing \( H_0^* \) against \( H_1^* \) in (4.8) is the distribution of

\[ \text{Trace}\left\{ \int_0^1 \mathbf{dW}_k(a) \mathbf{F}_k(a)' \left( \int_0^1 \mathbf{F}_k(a) \mathbf{F}_k(a)' \mathbf{d}a \right)^{-1} \int_0^1 \mathbf{F}_k(a) \mathbf{dW}_k(a)' \right\}, \]

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

\[ \Delta x_t = a_{x0} + \sum_{i=1}^{p-1} \Gamma_{xi} \Delta z_{t-i} + \alpha_{xy} \hat{\beta}_s z_{t-1}^* + \tilde{e}_{xt}, \]  

(4.9)

t = 1, \ldots, T, \] where \( \hat{\beta}_s \) denotes the ML estimator for the cointegrating vector \( \beta_s \) under \( H_r \) in Cases I–V of the earlier sections. Eq. (4.9) allows for the
alternative possibility that the lagged cointegrating relationships \( \{ \beta'_{it} z_{it}^s \}_{i=1}^\infty \) might enter (3.6).

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

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 et al. (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 \( \Psi \) associated with the short-run dynamics in the VECM (3.18):

\[
\text{vec}(\Psi') = S\varphi, \tag{5.1}
\]

where \( S \) is a \((nmp - n^2, s)\) matrix of full column rank \( s \) with known elements and \( \varphi \) an \( s \)-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,

\[
\Delta y'_j = c_0 y'_T + \varphi'_j S'_j \Delta Z_ - + \pi'_{y\theta_j} Z_{\theta j}^\theta + u'_j \tag{5.2}
\]

where \( \Delta y_j \equiv (\Delta y_{1j}, \ldots, \Delta y_{Tj})' \) and \( u_j \equiv (u_{1j}, \ldots, u_{Tj})' \), \( j = 1, \ldots, n \), and \( \Pi_{y\theta_j} = (\pi_{y\theta_1}, \ldots, \pi_{y\theta_n})' \). The formulation (5.2) allows certain elements of the stationary component \( \Delta Z_ - \) to be excluded from the \( j \)th equation and, thus, also the lag lengths to differ across equations. These constraints are summarized by \( \psi'_j = S'_j \varphi'_j \), with \( S'_j \) an appropriately defined selection matrix, \( j = 1, \ldots, n \), and \( \Psi = (\psi'_1, \ldots, \psi'_n)' \). Stacking these \( n \) equations yields

\[
\Delta Y = c_0 y'_T + \Psi' \Delta Z_ - + \Pi_{y\theta} Z_{\theta - 1}^\theta + U, \tag{5.3}
\]

\[\text{Boswijk (1995) provides a Lagrange multiplier test for } \alpha_{xy} = 0 \text{ in (4.9). See also Johansen (1992) and Harbo et al. (1998, Section 4.1).} \]
where $\Delta Y \equiv (\Delta y_1, \ldots, \Delta y_n)'$, subject to the constraints $vec(\Psi') = S\phi$ as in (5.1) with $S \equiv \text{diag}(S^1, \ldots, S^n)$ and $\phi = (\phi^1', \ldots, \phi^r')'$; cf. (3.18). Note that the constraints (5.1) also permit the possibility of cross-equation restrictions.

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

The next step 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

$$(\Delta Y - \hat{\Psi}\Delta Z) = c_01_T + \Pi_{y*}Z_{*-1} + \hat{U},$$

where $vec(\hat{\Psi}) = S\hat{\phi}$ and $\hat{U} \equiv U - (\hat{\Psi} - \Psi)\Delta Z$.

For this purpose consider the minimization with respect to $\beta_*$ under $H_r$ of (3.17) of the determinantal criterion function

$$|T^{-1}\Delta \tilde{Y}(I_T - \tilde{Z}_{*-1}^{*'}\beta_*(\beta_*'\tilde{Z}_{*-1}'\beta_*)^{-1}\beta_*'\tilde{Z}_{*-1})\Delta \tilde{Y}|,$$

cf. (3.20), where $\Delta \tilde{Y} \equiv (\Delta Y - \hat{\Psi}\Delta Z)\hat{P}_t$ and $\tilde{Z}_{*-1} = Z_{*-1}^*\hat{P}_t$ are respectively the de-meaned versions of $\Delta Y - \hat{\Psi}\Delta Z$ and $Z_{*-1}^*$, $\hat{P}_t \equiv I_T - t_t(t'_tt)'^{-1}t'_t$, which results from (5.4) by concentrating out $\alpha_y$ after substitution of $\Pi_{y*} = \alpha_y\beta_*'$. Defining the moment matrices

$$\hat{S}_{Yy} \equiv T^{-1}\Delta \tilde{Y}\Delta \tilde{Y}, \quad \hat{S}_{Yz} \equiv T^{-1}\Delta \tilde{Y}\tilde{Z}_{*-1}', \quad \hat{S}_{zz} \equiv T^{-1}\tilde{Z}_{*-1}'\tilde{Z}_{*-1},$$

cf. (3.21), the solution $\hat{\beta}_*$ to the above minimization problem is given by the eigenvectors corresponding to the $r$ largest eigenvalues $\lambda_1 > \cdots > \lambda_r > 0$ of (cf. (3.22))

$$|\tilde{T}\hat{S}_{zz} - \hat{S}_{zz}\hat{S}_{yy}^{-1}\hat{S}_{yz}| = 0.$$

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

\footnote{A proof of $T^{1/2}$-consistency of the SUR estimator of $\phi$ 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 and 5.4, an estimator for $\Omega_{uu}$ based on residuals from equation-by-equation ordinary least squares treating $\Pi_{y*}$ 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 $\phi$ is $T^{1/2}$-consistent.}
4.2 respectively for Cases I–IV, required for (5.7) and (5.6) to be applicable. Therefore, as the regression contains only variables which are asymptotically b-valued and subject to the constraints of (2.1, 2.2, 3.1 and 4.1, the limit distributions of the pseudo-likelihood ratio statistics \( \mathcal{R}(H_r|H_{r+1}) \) and \( \mathcal{R}(H_r|H_n) \) are identical to those of (4.5) and (4.7) stated in Theorems 4.1 and 4.2 respectively for Cases I–V, \( r = 0, \ldots, n - 1 \).

For Case IV, let \( \tilde{\beta}_* \) denote the \((m+1, r)\) matrix of just-identified eigenvectors of (5.6). Consider the multivariate regression

\[
\Delta Y = c_0' \tau_T + \Upsilon \Delta Z - + \alpha \tilde{\beta}_* Z_{n-1} + \tilde{\Upsilon},
\]

subject to the constraints \( \text{vec}(\Psi') = \mathbf{S} \Phi \) of (5.1); cf. (5.3). Asymptotically, the inclusion of the lagged estimated cointegrating relationships \( \tilde{\beta}_* Z_{n-1} \) is the same as the inclusion of the actual cointegrating relationships \( \beta_{*} T_{n-1}, t = 1, \ldots, T \). 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 \( \Phi \) and the \((n, r)\) loadings matrix \( \alpha_{*} \). 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, $y_t = (y_{1t}, y_{2t})$. Their VAR(2) model is augmented by a scalar exogenous $I(1)$ variable $x_t$, which results in the following DGP:

$$(I_3 - AL)(I_3 - BL)z_t = e_t,$$  \hspace{1cm} (6.1)

where $z_t = (y_t', x_t')$, 

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

and $e_t \sim \text{IN}(0, \Omega)$. 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, $\text{Rank}[P] = 1$, 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 $\mu$ and $\gamma$, we have set $\mu = 0$ and $\gamma = 0$ in (6.1).

Consequently, the structural VECM (3.5) is given by

$$\Delta y_t = \lambda \Delta x_t + \Psi_1 \Delta z_{t-1} + \Pi_y z_{t-1} + u_t,$$  \hspace{1cm} (6.3)

where $u_t \sim \text{IN}(0, \Omega_{uu})$,

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

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

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

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

$$\Pi_y = -\begin{pmatrix} 1 - a_{11} \\ -a_{21} \end{pmatrix} (1, -b_{12}, a).$$
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 \( \text{var}(\varepsilon_{xt}) = 1 \) yielding \( \lambda_1 \) and \( \lambda_2 \) as the covariances of \( y_{1t} \) and \( y_{2t} \) with \( x_t \), respectively.

The above DGP, characterized via \( \lambda \), \( \Psi_1 \), \( \Pi_y \) and \( \Omega_{uu} \), 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, \( \text{Rank}[\Pi_y] = 1 \), 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:

| Parameters | Experiments: Set I | Experiments: Set II | | Parameters | Experiments: Set I | Experiments: Set II |
|-----------|-------------------|-------------------| | Parameters | Experiments: Set I | Experiments: Set II |
| \( a_{11} \) | 1.0 | 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 |
| \( \rho \) | 0 | 0.5 | 0 | 0.5 | 0 | 0.5 | 0 | 0.5 |

| Implied \( R^2 \) | | | | | | | | |
| \( R^2_{y_{1t}} \) | 0.121 | 0.121 | 0.123 | 0.122 | 0.198 | 0.172 | 0.290 | 0.270 |
| \( R^2_{y_{2t}} \) | 0.163 | 0.152 | 0.198 | 0.188 | 0.087 | 0.083 | 0.138 | 0.134 |

\(^{17}\)Population \( R^2 \) values of the structural VECM equations are obtained as

\[ R^2_{\Delta y_{it}} = 1 - \frac{\text{Var}(\Delta y_{it}|\Delta x_t, \Delta z_{i-1}, z_{i-1})}{\text{Var}(\Delta y_{it})} \]

\[ = 1 - \frac{1}{\text{Var}(\Delta y_{it})}, \quad i = 1,2. \]

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

\[ \text{Cov}(\Delta z_i) = \sum_{j=0}^{a_i} C_f \Omega C_j; \]

see below (2.9) and fn. 3.

\(^{18}\)Other experiments with slightly larger values for these population \( R^2 \) were also conducted and yielded similar results to those described below.
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) $LR_1$: the statistics due to Johansen (1991) which treat all three variables in $z_t$ as endogenous;
(ii) $LR_2$: the statistics developed in Sections 4.1 and 4.2 which correctly treat $x_t$ as an exogenous $I(1)$ variable;
(iii) $LR_3$: the statistics of Section 5 which also incorporate the zero restrictions on the short-run dynamics in the matrix $\Psi_1$ which is estimated by equation-by-equation ordinary least squares treating $P_{yi}$ 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.\(^\text{19}\)

The rejection frequencies for testing $H_r$: $\text{Rank}(P_{yi}) = r$ 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) 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 $LR_3$ tests of Section 5 is significantly higher than their $LR_1$ and $LR_2$ counterparts with the power of the $LR_2$ tests of Sections 4.1 and 4.2 being slightly better than the $LR_1$ tests of Johansen (1991) in most cases. There seems little to choose between the $LR_3$ 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 $LR_3$ tests still perform best in terms of power with the $LR_2$ test also slightly more powerful than the $LR_1$ tests. In relative terms, the results are similar to those in Table 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 $R^2_{\text{yit}}, i = 1, 2$) 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 $\rho$ has little effect on the finite sample performance of these tests for the experiments in Set I.

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 $\rho$, although it does not result in a higher value for the population $R^2_{\text{yit}}, i = 1, 2$, improves the power of all tests significantly.

\(^{19}\) Similar results for Cases I–V to those for Case IV given below were also found and are available from the authors on request.
Table 1(a)
Rejection frequencies for experiment $E_{1a}$

<table>
<thead>
<tr>
<th>$T$</th>
<th>$H_r$</th>
<th>$LR1$</th>
<th>$LR2$</th>
<th>$LR3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$H_0$</td>
<td>$H_{r+1}$</td>
<td>$H_0$</td>
</tr>
<tr>
<td>30</td>
<td>$r = 0$</td>
<td>0.204</td>
<td>0.166</td>
<td>0.196</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.024</td>
<td>0.015</td>
<td>0.020</td>
</tr>
<tr>
<td>50</td>
<td>$r = 0$</td>
<td>0.151</td>
<td>0.123</td>
<td>0.164</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.017</td>
<td>0.009</td>
<td>0.015</td>
</tr>
<tr>
<td>75</td>
<td>$r = 0$</td>
<td>0.199</td>
<td>0.159</td>
<td>0.226</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.020</td>
<td>0.013</td>
<td>0.022</td>
</tr>
<tr>
<td>100</td>
<td>$r = 0$</td>
<td>0.280</td>
<td>0.248</td>
<td>0.342</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.029</td>
<td>0.019</td>
<td>0.031</td>
</tr>
<tr>
<td>150</td>
<td>$r = 0$</td>
<td>0.418</td>
<td>0.430</td>
<td>0.524</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.033</td>
<td>0.031</td>
<td>0.040</td>
</tr>
<tr>
<td>300</td>
<td>$r = 0$</td>
<td>0.957</td>
<td>0.988</td>
<td>0.988</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.050</td>
<td>0.051</td>
<td>0.053</td>
</tr>
</tbody>
</table>

Table 1(b)
Rejection frequencies for experiment $E_{1b}$

<table>
<thead>
<tr>
<th>$T$</th>
<th>$H_r$</th>
<th>$LR1$</th>
<th>$LR2$</th>
<th>$LR3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$H_0$</td>
<td>$H_{r+1}$</td>
<td>$H_0$</td>
</tr>
<tr>
<td>30</td>
<td>$r = 0$</td>
<td>0.193</td>
<td>0.164</td>
<td>0.184</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.021</td>
<td>0.014</td>
<td>0.018</td>
</tr>
<tr>
<td>50</td>
<td>$r = 0$</td>
<td>0.135</td>
<td>0.112</td>
<td>0.144</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.013</td>
<td>0.008</td>
<td>0.012</td>
</tr>
<tr>
<td>75</td>
<td>$r = 0$</td>
<td>0.162</td>
<td>0.133</td>
<td>0.187</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.017</td>
<td>0.010</td>
<td>0.017</td>
</tr>
<tr>
<td>100</td>
<td>$r = 0$</td>
<td>0.227</td>
<td>0.196</td>
<td>0.278</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.022</td>
<td>0.016</td>
<td>0.025</td>
</tr>
<tr>
<td>150</td>
<td>$r = 0$</td>
<td>0.419</td>
<td>0.429</td>
<td>0.519</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.036</td>
<td>0.035</td>
<td>0.036</td>
</tr>
<tr>
<td>300</td>
<td>$r = 0$</td>
<td>0.958</td>
<td>0.991</td>
<td>0.989</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.048</td>
<td>0.047</td>
<td>0.049</td>
</tr>
</tbody>
</table>
Table 1(c)
Rejection frequencies for experiment $E_{1c}$

<table>
<thead>
<tr>
<th>$T$</th>
<th>$H_r$</th>
<th>$LR1$</th>
<th>$LR2$</th>
<th>$LR3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$H_n$</td>
<td>$H_{r+1}$</td>
<td>$H_n$</td>
</tr>
<tr>
<td>30</td>
<td>$r = 0$</td>
<td>0.231</td>
<td>0.182</td>
<td>0.221</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.028</td>
<td>0.019</td>
<td>0.024</td>
</tr>
<tr>
<td>50</td>
<td>$r = 0$</td>
<td>0.203</td>
<td>0.159</td>
<td>0.221</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.024</td>
<td>0.015</td>
<td>0.022</td>
</tr>
<tr>
<td>75</td>
<td>$r = 0$</td>
<td>0.299</td>
<td>0.257</td>
<td>0.351</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.033</td>
<td>0.024</td>
<td>0.034</td>
</tr>
<tr>
<td>100</td>
<td>$r = 0$</td>
<td>0.443</td>
<td>0.431</td>
<td>0.535</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.044</td>
<td>0.036</td>
<td>0.047</td>
</tr>
<tr>
<td>150</td>
<td>$r = 0$</td>
<td>0.763</td>
<td>0.824</td>
<td>0.858</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.054</td>
<td>0.053</td>
<td>0.058</td>
</tr>
<tr>
<td>300</td>
<td>$r = 0$</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.056</td>
<td>0.055</td>
<td>0.057</td>
</tr>
</tbody>
</table>

Table 1(d)
Rejection frequencies for experiment $E_{1d}$

<table>
<thead>
<tr>
<th>$T$</th>
<th>$H_r$</th>
<th>$LR1$</th>
<th>$LR2$</th>
<th>$LR3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$H_n$</td>
<td>$H_{r+1}$</td>
<td>$H_n$</td>
</tr>
<tr>
<td>30</td>
<td>$r = 0$</td>
<td>0.226</td>
<td>0.182</td>
<td>0.219</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.027</td>
<td>0.017</td>
<td>0.022</td>
</tr>
<tr>
<td>50</td>
<td>$r = 0$</td>
<td>0.198</td>
<td>0.157</td>
<td>0.220</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.023</td>
<td>0.015</td>
<td>0.022</td>
</tr>
<tr>
<td>75</td>
<td>$r = 0$</td>
<td>0.289</td>
<td>0.252</td>
<td>0.345</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.033</td>
<td>0.023</td>
<td>0.036</td>
</tr>
<tr>
<td>100</td>
<td>$r = 0$</td>
<td>0.433</td>
<td>0.422</td>
<td>0.526</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.041</td>
<td>0.033</td>
<td>0.046</td>
</tr>
<tr>
<td>150</td>
<td>$r = 0$</td>
<td>0.751</td>
<td>0.811</td>
<td>0.848</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.054</td>
<td>0.052</td>
<td>0.057</td>
</tr>
<tr>
<td>300</td>
<td>$r = 0$</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.054</td>
<td>0.056</td>
<td>0.057</td>
</tr>
</tbody>
</table>
Table 2(a)
Rejection frequencies for experiment $E_{2a}$

<table>
<thead>
<tr>
<th>$T$</th>
<th>$H_r$</th>
<th>$LR1$</th>
<th>$LR2$</th>
<th>$LR3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$H_n$</td>
<td>$H_{r+1}$</td>
<td>$H_n$</td>
</tr>
<tr>
<td>30</td>
<td>$r = 0$</td>
<td>0.202</td>
<td>0.163</td>
<td>0.191</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.022</td>
<td>0.015</td>
<td>0.018</td>
</tr>
<tr>
<td>50</td>
<td>$r = 0$</td>
<td>0.156</td>
<td>0.128</td>
<td>0.166</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.015</td>
<td>0.009</td>
<td>0.014</td>
</tr>
<tr>
<td>75</td>
<td>$r = 0$</td>
<td>0.197</td>
<td>0.164</td>
<td>0.233</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.021</td>
<td>0.012</td>
<td>0.022</td>
</tr>
<tr>
<td>100</td>
<td>$r = 0$</td>
<td>0.276</td>
<td>0.251</td>
<td>0.340</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.027</td>
<td>0.022</td>
<td>0.029</td>
</tr>
<tr>
<td>150</td>
<td>$r = 0$</td>
<td>0.521</td>
<td>0.554</td>
<td>0.642</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.041</td>
<td>0.036</td>
<td>0.043</td>
</tr>
<tr>
<td>300</td>
<td>$r = 0$</td>
<td>0.986</td>
<td>0.998</td>
<td>0.998</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.054</td>
<td>0.054</td>
<td>0.053</td>
</tr>
</tbody>
</table>

Table 2(b)
Rejection frequencies for experiment $E_{2b}$

<table>
<thead>
<tr>
<th>$T$</th>
<th>$H_r$</th>
<th>$LR1$</th>
<th>$LR2$</th>
<th>$LR3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$H_n$</td>
<td>$H_{r+1}$</td>
<td>$H_n$</td>
</tr>
<tr>
<td>30</td>
<td>$r = 0$</td>
<td>0.258</td>
<td>0.197</td>
<td>0.247</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.035</td>
<td>0.023</td>
<td>0.028</td>
</tr>
<tr>
<td>50</td>
<td>$r = 0$</td>
<td>0.262</td>
<td>0.210</td>
<td>0.291</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.032</td>
<td>0.021</td>
<td>0.030</td>
</tr>
<tr>
<td>75</td>
<td>$r = 0$</td>
<td>0.408</td>
<td>0.377</td>
<td>0.474</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.047</td>
<td>0.034</td>
<td>0.046</td>
</tr>
<tr>
<td>100</td>
<td>$r = 0$</td>
<td>0.597</td>
<td>0.616</td>
<td>0.691</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.055</td>
<td>0.049</td>
<td>0.055</td>
</tr>
<tr>
<td>150</td>
<td>$r = 0$</td>
<td>0.897</td>
<td>0.942</td>
<td>0.951</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.060</td>
<td>0.057</td>
<td>0.057</td>
</tr>
<tr>
<td>300</td>
<td>$r = 0$</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.060</td>
<td>0.060</td>
<td>0.057</td>
</tr>
</tbody>
</table>
Table 2(c)
Rejection frequencies for experiment $E_{2c}$

<table>
<thead>
<tr>
<th>$T$</th>
<th>$H_r$</th>
<th>$H_n$</th>
<th>$H_{r+1}$</th>
<th>$H_n$</th>
<th>$H_{r+1}$</th>
<th>$H_n$</th>
<th>$H_{r+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>$r = 0$</td>
<td>0.285</td>
<td>0.209</td>
<td>0.274</td>
<td>0.229</td>
<td>0.348</td>
<td>0.298</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.044</td>
<td>0.026</td>
<td>0.033</td>
<td>0.033</td>
<td>0.057</td>
<td>0.057</td>
</tr>
<tr>
<td>50</td>
<td>$r = 0$</td>
<td>0.324</td>
<td>0.261</td>
<td>0.364</td>
<td>0.325</td>
<td>0.502</td>
<td>0.465</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.044</td>
<td>0.030</td>
<td>0.042</td>
<td>0.042</td>
<td>0.077</td>
<td>0.077</td>
</tr>
<tr>
<td>75</td>
<td>$r = 0$</td>
<td>0.524</td>
<td>0.501</td>
<td>0.600</td>
<td>0.600</td>
<td>0.768</td>
<td>0.781</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.058</td>
<td>0.047</td>
<td>0.061</td>
<td>0.061</td>
<td>0.096</td>
<td>0.096</td>
</tr>
<tr>
<td>100</td>
<td>$r = 0$</td>
<td>0.735</td>
<td>0.769</td>
<td>0.814</td>
<td>0.842</td>
<td>0.932</td>
<td>0.948</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.066</td>
<td>0.062</td>
<td>0.065</td>
<td>0.065</td>
<td>0.094</td>
<td>0.094</td>
</tr>
<tr>
<td>150</td>
<td>$r = 0$</td>
<td>0.965</td>
<td>0.986</td>
<td>0.984</td>
<td>0.993</td>
<td>0.998</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.064</td>
<td>0.060</td>
<td>0.060</td>
<td>0.060</td>
<td>0.079</td>
<td>0.079</td>
</tr>
<tr>
<td>300</td>
<td>$r = 0$</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.061</td>
<td>0.060</td>
<td>0.058</td>
<td>0.058</td>
<td>0.066</td>
<td>0.066</td>
</tr>
</tbody>
</table>

Table 2(d)
Rejection frequencies for experiment $E_{2d}$

<table>
<thead>
<tr>
<th>$T$</th>
<th>$H_r$</th>
<th>$H_n$</th>
<th>$H_{r+1}$</th>
<th>$H_n$</th>
<th>$H_{r+1}$</th>
<th>$H_n$</th>
<th>$H_{r+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>$r = 0$</td>
<td>0.392</td>
<td>0.288</td>
<td>0.388</td>
<td>0.328</td>
<td>0.496</td>
<td>0.435</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.071</td>
<td>0.044</td>
<td>0.056</td>
<td>0.056</td>
<td>0.093</td>
<td>0.093</td>
</tr>
<tr>
<td>50</td>
<td>$r = 0$</td>
<td>0.539</td>
<td>0.490</td>
<td>0.595</td>
<td>0.579</td>
<td>0.770</td>
<td>0.763</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.074</td>
<td>0.057</td>
<td>0.072</td>
<td>0.072</td>
<td>0.119</td>
<td>0.119</td>
</tr>
<tr>
<td>75</td>
<td>$r = 0$</td>
<td>0.814</td>
<td>0.837</td>
<td>0.875</td>
<td>0.893</td>
<td>0.964</td>
<td>0.972</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.085</td>
<td>0.073</td>
<td>0.080</td>
<td>0.080</td>
<td>0.113</td>
<td>0.113</td>
</tr>
<tr>
<td>100</td>
<td>$r = 0$</td>
<td>0.958</td>
<td>0.978</td>
<td>0.978</td>
<td>0.987</td>
<td>0.997</td>
<td>0.998</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.082</td>
<td>0.078</td>
<td>0.078</td>
<td>0.078</td>
<td>0.103</td>
<td>0.103</td>
</tr>
<tr>
<td>150</td>
<td>$r = 0$</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.068</td>
<td>0.066</td>
<td>0.066</td>
<td>0.066</td>
<td>0.079</td>
<td>0.079</td>
</tr>
<tr>
<td>300</td>
<td>$r = 0$</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>$r = 1$</td>
<td>0.063</td>
<td>0.061</td>
<td>0.061</td>
<td>0.061</td>
<td>0.068</td>
<td>0.068</td>
</tr>
</tbody>
</table>
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 oversized with the $LR1$ and $LR2$ tests oversized for $T \leq 100$, 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 $p$ 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 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 U.K. 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 U.K. wholesale price index, $p_t$, the logarithm of the trade-weighted foreign wholesale price index, $p_{tH}$, the logarithm of oil prices, $p_{ot}$, and the domestic and foreign interest rate variables, $r_t = \ln(1 + R_t/100)$ and $r_{tH} = \ln(1 + R_{tH}/100)$, where $R_t$ is the U.K. three-month Treasury Bill rate, and $R_{tH}$ 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 $p_0_t$, 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 U.K. 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 $p_0_t$, treated as $I(1)$ and exogenous, we estimated the following system of equations over the period 1972(3)–1987(2):

$$
\Delta y_t = c_0 + (\Pi_y y)_t + A \Delta x_t + \Psi_1 \Delta z_{t-1} + \Pi_y z_{t-1} + u_t,
$$

$$
\Delta x_t = a_{x0} + \Gamma_{x1} \Delta z_{t-1} + e_{xt},
$$

where $y_t = (p_t, e_t, r_t, r^*_t)'$, $x_t = (p^*_t, p_0_t)'$, $z_t = (y'_t, x'_t)$, $\Pi_y$ and $\Psi_1$ are each (4, 6) coefficient matrices, and $A$ a (4, 2) coefficient matrix. Note that in the above specification the trend coefficients ($\Pi_y y$) are restricted as in Case IV defined by (3.12) and, hence, the level of $y_t$ 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 $L(\mathcal{H}_r|\mathcal{H}_n)$ and $L(\mathcal{H}_r|\mathcal{H}_{r+1})$ 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 6(d) with $k = 2$ and $n - r = 4, 3, 2, 1$.

Neither statistic rejects the hypothesis of no cointegration at the 0.05 significance level but $L(\mathcal{H}_r|\mathcal{H}_n)$ 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.

---

20 These results are based on the assumption that $p^*$ and $p_0$ 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 $p_0$ 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.
Table 3
Cointegration rank statistics

<table>
<thead>
<tr>
<th>( H_r )</th>
<th>LR ( (H_r \mid H_n) ) Statistic</th>
<th>0.05 CV</th>
<th>0.10 CV</th>
<th>LR ( (H_r \mid H_{r+1}) ) Statistic</th>
<th>0.05 CV</th>
<th>0.10 CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r = 0 )</td>
<td>80.71</td>
<td>81.20</td>
<td>76.68</td>
<td>34.09</td>
<td>37.85</td>
<td>35.04</td>
</tr>
<tr>
<td>( r = 1 )</td>
<td>46.62</td>
<td>56.43</td>
<td>52.71</td>
<td>25.59</td>
<td>31.68</td>
<td>29.00</td>
</tr>
<tr>
<td>( r = 2 )</td>
<td>21.03</td>
<td>35.37</td>
<td>32.51</td>
<td>14.10</td>
<td>24.88</td>
<td>22.53</td>
</tr>
<tr>
<td>( r = 3 )</td>
<td>6.93</td>
<td>18.08</td>
<td>15.82</td>
<td>6.93</td>
<td>18.08</td>
<td>15.82</td>
</tr>
</tbody>
</table>

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 \( \Delta p_t, \Delta e_t, \Delta r_t \) and \( \Delta r_t^* \) suggest the following 24 zero restrictions on the short-run coefficients:

\[
A = \begin{bmatrix}
0 & * \\
0 & 0 \\
0 & 0 \\
0 & 0
\end{bmatrix}, \quad \Psi_1 = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & * & * & 0 & 0 \\
0 & 0 & * & 0 & 0 \\
0 & 0 & 0 & * & 0
\end{bmatrix},
\]

(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

<table>
<thead>
<tr>
<th>Unrestricted</th>
<th>Restricted</th>
</tr>
</thead>
<tbody>
<tr>
<td>LL</td>
<td>725.01</td>
</tr>
<tr>
<td>AIC</td>
<td>661.01</td>
</tr>
<tr>
<td>SBC</td>
<td>593.99</td>
</tr>
</tbody>
</table>

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 $\mathcal{LR}(H_r|H_n)$ and $\mathcal{LR}(H_r|H_{r+1})$ of Section 5 based on the above restricted model. 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 co-integrating 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 $\beta^* = (-\gamma, I_m)\beta$ associated with $\varepsilon^* = (t, p_t, e_t, r_t, r_n^*, p_t^*, p_o^*)$ 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 $\beta^*_1$ as explaining domestic prices and $\beta^*_2$ 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.

$$H_E: \beta^*_1 = \begin{pmatrix} * & 1 & * & * & 0 & * & * \\ * & 0 & * & 1 & * & * & * \end{pmatrix},$$

which yielded the estimate

$$\hat{\beta}^*_{1E} = \begin{pmatrix} 0.0195 & -0.701 & -6.005 & 0 \\ (0.0229) & (0.345) & (2.088) \end{pmatrix} - \begin{pmatrix} 2.564 & 0.206 \\ (1.943) & (0.274) \end{pmatrix},$$

$$\hat{\beta}^*_{2E} = \begin{pmatrix} -0.0062 & 0 & -0.091 \\ (0.0038) & (0.095) \end{pmatrix} - \begin{pmatrix} 1 & 0.791 & 0.493 & -0.054 \\ (1) & (0.326) & (0.334) & (0.046) \end{pmatrix},$$

$$LL_E = 714.50,$$
where $LL_E$ is the maximized value of the log-likelihood function for the just-identified case. Asymptotic standard errors are given in parentheses.

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:

$$H_{co}: \quad \beta^*_{co} = \begin{pmatrix} 0 & 1 & * & * & 0 & * \\ 0 & 0 & 1 & * & * & * \end{pmatrix}.$$

Under $H_{co}$, we obtained the estimate

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

$$LL_{co} = 713.30.$$

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 $H_{po}$. To save space, we only report the test of this hypothesis jointly with the co-trending hypothesis, $H_{co} \cap H_{po}$, which yields

$$\hat{\beta}^*_{co,po} = \begin{pmatrix} 0 & 1 & -0.834 & -5.646 & 0 & -0.925 & 0 \\ 0 & 0 & -0.062 & 1 & -0.803 & -0.005 & 0 \\ 0.327 & 1.850 & 0.227 & 0.102 & 0.432 & 0.063 \end{pmatrix},$$

$$LL_{co,po} = 711.49.$$

---

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

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

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

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

the estimate of which is

$$\hat{\beta}_{UIP} = \begin{pmatrix} 0 & 1 & -0.965 & -5.132 & 0 & -0.943 & 0 \\ 0.281 & 1.257 & 0.167 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$LL_{UIP} = 709.61.$$  

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 $H_{eo} \cap H_{po}$) is not rejected even at the 0.01 level.

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

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

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 $H_{eo} \cap H_{po}$ 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, consider the following modified version of the PPP hypothesis (denoted by $H_{PPP'}$) that allows the effect of domestic interest rate on

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

the real exchange rate to be unrestricted:

\[ H_{\text{PPP}^*} : \beta^* = \begin{pmatrix} 0 & 1 & -1 & * & 0 & -1 & 0 \\ 0 & 0 & * & 1 & * & * & 0 \end{pmatrix}. \]

Under these restrictions, we obtained the following estimate:

\[ \hat{\beta}^*_{\text{PPP}^*} = \begin{pmatrix} 0 & 1 & -1 & -5.363 & 0 & -1 & 0 \\ 0 & 0 & -0.022 & 1 & -0.778 & -0.002 & 0 \\ 0 & 0 & 0 & (0.97) & 1 & (0.325) & (0.045) \end{pmatrix}, \]

\[ LL_{\text{PPP}^*} = 709.81. \]

The (log-) likelihood ratio statistics for testing \( H_{\text{PPP}^*} \) 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 \( H_{\text{uo}} \cap H_{\text{po}} \)) yielding the following estimate:

\[ \hat{\beta}^*_{\text{PPP}^*,\text{UIP}} = \begin{pmatrix} 0 & 1 & -1 & -4.775 & 0 & -1 & 0 \\ 0 & 0 & 0 & (0.627) & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \end{pmatrix}, \]

\[ LL_{\text{PPP}^*,\text{UIP}} = 709.03. \] (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.

\[ ^{24} \text{Similar results also follow if instead of the domestic interest rate the coefficient of the foreign interest rate variable is left unrestricted.} \]
Table 5
Estimates of the error correction coefficients and diagnostic statistics

<table>
<thead>
<tr>
<th>Equation</th>
<th>$\alpha_{1y}$</th>
<th>$\alpha_{2y}$</th>
<th>$R^2$</th>
<th>$\chi^2_{SC}(4)$</th>
<th>$\chi^2_{FE}(1)$</th>
<th>$\chi^2_{JS}(2)$</th>
<th>$\chi^2_{HE}(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta p_t$</td>
<td>$-0.062$</td>
<td>$-0.063$</td>
<td>$0.784$</td>
<td>$3.35$</td>
<td>$3.31$</td>
<td>$3.15$</td>
<td>$0.02$</td>
</tr>
<tr>
<td>(0.016)</td>
<td>(0.066)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta e_t$</td>
<td>$0.124$</td>
<td>$0.049$</td>
<td>$0.154$</td>
<td>$0.76$</td>
<td>$0.72$</td>
<td>$0.13$</td>
<td>$0.06$</td>
</tr>
<tr>
<td>(0.070)</td>
<td>(0.285)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta r_t$</td>
<td>$0.044$</td>
<td>$-0.036$</td>
<td>$0.070$</td>
<td>$1.87$</td>
<td>$1.35$</td>
<td>$5.02$</td>
<td>$0.13$</td>
</tr>
<tr>
<td>(0.025)</td>
<td>(0.103)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta r_t^*$</td>
<td>$0.075$</td>
<td>$0.421$</td>
<td>$0.172$</td>
<td>$9.75$</td>
<td>$1.35$</td>
<td>$29.32$</td>
<td>$0.023$</td>
</tr>
<tr>
<td>(0.028)</td>
<td>(0.114)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

$$\hat{\beta}_1 z_{t-1} = p_{t-1} - e_{t-1} - p_{t-1}^* - 4.775 r_{t-1},$$
$$\hat{\beta}_2 z_{t-1} = r_{t-1} - r_{t-1}^*.$$ 

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

$$\Delta y_t = c_0 - \alpha_y \gamma y_t + \Delta \chi_t + \Psi_1 \Delta z_{t-1} + \alpha_y \beta z_{t-1} + u_t$$

$$= c_0 - \alpha_y \gamma y_t + \Delta \chi_t + \Psi_1 \Delta z_{t-1} + \alpha_{1y} \beta_1 z_{t-1} + \alpha_{2y} \beta_2 z_{t-1} + u_t,$$

where $\alpha_{1y}$ and $\alpha_{2y}$ are 4-dimensional vectors of adjustment (error correction) coefficients. The estimates of these adjustment coefficients together with a number of diagnostic test statistics are presented in Table 5 below.\textsuperscript{25}

The figures in ( ) are estimated asymptotic standard errors whereas those in [ ] 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 $\Delta r_t^*$, 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

\textsuperscript{25}Full estimation results are available on request. $\chi^2_{SC}(4)$ is the Lagrange multiplier statistic for testing the null of no serial correlation, $\chi^2_{FE}(1)$ is Ramsey’s RESET test statistic, $\chi^2_{JS}(2)$ is the Jarque–Bera statistic for testing the null of Gaussian errors, and $\chi^2_{HE}(1)$ is the statistic for testing the null of no heteroskedasticity. The number in ( ) indicates the degrees of freedom.
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 U.K. 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 \( \beta_\alpha = (\beta_0, \beta_1, \ldots, \beta_5, \beta_6)' \). Under the just-identifying normalization restriction \( \beta_1 = 1 \), we obtained the estimates

\[
\hat{\beta}_\alpha = \begin{pmatrix}
0.004 & 1 & -0.923 & -3.567 & -1.929 & -1.362 & 0.074 \\
0.018 & (0.323) & (1.891) & (1.434) & (1.553) & (0.207)
\end{pmatrix},
\]

\( LL = 701.70 \).

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 \( H_{co} \cap H_{po} \)), namely, \( \beta_3 = 1 \) and \( \beta_4 = \beta_1 = \beta_2 = \beta_5 = \beta_6 = 0 \), we obtain

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

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.

Acknowledgements

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).

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 \(\delta\) as

\[
\delta \equiv \begin{pmatrix} - \gamma' \\ I_{m-r} \end{pmatrix} \beta_{\perp},
\]
where \( \beta_1 \) is a \((m, m - r)\) matrix whose columns are a basis for the orthogonal complement of \( \beta \). Hence, \( (\beta, \beta_1) \) is a basis for \( \mathbb{R}^m \). Let \( \xi \) be the \((m + 1)\)-unit vector \((1, 0)\)' Then, \( (\beta_\#, \xi, \delta) \) is a basis for \( \mathbb{R}^{m+1} \). It therefore follows from (2.11) that for \((t - 1)/T \leq a < t/T, \ a \in [0, 1]\)

\[
T^{-1/2} \delta \xi(a) = T^{-1/2} \beta_1 \mu + \beta_1 C \delta \xi + \beta_1 T^{-1/2} C \xi e_t \Rightarrow \beta_1 CB_m(a) , \quad (A.1)
\]

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

\[
T^{-1} \xi \xi(a) = T^{-1} t \Rightarrow a. \quad (A.2)
\]

Similarly, we note that

\[
\beta_\# \xi(a) = \beta_\# \mu + \beta_\# C \xi e_t = O_p(1). \quad (A.3)
\]

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

\[
\beta_\# S_{ZZ} \beta_\# \rightarrow_p \Sigma_{\beta \beta},
\]

where \( \Sigma_{\beta \beta} \) is a \((r, r)\) positive-definite matrix. Similarly, as

\[
\beta_\# S_{ZY} = \beta_\# S_{ZZ} \Pi_{y^\#} + T^{-1} \beta_\# \hat{Z}_{-1}^* U,
\]

\[
\beta_\# S_{ZY} \rightarrow_p \Sigma_y = \Sigma_{\beta \beta} \mu_y,
\]

and, as

\[
S_{YY} = T^{-1} (\Pi_{y \#} \hat{Z}_{-1} + U)(\bar{P} - \bar{P}, \Delta Z - (\Delta Z - \bar{P}, \Delta Z - \bar{P}) - 1 \Delta Z - \bar{P})
\]

\[
(\Pi_{y \#} \hat{Z}_{-1} + U)^\prime,
\]

\[
S_{YY} \rightarrow_p \Sigma_{yy} = \Omega_{uu} + \mu_y \Sigma_{\beta \beta} \mu_y
\]


**Lemma A.1.** Let \( B_T \equiv (\delta, T^{-1/2} \xi) \) and define \( G(a) = (G_1(a), G_2(a))^\prime \), where

\[
G_1(a) = \beta_1 C \bar{B}_m(a), \quad \bar{B}_m(a) = [\bar{B}_n(a), \hat{B}_k(a)]^\prime = B_m(a) - \int_0^1 B_m(a) da, \quad \text{and}
\]

\[
G_2(a) = a - \frac{1}{2}, \quad a \in [0, 1]. \quad \text{Then}
\]

\[
T^{-1} B_T S_{ZZ} B_T = \int_0^1 G(a) G(a)^\prime da , \quad (A.4)
\]

\[
B_T^\prime (S_{ZY} - S_{ZZ} \Pi_{y^\#}) \Rightarrow \int_0^1 G(a) \bar{G}_n^\prime (a), \quad (A.5)
\]

where \( \bar{B}_n^\prime (a) = [\bar{B}_n(a) - \Omega_{yy} \Omega_{xx}^{-1} \hat{B}_k(a)], \quad a \in [0, 1], \) and \( S_{ZZ}, S_{ZY} \) and \( S_{YY} \) 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 \( \mathbb{R}^m \) and \( B_T = \delta \). For Case II, where \( Z_{\cdot 1}^\perp = (t_T, Z_{-1}^\perp) \), we have

\[
\beta_* = \left( -\mu' \right) \beta_{\perp} I_m
\]

and, consequently, we define \( \xi \) as in Case IV,

\[
\delta = \left( -\mu' \right) \beta_{\perp} \quad \text{and} \quad B_T = (\delta, \xi).
\]

Case III is similar to Case I as is Case V.

Proof of Theorems 4.1 and 4.2. See Pesaran et al. (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 \( H_r: \text{Rank}[P_{r}] = r \) of (3.8).

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

\[
\hat{\beta}_* = \beta_* \hat{P}_* \hat{P}_* + \delta \bar{\delta} \hat{\beta}_* + \bar{\xi} \bar{\xi} \hat{\beta}_*,
\]

where \( \hat{P}_* \equiv P_{(\delta, \xi)}(\beta_* (\beta_* P_{(\delta, \xi)} \beta_*)^{-1} \text{ with } P_{(\delta, \xi)} \text{ the orthogonal projector onto the orthogonal complement of } (\delta, \xi) \text{ and } \bar{\delta} \text{ and } \bar{\xi} \text{ defined similarly. We then define}

\[
\bar{\beta}_* \equiv \bar{P}_* \hat{P}_* \hat{P}_*^{-1} = \beta_* + \delta \bar{\delta} \hat{\beta}_* + \bar{\xi} \bar{\xi} \hat{\beta}_* = \beta_* + B_T U_T,
\]

where \( B_T \) is defined in Lemma A.1 and

\[
U_T \equiv (\delta, T^{1/2} \bar{\xi}) \hat{P}_*;
\]

cf. Johansen (1995, (13.1), (13.2), p. 179). The normalization \( \hat{\beta}_*, \hat{P}_* = I_r \) provides the \( r^2 \) restrictions required to just-identify \( \hat{\beta}_* \). The ML estimator for \( \Omega_{uu} \) is given from (3.20) by

\[
\hat{\Omega}_{uu} = T^{-1} \Delta \hat{Y}(I_T - \hat{Z}_{\cdot 1}^\perp \hat{\beta}_* (\hat{\beta}_* \hat{Z}_{\cdot 1}^\perp \hat{Z}_{\cdot 1}^\perp \hat{\beta}_*)^{-1} \hat{\beta}_* \hat{Z}_{\cdot 1}^\perp) \Delta \hat{Y}.
\]
Lemma A.2. Under \( H \), defined in (3.8) and Assumptions 2.1, 2.2, 3.1 and 4.1,
\[
\begin{align*}
\bar{\beta}_* - \beta_* &= o_p(T^{-1/2}), \\
\bar{\beta}_* S_{ZZ} \bar{\beta}_* &= \Sigma_{\beta\beta} + o_p(1), \\
\bar{\beta}_* S_{ZY} &= \Sigma_{\beta y} + o_p(1), \\
\hat{\Omega}_{uu} &= \Omega_{uu} + o_p(1).
\end{align*}
\]
Proof. Defining \( A_T = (\beta_*, T^{-1/2} B_T) \) and noting that \( A_T^{-1} = (\bar{\beta}_*, T^{1/2} \delta, T \bar{\xi}) \), these results are obtained by the application of an argument similar to that used in the proof of Lemma 13.1 of Johansen (1995, pp. 180–181).

Lemma A.3. Under \( H \), defined in (3.8) and Assumptions 2.1, 2.2, 3.1 and 4.1,
\[
TU_T = (T \delta, T^{3/2} \bar{\xi})(\bar{\beta}_* - \beta_*)
\Rightarrow \left( \int_{0}^{1} G(a)G(a)' \, da \right)^{-1} \int_{0}^{1} G(a) dB_n^*(a)' \Omega_{uu}^{-1} \mathbf{x}_y (\mathbf{x}_y' \Omega_{uu}^{-1} \mathbf{x}_y)^{-1},
\]
which is mixed normal with conditional variance given by
\[
(\mathbf{x}_y' \Omega_{uu}^{-1} \mathbf{x}_y)^{-1} \otimes \left( \int_{0}^{1} G(a)G(a)' \, da \right)^{-1},
\]
where \( B_n^*(a) \equiv [B_n(a) - \Omega_{yx} \Omega_{xx}^{-1} B_k(a)] \) and \( G(a) \) is defined in Lemma A.1, \( a \in [0, 1] \). Hence, \( (\bar{\beta}_* - \beta_*) = O_p(T^{-1}) \).

Proof. It is straightforward to show that \( \bar{\beta}_* \) satisfies the likelihood equations for \( \beta_* \) derived from (3.20); viz.
\[
\tilde{\mathbf{x}}_y \hat{\Omega}_{uu}^{-1} [S_{YZ} - \tilde{\mathbf{x}}_y \bar{\beta}_* S_{ZZ}] = 0,
\]
where \( \tilde{\mathbf{x}}_y = \tilde{\mathbf{x}}_y(\bar{\beta}_*) \bar{\beta}_* \) and, from Lemma A.2, \( \tilde{\mathbf{x}}_y - \mathbf{x}_y = o_p(1) \). Eq. (A.9) after post-multiplication by \( B_T \) may be rearranged as
\[
\tilde{\mathbf{x}}_y \hat{\Omega}_{uu}^{-1} [(S_{YZ} - \mathbf{x}_y \beta_* S_{ZZ}) - \tilde{\mathbf{x}}_y (\bar{\beta}_* - \beta_*) S_{ZZ}] = (\tilde{\mathbf{x}}_y - \mathbf{x}_y) \beta_* S_{ZZ} = 0.
\]
Now, \( \beta_* S_{ZZ} B_T = O_p(1); \) cf. Johansen (1995, Lemma 10.3, (10.18), p. 146). Substituting for \( \bar{\beta}_* - \beta_* \) from (A.6) yields
\[
TU_T = (T^{-1} B_T S_{ZZ} B_T)^{-1} B_T (S_{ZY} - S_{ZZ} \beta_* \mathbf{x}_y) \Omega_{uu}^{-1} \mathbf{x}_y (\mathbf{x}_y' \Omega_{uu}^{-1} \mathbf{x}_y)^{-1} + o_p(1)
\Rightarrow \left( \int_{0}^{1} G(a)G(a)' \, da \right)^{-1} \int_{0}^{1} G(a) dB_n^*(a)' \Omega_{uu}^{-1} \mathbf{x}_y (\mathbf{x}_y' \Omega_{uu}^{-1} \mathbf{x}_y)^{-1},
\]
by the consistency of \( \hat{\Omega}_{uu} \), Lemma A.1 and \( \tilde{\mathbf{x}}_y - \mathbf{x}_y = o_p(1) \). The first equality in (A.7) results from the definition of \( U_T \) and the representation (A.7) follows by noting that \( B_n^*(\cdot) \) may be replaced by \( B_n^*(\cdot) \) in (A.5). As \( G(\cdot) \) is independent of \( (\mathbf{x}_y' \Omega_{uu}^{-1} \mathbf{x}_y)^{-1} \mathbf{x}_y' \Omega_{uu}^{-1} B_n^*(\cdot) \), \( TU_T \) is asymptotically mixed normal with conditional variance (A.8). Cf. Johansen (1995, Lemma 13.2, pp. 181–183). □
Lemma A.4. Under $H$, defined in (3.8) and Assumptions 2.1, 2.2, 3.1 and 4.1, the likelihood ratio statistic

$$-2[\ell_T^*(\beta_*; r) - \ell_T^*(\hat{\beta}_*; r)]$$

$$\Rightarrow \text{Trace}\left\{ \left( x'_y \Omega_{uy}^{-1} x_y \right)^{-1} x'_y \Omega_{uy}^{-1} \int_0^1 dB_n^*(a) G(a) \left( \int_0^1 G(a) G(a)' da \right)^{-1} \right\}$$

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


$$-2[\ell_T^*(\beta_*; r) - \ell_T^*(\hat{\beta}_*; r)]$$

equals

$$T \times \text{Trace}\left\{ (\Sigma_{\beta y} - \Sigma_{\beta \beta} \Sigma_{yy}^{-1} \Sigma_{y \beta})^{-1} - \Sigma_{\beta \beta}^{-1} (\hat{\beta}_* - \beta_*) S_{ZZ} (\hat{\beta}_* - \beta_*) \right\}$$

$$+ O_p(T^{-1})$$

and is equal to

$$\text{Trace}\left\{ x'_y \Omega_{uy}^{-1} x_y T U_T^T (T^{-1} B_T^* S_{ZZ} B_T^* T U_T) \right\} + O_p(T^{-1});$$

cf. Johansen (1995, Eq. (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 $G(.)$, the random $(m - r + 1)r$-vector $\text{vec}(x'_y \Omega_{uy}^{-1} \int_0^1 dB_n^*(a) G(a))$ is normally distributed with mean zero and conditional variance matrix

$$\int_0^1 G(a) G(a)' da \otimes x'_y \Omega_{uy}^{-1} x_y.$$ 

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

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 $H_r$, analogous to (A.10), the likelihood ratio statistic $-2[\ell_T^*(\beta; r) - \ell_T^*(\hat{\beta}; r)]$, where $\hat{\beta}$ denotes the ML estimator for $\beta$ in Case III, has the limiting representation

$$\text{Trace}\left\{ (x'_y \Omega_{uy}^{-1} x_y)^{-1} x'_y \Omega_{uy}^{-1} \int_0^1 dB_n^*(a) G_1(a) \left( \int_0^1 G_1(a) G_1(a)' da \right)^{-1} \right\}$$

$$\int_0^1 G_1(a) dB_n^*(a)' \Omega_{uy}^{-1} x_y \right\},$$

which is $\chi^2[r(m - r)]$ distributed. The likelihood ratio statistic for $c_1 = 0$ is given by $-2[\ell_T^*(\hat{\beta}; r) - \ell_T^*(\beta_*; r)]$ which by partitioned inversion of
\[
(\int_0^1 G(a)G(a)^\prime \, da)^{-1}
\]
has the limiting representation under \(H_r\) and \(c_1 = 0\) given by
\[
\text{Trace}\left\{ (x_r^\prime \Omega_{uu}^{-1} x_r)^{-1} x_r^\prime \Omega_{uu}^{-1} \int_0^1 dB^*_n(a) G_{2,1}(a) \left( \int_0^1 G_{2,1}(a)^2 \, ds \right)^{-1} \int_0^1 G_{2,1}(a) dB^*_n(a)^\prime \Omega_{uu}^{-1} x_r \right\},
\]
where \(G_{2,1}(a) = G_2(a) - \int_0^1 G_2(a)G_1(a)^\prime \, da \int_0^1 G_1(a)G_1(a)^\prime \, da^{-1}G_1(a), \ a \in [0, 1].\) Now, conditional on \(G(),\) the random \(r\)-vector \(\int_0^1 G_{2,1}(a) dB^*_n(a)\) is normally distributed with mean zero and conditional variance matrix \(x_r^\prime \Omega_{uu}^{-1} x_r \left( \int_0^1 G_{2,1}(a)^2 \, da \right).\) Hence, in Case IV, under \(H_r\) and \(c_1 = 0,\) the likelihood ratio statistic \(-2[\ell^*_T(\hat{\beta}); r] - \ell^*_T(\hat{\beta}_n); r]\) has a limiting \(\chi^2(r)\) distribution. \(\square\)

**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
\[
12 \int_0^1 \left( a - \frac{1}{2} \right) dW_n(a)^\prime \int_0^1 \left( a - \frac{1}{2} \right) dW_n(a)
\]
\[
+ \text{Trace} \left\{ \int_0^1 dW_n(a) \tilde{W}_n(a)^\prime \left[ \int_0^1 \tilde{W}_n(a) \tilde{W}_n(a)^\prime \, da \right]^{-1} \times \int_0^1 \tilde{W}_n(a) dW_n(a) \right\},
\]
the latter being the limiting representation for the likelihood ratio statistic for \(H_r\) against \(H_n\) in Case V. As \(\int_0^1 (a - \frac{1}{2}) dW_n(a) \sim N(0, (1/12)I_{n-r}),\) the result follows. Cf. the proof of Theorem 2.3 of Johansen (1991, p. 1571). \(\square\)

**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 et al. (1997, Appendix A). \(\square\)

**Proof of Theorem 4.9.** By Lemma A.3, \(\tilde{\beta}_n - \beta_* = O_P(T^{-1}).\) Hence, asymptotically, the inclusion of the lagged estimated cointegrating relationships \(\beta^*_n z_{n-1}^*\) is the same as the inclusion of the actual cointegrating relationships \(\beta_* z_{n-1}^*, \ t = 1, \ldots, T.\) Therefore, as the regression contains only variables which are stationary, the result follows; see Phillips and Durlauf (1986) and Phillips and Solo (1992). \(\square\)
Proof of Theorem 5.1. By an argument similar to that above Lemma A.1, 
\( \mathbf{\beta}_r \mathbf{\tilde{S}}_{Z\mathbf{Z}} \mathbf{\beta}_r \to \mathbf{\tilde{S}}_{\mathbf{\beta}\mathbf{\beta}} \), where \( \mathbf{\tilde{S}}_{\mathbf{\beta}\mathbf{\beta}} \) is a \((r, r)\) positive-definite matrix. As 
\( \mathbf{\beta}_r \mathbf{\tilde{S}}_{Z\mathbf{Y}} = \mathbf{\tilde{S}}_{Z\mathbf{Z}} \Pi_{Y}^{*} + T^{-1} \mathbf{\beta}_r \mathbf{\tilde{Z}}^{-1}_r \mathbf{\tilde{U}} \), 
\( T^{-1} \mathbf{\beta}_r \mathbf{\tilde{Z}}^{-1}_r \mathbf{\tilde{U}} = T^{-1} \mathbf{\beta}_r \mathbf{\tilde{Z}}^{-1}_r \mathbf{U} - T^{-1} \mathbf{\beta}_r \mathbf{\tilde{Z}}^{-1}_r \Delta \mathbf{Z}^- (\mathbf{\tilde{V}} - \mathbf{\Psi}) \)
and 
\( T^{1/2}(\mathbf{\tilde{V}} - \mathbf{\Psi}) = \mathbf{O}_p(1), \mathbf{\beta}_r \mathbf{\tilde{S}}_{Z\mathbf{Y}} \to \mathbf{\tilde{S}}_{\mathbf{\beta}\mathbf{\beta}} = \mathbf{\tilde{S}}_{\mathbf{\beta}\mathbf{\beta}} \mathbf{\tilde{Z}}. \) Similarly, as 
\( \mathbf{\tilde{S}}_{Y\mathbf{Y}} = T^{-1} (\Pi_{\gamma \gamma} \mathbf{\tilde{Z}}^{-1}_r + \mathbf{\tilde{U}}) \mathbf{\tilde{P}} (\Pi_{\gamma \gamma} \mathbf{\tilde{Z}}^{* -1}_r + \mathbf{\tilde{U}}), \) 
\( \mathbf{\tilde{S}}_{Y\mathbf{Y}} \to \mathbf{\tilde{S}}_{\mathbf{\beta}\mathbf{\beta}} \mathbf{\tilde{Z}}. \) It also straightforwardly seen that 
\( T^{-1} \mathbf{B}_T \mathbf{\tilde{S}}_{Z\mathbf{Z}} \mathbf{B}_T = T^{-2} \mathbf{B}_T \mathbf{\tilde{Z}}^{* -1}_r \mathbf{\tilde{P}} \mathbf{\tilde{Z}}^{* -1}_r \mathbf{B}_T \to \int_0^1 \mathbf{G}(a) \mathbf{G}(a)^{\prime} \, da \) and 
\( \mathbf{\tilde{B}}_T (\mathbf{\tilde{S}}_{Z\mathbf{Z}} - \mathbf{\tilde{S}}_{Z\mathbf{Z}} \Pi_{Y}^{*}) = T^{-1} \mathbf{\tilde{B}}_T \mathbf{\tilde{Z}}^{* -1}_r \mathbf{\tilde{P}} \mathbf{\tilde{U}} = \int_0^1 \mathbf{G}(a) \mathbf{\tilde{B}}_T (a)^{\prime}, \) 
where \( \mathbf{B}_T \) is defined in Lemma A.1; cf. (A.4) and (A.5). Therefore, the result is obtained by employing similar arguments to those in the Proof of Theorems 4.1 and 4.2 given in Pesaran et al. (1997, Appendix A).

Appendix B. Computation of critical values

The calculations of asymptotic critical values for the (log-) likelihood ratio cointegration rank statistics \( \mathcal{L} \mathcal{R}(H_r|H_{r+1}) \) and \( \mathcal{L} \mathcal{R}(H_r|H_0) \), 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 
\( z_t = z_{t-1} + e_t, \quad t = 1, 2, \ldots, T, \)
with \( z_0 = 0 \) and \( e_t \sim \mathcal{N}(0, \mathbf{I}_m) \). We then partition \( z_t = (y_t', x_t') \) where \( y_t \) and \( x_t \) are respectively \((n - r)\)- and \( k\)-vectors of random walk variables and \( e_t = (w_t', u_t') \) is partitioned conformably with \( z_t = (y_t', x_t') \), \( t = 1, 2, \ldots, T \). For Cases I–V, we compute the following statistic:
\[
\sum_{t=1}^{T} w_t f_t' \left( \sum_{j=1}^{T} f_j f_j' \right)^{-1} \sum_{t=1}^{T} f_t w_t',
\]  
(B.1)

where
\[
\begin{aligned}
\{z_{t-1}\} & \quad \text{Case I} \\
\{z_{t-1}', 1\}' & \quad \text{Case II} \\
f_t & \equiv \{z_{t-1}, \mathbf{\hat{e}}_{t-1}\}' \quad \text{Case III,} \\
\{z_{t-1}', t - \frac{1}{2}\}' & \quad \text{Case IV} \\
\hat{z}_{t-1} & \quad \text{Case V}
\end{aligned}
\]
\( \mathbf{\hat{z}}_{t-1} \) is de-meaned \( z_{t-1} \) (or the OLS residual from the regression of \( z_{t-1} \) on 1) and \( \mathbf{\hat{z}}_{t-1} \) is de-meaned and de-trended \( z_{t-1} \) (or the OLS residual from the
### Table 6(a)

Case I: no intercepts and no trends

<table>
<thead>
<tr>
<th>$n - r \backslash k$</th>
<th>$\lambda_{\text{ave}}$</th>
<th>$\lambda_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>304.4</td>
<td>71.25</td>
</tr>
<tr>
<td>11</td>
<td>258.7</td>
<td>65.86</td>
</tr>
<tr>
<td>10</td>
<td>215.4</td>
<td>59.62</td>
</tr>
<tr>
<td>9</td>
<td>175.9</td>
<td>53.91</td>
</tr>
<tr>
<td>8</td>
<td>141.2</td>
<td>47.94</td>
</tr>
<tr>
<td>7</td>
<td>110.1</td>
<td>42.30</td>
</tr>
<tr>
<td>6</td>
<td>83.18</td>
<td>36.27</td>
</tr>
<tr>
<td>5</td>
<td>59.33</td>
<td>29.95</td>
</tr>
<tr>
<td>4</td>
<td>39.81</td>
<td>23.92</td>
</tr>
<tr>
<td>3</td>
<td>24.05</td>
<td>17.68</td>
</tr>
<tr>
<td>2</td>
<td>12.36</td>
<td>11.03</td>
</tr>
<tr>
<td>1</td>
<td>4.16</td>
<td>4.16</td>
</tr>
<tr>
<td>0.10</td>
<td>295.7</td>
<td>67.95</td>
</tr>
<tr>
<td>11</td>
<td>251.2</td>
<td>62.62</td>
</tr>
<tr>
<td>10</td>
<td>208.0</td>
<td>56.60</td>
</tr>
<tr>
<td>9</td>
<td>169.8</td>
<td>50.60</td>
</tr>
<tr>
<td>8</td>
<td>135.6</td>
<td>45.00</td>
</tr>
<tr>
<td>7</td>
<td>105.4</td>
<td>39.39</td>
</tr>
<tr>
<td>6</td>
<td>78.47</td>
<td>33.48</td>
</tr>
<tr>
<td>5</td>
<td>55.42</td>
<td>27.57</td>
</tr>
<tr>
<td>4</td>
<td>36.69</td>
<td>20.58</td>
</tr>
<tr>
<td>3</td>
<td>21.46</td>
<td>15.57</td>
</tr>
<tr>
<td>2</td>
<td>10.25</td>
<td>9.28</td>
</tr>
<tr>
<td>1</td>
<td>3.04</td>
<td>3.04</td>
</tr>
</tbody>
</table>

The critical values of $\lambda_{\text{max}}$ and $\lambda_{\text{ave}}$ 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 6(b)
Case II: restricted intercepts and no trends

<table>
<thead>
<tr>
<th>$n - r$</th>
<th>$\lambda_{case}$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>$\lambda_{max}$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>341.2</td>
<td>365.1</td>
<td>389.4</td>
<td>413.6</td>
<td>437.1</td>
<td>461.5</td>
<td>75.51</td>
<td>78.57</td>
<td>81.74</td>
</tr>
<tr>
<td>11</td>
<td>291.7</td>
<td>314.2</td>
<td>336.4</td>
<td>358.7</td>
<td>381.0</td>
<td>403.4</td>
<td>69.76</td>
<td>73.11</td>
<td>76.02</td>
</tr>
<tr>
<td>10</td>
<td>245.7</td>
<td>265.9</td>
<td>287.0</td>
<td>307.2</td>
<td>327.8</td>
<td>348.5</td>
<td>64.11</td>
<td>66.85</td>
<td>70.17</td>
</tr>
<tr>
<td>9</td>
<td>204.0</td>
<td>223.8</td>
<td>242.3</td>
<td>260.8</td>
<td>279.5</td>
<td>297.9</td>
<td>57.95</td>
<td>61.34</td>
<td>64.47</td>
</tr>
<tr>
<td>8</td>
<td>166.1</td>
<td>182.7</td>
<td>199.7</td>
<td>216.6</td>
<td>232.8</td>
<td>249.3</td>
<td>52.06</td>
<td>55.26</td>
<td>58.21</td>
</tr>
<tr>
<td>7</td>
<td>132.5</td>
<td>148.0</td>
<td>163.1</td>
<td>178.0</td>
<td>192.5</td>
<td>206.5</td>
<td>46.47</td>
<td>49.57</td>
<td>52.87</td>
</tr>
<tr>
<td>6</td>
<td>102.6</td>
<td>116.3</td>
<td>128.9</td>
<td>141.7</td>
<td>154.4</td>
<td>167.0</td>
<td>40.53</td>
<td>43.76</td>
<td>46.90</td>
</tr>
<tr>
<td>5</td>
<td>75.98</td>
<td>86.58</td>
<td>97.57</td>
<td>108.6</td>
<td>119.8</td>
<td>130.8</td>
<td>34.40</td>
<td>37.48</td>
<td>40.57</td>
</tr>
<tr>
<td>4</td>
<td>53.48</td>
<td>62.75</td>
<td>72.15</td>
<td>81.25</td>
<td>90.60</td>
<td>99.37</td>
<td>28.27</td>
<td>31.48</td>
<td>34.69</td>
</tr>
<tr>
<td>3</td>
<td>34.87</td>
<td>42.40</td>
<td>49.43</td>
<td>56.28</td>
<td>63.10</td>
<td>69.72</td>
<td>22.04</td>
<td>25.54</td>
<td>29.01</td>
</tr>
<tr>
<td>2</td>
<td>20.18</td>
<td>25.23</td>
<td>30.46</td>
<td>35.46</td>
<td>39.94</td>
<td>44.56</td>
<td>15.87</td>
<td>18.88</td>
<td>21.92</td>
</tr>
<tr>
<td>1</td>
<td>9.16</td>
<td>12.45</td>
<td>15.27</td>
<td>17.80</td>
<td>20.63</td>
<td>23.02</td>
<td>9.16</td>
<td>12.45</td>
<td>15.27</td>
</tr>
</tbody>
</table>

Case II is defined by (3.10). Also see the footnote to Table 6(a).
Table 6(c)
Case III: unrestricted intercepts and no trends

\[ \lambda_{\text{case}} \]
\[ \lambda_{\text{max}} \]

<table>
<thead>
<tr>
<th>n - r</th>
<th>k</th>
<th>\lambda_{\text{case}}</th>
<th>\lambda_{\text{max}}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0</td>
<td>328.5</td>
<td>74.62</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>352.4</td>
<td>77.72</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>376.5</td>
<td>80.87</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>400.8</td>
<td>83.93</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>424.3</td>
<td>87.13</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>448.8</td>
<td>90.23</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>531.4</td>
<td>74.62</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>555.9</td>
<td>77.72</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>580.4</td>
<td>80.87</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>606.0</td>
<td>83.93</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>632.5</td>
<td>87.13</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>659.5</td>
<td>90.23</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>687.4</td>
<td>74.62</td>
</tr>
</tbody>
</table>

| 0.10  |  0 |  319.6 |  71.36 |
|       |  1 |  343.5 |  74.61 |
|       |  2 |  367.5 |  77.65 |
|       |  3 |  391.1 |  80.60 |
|       |  4 |  415.0 |  83.55 |
|       |  5 |  438.0 |  86.49 |
|       |  6 |  466.1 |  74.61 |
|       |  7 |  490.0 |  77.65 |
|       |  8 |  513.8 |  80.60 |
|       |  9 |  537.7 |  83.55 |
|       | 10 |  561.6 |  86.49 |

*Case III is defined by (3.11). Also see the footnote to Table 6(a).
Case IV: unrestricted intercepts and restricted trends

<table>
<thead>
<tr>
<th>$n - r/k$</th>
<th>0.05</th>
<th>0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>364.8</td>
<td>355.9</td>
</tr>
<tr>
<td>1.0</td>
<td>389.2</td>
<td>403.3</td>
</tr>
<tr>
<td>2.0</td>
<td>322.4</td>
<td>292.9</td>
</tr>
<tr>
<td>3.0</td>
<td>315.6</td>
<td>265.3</td>
</tr>
<tr>
<td>4.0</td>
<td>307.7</td>
<td>242.4</td>
</tr>
<tr>
<td>5.0</td>
<td>300.1</td>
<td>239.3</td>
</tr>
<tr>
<td>6.0</td>
<td>293.7</td>
<td>287.7</td>
</tr>
<tr>
<td>7.0</td>
<td>286.8</td>
<td>282.4</td>
</tr>
<tr>
<td>8.0</td>
<td>279.3</td>
<td>262.4</td>
</tr>
<tr>
<td>9.0</td>
<td>272.9</td>
<td>239.7</td>
</tr>
<tr>
<td>10.0</td>
<td>265.3</td>
<td>215.8</td>
</tr>
<tr>
<td>11.0</td>
<td>257.8</td>
<td>205.9</td>
</tr>
<tr>
<td>12.0</td>
<td>250.3</td>
<td>193.9</td>
</tr>
</tbody>
</table>

Case IV is defined by (3.12). Also see the footnote to Table 6(a).
Table 6(e)
Case V: unrestricted intercepts and unrestricted trends

<table>
<thead>
<tr>
<th>$n - r$</th>
<th>$\lambda_{\text{case}}$</th>
<th>$\lambda_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>352.1</td>
<td>77.73</td>
</tr>
<tr>
<td>11</td>
<td>302.4</td>
<td>71.84</td>
</tr>
<tr>
<td>10</td>
<td>255.1</td>
<td>66.17</td>
</tr>
<tr>
<td>9</td>
<td>213.4</td>
<td>60.48</td>
</tr>
<tr>
<td>8</td>
<td>174.9</td>
<td>54.17</td>
</tr>
<tr>
<td>7</td>
<td>140.2</td>
<td>48.57</td>
</tr>
<tr>
<td>6</td>
<td>109.2</td>
<td>42.67</td>
</tr>
<tr>
<td>5</td>
<td>82.33</td>
<td>37.07</td>
</tr>
<tr>
<td>4</td>
<td>58.93</td>
<td>31.00</td>
</tr>
<tr>
<td>3</td>
<td>39.33</td>
<td>24.35</td>
</tr>
<tr>
<td>2</td>
<td>23.83</td>
<td>18.33</td>
</tr>
<tr>
<td>1</td>
<td>11.54</td>
<td>11.54</td>
</tr>
<tr>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>343.4</td>
<td>74.25</td>
</tr>
<tr>
<td>11</td>
<td>294.1</td>
<td>68.56</td>
</tr>
<tr>
<td>10</td>
<td>247.8</td>
<td>62.80</td>
</tr>
<tr>
<td>9</td>
<td>206.4</td>
<td>57.37</td>
</tr>
<tr>
<td>8</td>
<td>168.2</td>
<td>51.26</td>
</tr>
<tr>
<td>7</td>
<td>134.5</td>
<td>45.75</td>
</tr>
<tr>
<td>6</td>
<td>104.3</td>
<td>39.90</td>
</tr>
<tr>
<td>5</td>
<td>77.55</td>
<td>33.16</td>
</tr>
<tr>
<td>4</td>
<td>55.01</td>
<td>28.22</td>
</tr>
<tr>
<td>3</td>
<td>36.28</td>
<td>22.26</td>
</tr>
<tr>
<td>2</td>
<td>21.23</td>
<td>16.28</td>
</tr>
<tr>
<td>1</td>
<td>9.75</td>
<td>9.75</td>
</tr>
</tbody>
</table>

*Case V is defined by (3.13). Also see the footnote to Table 6(a).
regression of $z_{t-1}$ on $(1,t)$, $t = 1, 2, \ldots, T$. 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 6(a)–(e) provide the 0.05 and 0.10 upper critical values from these empirical distributions, denoted $\hat{\lambda}_{\text{max}}$ and $\hat{\lambda}_{\text{trace}}$ respectively, as estimates of the corresponding critical values of the limiting distributions of the cointegration rank statistics $\mathcal{L}(H_r | H_{r+1})$ and $\mathcal{L}(H_r | H_n)$ for Cases I–V with $n - r = 1, 2, \ldots, 12$ and $k = 0, \ldots, 5$.

For Cases I, II and IV, the asymptotic critical values reported in Tables 6(a), 6(b) and 6(d) are similar to those in Tables 3–6 of Harbo et al. (1998). However, for Cases III and V, Tables 6(c) and 6(e), 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:

<table>
<thead>
<tr>
<th>Intercept</th>
<th>Linear trend</th>
<th>Quadratic trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSS JOH</td>
<td>PSS JOH</td>
<td>PSS JOH</td>
</tr>
<tr>
<td>Case I</td>
<td>N N</td>
<td>N N</td>
</tr>
<tr>
<td>Case II</td>
<td>Restricted</td>
<td>Restricted</td>
</tr>
<tr>
<td>Case III</td>
<td>Unrestricted</td>
<td>N Y</td>
</tr>
<tr>
<td>Case IV</td>
<td>Unrestricted</td>
<td>Restricted</td>
</tr>
<tr>
<td>Case V</td>
<td>Unrestricted</td>
<td>N Y</td>
</tr>
</tbody>
</table>

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 $\hat{\tau}_\nu$ in Dickey and Fuller (1979) while those reported for Cases II and IV are multivariate equivalents of the $F$-statistics $\Phi_1$ and $\Phi_3$ in Dickey and Fuller (1981).

References


