UvA ECONOMETRICS

Mixed normal inference on multicointegration

H. Peter Boswijk

UvA UNIVERSITEIT VAN AMSTERDAM

Mixed Normal Inference on Multicointegration[*]

H. Peter Boswijk

Department of Quantitative Economics, University of Amsterdam[†]

October 26, 2009

Abstract

Asymptotic likelihood analysis of cointegration in $I(2)$ models, see Johansen (1997, 2006), Boswijk (2000) and Paruolo (2000), has shown that inference on most parameters is mixed normal, implying hypothesis test statistics with an asymptotic χ^2 null distribution. The asymptotic distribution of the multicointegration parameter estimator so far has been characterised by a Brownian motion functional, which has been conjectured to have a mixed normal distribution, based on simulations. The present paper proves this conjecture.

1 Introduction

The notion of multicointegration was introduced by Granger (1986) and Granger and Lee (1990). Although originally developed for processes integrated of order 1 ($I(1)$), it has subsequently become clear that the phenomenon occurs naturally in $I(2)$ cointegrated vector autoregressive (VAR) models, see Johansen (1992) and Engsted and Johansen (1999). With $\{X_t\}_{t\geq 1}$ a p-vector time series process, the $I(2)$ VAR model of order k is expressed as

$$\Delta^2 X_t \;=\; \alpha\beta' X_{t-1} + \Gamma\Delta X_{t-1} + \sum_{j=1}^{k-2}\Psi_j\Delta^2 X_{t-j} + \varepsilon_t, \tag{1}$$

$$\bar{\alpha}_\perp\alpha'_\perp\Gamma\beta_\perp\bar{\beta}'_\perp \;=\; \alpha_1\beta'_1, \tag{2}$$

where $\{\varepsilon_t\}_{t\geq 1}$ is assumed to be an independent and identically distributed (i.i.d.) $N(0,\Omega)$ sequence, and where α and β are $p\times r$ matrices ($0\leq r < p$), α_1 and β_1 are $p\times s$ matrices ($0\leq s < p - r$), and Γ, $\{\Psi_j\}_{j=1}^{k-2}$ and Ω are $p\times p$ matrices, with Ω positive definite.[1] The model can be extended to include deterministic components such as a constant and trend, see Rahbek *et al.* (1999), without qualitatively affecting the results to follow.

[†]Address for correspondence: Department of Quantitative Economics, University of Amsterdam, Roetersstraat 11, NL-1018 WB Amsterdam, The Netherlands. E-mail: `h.p.boswijk@uva.nl`.

[1]For an $n\times m$ matrix A of rank $m < n$, $A_\perp$ denotes an $n\times(n-m)$ matrix of rank $n-m$ satisfying $A'_\perp A = 0$; and $\bar{A} = A(A'A)^{-1}$, so that $\bar{A}'A = I_m$.

Paruolo and Rahbek (1999) show that the $I(2)$ restriction (2) implies

$$\Gamma = \alpha\delta\beta_2' + \zeta_1\beta' + \zeta_2\beta_1',$$

where $\delta = \bar{\alpha}'\Gamma\bar{\beta}_2$, $\beta_2 = (\beta, \beta_1)_\perp$, $\zeta_1 = \Gamma\bar{\beta}$ and $\zeta_2 = \Gamma\bar{\beta}_1$. Therefore, the model (1) under this restriction becomes

$$\Delta^2 X_t = \alpha(\beta' X_{t-1} + \delta\beta_2'\Delta X_{t-1}) + \zeta_1\beta'\Delta X_{t-1} + \zeta_2\beta_1'\Delta X_{t-1} + \sum_{j=1}^{k-2} \Psi_j\Delta^2 X_{t-j} + \varepsilon_t. \quad (3)$$

In this model, $(\beta, \beta_1)'X_t$ are $I(1)$ linear combinations of the $I(2)$ process X_t, and $\beta'X_t$ further cointegrates with the $I(1)$ process $\beta_2'\Delta X_t$ to the $I(0)$ linear combination

$$\beta'X_t + \delta\beta_2'\Delta X_t = (\beta + \beta_2\delta'\Delta)'X_t. \quad (4)$$

This phenomenon is known as *multicointegration*, and also as *polynomial cointegration*, because the right-hand-side expression in (4) is first-order vector lag polynomial operating on X_t. Various alternative parametrisations of the $I(2)$ model have been proposed in the literature, see Johansen (1997), Boswijk (2000) and Mosconi and Paruolo (2009). However, they do not affect inference on the multicointegration parameter δ, which is the subject of this paper.

Asymptotic likelihood-based inference on the parameters of (3) was studied by Johansen (1997, 2006), Boswijk (2000) and Paruolo (2000). They showed that under suitable identifying restrictions, the asymptotic distributions of the maximum likelihood estimators $\hat{\beta}$ and $\hat{\beta}_1$ are scale mixtures of normals, where the random scaling matrix is the distributional limit of the inverse observed information matrix. This implies that likelihood ratio test statistics for smooth hypotheses on β and β_1 have an asymptotic χ^2 null distribution, at least under particular conditions on the hypotheses, derived by Boswijk (2000) and Johansen (2006). The asymptotic distribution of the multicointegration parameter estimator $\hat{\delta}$, however, at first sight does not appear to be mixed normal. It can be written as the distribution of the sum of two mixed normal random variables, but there is no common conditioning set such that both are conditionally normally distributed, which complicates deriving a valid inference procedure for δ. Yet, as noted by Paruolo (1995) and Johansen (2006), Monte Carlo simulations of the Brownian motion functionals that characterise the asymptotic distribution of δ strongly suggest that $\hat{\delta}$ is in fact asymptotically mixed normal. The present paper provides a proof of this conjecture, implying that likelihood-based inference on multicointegration can be conducted using χ^2 critical values.

The outline of the remainder of this paper is as follows. The next section summarises the asymptotic distributions of $\hat{\beta}$, $\hat{\beta}_1$ and $\hat{\delta}$, as obtained by Johansen (1997, 2006) and Paruolo (2000) (and in a mixture of their notation). In Section 3 the main result is stated and proved. The final section discusses some extensions. An appendix contains proofs of some auxiliary lemmas.

2 Asymptotic results

The starting point of the asymptotic analysis is the multivariate invariance principle: as $n \to \infty$,

$$n^{-1/2} \sum_{t=1}^{\lfloor un \rfloor} \varepsilon_t \xrightarrow{\mathcal{L}} W(u), \qquad u \in [0,1],$$

where W is a p-vector Brownian motion with variance matrix Ω. The i.i.d. normality of $\{\varepsilon_t\}_{t \geq 1}$ is sufficient, but not necessary for this result to hold. From W, define

$$\begin{aligned}
W_1 &= (\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}W, \\
W_2 &= \left[\bar{\alpha}_1' - \bar{\alpha}_1'\Omega\alpha_2(\alpha_2'\Omega\alpha_2)^{-1}\alpha_2'\right]W,
\end{aligned}$$

with $\alpha_2 = (\alpha, \alpha_1)_\perp$. These are two independent vector Brownian motions, of dimensions r and s, and with variance matrices denoted Ω_1 and Ω_2, respectively. Furthermore, (W_1, W_2) is independent of the $(p-r-s)$-vector Brownian motion $W_3 = \alpha_2'W$.

Johansen (2006) shows that

$$n^{-1/2}\begin{pmatrix} \beta_2'\Delta X_{\lfloor un \rfloor} \\ \beta_1'\Delta X_{\lfloor un \rfloor} \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} H_0(u) \\ H_1(u) \end{pmatrix} = \begin{pmatrix} A_{03}W_3(u) \\ A_{12}W_2(u) + A_{13}W_3(u) \end{pmatrix}, \qquad u \in [0,1],$$

where A_{03}, A_{12} and A_{13} are conformable matrices, depending on the parameters, with A_{03} and A_{12} non-singular. Define

$$H_*(u) = \begin{pmatrix} H_0(u) \\ H_1(u) \\ H_2(u) \end{pmatrix} = \begin{pmatrix} H_0(u) \\ H_1(u) \\ \int_0^u H_0(s)ds \end{pmatrix}, \qquad u \in [0,1],$$

as well as

$$H_{**} = \int_0^1 H_*(u)H_*(u)'du, \qquad H_{ij} = \int_0^1 H_i(u)H_j(u)'du, \qquad i,j = 0,1,2. \tag{5}$$

Let $\psi' = (\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}\Gamma$; it can be shown that $\psi'\bar{\beta}_2 = \bar{\alpha}'\Gamma\bar{\beta}_2 = \delta$. Johansen (2006, Theorem 4) and Paruolo (2000, Theorems 4.1 and 4.2) prove the following results for the maximum likelihood estimators[2] $\hat{\beta}$, $\hat{\beta}_1$ and $\hat{\psi}$ based on a sample $\{X_t\}_{t=1}^n$, with starting values $\{X_{1-k}, \ldots, X_0\}$:

$$\begin{pmatrix} n\bar{\beta}_2'(\hat{\psi} - \psi) \\ n\bar{\beta}_1'(\hat{\beta} - \beta) \\ n^2\bar{\beta}_2'(\hat{\beta} - \beta) \\ n\bar{\beta}_2'(\hat{\beta}_1 - \beta) \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} B_0^\infty \\ B_1^\infty \\ B_2^\infty \\ C^\infty \end{pmatrix}, \tag{6}$$

where

$$B^\infty = \begin{pmatrix} B_0^\infty \\ B_1^\infty \\ B_2^\infty \end{pmatrix} = H_{**}^{-1}\int_0^1 H_* dW_1', \qquad C^\infty = H_{00}^{-1}\int_0^1 H_0 dW_2'. \tag{7}$$

---

[2]The cointegration parameters have been identified by $c'\beta = I_r$ and $c_1'\beta_1 = I_s$, where c and c_1 are known conformable matrices. The results given here are for $c = \bar{\beta}$ and $c_1 = \bar{\beta}_1$, from which the results for general (c, c_1) can be derived.

Because W_1 is independent of (W_2, W_3), and H_* is defined from (W_2, W_3), it follows that W_1 is independent of H_*. Similarly, W_2 is independent of W_3 and hence H_0. This implies

$$B^\infty | H_* \sim N(\Omega_1 \otimes H_{**}^{-1}), \qquad C^\infty | H_0 \sim N(\Omega_2 \otimes H_{00}^{-1}). \tag{8}$$

Thus both B^∞ and C^∞ have a conditionally normal and hence mixed normal distribution, but with a different conditioning set. Therefore, (B^∞, C^∞) is not *jointly* normal conditional on the same information: the distribution of $B^\infty | H_0$ is not normal, and $C^\infty | H_*$ has a degenerate distribution. This lack of joint mixed normality was analysed in more detail by Boswijk (2000).

The conditional variances in (8) are estimated consistently by the estimated variance matrix based on the inverse observed information matrix, in the sense that

$$\widehat{\mathrm{var}} \begin{pmatrix} n\bar{\beta}_2'(\hat{\psi} - \psi) \\ n\bar{\beta}_1'(\hat{\beta} - \beta) \\ n^2\bar{\beta}_2'(\hat{\beta} - \beta) \\ n\bar{\beta}_2'(\hat{\beta}_1 - \beta) \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} \Omega_1 \otimes H_{**}^{-1} & 0 \\ 0 & \Omega_2 \otimes H_{00}^{-1} \end{pmatrix}.$$

Letting θ denote the full vector of cointegration parameters, this means that likelihood ratio or Wald test statistics for smooth hypotheses $\mathcal{H}_0 : g(\theta) = 0$ have an asymptotic χ^2 null distribution, at least if for some suitable sequence of norming matrices D_n,

$$D_n^{-1}[g(\hat{\theta}) - g(\theta)] \xrightarrow{\mathcal{L}} \begin{pmatrix} G_B \, \mathrm{vec}\, B^\infty \\ G_C \, \mathrm{vec}\, C^\infty \end{pmatrix}, \tag{9}$$

with G_B and G_C matrices of full row rank. Thus the components of $D_n^{-1}[g(\hat{\theta}) - g(\theta)]$ should be asymptotically linear in either B^∞ or C^∞, but not both. See Boswijk (2000) and Johansen (2006) for a further discussion of this sufficient, but possibly not necessary condition for mixed normal inference.

The asymptotic distribution of the estimated multicointegration parameter $\hat{\delta} = \hat{\psi}'\hat{\beta}_2$ is obtained from (6), together with $n\beta_1'(\hat{\beta}_2 - \bar{\beta}_2) = -n(\hat{\beta}_1 - \beta_1)'\bar{\beta}_2 + o_p(1)$, which yields (Paruolo, 2000, Theorem 4.2)

$$n(\hat{\delta} - \delta)' \xrightarrow{\mathcal{L}} B_0^\infty - C^\infty A, \tag{10}$$

with $A = \bar{\beta}_1'\psi$. Its estimated variance matrix, based on the inverse observed information matrix, satisfies

$$\widehat{\mathrm{var}}\left[n(\hat{\delta} - \delta)' \right] \xrightarrow{\mathcal{L}} \Omega_1 \otimes (H_{**}^{-1})_{00} + (A'\Omega_2 A) \otimes H_{00}^{-1} =: V_B + V_C. \tag{11}$$

This implies that hypotheses on δ do not satisfy (9), unless the restriction $A = \bar{\beta}_1'\psi = 0$ is satisfied; in all other cases, the asymptotic distribution of $\hat{\delta}$ is characterised by the sum of two random variables that are marginally, but not jointly mixed normal. As noted by Paruolo (1995) and Johansen (2006), however, Monte Carlo simulation of (10)–(11) suggest that inference on δ is asymptotically mixed normal even if $A \neq 0$. In the next section, this result will be proved.

3 Mixed normality

This section studies asymptotic inference on the multicointegration parameter δ, based on the limit in distribution of the standardised estimator, as implied by (10)–(11):

$$\widehat{\text{var}}\left(\hat{\delta}'\right)^{-1/2} \text{vec}(\hat{\delta} - \delta)' \xrightarrow{\mathcal{L}} [V_B + V_C]^{-1/2} [\text{vec}\, B_0^\infty + \text{vec}(C^\infty A)] =: Z. \tag{12}$$

When δ is a scalar parameter, Z may be interpreted as the limit in distribution of the t-statistic of $\hat{\delta}$. More generally, a Wald or likelihood ratio test statistic for a simple hypothesis on δ will converge in distribution, under the null hypothesis, to $Z'Z$.

In order to prove the main result, we first need some auxiliary lemmas, proved in the appendix. The first lemma provides a convenient expression for $(H_{**}^{-1})_{00}$.

Lemma 1 *Let H_{**} and $H_{ij}, i, j = 0, 1, 2$, be as defined in (5), and define*

$$
\begin{aligned}
H_{ij|k} &= H_{ij} - H_{ik}H_{kk}^{-1}H_{kj}, & i, j, k = 0, 1, 2, \\
H_{ij|kl} &= H_{ij} - (H_{ik}, H_{il}) \begin{pmatrix} H_{kk} & H_{kl} \\ H_{lk} & H_{ll} \end{pmatrix}^{-1} \begin{pmatrix} H_{kj} \\ H_{lj} \end{pmatrix}, & i, j, k, l = 0, 1, 2.
\end{aligned}
$$

Then

$$(H_{**}^{-1})_{00} = H_{00|2}^{-1} + H_{00|2}^{-1}H_{01|2}H_{11|02}^{-1}H_{10|2}H_{00|2}^{-1}. \tag{13}$$

Note that a simpler expression for $(H_{**}^{-1})_{00}$ is $H_{00|12}^{-1}$, but the expression in Lemma 1 is more convenient for our purposes. In particular, using the fact that $H_1 = A_{12}W_2 + A_{13}W_3$, and H_0 and H_2 are defined from W_3, the lemma implies that W_2 appears in $(H_{**}^{-1})_{00}$ and hence V_B in the linear functional $H_{01|2}$, and in the quadratic functional $H_{11|02}$. The next lemma characterises conditions for conditional independence between stochastic integrals and such functionals of a vector Brownian motion.

A function or *kernel* K on $[0, 1]^2$ is said to be symmetric if $K(s, u) = K(u, s)$ for all $(u, s) \in [0, 1]^2$, and positive semi-definite if $\int_0^1 \int_0^1 K(u, s)g(u)g(s)duds \geq 0$ for all continuous functions g on $[0, 1]$.

Lemma 2 *Let W be a vector Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, independent of $\mathcal{G} \subset \mathcal{F}$. Let X and Y be $\mathcal{G}$-measurable vector processes satisfying $E\left(\int_0^1 (XX' + YY')du\right) < \infty$, and let K be a positive semi-definite $\mathcal{G}$-measurable kernel on $[0, 1]^2$. Then, conditionally on $\mathcal{G}$, $\int_0^1 X dW'$ is independent of $\int_0^1 YW'du$ and $\int_0^1 \int_0^1 K(u, s)dW(u)dW(s)'$ if and only if, with probability one,*

$$\int_0^1 \left(\int_0^u X(s)ds\right) Y(u)'du = 0, \qquad \int_0^1 K(u, s)X(u)du = 0, \quad s \in [0, 1].$$

We are now in a position to prove the main result.

Theorem 1 *Let $B_0^\infty, C^\infty, V_B, V_C$ and Z be as defined in (7), (11) and (12). Then we have*

$$\begin{pmatrix} B_0^\infty \\ C^\infty A \end{pmatrix} \bigg| (V_B, V_C) \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} V_B & 0 \\ 0 & V_C \end{pmatrix}\right), \tag{14}$$

so that inference on δ is asymptotically mixed normal, i.e.,

$$Z \,|(V_B, V_C) \sim N(0, I_{r(p-r-s)}). \tag{15}$$

Proof. Define
$$Z_B = V_B^{-1/2} \operatorname{vec} B_0^\infty, \qquad Z_C = V_C^{-1/2} \operatorname{vec}(C^\infty A).$$

From (8), it directly follows that $Z_B | H_* \sim N(0, I_q)$ and $Z_C | H_0 \sim N(0, I_q)$, where $q = r(p - r - s)$. Note that conditioning on a process X in fact means conditioning on the σ-field generated by $\{X(u)\}_{u \in [0,1]}$. We will use the notation $X \equiv Y$ if both processes or random variables generate the same σ-field, and $X \subset Y$ if the σ-field generated by X is contained in the σ-field generated by Y.

The result $Z_B | H_* \sim N(0, I_q)$ implies that Z_B is independent of H_*, and hence also of $(Z_C, V_B, H_0) \subset (H_0, W_1) \equiv H_*$, so that
$$Z_B | (Z_C, V_B, H_0) \sim N(0, I_q). \tag{16}$$

We will show that conditionally on H_0, Z_C is independent of V_B. This implies $Z_C | (V_B, H_0) \sim N(0, I_q)$, and together with (16), this implies
$$\begin{pmatrix} Z_B \\ Z_C \end{pmatrix} \Bigg| (V_B, H_0) \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} I_q & 0 \\ 0 & I_q \end{pmatrix} \right).$$

Because this conditional distribution of (Z_B, Z_C) does not depend on (V_B, H_0), the same joint $N(0, I_{2q})$ distribution applies conditionally on $(V_B, V_C) \subset (V_B, H_0)$. This directly implies (14) and (15).

Recall that $Z_C = \operatorname{vec} \left(H_{00}^{-1/2} \int_0^1 H_0 dW_2' A (A' \Omega_2 A)^{-1/2} \right)$ and $V_B = \Omega_1 \otimes (H_{**}^{-1})_{00}$, where $(H_{**}^{-1})_{00}$ is given by (13). This means that, conditionally on H_0, Z_C is independent of V_B if $\int_0^1 H_0 dW_2'$ is independent of $H_{01|2}$ and $H_{11|02}$; the other ingredients of $(H_{**}^{-1})_{00}$ are fixed conditional on H_0. Using $H_1 = A_{12} W_2 + A_{13} W_3$ for fixed matrices A_{12} and A_{13}, with $|A_{12}| \neq 0$, it follows that
$$
\begin{aligned}
H_{01|2} &= \int_0^1 (H_0 - H_{02} H_{22}^{-1} H_2) H_1' du \\
&= \int_0^1 H_{0|2} H_1' du \\
&= \int_0^1 H_{0|2} W_2' du A_{12}' + \int_0^1 H_{0|2} W_3' du A_{13}',
\end{aligned}
$$

where $H_{0|2}(u) = H_0(u) - H_{02} H_{22}^{-1} H_2(u)$. By Lemma 2, this implies that $\int_0^1 H_0 dW_2'$ is conditionally independent of $H_{01|2}$, because
$$\int_0^1 \left(\int_0^u H_0 ds \right) H_{0|2}(u)' du = \int_0^1 H_2(u) H_{0|2}(u)' du = 0.$$

Next, let $H_3 = (H_0', H_2')'$ and $H_{33} = \int_0^1 H_3 H_3' du$, so that
$$
\begin{aligned}
H_{11|02} &= \int_0^1 H_1 H_1' du - \int_0^1 H_1 H_3' du H_{33}^{-1} \int_0^1 H_3 H_1' du \\
&= A_{12} \left(\int_0^1 W_2 W_2' du - \int_0^1 W_2 H_3' du H_{33}^{-1} \int_0^1 H_3 W_2' du \right) A_{12}',
\end{aligned}
$$

6

where the final equality follows from $H_1 = A_{12}W_2 + A_{13}W_3 = A_{12}W_2 + A_{13}A_{03}^{-1}H_0$. From this, we find

$$A_{12}^{-1}H_{11|02}A_{12}'^{-1} = \int_0^1 \left(\int_0^u dW_2\right)\left(\int_0^u dW_2\right)' du$$
$$- \int_0^1 \left(\int_0^u dW_2\right) H_3' du H_{33}^{-1} \int_0^1 H_3 \left(\int_0^u dW_2\right)' du$$
$$= \int_0^1 \int_0^1 K(u,s) dW_2(u) dW_2(s),$$

where

$$K(u,s) = 1 - u \vee s - \tilde{H}_3(u)' H_{33}^{-1} \tilde{H}_3(s),$$

with $\tilde{H}_i(u) = \int_u^1 H_i ds$, $i = 0,1,2,3$. Applying again Lemma 2, we find that conditionally on H_0, $\int_0^1 H_0 dW_2'$ is independent of $H_{11|02}$, because

$$\int_0^1 K(u,s)H_0(u)du = \int_s^1 \left(\int_0^u H_0 dt\right) du - \int_0^1 \left(\int_0^u H_0 dt\right) H_3(u)' du H_{33}^{-1} \tilde{H}_3(s)$$
$$= \tilde{H}_2(s) - \int_0^1 H_2 H_3(u)' du H_{33}^{-1} \tilde{H}_3(s)$$
$$= 0.$$

The final equality follows from $H_2 = (I_s, 0)H_3$, and hence $\tilde{H}_2 = (I_s, 0)\tilde{H}_3$. Thus we have shown that $\int_0^1 H_0 dW_2'$ is independent of both $H_{01|2}$ and $H_{11|02}$, and hence of V_B. $\quad\square$

4 Discussion

Theorem 1 states that the maximum likelihood estimator of the multicointegration parameter δ has an asymptotically mixed normal distribution. This means that a likelihood ratio or Wald test statistic of a simple hypothesis $\mathcal{H}_0 : \delta = \delta_0$ will have an asymptotic $\chi^2_{r(p-r-s)}$ null distribution, arising as the distribution of $Z'Z$. More generally, it is not hard to prove that test statistics of smooth hypotheses $\mathcal{H}_0 : g(\delta) = 0$, with g a continuously differentiable function with derivative $G(\delta)$ of full row rank, will have an asymptotic χ^2 null distribution.

A further extension is to consider hypotheses on β, β_1 and δ together. For example, Mosconi and Paruolo (2009) consider possibly over-identifying restrictions of the form $(\beta', \delta)' = h(\phi)$, where h is a linear function of a parameter vector ϕ. Extending Johansen's (2006) Theorem 5.1, we may obtain conditions on h such that the restricted log-likelihood is locally asymptotically quadratic. As indicated by Johansen (2006), these conditions entail that ϕ can be partitioned as (ϕ_1, ϕ_2), with $n\hat{\phi}_1$ and $n^2\hat{\phi}_2$ converging in distribution to linear functions of $(B_0^\infty, B_1^\infty, C^\infty)$ and B_2^∞, respectively. Theorem 1 can be extended to show that $(B_0^\infty, B_1^\infty, C^\infty)$ is jointly mixed normal, but not independent of B_2^∞ and its conditional variance $\Omega_1 \otimes (H_{**}^{-1})_{22}$. This implies that hypotheses that only restrict $B_0 = \bar{\beta}_2^{0'}(\psi - \psi^0)$, $B_1 = \bar{\beta}_1^{0'}\beta$ and $C = \bar{\beta}_2^{0'}\beta_1$, but leave $B_2 = \bar{\beta}_2^{0'}\beta$ unrestricted (where θ^0 denotes the true value of θ), can be tested based on asymptotically χ^2 likelihood ratio statistics. In other words, hypotheses that involve β and δ only allow for mixed normal inference if they do not restrict B_2.

Appendix

Proof of Lemma 1. We use the following well-known result for partitioned inverses:

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11|2}^{-1} & -A_{11|2}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}A_{11|2}^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}A_{11|2}^{-1}A_{12}A_{22}^{-1} \end{pmatrix}, \quad (17)$$

where $A_{11|2} = A_{11} - A_{12}A_{22}^{-1}A_{21}$, and where A_{11}, A_{22} and $A_{11|2}$ are assumed to be non-singular. It is convenient to define a re-ordered version of H_{**}:

$$H_{\dagger\dagger} = \begin{pmatrix} H_{11} & \begin{pmatrix} H_{10} & H_{12} \end{pmatrix} \\ \begin{pmatrix} H_{01} \\ H_{21} \end{pmatrix} & \begin{pmatrix} H_{00} & H_{02} \\ H_{20} & H_{22} \end{pmatrix} \end{pmatrix},$$

so that $(H_{**}^{-1})_{00}$ is identical to the middle diagonal block of $H_{\dagger\dagger}^{-1}$. Applying (17) to $H_{\dagger\dagger}$ with $A_{11} = H_{11}$ (and the implied choice for A_{12} and A_{22}), leads to

$$\begin{aligned} (H_{**}^{-1})_{00} = & \left[\begin{pmatrix} H_{00} & H_{02} \\ H_{20} & H_{22} \end{pmatrix}^{-1} \right]_{00} \\ & + \left[\begin{pmatrix} H_{00} & H_{02} \\ H_{20} & H_{22} \end{pmatrix}^{-1} \begin{pmatrix} H_{01} \\ H_{21} \end{pmatrix} H_{11|02}^{-1} \begin{pmatrix} H_{10} & H_{12} \end{pmatrix} \begin{pmatrix} H_{00} & H_{02} \\ H_{20} & H_{22} \end{pmatrix}^{-1} \right]_{00}. \end{aligned}$$

Next, applying (17) again to

$$\begin{pmatrix} H_{00} & H_{02} \\ H_{20} & H_{22} \end{pmatrix}^{-1} = \begin{pmatrix} H_{00|2}^{-1} & -H_{00|2}^{-1}H_{02}H_{22}^{-1} \\ -H_{22}^{-1}H_{20}H_{00|2}^{-1} & H_{22}^{-1} + H_{22}^{-1}H_{20}H_{00|2}^{-1}H_{02}H_{22}^{-1} \end{pmatrix},$$

yields

$$\begin{aligned} (H_{**}^{-1})_{00} = & \; H_{00|2}^{-1} \\ & + H_{00|2}^{-1} \begin{pmatrix} I & -H_{02}H_{22}^{-1} \end{pmatrix} \begin{pmatrix} H_{01} \\ H_{21} \end{pmatrix} H_{11|02}^{-1} \begin{pmatrix} H_{10} & H_{12} \end{pmatrix} \begin{pmatrix} I \\ -H_{22}^{-1}H_{20} \end{pmatrix} H_{00|2}^{-1} \\ = & \; H_{00|2}^{-1} + H_{00|2}^{-1}H_{01|2}H_{11|02}^{-1}H_{10|2}H_{00|2}^{-1}, \end{aligned}$$

which is the required result. $\qquad\square$

Proof of Lemma 2. First, note that integration by parts yields

$$\int_0^1 Y(u)W(u)'du = \int_0^1 Y(u)\left(\int_0^u dW(s)'\right)du = \int_0^1 \tilde{Y}(u)dW(u)',$$

where $\tilde{Y}(u) = \int_u^1 Y(s)ds$. Next, for a positive definite K, Mercer's theorem, see Tanaka (1996), states that

$$K(u,s) = \sum_{i=1}^{\infty} \lambda_i f_i(u)f_i(s),$$

8

where $\{\lambda_i\}_{i\geq 1}$ and $\{f_i\}_{i\geq 1}$ are the eigenvalues[3] and orthonormal eigenfunctions of K, solving the integral equation

$$\int_0^1 K(u,s)f(u)du = \lambda f(s).$$

This implies that

$$\int_0^1 \int_0^1 K\,dW\,dW' = \sum_{i=1}^\infty \lambda_i \left(\int_0^1 f_i dW\right)\left(\int_0^1 f_i dW\right)'.$$

The basic properties of the Itô integral imply

$$\left.\begin{pmatrix} \int_0^1 X dW' \\ \int_0^1 \tilde{Y} dW' \\ \int_0^1 f_i dW' \end{pmatrix}\right| \mathcal{G} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \int_0^1 \begin{pmatrix} XX' & X\tilde{Y}' & Xf_i' \\ \tilde{Y}X' & \tilde{Y}\tilde{Y}' & \tilde{Y}f_i' \\ f_iX' & f_i\tilde{Y}' & f_if_i' \end{pmatrix} du\right).$$

Therefore $\int_0^1 X dW'$ is conditionally independent of $\int_0^1 YW'du$ and $\int_0^1 f_i dW'$ if and only if

$$\int_0^1 X(u)\tilde{Y}(u)du = \int_0^1 \left(\int_0^u X(s)ds\right) Y'du = 0, \tag{18}$$

(the first equality follows from integration by parts), and

$$\int_0^1 X(u)f_i(u)du = 0. \tag{19}$$

This in turn implies that $\int_0^1 X dW'$ is conditionally independent of $\int_0^1 YW'du$ and $\int_0^1 \int_0^1 K\,dW\,dW'$ if and only if both (18) holds and (19) holds for all eigenfunctions f_i corresponding tot non-zero eigenvalues. The latter condition is equivalent to

$$\int_0^1 K(u,s)X(u)du = \sum_{i=1}^\infty \lambda_i f_i(s) \int_0^1 f_i(u)X(u)du = 0, \qquad s \in [0,1].$$

Hence the components of X are eigenfunctions of K corresponding to zero eigenvalues. $\qquad\square$

References

Boswijk, H. P. (2000), "Mixed normality and ancillarity in $I(2)$ systems," *Econometric Theory*, 16, 878–904.

Engsted, T. and S. Johansen, "Granger's representation theorem and multicointegration," pp. 200–212 in: R. F. Engle and H. White (eds.), *Cointegration, Causality and Forecasting: Festschrift in Honour of Clive Granger*. Oxford: Oxford University Press.

Granger, C. W. J. (1986), "Developments in the study of cointegrated economic variables," *Oxford Bulletin of Economics and Statistics*, 48, 213–228.

Granger, C. W. J. and T.-H. Lee (1990), "Multicointegration," pp. 71–84 in: G. F. Rhodes, Jr. and T. B. Fomby (eds.), *Advances in Econometrics: Cointegration, Spurious Regressions and Unit Roots*. Greenwich, CT: JAI Press.

---

[3]Tanaka (1996) refers to $1/\lambda_i$ as the eigenvalues.

Johansen, S. (1992), "A representation of vector autoregressive processes integrated of order 2," *Econometric Theory*, 8, 188–202.

Johansen, S. (1997), "Likelihood analysis of the $I(2)$ model," *Scandinavian Journal of Statistics*, 24, 433–462.

Johansen, S. (2006), "Statistical analysis of hypotheses on the cointegrating relations in the $I(2)$ model," *Journal of Econometrics*, 132, 81–115.

Mosconi, R. and P. Paruolo (2009), "Identification of cointegrating relations in $I(2)$ vector autoregressive models," Working paper.

Paruolo, P. (1995), *Topics in statistical inference for vector autoregressive processes integrated of order 1 and 2.* PhD Thesis, University of Copenhagen.

Paruolo, P. (2000), "Asymptotic efficiency of the two stage estimator in $I(2)$ systems," *Econometric Theory*, 16, 524–550.

Paruolo, P. and A. Rahbek (1999), "Weak exogeneity in $I(2)$ VAR systems," *Journal of Econometrics*, 93, 281–308.

Rahbek, A., H. C. Kongsted and C. Jørgensen (1999), "Trend-stationarity in the $I(2)$ cointegration model," *Journal of Econometrics*, 90, 265–289.

Tanaka, K. (1996), *Time Series Analysis. Nonstationary and Noninvertible Distribution Theory.* New York: Wiley.