

Distribution Approximations for Cointegration Tests with Stationary Exogenous Regressors

H. Peter Boswijk

*Department of Quantitative Economics, Universiteit van Amsterdam**

Jurgen A. Doornik

Nuffield College, University of Oxford

10th February 1999

Abstract

The distribution of a functional of two correlated vector Brownian motions is approximated by a Gamma distribution. This functional represents the limiting distribution for cointegration tests with stationary exogenous regressors, but also for cointegration tests based on a non-Gaussian likelihood. The approximation is accurate, fast, and easy to use in comparison to both tabulated critical values and simulated p -values.

1 Introduction

Consider the vector error correction model (VECM)

$$\Delta X_t = \Pi^* X_{t-1}^* + \sum_{j=1}^{k-1} \Gamma_j \Delta X_{t-j} + \Phi q_t + \varepsilon_t, \quad t = 1, \dots, T, \quad (1)$$

where $\{X_t\}$ is a p -vector time series, the starting values (X_{1-k}, \dots, X_0) are fixed, $\{\varepsilon_t\}$ is i.i.d. $N(0, \Omega)$, and $X_{t-1}^* = (X'_{t-1}, d'_t)'$, where d_t and q_t are deterministic regressors. The three deterministic specifications that are most commonly used, are, in the notation of Doornik *et al.* (1998) (also see Johansen, 1995, Section 5.7):

- H_z : both d_t and q_t are void; no deterministic;
- H_c : $d_t = 1$, q_t is void; restricted constant;
- H_l : $d_t = t$, $q_t = 1$; restricted linear trend.

*Correspondence to: H. Peter Boswijk, Department of Quantitative Economics, Universiteit van Amsterdam, Roetersstraat 11, NL-1018 WB Amsterdam, The Netherlands. E-mail: peterb@fee.uva.nl.

When $\text{rank}\Pi^* = r \leq p$ and some further restrictions on $(\Pi^*, \{\Gamma_j\})$ are satisfied, the model implies that X_t is cointegrated. The likelihood ratio (LR) statistic for the hypothesis $H(r) : \text{rank}\Pi^* \leq r$ has been derived by Johansen (1988, 1995), and its limiting distribution (when $\text{rank}\Pi^* = r$) is characterized by the functional

$$\mathbb{T} = \text{tr} \left\{ \int_0^1 dB F' \left(\int_0^1 FF' du \right)^{-1} \int_0^1 F dB' \right\}, \quad (2)$$

where $B(u)$ is a standard $(p-r)$ -vector Brownian motion, and where either $F(u) = B(u)$ (model H_z), or $F(u) = \{B(u)', 1\}'$ (model H_c), or $F(u) = \{[B(u) - \int_0^1 B du]', [u - \frac{1}{2}]\}'$ (model H_l).

Recently, Seo (1998) and Rahbek and Mosconi (1998) considered an extension of (1), where some exogenous stationary vector process Z_t is added to the regressors (together with some of its lags):

$$\Delta X_t = \Pi^* X_{t-1}^* + \sum_{j=1}^{k-1} \Gamma_j \Delta X_{t-j} + \Phi q_t + \sum_{j=0}^m D_j Z_{t-j} + \varepsilon_t, \quad t = 1, \dots, T. \quad (3)$$

The limiting distribution under the null hypothesis of the LR statistic for $H(r)$ in this case turns out to be characterized by

$$\mathbb{Q} = \text{tr} \left\{ \int_0^1 dW F' \left(\int_0^1 FF' du \right)^{-1} \int_0^1 F dW' \right\}, \quad (4)$$

where $F(u)$ is the same as before, and $W(u)$ is a standard $(p-r)$ -vector Brownian motion, with $E[W(1)B(1)'] = P = \text{diag}(\rho_1, \dots, \rho_{p-r})$, where $\rho_i \in [0, 1]$ are correlation coefficients. Note that $\mathbb{Q} = \mathbb{T}$ when $P = I_{p-r}$.

In order to save notation, but without loss of generality, we shall only consider the case $r = 0$ henceforth, so that W and B are of dimension p .

When $p = 1$ (and $P = \rho$), the random variable \mathbb{Q} in model H_z is the square of

$$\mathbb{R} = \left(\int_0^1 B^2 du \right)^{-1/2} \int_0^1 B dW.$$

Its distribution was obtained by Kremers *et al.* (1992) as the limiting distribution of a t -statistic for cointegration with known cointegrating vector. Because we may decompose W as $\rho B + (1 - \rho^2)^{1/2} U$, with U a standard Brownian motion, independent of B , it follows that

$$\begin{aligned} \mathbb{R} &= \rho \frac{\int_0^1 B dB}{\left(\int_0^1 B^2 du \right)^{1/2}} + (1 - \rho^2)^{1/2} \frac{\int_0^1 B dU}{\left(\int_0^1 B^2 du \right)^{1/2}} \\ &= \rho X + (1 - \rho^2)^{1/2} Z, \end{aligned} \quad (5)$$

where X corresponds to the limiting distribution of the Dickey-Fuller t -statistic, and Z is a standard normal random variable, independent of B and hence X . Kremers *et al.* (1992) suggested to use critical values from the standard normal distribution by a small- σ (in this case small- ρ) asymptotic argument.

The same distribution of \mathbb{R} was also obtained by Hansen (1995), in the context of testing for a unit root with stationary exogenous regressors. Hansen tabulated the distribution of \mathbb{R} for $\rho^2 \in \{0.1, 0.2, \dots, 1\}$, and suggested to interpolate these critical values for other values of ρ^2 . This approach was extended

by Seo (1998) to the multivariate case, leading to Q . With $p > 1$, however, one has to construct tables for different values of (ρ_1, \dots, ρ_p) which is rather impractical. Seo provided tables for $p \leq 5$ and $(\rho_1, \dots, \rho_p) \in \mathcal{P} \times \dots \times \mathcal{P}$, where $\mathcal{P} = \{0, 0.2, 0.4, 0.6, 0.8, 1\}$, which already resulted in 20 pages of tables. Requiring from practitioners to obtain appropriate critical values by interpolation in several dimensions may be too much to ask, and definitely does not add to the ease-of-use of the proposed test. Therefore, we suggest in the next section an alternative approach to obtaining critical values or (preferably) p -values, based on Doornik's (1998a) approach to approximate the distribution of T by a Gamma distribution with the same mean and variance as T , also see Johansen (1988) and Nielsen (1997). This is closely related to Abadir and Lucas' (1996) approximation of the distribution of R by a normal distribution with non-zero mean and non-unit variance.

The distribution of R and Q also arises when $H(r)$ is tested in model (1) with non-Gaussian $\{\varepsilon_t\}$ and corresponding non-Gaussian likelihood function. See, e.g., Lucas (1997), who considers cointegration testing based on a Student- t likelihood function, and Boswijk and Lucas (1997), who use a semi-nonparametric likelihood function. Furthermore, the distribution of R also emerges when $\{\varepsilon_t\}$ is assumed to follow a GARCH process, and an LR test for a unit root is based on the corresponding likelihood function, see Ling and Li (1997, 1998).

The plan of the rest of this paper is as follows. In Section 2, we show that the mean and variance of Q can be expressed as a function of the mean and variance of T , a covariance parameter, and the correlations (ρ_1, \dots, ρ_p) . It is then suggested to use a Gamma distribution with the same mean and variance as an approximation to the true distribution of Q . In Section 3, this approximation is shown to be very accurate, at least for quantiles and p -values where accuracy is required (in the right-hand tail of the distribution). Section 4 applies the result to the purchasing-power parity model of Johansen and Juselius (1992). Section 5 concludes.

2 Mean and Variance of Q

The Gamma distribution function $\Gamma(x; b, a)$ is defined here as:

$$\Gamma(x; b, a) = \int_0^x \frac{a^b}{\Gamma(b)} t^{b-1} e^{-at} dt, \quad x > 0, b > 0, a > 0, \quad (6)$$

with $\Gamma(b) = \int_0^\infty t^{b-1} e^{-t} dt$, the Gamma function. A random variable X with this distribution has mean $E(X) = b/a$ and variance $\text{var}(X) = b/a^2$.

Doornik (1998a) shows that the Gamma distribution with $b = E(T)^2/\text{var}(T)$ and $a = E(T)/\text{var}(T)$ provides an accurate approximation of the distribution of T . The mean and variance of T could in principle be simulated for the three different deterministic models and many values of p . However, Doornik shows, using estimated response surfaces based on Monte Carlo simulation, that the following approximations are sufficiently accurate:

$$E(T) \approx \begin{cases} 2p^2 - p + 0.07 + 0.07 \cdot \mathbf{1}_{\{p=1\}} & \text{for } H_z, \\ 2p^2 + 2.01p + 0.06 \cdot \mathbf{1}_{\{p=1\}} + 0.05 \cdot \mathbf{1}_{\{p=2\}} & \text{for } H_c, \\ 2p^2 + 4.05p + 0.5 - 0.23 \cdot \mathbf{1}_{\{p=1\}} - 0.07 \cdot \mathbf{1}_{\{p=2\}} & \text{for } H_l, \end{cases} \quad (7)$$

and¹

$$\text{var}(\mathbb{T}) \approx \begin{cases} 3p^2 - 0.33p - 0.55 & \text{for } H_z, \\ 3p^2 + 3.60p + 0.75 - 0.40 \cdot \mathbf{1}_{\{p=1\}} - 0.30 \cdot \mathbf{1}_{\{p=2\}} & \text{for } H_c, \\ 3p^2 + 5.70p + 3.20 - 1.30 \cdot \mathbf{1}_{\{p=1\}} - 0.50 \cdot \mathbf{1}_{\{p=2\}} & \text{for } H_l. \end{cases} \quad (8)$$

Doornik also analyzes

$$\mathbb{T}_i = \int_0^1 dB_i F' \left(\int_0^1 FF' du \right)^{-1} \int_0^1 F dB_i,$$

where B_i is the i th component of B , so that $\mathbb{T} = \sum_{i=1}^p \mathbb{T}_i$. Since \mathbb{T}_i and \mathbb{T}_j have the same distribution, we have $E(\mathbb{T}_i) = E(\mathbb{T})/p$. Furthermore, he finds (by simulation) that for $i \neq j$,

$$\text{cov}(\mathbb{T}_i, \mathbb{T}_j) \approx \begin{cases} -1.270 & \text{for } H_z, \\ -1.066 & \text{for } H_c, \\ -1.35 & \text{for } H_l. \end{cases} \quad (9)$$

This can be used to evaluate $\text{var}(\mathbb{T}_i) = \text{var}(\mathbb{T})/p - (p-1)\text{cov}(\mathbb{T}_i, \mathbb{T}_j)$.

Here we adopt a similar approach for Q . Theorem 1 provides an expression for the mean and variance of Q in terms of $E(\mathbb{T})$, $\text{var}(\mathbb{T})$, $\text{cov}(\mathbb{T}_i, \mathbb{T}_j)$ and (ρ_1, \dots, ρ_p) . Subsequent substitution of the approximations (7)–(9) provides the first two moments of Q . This, in turn, may be used to obtain a Gamma approximation of its distribution.

Theorem 1 *Let $q = \dim(F)$. Then*

$$E(Q) = \frac{\sum_{i=1}^p \rho_i^2}{p} E(\mathbb{T}) + \left(1 - \frac{\sum_{i=1}^p \rho_i^2}{p}\right) pq, \quad (10)$$

and

$$\text{var}(Q) = \sum_{i=1}^p \rho_i^4 \text{var}(\mathbb{T}_i) + 2 \sum_{i=2}^p \sum_{j=1}^{i-1} \rho_i^2 \rho_j^2 \text{cov}(\mathbb{T}_i, \mathbb{T}_j) + 4 \frac{\sum_{i=1}^p \rho_i^2 (1 - \rho_i^2)}{p} E(\mathbb{T}) + 2q \sum_{i=1}^p (1 - \rho_i^2)^2. \quad (11)$$

Proof. Decompose W as $PB + RU$, where $R = \text{diag}\{(1 - \rho_1^2)^{1/2}, \dots, (1 - \rho_p^2)^{1/2}\} = (I - P^2)^{1/2}$, and where U is a standard p -vector Brownian motion, independent of B . This implies that

$$Z = \left(\int_0^1 FF' du \right)^{-1/2} \int_0^1 F dU' \sim N(0, I_{qp}),$$

independently of B . Defining

$$X = \left(\int_0^1 FF' du \right)^{-1/2} \int_0^1 F dB',$$

¹The variance entries in Table 7 of Doornik (1998a) should be labelled $(n-p)$, $1, n-p=1, n-p=2$.

it follows that (remembering that both P and Q are diagonal)

$$\begin{aligned}
\mathbf{Q} &= \text{tr} \left([XP + ZR]' [XP + ZR] \right) \\
&= \text{tr} \left(PX'XP + PX'ZR + RZ'XP + RZ'ZR \right), \\
&= \text{tr} \left(PX'XP \right) + 2\text{tr} \left(PX'ZR \right) + \text{tr} \left(RZ'ZR \right) \\
&= \sum_{i=1}^p \rho_i^2 \mathbb{T}_i + 2\text{tr} \left(PX'ZR \right) + \sum_{i=1}^p (1 - \rho_i^2) \xi_i,
\end{aligned} \tag{12}$$

where $\xi_i = \sum_{j=1}^q Z_{ji}^2$ are independent $\chi^2(q)$ random variables. Because Z is independent of X , $E(2\text{tr}[PX'ZR]) = 0$. Thus we find

$$\begin{aligned}
E(\mathbf{Q}) &= \sum_{i=1}^p \rho_i^2 E(\mathbb{T}_i) + \sum_{i=1}^p (1 - \rho_i^2) q \\
&= \frac{\sum_{i=1}^p \rho_i^2}{p} E(\mathbb{T}) + \left(1 - \frac{\sum_{i=1}^p \rho_i^2}{p} \right) pq.
\end{aligned}$$

To obtain the variance of \mathbf{Q} , we first note that the first and third term in (12) are independent, and hence uncorrelated. Furthermore, the second term is uncorrelated with the first term, because it has mean zero conditionally on X . Next, the covariance between the second and third term in (12) is zero (conditionally on X), because elements of Z are uncorrelated with squared elements of Z . Hence all covariances are zero, and the variance of \mathbf{Q} can be reduced to

$$\text{var}(\mathbf{Q}) = \text{var} \left(\sum_{i=1}^p \rho_i^2 \mathbb{T}_i \right) + 4\text{var}(\text{tr}[PX'ZR]) + \text{var} \left(\sum_{i=1}^p (1 - \rho_i^2) \xi_i \right). \tag{13}$$

For the first term of (13) we find

$$\text{var} \left(\sum_{i=1}^p \rho_i^2 \mathbb{T}_i \right) = \sum_{i=1}^p \rho_i^4 \text{var}(\mathbb{T}_i) + 2 \sum_{i=2}^p \sum_{j=1}^{i-1} \rho_i^2 \rho_j^2 \text{cov}(\mathbb{T}_i, \mathbb{T}_j).$$

To evaluate the second variance term, we use

$$\text{tr}(PX'ZR) = \text{vec}(P)'(X' \otimes R)\text{vec}(Z').$$

Hence

$$\begin{aligned}
\text{var}(\text{tr}[PX'ZR]) &= E \left\{ \text{var} [\text{vec}(P)'(X' \otimes R)\text{vec}(Z') | X] \right\} \\
&= E [\text{vec}(P)'(X'X \otimes R^2)\text{vec}(P)] \\
&= E [\text{tr}(X'XP^2R^2)] \\
&= \sum_{i=1}^p \rho_i^2 (1 - \rho_i^2) E(\mathbb{T}_i) \\
&= \frac{\sum_{i=1}^p \rho_i^2 (1 - \rho_i^2)}{p} E(\mathbb{T}).
\end{aligned}$$

The final term in (13) follows immediately from the fact that $\xi_i \sim \text{i.i.d. } \chi^2(q)$:

$$\text{var} \left(\sum_{i=1}^p (1 - \rho_i^2) \xi_i \right) = 2q \sum_{i=1}^p (1 - \rho_i^2)^2.$$

This completes the proof. \square

3 Evaluation of the Approximation

The accuracy of the Gamma approximation based on the results from Theorem 1 is assessed using simulation. The experimental design is as follows. Following Doornik (1998a, Section 4), the distribution of Q is simulated from $-T \sum \log(1 - \lambda_i)$. Here, λ_i are the eigenvalues of $T^{-1}E'R(R'R)^{-1}R'E$, which is the discrete approximation to the expression inside the trace of (4). The dimensions and correlations are respectively: $p = 1, \dots, 5$, and $(\rho_1, \dots, \rho_p) \in \mathcal{P} \times \dots \times \mathcal{P}$, where $\mathcal{P} = \{0, 0.2, 0.4, 0.6, 0.8, 1\}$ (the ordering of the correlations is irrelevant). The design consists of 461 specifications for each of the three treatments of deterministic terms, so 1383 experiments in total. This corresponds to the tables in Seo (1998). The number of Monte Carlo replications was chosen as $M = 10\,000$, and the sample size in the discretization as $T = 2000$.²

Table 1 compares the absolute relative error in the mean and standard deviation of Q when using Theorem 1 together with (7)–(9) to that found in the simulations. The table reports the mean of the absolute relative errors for each dimension separately as a percentage. The proposed procedure is very accurate indeed. The discrepancy is somewhat higher at dimension one. Here, the distribution is very skewed (as discussed in Doornik, 1998a), requiring more precise estimates of the mean and variance, corresponding to the dummies for low dimensions needed in (7) and (8).

Table 1: Mean absolute relative errors of $E(Q)$ and $s.d.(Q)$

	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
mean	0.83%	0.35%	0.37%	0.23%	0.21%
std.deviation	1.27%	0.75%	0.75%	0.70%	0.71%
count	18	63	168	378	756

To evaluate the whole procedure, we contrast the p -values obtained from the Gamma approximation based on (7)–(9) and Theorem 1 with the empirical rejection frequencies. We record the absolute difference between the p -value from the Gamma approximation and the simulated values. Table 2 reports the percentage of experiments where this absolute difference exceeds 0.0025, 0.005, 0.01 respectively. In no experiment did the difference exceed 0.02. For example, at 0.95 (the most commonly used p -value), 27% of the experiments (379 out of 1383) have a difference exceeding 0.0025, and only 2% a difference greater than 0.005 (but never exceeding 0.01). If the simulated distribution were exact, this implies that only 2% from the Gamma approximation are outside the range 0.945 – 0.955. Actually, we cannot rule out that the Gamma approximation is more accurate than the simulated values, because the latter uses $T = 2000$, rather than the range of sample sizes which were used to obtain (7) and (8).

²All experiments are done using Ox 2.0, see Doornik (1998b).

Table 2: Difference between Gamma-based and simulated distribution of Q

probability:	0.01	0.05	0.1	0.2	0.4	0.5	0.6	0.8	0.9	0.95	0.99
> 0.0025 difference	6%	30%	43%	55%	70%	73%	70%	63%	45%	27%	2%
> 0.005 difference	1%	5%	12%	27%	43%	44%	43%	29%	11%	2%	0%
> 0.01 difference	0%	1%	1%	3%	9%	11%	11%	3%	0%	0%	0%
> 0.02 difference	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

4 Application

To illustrate the proposed procedure, we use a variant of the model estimated by Johansen and Juselius (1992). They estimate a VAR(2) with unrestricted constant (model H_{lc}) and seasonal dummies using the UK wholesale price index (p_1), the trade-weighted foreign wholesale price index (p_2), the UK effective exchange rate (e_{12}), the three-month treasury bill rate in the UK (i_1), and the three-month Eurodollar interest rate in the UK (i_2). In addition, the change in oil price (Δp_{oil}) and its lag were used as conditioning variables. Johansen and Juselius (1992) found two cointegrating vectors. Hansen and Juselius (1995) use a transformed version in terms of $p_1 - p_2, \Delta p_1$ to avoid I(2)-ness. Following Rahbek and Mosconi (1998) we adopt model H_l by allowing a trend to enter the cointegrating space. In terms of (3) our specification is a VAR(2):

$$\begin{aligned} X_t &= (p_{1t} - p_{2t}, \Delta p_{1t}, e_{12,t}, i_{1t}, i_{2t})', & X_t^* &= (X_t', t)', \\ q_t &= (1, S_{1t}, S_{2t}, S_{3t})', \\ Z_t &= \Delta p_{oil,t}, \end{aligned}$$

with $k - 1 = m = 1$, and where S_{it} are seasonal dummy variables. The effective sample size, after taking all lags into account is 1972 (4) – 1987 (2).

Table 3: p -values for the tests T and Q

r	$p - r$	trace test	$1 - P(T)$	$1 - P(Q)$	$\hat{\rho}$
0	5	95.3	0.014	0.002	1, 1, 1, 0.842, 0.378
1	4	61.4	0.077	0.016	1, 1, 0.881, 0.396,
2	3	37.8	0.150	0.084	1, 0.918, 0.813
3	2	16.7	0.445	0.303	0.968, 0.826,
4	1	5.27	0.567	0.515	0.960,

Table 3 lists the test values for each rank, together with p -values. The fourth column, labelled $1 - P(T)$, gives the asymptotic p -value under the assumption that the presence of Z_t does not affect the distribution. The second cointegrating vector is only present if we are willing to adopt a 10% significance level. However, the sample is very small: 59 observations with 17 regressors in each equation. There

may be a tendency to overreject the true rank in small samples. Indeed, when adopting the sample-size adjusted p -value for $H_l(1)$ (using the formulae in Doornik, 1998a), we find that it changes from 7.7% to 14%. Correcting for the stationary exogenous regressors shrinks the distribution towards zero, so that the p -values will always decrease. Now rank one is firmly rejected with a p -value of 1.6%. If we use the small sample mean and variance of T in the formulae for Q (which is somewhat ad hoc), the p -values for ranks 0–2 change to 1%, 3%, 11% respectively.

The canonical correlations $\hat{\rho}_i$ are computed using the kernel method of Andrews (1991), as suggested in Seo (1998). We use the quadratic spectral kernel with automatic bandwidth and an AR(1) for each component. However, we found that the standard long-run covariance matrix gave nearly identical results.

To illustrate the procedure to obtain the distribution, consider $H_l(r = 4)$, which has only one canonical correlation. The following steps are involved:

- Use $p - r = 1$ for p in (7) and (8) to compute $E(T)$ and $\text{var}(T)$. This yields 6.32 and 10.6. The next step requires $\text{var}(T_i)$ which is also 10.6 in this case.
- To compute $E(Q)$ and $\text{var}(Q)$, again use $p - r = 1$ for p ; q is one more, corresponding to the trend which has been added to the cointegrating vector. With $\hat{\rho} = 0.96$, the result is 5.98 and 10.85 respectively.
- The approximating distribution is $\Gamma(\cdot; 5.98^2/10.85, 5.98/10.85)$. Or roughly: 1.1×5.27 comes from a $\chi^2(6.6)$.

Rahbek and Mosconi (1998) suggest to add p_{oil} to the cointegrating space to avoid the need to compute the nuisance parameters. In that case, the analysis is conditional on an I(1) variable, and the analysis of Harbo *et al.* (1998) pertains. This test statistic, denoted S here, was considered by Doornik (1998a), and the p -values are listed in Table 4. As Rahbek and Mosconi (1998) note, these results barely support the hypothesis that the rank is one.

Table 4: p -values for the test conditional on the I(1) variable p_{oil}

r	$p - r$	trace test	$1 - P(S)$
0	5	99.1	0.059
1	4	65.2	0.179
2	3	41.5	0.246
3	2	18.8	0.570
4	1	6.10	0.682

This discrepancy merits further investigation. The small sample size is a possible limitation to the power of the tests, and therefore we shall work with a model which is more parsimonious. However, we first note that, although Δp_{oil} was added by Johansen and Juselius (1992) to avoid non-normality, there

is still strong non-normality in the i_1 equation, caused by a single outlier. Therefore we add a dummy for 1980(2). With this adjustment, all the vector and univariate misspecification tests of PcFiml (Doornik and Hendry, 1997) are passed. A joint tests on the seasonals supports their deletion; the same holds for the second lags, with the exception of $i_{2,t-2}$. The seasonals and second lags are deleted, but $\Delta i_{2,t-1}$ is entered unrestrictedly.

Table 5: p -values for the tests T, Q, and S

r	$p - r$	trace test	$1 - P(\text{T})$	$1 - P(\text{Q})$	trace test	$1 - P(\text{S})$
0	5	171.6	0.000	0.000	176.3	0.000
1	4	73.5	0.005	0.002	78.3	0.018
2	3	41.2	0.072	0.039	45.0	0.137
3	2	16.4	0.468	0.348	18.9	0.566
4	1	3.22	0.841	0.816	3.47	0.932

The new test results, using the same sample period, are in Table 5. (Doornik *et al.*, 1998, noted that an impulse dummy, when entered unrestrictedly, does not affect the distribution.) The outcomes are no longer contradictory: a rank of two or more (the original conclusion of Johansen and Juselius), is clearly supported. There is some evidence of a third cointegrating vector, but the small sample argument leads us to reject this. Accepting $r = 2$, we can test whether the oil price can be deleted in the model corresponding to S. The test supports this: $\chi^2(2) = 0.92$ [0.63]. A weak form of purchasing power parity, namely that $p_1 - p_2$ and e_{12} have equal but opposite coefficients in both cointegrating vectors, is not rejected: $\chi^2(2) = 4.38$ [0.11] (without the trend, it would be strongly rejected).

5 Conclusion

We have derived a convenient way to tabulate the distribution of cointegration tests in the presence of additional stationary regressors. The proposed method is very accurate, and avoids the need for interpolation required with previous tabulations. In addition, the method is compact and easy-to-use, making it suitable for application in computer programs.

References

- Abadir, K.M. and A. Lucas (1996), "Approximate Normality of t -statistics Based on M-estimators of Unit Roots", Discussion Papers in Economics No. 96/32, University of York.
- Andrews, D.W.K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation", *Econometrica*, 59, 817–858.
- Boswijk, H.P. and A. Lucas (1997), "Semi-nonparametric Cointegration Testing", Research Memorandum 1997–41, Vrije Universiteit, Amsterdam.

- Doornik, J.A. (1998a), "Approximations to the Asymptotic Distribution of Cointegration Tests", *Journal of Economic Surveys*, 12, 573–593.
- Doornik, J.A. (1998b), *Ox – An Object-Oriented Matrix Programming Language*. London: Timberlake Consultants Press.
- Doornik, J.A. and D.F. Hendry (1997), *Modelling Dynamic Systems using PcFiml 9 for Windows*. London: Timberlake Consultants Press.
- Doornik, J.A., D.F. Hendry and B. Nielsen (1998), "Inference in Cointegrating Models: UK M1 Revisited", *Journal of Economic Surveys*, 12, 533–572.
- Hansen, B. (1995), "Rethinking the Univariate Approach to Unit Root Testing: Using Covariates to Increase Power", *Econometric Theory*, 11, 1148–1171.
- Hansen, H. and K. Juselius (1995), *CATS in RATS: Cointegration Analysis of Time Series*. Evanston (IL): Estima.
- Harbo, I., S. Johansen, B. Nielsen and A. Rahbek (1998), "Asymptotic Inference on Cointegrating Rank in Partial Systems", *Journal of Business Economics and Statistics*, 16, 388–399.
- Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 12, 231–254.
- Johansen, S. (1995), *Likelihood based Inference on Cointegration in the Vector Autoregressive Model*. Oxford: Oxford University Press.
- Johansen, S. and K. Juselius (1992), "Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and the UIP for UK". *Journal of Econometrics*, 53, 211–244.
- Kremers, J., N.R. Ericsson and J.J. Dolado (1992), "The Power of Cointegration Tests", *Oxford Bulletin of Economics and Statistics*, 54, 325–348.
- Ling, S. and W.K. Li (1997), "Estimation and Testing for Unit Root Processes with GARCH(1,1) Errors", Technical report, University of Hong Kong.
- Ling, S. and W.K. Li (1998), "Limiting Distributions of Maximum Likelihood Estimators for Unstable Autoregressive-Moving Average Time Series with General Autoregressive Heteroskedastic Errors", *Annals of Statistics*, 26, 84–125.
- Lucas, A. (1997), "Cointegration Testing Using Pseudo Likelihood Ratio Tests", *Econometric Theory*, 13, 149–169.
- Nielsen, B. (1997), "Bartlett Correction of the Unit Root Test in Autoregressive Models", *Biometrika*, 84, 500–504.
- Rahbek, A. and R. Mosconi (1998), "The Role of Stationary Regressors in the Cointegration Test", Working paper, University of Copenhagen.
- Seo, B. (1998), "Statistical Inference on Cointegration Rank in Error Correction Models with Stationary Covariates", *Journal of Econometrics*, 85, 339–385.