Econometric Theory, **16**, 2000, 176–199. Printed in the United States of America.

TESTS OF COMMON STOCHASTIC TRENDS

JUKKA NYBLOM
University of Joensuu

ANDREW HARVEY
University of Cambridge

This paper is concerned with tests in multivariate time series models made up of random walk (with drift) and stationary components. When the stationary component is white noise, a Lagrange multiplier test of the hypothesis that the covariance matrix of the disturbances driving the multivariate random walk is null is shown to be locally best invariant, something that does not automatically follow in the multivariate case. The asymptotic distribution of the test statistic is derived for the general model. The test is then extended to deal with a serially correlated stationary component. The main contribution of the paper is to propose a test of the validity of a specified value for the rank of the covariance matrix of the disturbances driving the multivariate random walk. This rank is equal to the number of common trends, or levels, in the series. The test is very simple insofar as it does not require any models to be estimated, even if serial correlation is present. Its use with real data is illustrated in the context of a stochastic volatility model, and the relationship with tests in the cointegration literature is discussed.

1. INTRODUCTION

The random walk plus noise model has been used in a wide variety of situations. The test of Nyblom and Mäkeläinen (1983) is the locally best invariant (LBI) test of the hypothesis that the disturbance term generating the random walk is equal to zero in the special case of a random walk plus white noise model, but it may be modified as in Kwiatkowski, Phillips, Schmidt, and Shin (1992) to allow the random walk to be embedded within a more general stationary process.

The Nyblom and Mäkeläinen and Kwiatkowski et al. tests are of the null hypothesis that the series is stationary, or stationary around a deterministic trend, against the alternative that a random walk component is present. The standard unit root tests, such as Dickey–Fuller, are set up in an autoregressive framework and test the null hypothesis that the series is nonstationary against the alternative that it is stationary. Thus they are designed for rather different situations.

We thank Juha Alho, Bruce Hansen, David Harris, Grant Hillier, Pentti Saikkonen, Neil Shephard, and two anonymous referees for helpful comments. Of course we are responsible for any errors. Address correspondence to: A.C. Harvey, Faculty of Economics and Politics, University of Cambridge, Sidgwick Ave., Cambridge CB3 9DD, United Kingdom.

 © 2000 Cambridge University Press 0266-4666/00 $9.50

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

This paper is concerned with tests in multivariate models. Using the Lagrange multiplier (LM) principle Nyblom (1989) proposed Nyblom and Mäkeläinen type tests for the parameter constancy against their changing in fairly general models. Our first result is to show that this test applied to the multivariate random walk plus noise model is LBI, something that does not immediately follow in the multivariate case (cf. King and Hillier, 1985). Under the null hypothesis that no random walk component is present, it is known that, asymptotically, the test statistic has a generalized Cramér–von Mises distribution. Here we also derive its asymptotic distribution for the general model and show that, somewhat surprisingly, this distribution depends only on the rank, denoted K, of the covariance matrix of the disturbances driving the random walk. We then move on to discuss ways in which the test may be generalized to deal with time trends and serial correlation. This completes Section 2.

When K is less than the number of series, N, the model contains K common random walk components or common trends. Common trends imply cointegration and vice versa. When economic theory suggests a given cointegrating vector, we can try to find evidence in its favor by rejecting the null hypothesis that there is a unit root in the supposed cointegrating relationship and applying a test such as augmented Dickey–Fuller. Alternatively, we can test whether the data are consistent with such a cointegrating relationship by using the Nyblom and Mäkeläinen or Kwiatkowski et al. test to test for the presence of a random walk in an otherwise stationary time series. If the coefficients in the cointegrating relationship are not prespecified, they must be estimated by an efficient estimation procedure and the tests modified as in Harris and Inder (1994) and Shin (1994). The multivariate generalizations of Nyblom and Mäkeläinen and Kwiatkowski et al. may be used to test the joint validity of several prespecified cointegrating vectors. Again, an efficient estimation procedure must be used if the coefficients in the cointegrating relationship are not prespecified (see Choi and Ahn, 1995). Section 3 discusses these issues, primarily to set the scene for the development of the new test in Section 4.

The test proposed in Section 4 is for the null hypothesis that K takes a given value, less than N, against the alternative that it has a greater value. Thus the test can be regarded as a test of the validity of the hypothesis that there are a certain number of common trends. Recognition of common trends yields models that may not only have an interesting interpretation but may also provide more efficient inferences and forecasts. The test generalizes the multivariate Nyblom and Mäkeläinen test of Section 2 where K is taken to be zero. It does not actually require that a model be estimated, and this remains the case even when we make a correction for serial correlation along the lines of Kwiatkowski et al. The asymptotic distribution of the test statistic, which depends only on the number of series and the hypothesized number of common trends, is derived, and critical values are tabulated. Critical values are also given for models containing time trends. The use of the test is illustrated in the context of the multivariate stochastic volatility model of Harvey, Ruiz, and Shephard (1994). The final subsection shows how the test relates to the proce-

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

dures described in Section 3 and to the tests of Johansen (1988) and Stock and Watson (1988). Section 5 presents the conclusions.

2. TESTS FOR THE PRESENCE OF RANDOM WALK COMPONENTS

2.1. Univariate Random Walk plus Noise Model

Consider a univariate unobserved components model consisting of a random walk plus noise for a set of observations, y_t

$$y_t = \mu_t + \varepsilon_t, \qquad \mu_t = \mu_{t-1} + \eta_t, \qquad t = 1,\ldots,T, \tag{1}$$

where the η_t''s and ε_t''s are mutually and serially uncorrelated Gaussian disturbances with variances σ_η^2 and σ_ε^2, respectively. Assume μ_0 is fixed. Then when $\sigma_\eta^2 = 0$ the random walk becomes a constant level. Testing this hypothesis is nonstandard, and the distribution of the likelihood ratio (LR) statistic is unknown. This motivated Nyblom and Mäkeläinen (1983) to consider most powerful invariant (MPI) tests. These tests depend on the signal to noise ratio, $q = \sigma_\eta^2/\sigma_\varepsilon^2$, but maximizing power near $q = 0$ leads to the LBI test, and this does not depend on q.

The LBI test of the null hypothesis that $\sigma_\eta^2 = 0$, against the alternative that $\sigma_\eta^2 > 0$, can be formulated as

$$\xi = T^{-2}\sum_{i=1}^{T}\left[\sum_{t=1}^{i}(y_t - \bar{y})\right]^2 \Big/ s^2 > c, \tag{2}$$

where s^2 is the sample variance of the observations and c is a critical value. In fact, one initially obtains a form of the statistic with the summations running in reverse, that is, from $t = i$ to T, but it is easily seen that the two statistics are identical. Nyblom and Mäkeläinen (1983, p. 859) give a table of critical values for the LBI statistic. Bearing in mind that they use a normalizing factor of $T - 1$ instead of T, the 5% critical value for ξ is 0.462 for sample sizes of both 50 and 100. Asymptotically, ξ has the Cramér–von Mises distribution for which the 5% critical value is 0.461. Invoking the invariance principle for cumulative sums of independently and identically distributed (i.i.d.) variables guarantees that the same asymptotic distribution is obtained without the normality assumption (see Nabeya and Tanaka, 1988). When a time trend is included in (1), ξ is formed from the residuals from a linear trend regression, and it follows from Nyblom (1986) that the 5% critical value is 0.148 when $T = 100$, whereas the asymptotic distribution yields 0.148 (Nyblom [1986] gave 0.149, but 0.148 seems to be correct as given already by MacNeill [1978]).

2.2. Multivariate Random Walk plus Noise Model

If $\mathbf{y}_t$ is a vector containing N time series the Gaussian multivariate local level model is

$$\begin{aligned}\mathbf{y}_t &= \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathrm{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon),\\ \boldsymbol{\mu}_t &= \boldsymbol{\mu}_{t-1} + \boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t \sim \mathrm{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\eta), \qquad t = 1,\ldots,T,\end{aligned} \tag{3}$$

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

where $\mathbf{\Sigma}_\varepsilon$ is an $N \times N$ positive definite matrix. We seek a test of the null hypothesis that $\mathbf{\Sigma}_\eta = \mathbf{0}$; in other words a test of whether there is any nonstationarity in the system.

The LR test is automatically invariant with respect to affine linear transformations, that is,

$$\mathbf{y}_t \mapsto \mathbf{P}\mathbf{y}_t + \mathbf{a}, \tag{4}$$

but because its limiting distribution cannot be obtained we would like to find a viable invariant test that is in some sense optimal. Now from the theory of matrices it is known that $\mathbf{\Sigma}_\varepsilon$ and $\mathbf{\Sigma}_\eta$ can be simultaneously diagonalized (see Rao, 1973, p. 41). To be precise, there exists a nonsingular matrix $\mathbf{P}$ such that, for any $\mathbf{\Sigma}_\eta$, and any $\mathbf{\Sigma}_\varepsilon$ that is positive definite, $\mathbf{P}\mathbf{\Sigma}_\varepsilon\mathbf{P}' = \mathbf{I}$ and $\mathbf{P}\mathbf{\Sigma}_\eta\mathbf{P}' = \mathbf{Q} = \text{diag}[q_1, \ldots, q_N]$, where the q_j''s are roots obtained from

$$|\mathbf{\Sigma}_\eta - q_j \mathbf{\Sigma}_\varepsilon| = 0, \qquad j = 1, \ldots, N. \tag{5}$$

The distribution of any invariant test statistic depends only on these roots. Our null hypothesis corresponds to all of these roots being zero, and, if we have no prior beliefs about the relative magnitudes of the roots under the alternative hypothesis, it seems reasonable to seek tests that maximize power against equal roots, that is, against $q_1 = q_2 = \cdots = q_N = q > 0$. In terms of the original covariance matrices, this alternative corresponds to $\mathbf{\Sigma}_\eta = q\mathbf{\Sigma}_\varepsilon$, and such a model is said to be homogeneous (see Harvey, 1989, Ch. 8). Theorem A.1 in Appendix A gives the form of the MPI test against homogeneous alternatives. The LBI test, which does not depend on q, is obtained as a corollary, but we state it here as a theorem in its own right.

THEOREM 1. *In the multivariate local level model*, (3), *the LBI test*, *invariant under* (4), *against the homogeneous alternative*, $\mathbf{\Sigma}_\eta = q\mathbf{\Sigma}_\varepsilon$, *has the rejection region*

$$\xi_N = \text{tr}[\mathbf{S}^{-1}\mathbf{C}] > c, \tag{6}$$

where

$$\mathbf{C} = T^{-2}\sum_{i=1}^{T}\left[\sum_{t=1}^{i}(\mathbf{y}_t - \bar{\mathbf{y}})\right]\left[\sum_{t=1}^{i}(\mathbf{y}_t - \bar{\mathbf{y}})\right]' \quad \textit{and} \quad \mathbf{S} = T^{-1}\sum_{t=1}^{T}(\mathbf{y}_t - \bar{\mathbf{y}})(\mathbf{y}_t - \bar{\mathbf{y}})'. \tag{7}$$

It follows from Nyblom (1989) that (6) is the LM test under homogeneity, but unlike in the univariate case, the LM test is not automatically LBI.

Under the null hypothesis, the limiting distribution of (6) is that which is sometimes called the Cramér–von Mises distribution with N degrees of freedom, denoted $CvM(N)$. Following Nyblom and Mäkeläinen (1983) this distribution may be represented by a series expansion. The following theorem is proved in Appendix B, Corollary B.2, with the series expansion derived in Appendix C.

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

TABLE 1. Critical values for ξ_N test ($K = 0$) and $\zeta_{K,N}$ test

N	Significance level (%)	K			
		0	1	2	3
2	10	.607	.162		
	5	.748	.218		
	1	1.074	.383		
3	10	.841	.297	.094	
	5	1.000	.382	.120	
	1	1.359	.614	.200	
4	10	1.063	.427	.170	.063
	5	1.237	.538	.208	.078
	1	1.623	.830	.317	.121

THEOREM 2. *When* $\mathbf{\Sigma}_\eta = \mathbf{0}$ *in* (3) *the limiting distribution of* (6) *is*

$$\xi_N \xrightarrow{d} \int_0^1 \mathbf{B}(u)' \mathbf{B}(u)\, du = \sum_{k=1}^{\infty} (\pi k)^{-2} x_k, \tag{8}$$

where $\mathbf{B}$ *is a standard vector Brownian bridge of dimension* N *and* $x_k \sim \chi^2(N)$.

The series expansion was given in Nyblom (1989). It can be used to calculate critical values by simulation or by numerical inversion of the Laplace transformation of the distribution function. Selected percentage points for ξ_N are given in Table 1 in the $K = 0$ column (see also Canova and Hansen, 1995).

Although the test maximizes the power against homogeneous alternatives, it is consistent against all nonnull $\mathbf{\Sigma}_\eta'$'s. This follows from the next result where we show that $T^{-1}\xi_N$ has a nondegenerate limiting distribution.

The remarkable feature about the limiting distribution of $T^{-1}\xi_N$ is that it depends only on the number of nonzero q_j''s, that is, the rank of $\mathbf{\Sigma}_\eta$. The following theorem is proved in Appendix B, Corollary B.2, with the series expansion derived in Appendix C.

THEOREM 3. *When the rank of* $\mathbf{\Sigma}_\eta$ *in model* (3) *is* $K \geq 1$, *then*

$$T^{-1}\xi_N = T^{-1} \operatorname{tr}(\mathbf{S}^{-1}\mathbf{C}) \xrightarrow{d} \operatorname{tr}(\mathbf{S}_{11}^{*-1}\mathbf{C}_{11}^{*}), \tag{9}$$

where

$$\mathbf{S}_{11}^{*} = \int_0^1 \mathbf{W}^{*}(u)\mathbf{W}^{*}(u)'\, du, \tag{10}$$

$$\mathbf{C}_{11}^{*} = \int_0^1 \left[\int_0^u \mathbf{W}^{*}(s)\, ds\right]\left[\int_0^u \mathbf{W}^{*}(s)\, ds\right]' du \tag{11}$$

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

TABLE 2. Critical values for ξ_N and $\zeta_{K,N}$ tests with time trends

N	Significance level (%)	K			
		0	1	2	3
2	10	.211	.085		
	5	.247	.105		
	1	.329	.160		
3	10	.296	.151	.061	
	5	.332	.180	.075	
	1	.428	.245	.113	
4	10	.377	.215	.110	.046
	5	.423	.246	.128	.055
	1	.521	.321	.176	.081

are $K \times K$ matrices with

$$\mathbf{W}^*(u) = \mathbf{W}(u) - \int_0^1 \mathbf{W}(s)\, ds, \qquad 0 \le u \le 1 \tag{12}$$

and $\mathbf{W}$ *being a standard vector Brownian motion process of dimension K.*

The series expansion is

$$T^{-1}\xi_N \xrightarrow{d} \operatorname{tr}\left(\sum_{k=1}^{\infty} (\pi k)^{-2} \boldsymbol{v}_k \boldsymbol{v}_k'\right)^{-1} \left(\sum_{k=1}^{\infty} (\pi k)^{-4} \boldsymbol{v}_k \boldsymbol{v}_k'\right), \qquad K = 1, \ldots, N, \tag{13}$$

where $\boldsymbol{v}_k$ is a $K \times 1$ vector such that $\boldsymbol{v}_k \sim \mathrm{NID}(\mathbf{0}, \mathbf{I}_K)$.

If a linear time trend is included in (3) so that

$$\mathbf{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\beta} t + \boldsymbol{\varepsilon}_t, \qquad t = 1, \ldots, T, \tag{14}$$

where $\boldsymbol{\beta}$ is an $N \times 1$ vector of slope coefficients, the test statistic, (6), is formed from the ordinary least squares residuals from a set of regressions on a constant and time. As shown in Nyblom (1986, Sect. 3), the critical values are obtained by changing the weight $(\pi k)^{-2}$ to λ_k^{-2}, where $\lambda_{2k-1} = 2k\pi$ and λ_{2k} is the root of $\tan(\lambda/2) = \lambda/2$ on $(2k\pi, 2(k+1)\pi)$, $k = 1, 2, \ldots$ (see Table 2).

2.3. Nonparametric Adjustment for Serial Correlation

Introducing a simple modification to s^2 allows the univariate test in (2) to be valid when the noise, within which the random walk is embedded, has some degree of autocorrelation and even heteroskedasticity. The regularity conditions allow for finite order autoregressive moving average (ARMA) processes with i.d.d. disturbances as a special case. The key argument for the validity of the asymptotic distribution of the modified test statistic under the null hypothesis is that the

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

integrated noise process (the cumulative sums of the noise) converges weakly to a Brownian motion and that the scale of this Brownian motion is consistently estimated. The scale of the limiting Brownian motion depends on the "long-run" variance, which is equal to the spectral density at frequency zero. Kwiatkowski et al. proposed estimating the long-run variance by

$$s^2(m) = c(0) + 2\sum_{\tau=1}^{m} w_{\tau m} c(\tau), \tag{15}$$

where $c(\tau)$ is the sample autocovariance at lag τ and $w_{\tau m}$ is a weighting function such as $w_{\tau m} = 1 - \tau/(m+1), \tau = 1, \ldots, m$. The somewhat complicated interplay between various weighting functions, and also the lag window m and different moment conditions, is studied in Andrews (1991), but we will not delve into these issues here. When $m \to \infty$, but at a slower rate than T, say, $m = o(T^{1/4})$, expression (15) estimates the long-run variance consistently if the random walk component is absent. Thus the Kwiatkowski et al. test is a modification of ξ in which s^2 is replaced by $s^2(m)$. Of course, it is not without its drawbacks. For example, Kwiatkowski et al. demonstrated that if the stationary component is an AR(1) with strong serial correlation, m must be high to get a test size close to the nominal and this weakens the power (see also Leybourne and McCabe, 1994).

Now consider a generalization of the multivariate model in (3), in which the process $\boldsymbol{\varepsilon}_t$ satisfies the multivariate version of the conditions in Phillips and Perron (1988); joint stationarity is a special case. The ξ_N test can be generalized along the lines of the Kwiatkowski et al. test quite straightforwardly by replacing $\mathbf{S}$ in (6) by

$$\mathbf{S}(m) = \sum_{\tau=-m}^{\tau=m} w_{\tau m} \hat{\boldsymbol{\Gamma}}(\tau), \tag{16}$$

where $\hat{\boldsymbol{\Gamma}}(\tau)$ is the sample autocovariance matrix at lag τ, that is,

$$\hat{\boldsymbol{\Gamma}}(\tau) = T^{-1} \sum_{t=\tau+1}^{T} (\mathbf{y}_t - \bar{\mathbf{y}})(\mathbf{y}_{t-\tau} - \bar{\mathbf{y}})' \tag{17}$$

and $w_{\tau m}$ is a weighting function as in (15). Note that $\hat{\boldsymbol{\Gamma}}(-\tau) = \hat{\boldsymbol{\Gamma}}(\tau)'$. We will refer to this test as $\xi_N(m)$. The matrix in (16) now estimates the long-run covariance matrix, or spectral density matrix at frequency zero, of the noise process. If $m = o(T^{1/4})$, the limiting distribution of $\xi_N(m)$ is $CvM(N)$ under the null hypothesis. Under the alternative, the limiting distribution of Theorem 3 holds for $(m/T)\xi_N(m)$. The results are stated formally in Corollary B.5 of Appendix B.

3. COMMON TRENDS AND COINTEGRATION

If $\boldsymbol{\Sigma}_\eta$ is of full rank, the multivariate local level model of (3) can be written in the following way:

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

$$\mathbf{y}_t = \begin{bmatrix} \mathbf{y}_{1t} \\ \mathbf{y}_{2t} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_K & \mathbf{0} \\ \boldsymbol{\Theta} & \mathbf{I}_r \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}_t^\dagger \\ \bar{\boldsymbol{\mu}}_t \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_{1t} \\ \boldsymbol{\varepsilon}_{2t} \end{bmatrix}, \qquad t = 1, \ldots, T, \tag{18}$$

where $\mathbf{y}_t$ is partitioned into a $K \times 1$ vector $\mathbf{y}_{1t}$ and an $r \times 1$ vector $\mathbf{y}_{2t}$, with $r = N - K$, $\boldsymbol{\varepsilon}_t$ is similarly partitioned, $\boldsymbol{\Theta}$ is an $r \times K$ matrix of coefficients, and the $K \times 1$ vector $\boldsymbol{\mu}_t^\dagger$ follows a multivariate random walk

$$\boldsymbol{\mu}_t^\dagger = \boldsymbol{\mu}_{t-1}^\dagger + \boldsymbol{\eta}_t^\dagger, \qquad \boldsymbol{\eta}_t^\dagger \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\eta^\dagger), \tag{19}$$

where $\boldsymbol{\eta}_t^\dagger$ is a $K \times 1$ vector and $\boldsymbol{\Sigma}_\eta^\dagger$ is a $K \times K$ positive definite matrix. The $r \times 1$ vector $\bar{\boldsymbol{\mu}}_t$ also follows a multivariate random walk with the disturbance vector $\bar{\boldsymbol{\eta}}_t$ having a positive definite covariance matrix $\bar{\boldsymbol{\Sigma}}_\eta$.

If the rank of the covariance matrix $\boldsymbol{\Sigma}_\eta$ is $K < N$, an appropriate ordering of the series enables the model to be written in the form (18) with $\boldsymbol{\mu}_t^\dagger$ containing K common levels, or trends, and $\boldsymbol{\Sigma}_\eta^\dagger$ continuing to be positive definite, whereas $\bar{\boldsymbol{\Sigma}}_\eta = \mathbf{0}$ and thus $\bar{\boldsymbol{\mu}}_t \equiv \bar{\boldsymbol{\mu}}_0$. This means that $\mathbf{y}_{2t}$ is cointegrated with $\mathbf{y}_{1t}$, the latter being a set of variables that are not themselves cointegrated. Now consider a set of r linear combinations of the observations, $\mathbf{A}\mathbf{y}_t$, where $\mathbf{A}$ is an $r \times N$ matrix partitioned as $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2)$ with $\mathbf{A}_2$ being $r \times r$. Then

$$\begin{aligned} \mathbf{A}\mathbf{y}_t &= \mathbf{A}_1 \mathbf{y}_{1t} + \mathbf{A}_2 \mathbf{y}_{2t} \\ &= (\mathbf{A}_1 + \mathbf{A}_2 \boldsymbol{\Theta}) \boldsymbol{\mu}_t^\dagger + \mathbf{A}_2 \bar{\boldsymbol{\mu}}_t + \mathbf{A}_1 \boldsymbol{\varepsilon}_{1t} + \mathbf{A}_2 \boldsymbol{\varepsilon}_{2t}, \qquad t = 1, \ldots, T. \end{aligned} \tag{20}$$

If $\mathbf{A}$ satisfies $\mathbf{A}_1 + \mathbf{A}_2 \boldsymbol{\Theta} = \mathbf{0}$ and $\bar{\boldsymbol{\Sigma}}_\eta = \mathbf{0}$, then $\mathbf{A}\mathbf{y}_t$ is stationary and the rows of $\mathbf{A}$ constitute a set of r cointegrating vectors. The assumption that $\boldsymbol{\Sigma}_\eta^\dagger$ is positive definite implies that $\mathbf{A}_2$ must be nonsingular (see Saikkonen, 1993, p. 21).

Suppose that we wish to test the null hypothesis that $\bar{\boldsymbol{\Sigma}}_\eta = \mathbf{0}$, that is, $\mathbf{y}_{2t}$ is cointegrated with $\mathbf{y}_{1t}$, in the model (18). If we choose $\mathbf{A}$ such that $\mathbf{A}_1 + \mathbf{A}_2 \boldsymbol{\Theta} = \mathbf{0}$, then the ξ_N test applied to $\mathbf{A}\mathbf{y}_t$ is LBI against alternatives in which $\bar{\boldsymbol{\Sigma}}_\eta$ is proportional to $\boldsymbol{\Theta}\boldsymbol{\Sigma}_{11}\boldsymbol{\Theta}' + \boldsymbol{\Theta}\boldsymbol{\Sigma}_{12} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Theta}' + \boldsymbol{\Sigma}_{22}$, where the $\boldsymbol{\Sigma}_{ij}'$s are blocks of $\boldsymbol{\Sigma}_\varepsilon$. The form of $\mathbf{A}$ may come directly from economic theory or indirectly, through $\boldsymbol{\Theta}$, because of knowledge of common trends. For example, suppose there is a single common trend in the logarithms of income (y_1), consumption (y_2), and investment (y_3), so that $N = 3$, $K = 1$. With balanced growth $\boldsymbol{\Theta} = (1,1)'$, and if the cointegrating relationships are taken to be $y_{2t} - y_{1t}$ and $y_{3t} - y_{1t}$, they correspond to the "great ratios" of consumption and investment to income (see King, Plosser, Stock, and Watson, 1991).

The test statistic can be written as

$$\xi_r(\mathbf{A}) = \text{tr}(\mathbf{A}\mathbf{S}\mathbf{A}')^{-1}\mathbf{A}\mathbf{C}\mathbf{A}', \tag{21}$$

and its limiting distribution under the null hypothesis is, of course, $CvM(r)$. The test will tend to reject the null hypothesis not only if $\bar{\boldsymbol{\Sigma}}_\eta \neq \mathbf{0}$ but also if the supposed cointegrating vectors are incorrectly specified so that $\mathbf{A}_1 + \mathbf{A}_2 \boldsymbol{\Theta} \neq \mathbf{0}$. Thus the test can be regarded as a test of the validity of a set of prespecified

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

cointegrating vectors; in other words it is a test of whether $\mathbf{A}\mathbf{y}_t$ is stationary. Following (5), its power depends on the r roots obtained from

$$|\mathbf{A}\boldsymbol{\Sigma}_\eta\mathbf{A}' - q_j\mathbf{A}\boldsymbol{\Sigma}_\varepsilon\mathbf{A}'| = 0, \qquad j = 1,\ldots,r. \tag{22}$$

If $\mathbf{A}\mathbf{y}_t$ is stationary then $\mathbf{A}\boldsymbol{\Sigma}_\eta\mathbf{A}' = \mathbf{0}$ and all the roots are zero. If the test is modified to allow for serial correlation, as described toward the end of Section 2, it can be used as a general test for the validity of a specified cointegrating vector.

Suppose now that the coefficients of the variables in the hypothesized cointegrating relationships are not prespecified. Because $\mathbf{A}_2$ must be nonsingular, (20) can be multiplied through by $\mathbf{A}_2^{-1}$. Then, if there are no overidentifying restrictions in (20), it is tempting to test for the validity of the supposed cointegrating relationships by regressing $\mathbf{y}_{2t}$ on $\mathbf{y}_{1t}$. However, the $\xi_r(\mathbf{A})$ test described previously will not, in general, have its usual $CvM(r)$ asymptotic distribution if formed from the OLS residuals. Ways of constructing valid tests are described by Choi and Ahn (1995), Harris and Inder (1994), Phillips and Ouliaris (1990), and Shin (1994). The SBDH test proposed by Choi and Ahn (1995) is a generalization of the $\xi_N(m)$ test described in Section 2.3.

4. TESTING FOR A SPECIFIED NUMBER OF COMMON TRENDS

We now consider how, in (3), we might test

$$H_0: \text{rank}(\boldsymbol{\Sigma}_\eta) = K \quad \text{against } H_1: \text{rank}(\boldsymbol{\Sigma}_\eta) > K$$

for $K < N$.

Let $\ell_1 \geq \cdots \geq \ell_N$ be the ordered eigenvalues of $\mathbf{S}^{-1}\mathbf{C}$, given by

$$|\mathbf{C} - \ell_j\mathbf{S}| = 0, \qquad j = 1,\ldots,N. \tag{23}$$

The ξ_N test statistic is the sum of these eigenvalues, but in Theorem B.1 and Corollary B.2 of Appendix B it is shown that when the rank of $\boldsymbol{\Sigma}_\eta$ is $K^\dagger$, the limiting distribution of $T^{-1}\xi_N$ is the limiting distribution of T^{-1} times the sum of the $K^\dagger$ largest eigenvalues. This suggests basing a test of the hypothesis that $\text{rank}(\boldsymbol{\Sigma}_\eta) = K$ on the sum of the $N - K$ smallest eigenvalues, that is,

$$\zeta_{K,N} = \ell_{K+1} + \cdots + \ell_N, \qquad K = 1,\ldots,N-1. \tag{24}$$

Then if $K^\dagger > K$ the relatively large values taken by the first $K^\dagger - K$ of these eigenvalues will tend to lead to the null hypothesis being rejected. We will call this the *common trends* test. Of course if we allow K to be zero in (24), then $\zeta_{0,N} = \xi_N$.

The asymptotic distribution of the common trends test is given in the theorems that follow. Before stating these, however, we present another rationalization for the test in terms of cointegration. As was demonstrated in the previous section, if the rank of $\boldsymbol{\Sigma}_\eta$ is K then there are K common trends and $r = N - K$ linear combinations of $\mathbf{y}_t$, denoted $\mathbf{A}\mathbf{y}_t$, that are white noise. (The second of these proposi-

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

tions can be shown to hold directly by noting that if $\text{rank}(\boldsymbol{\Sigma}_\eta) = K$, then there exist $r \times K$ matrices such that $\mathbf{A}\boldsymbol{\Sigma}_\eta = \mathbf{0}$.) Thus under H_0 we have $\mathbf{A}\boldsymbol{\Sigma}_\eta\mathbf{A}' = \mathbf{0}$, and if $\mathbf{A}$ can be specified, the test (6) is applicable to $\mathbf{A}\mathbf{y}_t$. If $\mathbf{A}$ cannot be specified, a natural test statistic for the null hypothesis that there are K common trends, against the alternative that there are more, arises by minimizing (21) over $\mathbf{A}$, with $\mathbf{A}$ being $r \times N$ and of rank r. This minimization problem can be solved as follows. Because $\text{tr}[(\mathbf{ASA}')^{-1}\mathbf{ACA}']$ remains unchanged when $\mathbf{A}$ is premultiplied by a nonsingular matrix, the minimization can be carried out under the condition $\mathbf{ASA}' = \mathbf{I}$. Then writing $\mathbf{D} = \mathbf{AS}^{1/2}$, where $\mathbf{S}^{1/2}$ is the positive definite square root of $\mathbf{S}$, we can appeal to the Poincaré separation theorem, as given in Rao (1973, p. 64), to deduce

$$\begin{aligned}\min_{\mathbf{ASA}'=\mathbf{I}} \text{tr}(\mathbf{ACA}') &= \min_{\mathbf{DD}'=\mathbf{I}} \text{tr}(\mathbf{DS}^{-1/2}\mathbf{CS}^{-1/2}\mathbf{D}') \\ &= \sum_{j=0}^{r-1} \lambda_{N-j}(\mathbf{S}^{-1/2}\mathbf{CS}^{-1/2}),\end{aligned}$$

where $\lambda_i(\cdot)$ denotes the ith largest eigenvalue of the matrix in the parentheses. Because $\lambda_i(\mathbf{S}^{-1/2}\mathbf{CS}^{-1/2}) = \lambda_i(\mathbf{S}^{-1}\mathbf{C})$, we can write

$$\zeta_{K,N} = \min_{\mathbf{A}} \text{tr}[(\mathbf{ASA}')^{-1}\mathbf{ACA}']. \tag{25}$$

The distribution of the common trends test statistic under the null hypothesis is given by the following theorem, which is stated as Corollary B.3 in Appendix B.

THEOREM 4. *When the rank of $\boldsymbol{\Sigma}_\eta$ in model* (3) *is $K < N$, the limiting distribution of $\zeta_{N,K}$ is*

$$\zeta_{K,N} \xrightarrow{d} \text{tr}[\mathbf{C}^*_{22} - \mathbf{C}^{*\prime}_{12}(\mathbf{C}^*_{11})^{-1}\mathbf{C}^*_{12}], \tag{26}$$

where $\mathbf{C}^*_{11}$ is $K \times K$, $\mathbf{C}^*_{12}$ is $K \times r$, and $\mathbf{C}^*_{22}$ is $r \times r$, with

$$\mathbf{C}^*_{11} = \int_0^1 \left[\int_0^u \mathbf{W}^*(s)\,ds\right]\left[\int_0^u \mathbf{W}^*(s)\,ds\right]'\,du,$$

$$\mathbf{C}^*_{12} = \mathbf{C}^{*\prime}_{21} = \int_0^1 \left[\int_0^u \mathbf{W}^*(s)\,ds\right]\mathbf{B}(u)\,du,$$

$$\mathbf{C}^*_{22} = \int_0^1 \mathbf{B}(u)\mathbf{B}(u)'\,du,$$

where $\mathbf{B}$ and $\mathbf{W}^*$ are defined as in Theorems 2 *and* 3 *but are now of dimension r and K, respectively.*

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

The series expansion is

$$\zeta_{K,N} \xrightarrow{d} \sum_{k=1}^{\infty} (\pi k)^{-2} \mathbf{u}_k' \mathbf{u}_k - \operatorname{tr}\left(\sum_{k=1}^{\infty} (\pi k)^{-3} \mathbf{u}_k \boldsymbol{v}_k' \right) \left(\sum_{k=1}^{\infty} (\pi k)^{-4} \boldsymbol{v}_k \boldsymbol{v}_k' \right)^{-1} \times \left(\sum_{k=1}^{\infty} (\pi k)^{-3} \boldsymbol{v}_k \mathbf{u}_k' \right),$$

where $\boldsymbol{v}_k$ *and* $\mathbf{u}_k$ *are, respectively,* $K \times 1$ *and* $r \times 1$ *vectors that are mutually independent* NID$(\mathbf{0},\mathbf{I})$.

It is, perhaps, surprising that such a limiting distribution is obtained in view of the fact that $\mathbf{S}$ is not a consistent estimator of $\boldsymbol{\Sigma}_\varepsilon$ when $K > 0$.

The 10%, 5%, and 1% significance points of the distribution in (26) were found by simulation, a sufficiently accurate approximation being obtained by truncating the summations in (27) at $k = 50$ and replacing nonnegligible remainders by their expectations. The significance points for $\zeta_{K,N}$ depend on both K and N and are shown in Table 1. Note that although the ξ_N and $\zeta_{K,N}$ tests were derived for Gaussian models, their limiting distributions only require that the disturbances be i.i.d.

It is proved in Appendix B, Corollary B.3, and stated in Theorem 5, which follows, that under the alternative hypothesis, $T^{-1}\zeta_{K,N}$ has a limiting distribution that depends only on $K^\dagger - K$. Hence the test based on $\zeta_{K,N}$ is consistent.

THEOREM 5. *If the rank of* Σ_η *in model* (3) *is* $K^\dagger > K$, *the limiting distribution of* $T^{-1}\zeta_{K,N}$ *is the same as the limiting distribution of* $T^{-1}(\ell_{K+1} + \cdots + \ell_{K^\dagger})$. *This is the distribution of the sum of the* $K^\dagger - K$ *smallest eigenvalues of* $\mathbf{S}_{11}^{*-1}\mathbf{C}_{11}^{*}$.

As with the ξ_N test, the model may be extended to include time trends as in (14). Critical values are in Table 2. The test can make no allowance for the fact that there may be common slopes corresponding to the common levels. In the $\xi_r(\mathbf{A})$ test this possibility can be accommodated quite naturally because it implies that $\mathbf{A}\boldsymbol{\beta} = \mathbf{0}$, and so $\mathbf{A}\mathbf{y}_t$ includes no time trend.

4.1. Application to Stochastic Volatility

Harvey et al. (1994) considered a multivariate stochastic volatility (SV) model of the form

$$r_{it} = \sigma_i \varepsilon_{it} \exp(h_{it}/2), \qquad i = 1,\ldots,N, \qquad t = 1,\ldots,T,$$

where $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t},\ldots,\varepsilon_{Nt})'$ is i.i.d. $(\mathbf{0},\boldsymbol{\Sigma}_\varepsilon)$ and $\mathbf{h}_t = (h_{1t},\ldots,h_{Nt})'$ is a multivariate random walk

$$\mathbf{h}_t = \mathbf{h}_{t-1} + \boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t \sim \text{NID}(\mathbf{0},\boldsymbol{\Sigma}_\eta).$$

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

Taking the logarithms of the squared observations gives

$$\mathbf{y}_t = \boldsymbol{\omega} + \mathbf{h}_t + \boldsymbol{\varepsilon}_t^*, \qquad t = 1, \ldots, T,$$

where the ith elements in the $N \times 1$ vectors $\mathbf{y}_t$, $\boldsymbol{\omega}$, and $\boldsymbol{\varepsilon}_t^*$ are $\log r_{it}^2$, $\log \sigma_i^2$, and $\log \varepsilon_{it}^2$, respectively. The model is now of the form (3) except that it is not Gaussian. However, this makes no difference to the asymptotic distribution of the common trends test statistic.

In the application in Harvey et al. (1994), the r_{it}''s are the differences of 946 logged daily exchange rates of the dollar against four other currencies starting on November 1, 1981. Harvey et al. proposed a model based on two common trends. Computing the $\zeta_{2,4}$ statistic for $\mathbf{y}_t$ gives a value of 0.17. This is less than the 5% critical value in Table 1, and its probability value is 10%. Thus the assumption of two common trends is not rejected. On the other hand, the hypothesis of one common trend is rejected because $\zeta_{1,4} = 1.21$ and the 5% critical value is 0.54.

4.2. Serial Correlation

A nonparametric adjustment for serial correlation may be made in the common trends test in much the same way as was suggested for the ξ_N test. All that needs to be done is to replace the eigenvalues in (24) by those of $\mathbf{S}(m)^{-1}\mathbf{C}$, leading to a test statistic that we will denote as $\zeta_{K,N}(m)$. Under the null hypothesis, the asymptotic result stated in Theorem 4 continues to hold, whereas under the alternative the limiting distributions of Theorem 5 apply to $(m/T)\zeta_{K,N}(m)$ (see Appendix B, Corollaries B.5 and B.6).

4.3. Common Trends and Cointegration

The common trends test tests the null hypothesis that there are K common trends against the alternative that there are more. Equivalently, it tests the null hypothesis that there are $N - K$ cointegrating vectors against the alternative that there are less. Thus it is similar to the tests described at the end of Section 3. However, when $K > 1$, the common trends test tests a slightly different hypothesis than the one maintained in the tests of Section 3. This is because the latter are based on knowing that all the common trends can go in $\mathbf{y}_{1t}$, that is, the series in $\mathbf{y}_{1t}$ are not cointegrated. In terms of (18) and (19) this means that $\boldsymbol{\Sigma}_\eta^\dagger$ is positive definite. Thus suppose that $N = 3$ and that y_1 and y_2 depend on a single common trend driven by a disturbance that is uncorrelated with the disturbance of the trend in y_3. Then $K = 2$ and $r = 1$, but regressing y_3 on y_1 and y_2 will not provide a test of the null hypothesis that $r = 1$ against the alternative that $r = 0$. (Another way of seeing this problem is that the dependent variable is not in the cointegrating relationship.) A situation of this kind actually arises in the preceding SV example, where it can be argued that the trend in the exchange rate for the yen is uncorrelated with the common trend that appears in the three European currencies.

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

The procedures in Johansen (1988) and Stock and Watson (1988) also test for the number of cointegrating vectors/common trends. However, these are vector autoregression based tests, and our common trends test differs from them just as the Nyblom and Mäkeläinen and Kwiatkowski et al. tests differ from augmented Dickey–Fuller. Thus the alternative for $\zeta_{K,N}$ is in the direction of fewer cointegrating vectors rather than more. Furthermore no models are being estimated, either under the null or the alternative, and we are not testing against the alternative that there should specifically be one more common trend (or one less cointegrating vector).

Recent papers by Harris (1997) and Snell (1999) proposed tests of a similar hypothesis to the one addressed by our common trends tests. These tests are different from ours because they are based on estimating cointegrating relationships by principal components. In our approach no explicit estimators are needed. However, the eigenvectors corresponding to the eigenvalues in (26) may be interpreted as implicitly defining the hypothesized cointegration space. The question of whether these eigenvectors have some merit as estimators is not pursued here.

5. CONCLUSIONS

This paper has examined various aspects of testing for unobserved random walk components in multivariate time series models and contrasted them with existing tests for cointegration. The main contribution is to derive a test of the validity of a specified value for the rank of the covariance matrix of the disturbances driving a multivariate random walk. The alternative hypothesis is that the rank is higher; in other words there are more common trends. The rest of the model is assumed to be stationary, and if it is not white noise a correction can be made for serial correlation. This corrected test can be regarded as a generalization of the univariate one in Kwiatkowski et al. (1992). The test against the presence of a multivariate random walk component is a special case of the aforementioned common trends test in which the rank of the covariance matrix is zero under the null hypothesis. The tests are relatively simple to carry out because they do not require a model to be estimated and their asymptotic distributions depend only on the number of series and the hypothesized number of common trends.

REFERENCES

Anderson, T.W. (1984) *An Introduction to Multivariate Statistical Analysis*, 2nd ed. New York: Wiley.

Andrews, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–858.

Bhatia, R. (1997) *Matrix Analysis.* New York: Springer.

Billingsley, P. (1968) *Convergence of Probability Measures.* New York: Wiley.

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

Canova, F. & B.E. Hansen (1995) Are seasonal patterns constant over time? A test for seasonal stability. *Journal of Business and Economic Statistics* 13, 237–252.
Choi, I. & B.C. Ahn (1995) Testing for cointegration in systems of equations. *Econometric Theory* 11, 952–983.
Hall, P. & C.C. Heyde (1980) *Martingale Limit Theory and Its Applications.* New York: Academic Press.
Harris, D. (1997) Principal components analysis of cointegrated time series. *Econometric Theory* 13, 529–557.
Harris, D. & B. Inder (1994) A test of the null hypothesis of co-integration. In C. Hargreaves (ed.), *Non-stationary Time Series Analysis and Co-integration.* Oxford: Oxford University Press.
Harvey, A.C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.
Harvey, A.C., E. Ruiz, & N. Shephard (1994) Multivariate stochastic variance models. *Review of Economic Studies* 61, 247–264.
Herrndorf, N. (1984) A functional central limit theorem for weakly dependent sequences of random variables. *Annals of Probability* 12, 141–153.
Johansen, S. (1988) Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12, 13–54.
King, M.L. & G.H. Hillier (1985) Locally best invariant tests of the error covariance matrix of the linear regression model. *Journal of the Royal Statistical Society*, *Series B* 47, 98–102.
King, R.G., C.I. Plosser, J.H. Stock, & M.W. Watson (1991) Stochastic trends and economic fluctuations. *American Economic Review* 81, 819–840.
Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, & Y. Shin (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics* 44, 159–178.
Lehmann, E.L. (1994) *Testing Statistical Hypotheses*, 2nd ed. New York: Wiley.
Leybourne, S.J. & B.P.M. McCabe (1994) A consistent test for a unit root. *Journal of Business and Economic Statistics* 12, 157–166.
MacNeill, I.B. (1978) Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times. *Annals of Statistics* 6, 422–433.
Nabeya, S. & K. Tanaka (1988) Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative. *Annals of Statistics* 16, 218–235.
Nyblom, J. (1986) Testing for deterministic linear trend in time series. *Journal of the American Statistical Association* 81, 545–549.
Nyblom, J. (1989) Testing for the constancy of parameters over time. *Journal of the American Statistical Association* 84, 223–230.
Nyblom, J. & T. Mäkeläinen (1983) Comparison of tests for the presence of random walk coefficients in a simple linear model. *Journal of the American Statistical Association* 78, 856–864.
Phillips, P.C.B. (1987) Time series regression with a unit root. *Econometrica* 55, 277–301.
Phillips, P.C.B. & S. Durlauf (1986) Statistical inference in instrumental variables regression with $I(1)$ processes. *Review of Economic Studies* 57, 99–125.
Phillips, P.C.B. & S. Ouliaris (1990) Asymptotic properties of residual based tests for cointegration. *Econometrica* 58, 168–193.
Phillips, P.C.B. & P. Perron (1988) Testing for a unit root in time series regression. *Biometrika* 75, 335–346.
Rao, C.R. (1973) *Linear Statistical Inference and Its Applications*, 2nd ed. New York: Wiley.
Saikkonen, P. (1993) Estimation of cointegration vectors with linear restrictions. *Econometric Theory* 9, 19–35.
Shin, Y. (1994) A residual based test of the null of cointegration against the alternative of no cointegration. *Econometric Theory* 10, 91–115.
Snell, A. (1999) Testing for r versus $r-1$ cointegrating vectors. *Journal of Econometrics* 88, 151–192.
Stock, J.H. & M.W. Watson (1988) Testing for common trends. *Journal of the American Statistical Association* 83, 1097–1107.

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

APPENDIX A. LBI AND MPI TESTS IN MULTIVARIATE REGRESSION MODEL WITH RANDOM WALK PLUS NOISE DISTURBANCES

Here we prove that the test in (6) is LBI but as a corollary to a more general result. We consider the multivariate linear model

$$\mathbf{Y} = \mathbf{XM} + \mathbf{E}, \tag{A.1}$$

where $\mathbf{Y}$ is a $T \times N$ matrix of dependent variables, $\mathbf{X}$ is a $T \times p$ matrix of regressors, the $p \times N$ matrix $\mathbf{M}$ contains the regression coefficients, and $\mathbf{E}$ is a $T \times N$ matrix of regression errors. Naturally, we assume that $T - p > N$. Further, we assume that $\mathbf{E}$ has a special covariance structure: if $\mathbf{E}$ is vectorized columnwise its covariance matrix has the form $\mathbf{\Sigma}_\varepsilon \otimes \mathbf{G} + \mathbf{\Sigma}_\eta \otimes \mathbf{H}$, where $\mathbf{G}$ and $\mathbf{H}$ are arbitrary known positive definite $T \times T$ matrices. In terms of our model (3) the rows of $\mathbf{Y}$ consist of $\mathbf{y}'_1, \ldots, \mathbf{y}'_T$ and the columns of $\mathbf{Y}$ are the N time series of length T, $\mathbf{X}$ is a column of ones $\mathbf{i}$, $\mathbf{M} = \boldsymbol{\mu}'_0$, and $\mathbf{E}$ has rows $\boldsymbol{\varepsilon}'_t + \sum_{s=1}^{t} \boldsymbol{\eta}'_s$, $t = 1, \ldots, T$. Thus in (3) $\mathbf{G} = \mathbf{I}$ and $\mathbf{H} = [\min(s, t)]$. The hypotheses are as follows:

$$H_0 : \mathbf{\Sigma}_\eta = \mathbf{0}, \tag{A.2}$$

$$H_A : \mathbf{\Sigma}_\eta = q\mathbf{\Sigma}_\varepsilon, \qquad q > 0, \tag{A.3}$$

that is, we are testing against homogeneous alternatives. The pair of hypotheses H_0 and H_A is invariant under the group of affine linear transformations

$$\mathbf{Y} \mapsto \mathbf{YP} + \mathbf{XA}, \tag{A.4}$$

where $\mathbf{P}$ is an arbitrary nonsingular $N \times N$ matrix and $\mathbf{A}$ is an arbitrary $p \times N$ matrix. The next theorem gives the MPI and the LBI test for H_0 against H_A under the assumption that $\mathbf{E}$ follows a multinormal distribution. The best invariant test depends on the fixed $q > 0$, and the LBI test maximizes the derivative of the power function at $q = 0$. Our theorem generalizes the tests given by King and Hillier (1985) to a multivariate setting. A standard reference to the theory for invariant tests is Lehmann (1994, Ch. 6).

THEOREM A.1. *In model* (A.1) *when testing for* (A.2) *against* (A.3) *the best* (*most powerful*) *invariant test under* (A.4) *rejects when*

$$\frac{|\tilde{\mathbf{E}}'(\mathbf{G} + q\mathbf{H})^{-1}\tilde{\mathbf{E}}|}{|\hat{\mathbf{E}}'\mathbf{G}^{-1}\hat{\mathbf{E}}|} < c, \tag{A.5}$$

where $\tilde{\mathbf{E}}$ *and* $\hat{\mathbf{E}}$ *are the residual matrices under* H_A *and* H_0, *respectively*, *that is*,

$$\hat{\mathbf{E}} = \mathbf{Y} - \mathbf{X}[\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}]^{-1}\mathbf{X}'\mathbf{G}^{-1}\mathbf{Y}, \tag{A.6}$$

$$\tilde{\mathbf{E}} = \mathbf{Y} - \mathbf{X}[\mathbf{X}'(\mathbf{G} + q\mathbf{H})^{-1}\mathbf{X}]^{-1}\mathbf{X}'(\mathbf{G} + q\mathbf{H})^{-1}\mathbf{Y}. \tag{A.7}$$

Further, *the* LBI *test under* (A.4) *rejects when*

$$\mathrm{tr}(\hat{\mathbf{E}}'\mathbf{G}^{-1}\mathbf{H}\mathbf{G}^{-1}\hat{\mathbf{E}})(\hat{\mathbf{E}}'\mathbf{G}^{-1}\hat{\mathbf{E}})^{-1} > c. \tag{A.8}$$

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

Remark 1. Note that the homogeneity allows the regressions to be run separately on each variable (see Harvey, 1989, Sect. 8.3), allowing the formulas (A.6) and (A.7).

Remark 2. In model (3) (A.8) reduces to (6) because then $\mathbf{G} = \mathbf{I}$, $\mathbf{H} = [\min(s,t)]$, and $\hat{\mathbf{E}}$ consists of the rows $\mathbf{y}_t' - \bar{\mathbf{y}}'$.

Proof. When applying the theory of invariant testing to the hypotheses (A.2) and (A.3) the first task is to find the maximal invariant under (A.4). Because it is generated by two subgroups of transformations, namely, translations and postmultiplications by nonsingular matrices, we can proceed in two steps. First, we find a maximal invariant under translations $\mathbf{Y} \mapsto \mathbf{Y} + \mathbf{XA}$, say, $s(\mathbf{Y})$, such that

$$\text{if } s(\mathbf{Y}_1) = s(\mathbf{Y}_2) \quad \text{then also} \quad s(\mathbf{Y}_1\mathbf{P}) = s(\mathbf{Y}_2\mathbf{P}) \quad \text{for all nonsingular } \mathbf{P}. \tag{A.9}$$

Second, we find a maximal invariant in the space of $s(\mathbf{Y})$ under the group of transformations $s(\mathbf{Y}) \mapsto s(\mathbf{YP})$ with $\mathbf{P}$ being nonsingular. By Theorem 2 of Lehmann (1994, p. 288) the latter maximal invariant is also a maximal invariant under (A.4).

Translations induce a maximal invariant $\mathbf{Z}'\mathbf{Y}$ with $\mathbf{Z}$ being of a maximum rank such that $\mathbf{Z}'\mathbf{X} = \mathbf{0}$. The matrix $\mathbf{Z}$ is not unique, but any such $\mathbf{Z}$ will do. Fix any suitable $\mathbf{Z}$. Then another choice is obtained by postmultiplying $\mathbf{Z}$ by a nonsingular square matrix $\mathbf{F}$. Now, we choose $\mathbf{F}$ such that (see Rao, 1973, p. 41) $\mathbf{F}'(\mathbf{Z}'\mathbf{GZ})\mathbf{F} = \mathbf{I}$ and $\mathbf{F}'(\mathbf{Z}'\mathbf{HZ})\mathbf{F} = \mathbf{D} = \text{diag}[d_1,\ldots,d_{T-p}]$ and denote, for brevity, this $\mathbf{ZF}$ simply by $\mathbf{Z}$. Then

$$\mathbf{Z}'\mathbf{GZ} = \mathbf{I} \quad \text{and} \quad \mathbf{Z}'\mathbf{HZ} = \mathbf{D}. \tag{A.10}$$

A maximal invariant is $\mathbf{Z}'\mathbf{Y} = \mathbf{U}$. Clearly (A.9) is satisfied.

Denote the rows of $\mathbf{U}$ as $\mathbf{u}_1',\ldots,\mathbf{u}_{T-p}'$ and let $\mathbf{U}_0'$ be the submatrix of the first N rows of $\mathbf{U}$. The nonsingular linear transformations applied to the rows of $\mathbf{U}$, that is, $\mathbf{U} \mapsto \mathbf{UP}$, lead to a set of maximal invariant vectors

$$\mathbf{w}_t = \mathbf{U}_0^{-1}\mathbf{u}_t, \qquad t = N+1,\ldots,T-p, \tag{A.11}$$

where the inverse $\mathbf{U}_0^{-1}$ exists with probability one because $T - p > N$. The previously mentioned Theorem 2 of Lehmann now guarantees that the set (A.11) is maximal invariant also under (A.4).

From (A.10) we find that the rows of $\mathbf{U}$ are independent and $\text{cov}(\mathbf{u}_t) = (1 + qd_t)\boldsymbol{\Sigma}_\varepsilon$. Because the distribution of the maximal invariant (A.11) is independent of $\boldsymbol{\Sigma}_\varepsilon$, we can without loss of generality put $\boldsymbol{\Sigma}_\varepsilon = \mathbf{I}$. Then, conditionally on $\mathbf{U}_0$, $\mathbf{w}_{N+1},\ldots,\mathbf{w}_{T-p}$ are independently distributed with $\mathbf{w}_t \sim N(\mathbf{0},(1 + qd_t)(\mathbf{U}_0'\mathbf{U}_0)^{-1})$. Thus their joint marginal density is proportional to

$$E\{|\mathbf{U}_0'\mathbf{U}_0|^{1/2(T-N-p)} e^{-\sum_{t=N+1}^{T-p}\mathbf{w}_t'\mathbf{U}_0'\mathbf{U}_0\mathbf{w}_t/2(1+qd_t)}\}, \tag{A.12}$$

where the expectation is over $\mathbf{U}_0$, other quantities being fixed. Denote

$$\mathbf{D}_0 = \text{diag}[d_1,\ldots,d_N],$$

$$\mathbf{T} = (\mathbf{I} + q\mathbf{D}_0)^{-1/2}\mathbf{U}_0'\mathbf{U}_0(\mathbf{I} + q\mathbf{D}_0)^{-1/2}, \tag{A.13}$$

$$\mathbf{W} = (\mathbf{I} + q\mathbf{D}_0)^{1/2}\left(\sum_{t=N+1}^{T-p}(1 + qd_t)^{-1}\mathbf{w}_t\mathbf{w}_t'\right)(\mathbf{I} + q\mathbf{D}_0)^{1/2}. \tag{A.14}$$

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

With these notations (A.12) becomes proportional to

$$E[|\mathbf{T}|^{1/2(T-N-p)}e^{-1/2\mathrm{tr}\mathbf{TW}}], \tag{A.15}$$

where $\mathbf{W}$ is fixed and the expectation is over $\mathbf{T}$. Because the columns of $\mathbf{U}_0(\mathbf{I}+q\mathbf{D}_0)^{-1/2}$ are NID$(\mathbf{0},\mathbf{I})$, $\mathbf{T}$ defined by (A.13) follows a standard Wishart distribution with N degrees of freedom. Using this and invoking the general Wishart density (Anderson 1984, p. 249) the expectation in (A.15) is proportional to

$$\int |\mathbf{T}|^{1/2(T-N-p-1)}e^{-1/2\,\mathrm{tr}\,\mathbf{T}(\mathbf{I}+\mathbf{W})}\,d\mathbf{T} \propto |\mathbf{I}+\mathbf{W}|^{-1/2(T-p)}, \tag{A.16}$$

where the integral is taken over symmetric positive definite matrices. Substituting $\mathbf{W}$ from (A.14) to (A.16) we obtain after some algebra that

$$|\mathbf{I}+\mathbf{W}| = \frac{|\mathbf{U}'(\mathbf{I}+q\mathbf{D})^{-1}\mathbf{U}|}{|\mathbf{U}_0'(\mathbf{I}+q\mathbf{D}_0)^{-1}\mathbf{U}_0|} = \frac{|\mathbf{I}+q\mathbf{D}_0|}{|\mathbf{U}_0|^2}\,|\mathbf{U}'(\mathbf{I}+q\mathbf{D})^{-1}\mathbf{U}|. \tag{A.17}$$

The most powerful test rejects when the ratio of the densities (A.16) at q and at $q=0$ is sufficiently large. The formula (A.17) with $\mathbf{U}=\mathbf{Z}'\mathbf{Y}$ and (A.10) implies that the test has the rejection region

$$\frac{|\mathbf{U}'(\mathbf{I}+q\mathbf{D})^{-1}\mathbf{U}|}{|\mathbf{U}'\mathbf{U}|} = \frac{|\mathbf{Y}'\mathbf{Z}(\mathbf{Z}'\mathbf{G}\mathbf{Z}+q\mathbf{Z}'\mathbf{H}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}|}{|\mathbf{Y}'\mathbf{Z}(\mathbf{Z}'\mathbf{G}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}|} < c. \tag{A.18}$$

A similar treatment of the log derivative of (A.16) at $q=0$ leads to the LBI test

$$\mathrm{tr}[\mathbf{Y}'\mathbf{Z}(\mathbf{Z}'\mathbf{G}\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{H}\mathbf{Z})(\mathbf{Z}'\mathbf{G}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}][\mathbf{Y}'\mathbf{Z}(\mathbf{Z}'\mathbf{G}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}]^{-1} > c. \tag{A.19}$$

The formulas (A.5) and (A.8) involving residual matrices are found by noting that $\mathbf{Z}'(\mathbf{G}+q\mathbf{H})(\mathbf{G}+q\mathbf{H})^{-1}\mathbf{X}=\mathbf{0}$ and that $\mathbf{Z}'(\mathbf{G}+q\mathbf{H})$ is of maximum rank. Hence the residuals under H_A are obtained by regressing $\mathbf{Y}$ on the columns of $(\mathbf{G}+q\mathbf{H})\mathbf{Z}$ using the generalized least squares method with the covariance matrix $\mathbf{G}+q\mathbf{H}$. This regression can be run separately on each column of $\mathbf{Y}$ because of homogeneity (cf. Remark 1). A simple matrix calculation shows then that

$$\tilde{\mathbf{E}} = (\mathbf{G}+q\mathbf{H})\mathbf{Z}[\mathbf{Z}'(\mathbf{G}+q\mathbf{H})\mathbf{Z}]^{-1}\mathbf{Z}'\mathbf{Y}.$$

The null case is obtained with $q=0$, so

$$\hat{\mathbf{E}} = \mathbf{G}\mathbf{Z}(\mathbf{Z}'\mathbf{G}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}.$$

Although these expressions for $\hat{\mathbf{E}}$ and $\tilde{\mathbf{E}}$ look different from (A.7) and (A.6) they are, in fact, numerically the same (see Rao, 1973, problem 33, p. 77). The proof is completed by inserting these new formulas of $\hat{\mathbf{E}}$ and $\tilde{\mathbf{E}}$ into (A.5) and (A.8)and comparing with (A.18) and (A.19). ■

APPENDIX B. ASYMPTOTIC DISTRIBUTIONS

Although the tests are derived under multinormality the limiting distributions do not rely on this assumption. In the following we first assume that the errors in (3) are all mutually

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

independent (Theorem B.1) with finite covariance matrices $\mathbf{\Sigma}_\varepsilon$ and $\mathbf{\Sigma}_\eta$. Then we extend our results to allow for the serial dependence of the $\boldsymbol{\varepsilon}_t$'s (Theorem B.4).

A standard vector valued Brownian motion with N independent components is denoted by $\mathbf{W}$. Thus its covariance matrix function is $\operatorname{cov}(\mathbf{W}(s), \mathbf{W}(u)) = \min(s,u)\mathbf{I}$, $0 \le s, u \le 1$. Let us partition it into $K^\dagger$ and $N - K^\dagger$ components as $\mathbf{W}(s)' = (\mathbf{W}_1(s)', \mathbf{W}_2(s)')$. Define the matrices

$$\mathbf{S}_{11}^* = \int_0^1 \mathbf{W}^*(u)\mathbf{W}^*(u)'\,du,$$

$$\mathbf{C}_{11}^* = \int_0^1 \left[\int_0^u \mathbf{W}^*(s)\,ds\right]\left[\int_0^u \mathbf{W}^*(s)\,ds\right]'\,du,$$

$$\mathbf{C}_{12}^* = \int_0^1 \left[\int_0^u \mathbf{W}^*(s)\,ds\right]\mathbf{B}(u)'\,du = \mathbf{C}_{21}^{*\prime},$$

$$\mathbf{C}_{22}^* = \int_0^1 \mathbf{B}(u)\mathbf{B}(u)'\,du,$$

where

$$\mathbf{W}^*(u) = \mathbf{W}_1(u) - \int_0^1 \mathbf{W}_1(s)\,ds, \qquad 0 \le u \le 1$$

is the so-called demeaned Brownian motion and $\mathbf{B}(u) = \mathbf{W}_2(u) - u\mathbf{W}_2(1)$, $0 \le u \le 1$, a standard Brownian bridge.

We use the shorthand notation for the nonincreasingly ordered eigenvalues

$$\ell_j = \lambda_j(\mathbf{S}^{-1}\mathbf{C}), \qquad j = 1,\ldots,N,$$

$$\ell_j^* = \lambda_j(\mathbf{S}_{11}^{*-1}\mathbf{C}_{11}^*), \qquad j = 1,\ldots,K^\dagger,$$

$$\ell_{K^\dagger+j}^{**} = \lambda_j(\mathbf{C}_{22}^* - \mathbf{C}_{21}^*\mathbf{C}_{11}^{*-1}\mathbf{C}_{12}^*) \qquad j = 1,\ldots,N-K^\dagger.$$

We allow also the cases $K^\dagger = 0$ and $K^\dagger = N$. Then one or the other set of eigenvalues is empty. If $K^\dagger = 0$, the eigenvalues ℓ_j^{**}, $j = 1,\ldots,N$ are those of $\mathbf{C}_{22}^*$, which is then $N \times N$. On the other hand, if $K^\dagger = N$, then $\mathbf{S}_{11}^*$ and $\mathbf{C}_{11}^*$ are $N \times N$.

In the following it is understood that all limits are taken as $T \to \infty$, although the dependence on T is usually omitted. The main tools in the derivations are the functional central limit theorem and the continuous mapping theorem (CMT). A basic reference is Billingsley (1968). Phillips and Durlauf (1986) derived results for vector processes that are particularly relevant in our case.

THEOREM B.1. *Assume that the errors* $\{\boldsymbol{\varepsilon}_t\}$ *and* $\{\boldsymbol{\eta}_t\}$ *in* (3) *are i.i.d. and mutually independent with finite covariance matrices* $\mathbf{\Sigma}_\varepsilon$ *and* $\mathbf{\Sigma}_\eta$, *respectively. Then*

$$(T^{-1}\ell_1,\ldots,T^{-1}\ell_{K^\dagger},\ell_{K^\dagger+1},\ldots,\ell_N) \xrightarrow{d} (\ell_1^*,\ldots,\ell_{K^\dagger}^*,\ell_{K^\dagger+1}^{**},\ldots,\ell_N^{**}). \tag{B.1}$$

Proof. Because the eigenvalues of $\mathbf{S}^{-1}\mathbf{C}$ are invariant under nonsingular linear transformations we can assume without loss of generality that $\mathbf{\Sigma}_\varepsilon = \mathbf{I}$ and $\mathbf{\Sigma}_\eta = \operatorname{diag}[q_1,\ldots,q_N]$ with $q_1 \ge \cdots \ge q_N \ge 0$ (see the discussion around (5)). Then the assumption $\operatorname{rank}(\mathbf{\Sigma}_\eta) = K^\dagger$ implies that $q_{K^\dagger} > 0$ and $q_{K^\dagger+1} = \cdots = q_N = 0$.

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

Let us partition $\mathbf{y}_t'$ and $\boldsymbol{\varepsilon}_t'$ into $K^\dagger$ and $N - K^\dagger$ components as $(\mathbf{y}_{1t}', \mathbf{y}_{2t}')$ and $(\boldsymbol{\varepsilon}_{1t}', \boldsymbol{\varepsilon}_{2t}')$. Because only $\mathbf{y}_{1t}$ contains the random walk term, we can write

$$\begin{bmatrix} \mathbf{y}_{1t} - \bar{\mathbf{y}}_1 \\ \mathbf{y}_{2t} - \bar{\mathbf{y}}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_t - \bar{\boldsymbol{\mu}} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_{1t} - \bar{\boldsymbol{\varepsilon}}_1 \\ \boldsymbol{\varepsilon}_{2t} - \bar{\boldsymbol{\varepsilon}}_2 \end{bmatrix}. \tag{B.2}$$

Set $\boldsymbol{\mu}_0 = 0$ without loss of generality. As a random walk of i.i.d. errors, $T^{-1/2}\boldsymbol{\mu}_{[Tu]}$ converges weakly to a $K^\dagger$ dimensional Brownian motion $\mathbf{Q}_{K^\dagger}^{1/2}\mathbf{W}_1(u)$ with $\mathbf{Q}_{K^\dagger}^{1/2} = \mathrm{diag}[\sqrt{q_1}, \ldots, \sqrt{q_{K^\dagger}}]$. Then CMT and weak convergence of partial sums of $\boldsymbol{\varepsilon}_{2t}$'s give

$$\begin{bmatrix} T^{-3/2}\sum_{t=1}^{[Tu]}(\mathbf{y}_{1t} - \bar{\mathbf{y}}_1) \\ T^{-1/2}\sum_{t=1}^{[Tu]}(\mathbf{y}_{2t} - \bar{\mathbf{y}}_2) \end{bmatrix} = \begin{bmatrix} T^{-3/2}\sum_{t=1}^{[Tu]}(\boldsymbol{\mu}_t - \bar{\boldsymbol{\mu}}) \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} T^{-3/2}\sum_{t=1}^{[Tu]}(\boldsymbol{\varepsilon}_{1t} - \bar{\boldsymbol{\varepsilon}}_1) \\ T^{-1/2}\sum_{t=1}^{[Tu]}(\boldsymbol{\varepsilon}_{2t} - \bar{\boldsymbol{\varepsilon}}_2) \end{bmatrix}$$

$$\xrightarrow{d} \begin{bmatrix} \mathbf{Q}_{K^\dagger}^{1/2}\int_0^u \mathbf{W}^*(s)\,ds \\ \mathbf{B}(u) \end{bmatrix}, \qquad 0 \le u \le 1. \tag{B.3}$$

The matrices

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{12}' & \mathbf{S}_{22} \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{bmatrix}, \tag{B.4}$$

and

$$\mathbf{D} = \begin{bmatrix} T^{-1/2}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

are partitioned conformably with $(\mathbf{y}_{1t}', \mathbf{y}_{2t}')$.

It will be proved later that

$$\mathbf{S}_{12} = \mathbf{O}_p(1). \tag{B.5}$$

From (B.3) and (B.5) we easily obtain that

$$\mathbf{DSD} = \begin{bmatrix} T^{-1}\mathbf{S}_{11} & T^{-1/2}\mathbf{S}_{12} \\ T^{-1/2}\mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \mathbf{Q}_{K^\dagger}^{1/2}\mathbf{S}_{11}^*\mathbf{Q}_{K^\dagger}^{1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \tag{B.6}$$

$$T^{-1}\mathbf{DCD} = \begin{bmatrix} T^{-2}\mathbf{C}_{11} & T^{-3/2}\mathbf{C}_{12} \\ T^{-3/2}\mathbf{C}_{21} & T^{-1}\mathbf{C}_{22} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \mathbf{Q}_{K^\dagger}^{1/2}\mathbf{C}_{11}^*\mathbf{Q}_{K^\dagger}^{1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{B.7}$$

$$\mathbf{D}^2\mathbf{SD}^2 = \begin{bmatrix} T^{-2}\mathbf{S}_{11} & T^{-1}\mathbf{S}_{12} \\ T^{-1}\mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} \xrightarrow{P} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \tag{B.8}$$

$$\mathbf{D}^2\mathbf{CD}^2 = \begin{bmatrix} T^{-2}\mathbf{C}_{11} & T^{-1}\mathbf{C}_{12} \\ T^{-1}\mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \mathbf{Q}_{K^\dagger}^{1/2}\mathbf{C}_{11}^*\mathbf{Q}_{K^\dagger}^{1/2} & \mathbf{Q}_{K^\dagger}^{1/2}\mathbf{C}_{12}^* \\ \mathbf{C}_{21}^*\mathbf{Q}_{K^\dagger}^{1/2} & \mathbf{C}_{22}^* \end{bmatrix}. \tag{B.9}$$

These convergences are jointly valid.

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

Note that each ℓ_j/T is an eigenvalue of

$$T^{-1}(\mathbf{DSD})^{-1}(\mathbf{DCD}) \xrightarrow{d} \begin{bmatrix} \mathbf{S}_{11}^{*-1}\mathbf{C}_{11}^{*} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

and each $1/\ell_j$ is an eigenvalue of

$$(\mathbf{D}^2\mathbf{C}\mathbf{D}^2)^{-1}(\mathbf{D}^2\mathbf{S}\mathbf{D}^2) \xrightarrow{d} \begin{bmatrix} \mathbf{Q}_{K^\dagger}^{1/2}\mathbf{C}_{11}^{*}\mathbf{Q}_{K^\dagger}^{1/2} & \mathbf{Q}_{K^\dagger}^{1/2}\mathbf{C}_{12}^{*} \\ \mathbf{C}_{21}^{*}\mathbf{Q}_{K^\dagger}^{1/2} & \mathbf{C}_{22}^{*} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0} & -\mathbf{F}\mathbf{E}^{-1} \\ \mathbf{0} & \mathbf{E}^{-1} \end{bmatrix},$$

where $\mathbf{E} = \mathbf{C}_{22}^{*} - \mathbf{C}_{21}^{*}\mathbf{C}_{11}^{*-1}\mathbf{C}_{12}^{*}$ and $\mathbf{F} = \mathbf{Q}_{K^\dagger}^{-1/2}\mathbf{C}_{11}^{*-1}\mathbf{C}_{12}^{*}$ using the formula for the inverse of a partitioned matrix (Rao, 1973, p. 33). Because the ordered eigenvalues depend continuously on their matrix argument (e.g., see Bhatia, 1997, pp. 154, 227–228),

$$\begin{bmatrix} \ell_1/T & \cdots & \ell_{K^\dagger}/T & \ell_{K^\dagger+1}/T & \cdots & \ell_N/T \\ 1/\ell_1 & \cdots & 1/\ell_{K^\dagger} & 1/\ell_{K^\dagger+1} & \cdots & 1/\ell_N \end{bmatrix}$$

$$\xrightarrow{d} \begin{bmatrix} \ell_1^{*} & \cdots & \ell_{K^\dagger}^{*} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1/\ell_{K^\dagger+1}^{**} & \cdots & 1/\ell_N^{**} \end{bmatrix}.$$

The claim (B.1) now follows.

The proof is completed by showing (B.5). Take arbitrary vectors $\mathbf{a}$, $K^\dagger \times 1$ and $\mathbf{b}$, $(N - K^\dagger) \times 1$. Then

$$|\mathbf{a}'\mathbf{S}_{12}\mathbf{b}| = T^{-1}\left|\sum_{t=1}^{T} \mathbf{a}'(\mathbf{y}_{1t} - \bar{\mathbf{y}}_1)(\mathbf{y}_{2t} - \bar{\mathbf{y}}_2)'\mathbf{b}\right|$$

$$\leq T^{-1/2}\max|\mathbf{a}'(\mathbf{y}_{1t} - \bar{\mathbf{y}}_1)|\left|T^{-1/2}\sum_{t=1}^{T}(\mathbf{y}_{2t} - \bar{\mathbf{y}}_2)'\mathbf{b}\right|. \qquad \textbf{(B.10)}$$

The factors in (B.10) are clearly both $O_p(1)$, proving (B.5). ■

The following corollary is immediately deduced from Theorem B.1 and gives the limiting distribution of ξ_N both under the null and under the alternative hypothesis.

COROLLARY B.2. *Let* $\operatorname{rank}(\mathbf{\Sigma}_\eta) = K^\dagger$.

1. *If* $K^\dagger = 0$, *then*

$$\xi_N = \operatorname{tr}(\mathbf{S}^{-1}\mathbf{C}) \xrightarrow{d} \int_0^1 \mathbf{B}(u)'\mathbf{B}(u)\,du.$$

2. *If* $K^\dagger \geq 1$, *then*

$$T^{-1}\xi_N = T^{-1}\operatorname{tr}(\mathbf{S}^{-1}\mathbf{C}) \xrightarrow{d} \sum_{j=1}^{K^\dagger} \ell_j^{*} = \operatorname{tr}(\mathbf{S}_{11}^{*-1}\mathbf{C}_{11}^{*}).$$

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

The corresponding result for the common trends test,

$$\zeta_{K,N} = \sum_{j=K+1}^{N} \ell_j,$$

is given in the next corollary.

COROLLARY B.3. *Let* $\operatorname{rank}(\boldsymbol{\Sigma}_\eta) = K^\dagger$.

1. *If* $K^\dagger < K$, *then* $\zeta_{K,N} \xrightarrow{d} \sum_{j=K+1}^{N} \ell_j^{**}$
2. *If* $K^\dagger = K$, *then* $\zeta_{K,N} \xrightarrow{d} \sum_{j=K^\dagger+1}^{N} \ell_j^{**} = \operatorname{tr}(\mathbf{C}_{22}^* - \mathbf{C}_{21}^* \mathbf{C}_{11}^{*-1} \mathbf{C}_{12}^*)$.
3. *If* $K^\dagger > K$, *then* $T^{-1}\zeta_{K,N} \xrightarrow{d} \sum_{j=K+1}^{K^\dagger} \ell_j^*$.

Theorem B.1 can be generalized to allow a large class of weakly stationary errors $\boldsymbol{\varepsilon}_t$ including finite order vector valued ARMA processes. We follow Phillips and Perron (1988) and base our regularity conditions on the results of Herrndorf (1984, Corollary 1). We assume that the process $\{\boldsymbol{\varepsilon}_t\}$ satisfies the following regularity conditions:

(i) $E(\varepsilon_{it}) = 0, \qquad i = 1,\ldots,N; \qquad t = 1,2,\ldots,$
(ii) $\sup_t E|\varepsilon_{it}|^{2\beta} < \infty, \qquad i = 1,\ldots,N, \qquad \beta > 2,$
(iii) the vector process $\{\boldsymbol{\varepsilon}_t\}$ is strong mixing with mixing coefficients α_m that satisfy $\sum_{m=1}^{\infty} \alpha_m^{1-2/\beta} < \infty$ (for the definition of strong mixing coefficients, see Hall and Heyde, 1980, p. 132),
(iv) $\lim_{T\to\infty} T^{-1} E[(\sum_{t=1}^{T} \boldsymbol{\varepsilon}_t)(\sum_{t=1}^{T} \boldsymbol{\varepsilon}_t')] = \boldsymbol{\Omega}_\varepsilon$ exists and is positive definite.

The condition (ii) allows some degree of nonstationarity. In the stationary case there is no need to take the supremum. The conditions (i)–(iv) guarantee that the partial sums of any linear combination $\mathbf{a}'\boldsymbol{\varepsilon}_t$ divided by $\sqrt{T}$ converge to a Brownian motion with a variance $\mathbf{a}'\boldsymbol{\Omega}_\varepsilon\mathbf{a}$. This yields weak convergence for the partial sums $\sum_{t=1}^{[Tu]} \boldsymbol{\varepsilon}_t$. The moment condition (ii) is too strict to prove merely this convergence (cf. Phillips and Perron, 1988, condition (ii), p. 336). The existence of the moments slightly higher than the fourth in (ii) is for the consistent estimation of $\boldsymbol{\Omega}_\varepsilon$.

In the stationary case we can write

$$\boldsymbol{\Omega}_\varepsilon = \sum_{\tau=-\infty}^{\infty} \boldsymbol{\Gamma}(\tau), \tag{B.11}$$

where $\boldsymbol{\Gamma}(\tau) = E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_{t-\tau}')$, $\tau = 0,1,\ldots$ and $\boldsymbol{\Gamma}(-\tau) = \boldsymbol{\Gamma}(\tau)'$. These autocovariance matrices can be estimated by their sample analogues, $\hat{\boldsymbol{\Gamma}}(\tau)$, as defined in (17). Define now a truncated estimate of $\boldsymbol{\Omega}_\varepsilon$ as

$$\mathbf{S}(m) = \sum_{\tau=-m}^{m} \hat{\boldsymbol{\Gamma}}(\tau).$$

Then, in the absence of the random walk component, Theorem 4.2 of Phillips (1987) by the condition (ii) implies that, if

(v) $m = m(T) = o(T^{1/4})$,

$\mathbf{a}'\mathbf{S}(m)\mathbf{a} \to \mathbf{a}'\boldsymbol{\Omega}_\varepsilon\mathbf{a}$ in probability for all $\mathbf{a}$. From this we can deduce that

$$\mathbf{S}(m) \xrightarrow{P} \boldsymbol{\Omega}_\varepsilon. \tag{B.12}$$

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

Note that $\mathbf{S}(m)$ is still consistent for $\mathbf{\Omega}_\varepsilon$ in the presence of mild nonstationarity controlled by (ii) even though the representation (B.11) no longer exists.

Denote now the ordered eigenvalues of $\mathbf{S}(m)^{-1}\mathbf{C}$ by $\ell_j(m), j = 1,\ldots,N$. The definitions of the ℓ_j^*'s and ℓ_k^{**}'s remain unchanged.

THEOREM B.4. *Assume that the errors* $\{\boldsymbol{\varepsilon}_t\}$ *in* (3) *satisfy the conditions* (i)–(iv) *and* m *satisfies* (v). *Assume also that the process* $\{\boldsymbol{\eta}_t\}$ *is independent of* $\{\boldsymbol{\varepsilon}_t\}$ *and is such that its partial sums* $T^{-1/2}\sum_{t=1}^{[Tu]}\boldsymbol{\eta}_t$ *converge weakly to a Brownian motion with covariance matrix* $\mathbf{\Omega}_\eta$. *Then*

$$[(m/T)\ell_1(m),\ldots,(m/T)\ell_{K^\dagger}(m),\ell_{K^\dagger+1}(m),\ldots,\ell_N(m)] \xrightarrow{d} [\ell_1^*,\ldots,\ell_{K^\dagger}^*,\ell_{K^\dagger+1}^{**},\ldots,\ell_N^{**}].$$

Proof. Again by the invariance argument it is clear that we can assume $\mathbf{\Omega}_\varepsilon = \mathbf{I}$ and $\mathbf{\Omega}_\eta = \operatorname{diag}[q_1,\ldots,q_{K^\dagger},0,\ldots,0]$. Recall the partition (B.2) and denote the conformable partition of $\mathbf{S}(m)$ as $\mathbf{S}_{ij}(m)$, $i,j = 1,2$. Our assumptions and regularity conditions guarantee that (B.3) holds also now. Hence the asymptotic behavior of $\mathbf{C}$ is as established in the proof of Theorem B.1. Further, the limit (B.12) now can be applied to $\mathbf{S}_{22}(m)$. In contrast to these facts the matrices $\mathbf{S}_{11}(m)$ and $\mathbf{S}_{12}(m)$ require a more careful treatment.

Let us first consider $\mathbf{S}_{11}(m)$. The assumption (v) implies

$$(mT)^{-1}\sum_{\tau=0}^{m}\sum_{t=\tau+1}^{T}(\mathbf{h}[t/T]-\bar{\mathbf{h}}_T)(\mathbf{h}[(t-\tau)/T]-\bar{\mathbf{h}}_T)' \to \int_0^1 [\mathbf{h}(u)-\bar{\mathbf{h}}][\mathbf{h}(u)-\bar{\mathbf{h}}]'\,du$$

for any continuous function $\mathbf{h}$ on $[0,1]$ with

$$\bar{\mathbf{h}}_T = T^{-1}\sum \mathbf{h}(t/T) \quad \text{and} \quad \bar{\mathbf{h}} = \int_0^1 \mathbf{h}(u)\,du.$$

This is sufficient for

$$(mT)^{-1}\mathbf{S}_{11}(m) \xrightarrow{d} \mathbf{Q}_{K^\dagger}^{1/2}\mathbf{S}_{11}^*\mathbf{Q}_{K^\dagger}^{1/2}$$

by Theorem 5.5 of Billingsley (1968, p. 34).

For $\mathbf{S}_{12}(m)$ we find (cf. (B.10)) that

$$|\mathbf{a}'\mathbf{S}_{12}(m)\mathbf{b}| = T^{-1}\left|\sum_{k=0}^{m}\sum_{t=\tau+1}^{T}\mathbf{a}'(\mathbf{y}_{1t}-\bar{\mathbf{y}}_1)(\mathbf{y}_{2,t-\tau}-\bar{\mathbf{y}}_2)'\mathbf{b}\right|$$

$$\le T^{-1/2}\max|\mathbf{a}'(\mathbf{y}_{1t}-\bar{\mathbf{y}}_1)|\left|T^{-1/2}\sum_{k=0}^{m}\sum_{t=\tau+1}^{T}(\mathbf{y}_{2,t-\tau}-\bar{\mathbf{y}}_2)'\mathbf{b}\right|,$$

where the first factor is $O_p(1)$ and the second $O_p(m)$. Hence $\mathbf{S}_{12}(m)$ is $\mathbf{O}_p(m)$.

After these preparations we find that the limits (B.8) and (B.9) continue to hold. Defining

$$\mathbf{D}(m) = \begin{bmatrix}(mT)^{-1/2}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}$$

it is easily deduced that $\mathbf{D}(m)\mathbf{S}\mathbf{D}(m)$ and $(m/T)\mathbf{D}(m)\mathbf{C}\mathbf{D}(m)$ have the weak limits (B.6) and (B.7). The rest of the proof is like that of Theorem B.1. ■

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

Hence, under the assumptions (i)–(v) we can state the following corollaries.

COROLLARY B.5. *Let* $\operatorname{rank}(\mathbf{\Omega}_\eta) = K^\dagger$.

1. *If* $K^\dagger = 0$, *then*

$$\xi_N(m) = \operatorname{tr}(\mathbf{S}(m)^{-1}\mathbf{C}) \xrightarrow{d} \int_0^1 \mathbf{B}(u)'\mathbf{B}(u)\,du.$$

2. *If* $K^\dagger \geq 1$, *then*

$$(m/T)\xi_N(m) = \operatorname{tr}(\mathbf{S}(m)^{-1}\mathbf{C}) \xrightarrow{d} \operatorname{tr}(\mathbf{S}_{11}^{*-1}\mathbf{C}_{11}^*).$$

The modified test for a specified number of common trends is defined as

$$\zeta_{K,N}(m) = \sum_{j=K+1}^{N} \ell_j(m).$$

Its asymptotic distributions depending on the true rank of $\mathbf{\Sigma}_\eta$ are covered by the following corollary.

COROLLARY B.6. *Let* $\operatorname{rank}(\mathbf{\Sigma}_\eta) = K^\dagger$.

1. *If* $K^\dagger < K$, *then* $\zeta_{K,N}(m) \xrightarrow{d} \sum_{j=K+1}^{N} \ell_j^{**}$.
2. *If* $K^\dagger = K$, *then* $\zeta_{K,N}(m) \xrightarrow{d} \sum_{j=K^\dagger+1}^{N} \ell_j^{**} = \operatorname{tr}(\mathbf{C}_{22}^* - \mathbf{C}_{21}^*\mathbf{C}_{11}^{*-1}\mathbf{C}_{12}^*)$.
3. *If* $K^\dagger > K$, *then* $(m/T)\zeta_{K,N}(m) \xrightarrow{d} \sum_{j=K+1}^{K^\dagger} \ell_j^*$.

In practice when estimating $\mathbf{\Omega}_\varepsilon$, we may want to use some weighting system as in (16). If the weighting system is such that

$$w_{\tau m} = w(\tau/m) \quad \text{and} \quad \int_{-1}^{1} w(u)\,du = 1,$$

then the proof of Theorem B.4 holds with minor modifications. Although asymptotically $m = o(T^{1/4})$ there is the practical problem of choosing the lag window m in finite samples.

APPENDIX C. SERIES EXPANSIONS

Here we derive the series expansions (8), (13), and (27). They give convenient expressions for simulation, which is our tool for numerical calculation. The series expansions are established easily by assuming multinormality in (3). The invariance principle then implies that these expansions are valid under any distributional assumptions with finite second moments. The formula (3.5) of Nyblom and Mäkeläinen (1983) is easily generalized to the multivariate case: the (i,j) element of $\mathbf{S}$ may be characterized as

$$T^{-1}\sum_{t=1}^{T}(y_{it}-\bar{y}_i)(y_{jt}-\bar{y}_j) \stackrel{d}{=} T^{-1}\sum_{k=1}^{T-1}(1+q_i\lambda_{kT})^{1/2}(1+q_j\lambda_{kT})^{1/2}u_{ik}u_{jk} \qquad \textbf{(C.1)}$$

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press

and the corresponding element of $\mathbf{C}$ as

$$T^{-2}\sum_{t=1}^{T}\left(\sum_{s=1}^{t}(y_{is}-\bar{y}_i)\right)\left(\sum_{s=1}^{t}(y_{js}-\bar{y}_j)\right)$$
$$\stackrel{d}{=} T^{-2}\sum_{k=1}^{T-1}\lambda_{kT}(1+q_i\lambda_{kT})^{1/2}(1+q_j\lambda_{kT})^{1/2}u_{ik}u_{jk}. \qquad \textbf{(C.2)}$$

The notation $\stackrel{d}{=}$ means "equal in distribution." All u_{ik}, $i=1,\ldots,N$, $k=1,\ldots,T-1$ are mutually independent standard normal variables, and $\lambda_{kT}=1/[2(1-\cos(\pi k/T))]$, $k=1,\ldots,T-1$. Now $T^{-2}\lambda_{kT}\to(\pi k)^{-2}$ uniformly in $k=1,2,\ldots$. From (C.1) and (C.2) the following results are easily obtained:

(i) If $q_1=\cdots=q_N=0$, then

$$\mathbf{C}\xrightarrow{d}\sum_{k=1}^{\infty}(\pi k)^{-2}\mathbf{u}_k\mathbf{u}_k',$$

where $\mathbf{u}_k\sim \mathrm{NID}(\mathbf{0},\mathbf{I})$. Because $\mathbf{S}$ tends to $\mathbf{I}$ in probability, we get

$$\operatorname{tr}\mathbf{S}^{-1}\mathbf{C}\xrightarrow{d}\sum_{k=1}^{\infty}(\pi k)^{-2}\mathbf{u}_k'\mathbf{u}_k.$$

Note that $\mathbf{u}_k'\mathbf{u}_k\sim\chi^2(N)$. This proves (8).

(ii) If $q_j>0$, for $j=1,\ldots,K^\dagger$, and $q_j=0$, for $j=K^\dagger+1,\ldots,N$, and the partitions are as in (B.4), then with $\mathbf{Q}_{K^\dagger}^{1/2}=\operatorname{diag}[\sqrt{q_1},\ldots,\sqrt{q_K}]$ we now get

$$T^{-2}\mathbf{C}_{11}\xrightarrow{d}\mathbf{Q}_{K^\dagger}^{1/2}\left(\sum_{k=1}^{\infty}(\pi k)^{-4}\boldsymbol{v}_k\boldsymbol{v}_k'\right)\mathbf{Q}_{K^\dagger}^{1/2},$$

$$T^{-1}\mathbf{S}_{11}\xrightarrow{d}\mathbf{Q}_{K^\dagger}^{1/2}\left(\sum_{k=1}^{\infty}(\pi k)^{-2}\boldsymbol{v}_k\boldsymbol{v}_k'\right)\mathbf{Q}_{K^\dagger}^{1/2},$$

$$T^{-1}\mathbf{C}_{12}\xrightarrow{d}\mathbf{Q}_{K^\dagger}^{1/2}\left(\sum_{k=1}^{\infty}(\pi k)^{-3}\boldsymbol{v}_k\mathbf{u}_k'\right),$$

$$\mathbf{C}_{22}\xrightarrow{d}\sum_{k=1}^{\infty}(\pi k)^{-2}\mathbf{u}_k\mathbf{u}_k',$$

where $\boldsymbol{v}_k$ and $\mathbf{u}_k$ are mutually independent $N(\mathbf{0},\mathbf{I})$ random vectors of dimension $K^\dagger$ and $(N-K^\dagger)$, respectively. Thus

$$T^{-1}\xi_N=T^{-1}\operatorname{tr}\mathbf{S}^{-1}\mathbf{C}=T^{-1}\operatorname{tr}\mathbf{S}_{11}^{-1}\mathbf{C}_{11}+o_p(1)$$

converges in distribution to (13), whereas the common trends test statistic,

$$\zeta_{K,N}=\operatorname{tr}(\mathbf{C}_{22}-\mathbf{C}_{12}'\mathbf{C}_{11}^{-1}\mathbf{C}_{12})+o_p(1),$$

converges to (27) when $K=K^\dagger$.

https://doi.org/10.1017/S0266466600162024 Published online by Cambridge University Press