Economics Department

Multivariate Tests of Fractionally Integrated Hypotheses

LUIS A. GIL-ALAÑA

ECO No. 98/19

EUI WORKING PAPERS

EUROPEAN UNIVERSITY INSTITUTE

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE

ECONOMICS DEPARTMENT

EUI Working Paper **ECO** No. 98/19

Multivariate Tests of Fractionally Integrated Hypotheses

LUIS A. GIL-ALAÑA

BADIA FIESOLANA, SAN DOMENICO (FI)

Printed in Italy in July 1998
European University Institute
Badia Fiesolana
I – 50016 San Domenico (FI)
Italy

MULTIVARIATE TESTS OF FRACTIONALLY INTEGRATED HYPOTHESES

L. A. Gil-Alaña

European University Institute, Florence
and
London School of Economics, London

March 1998

ABSTRACT

Multivariate tests of fractionally integrated hypotheses are proposed in this article. They are a natural generalization of the univariate tests of Robinson (1994) for testing unit roots and other nonstationary hypotheses. The functional forms of the tests, based on the score, Wald and likelihood ratio principles are calculated in both, the time and the frequency domain. Some simulations based on Monte Carlo experiments are also carried out at the end of the article.

Key words: Multivariate tests, score tests, Wald tests, LR tests, fractional differencing, unit roots.

1. Introduction

In this article I extend the univariate tests of Robinson (1994) to a general multivariate context, testing the presence of unit roots and other nonstationarities on the residuals in a multiple time series system. The multivariate case is relevant in order to analyze the interrelationships between different variables, and it can provide a more detailed insight into properties and stochastic behaviour than the univariate work. In particular, we will initially take the underlying I(0) sequence to be contemporaneously correlated but uncorrelated in time, then going on to extend the treatment to a general case of I(0) parametric autocorrelation. Multivariate tests for unit roots have been widely analyzed in the literature, and they are commonly related to the problem of cointegration, testing the number of common unit roots in a system of equations, (e.g., Johansen (1988)). The test statistics proposed here go beyond that in the sense that they will allow us to test not only unit roots, but also fractional roots of any order for each one of the time series analyzed.

We consider a multivariate regression model of form

$$Y_t = Z_t(\delta) + X_t, \qquad t = 1,2,..., \tag{1}$$

with

$$X_t = 0, \qquad t \leq 0, \tag{2}$$

where the column vectors Y_t and X_t each has N components, and by δ we mean a (K x 1) vector of real parameters, and $Z_t(\delta)$ is a (N x 1) vector of possibly non-linear functions of δ and, in general a number of predetermined variables. We will assume that under the null hypothesis to be tested and described below, X_t in (1) and (2) satisfies

$$\Phi(L) X_t = U_t, \qquad t = 1,2,..., \tag{3}$$

where $\Phi(L)$ is a (N x N) diagonal matrix function of the backshift operator L, and U_t is a (Nx1) I(0) vector process[1] with mean zero and weak parametric correlation. We consider a given matrix function $\Phi(z;\theta)$ of the complex variate z and the p-dimensional vector θ of real-valued parameters, where $\Phi(z;\theta) = \Phi(z)$ for all z such that $|z| = 1$ if and only if the null hypothesis defined by

$$\mathbf{H_o}: \theta = 0 \tag{4}$$

[1] We define an I(0) vector process U_t, t = 0,±1,..., as a covariance stationary vector process with spectral density matrix $f(\lambda)$ that is finite and positive definite.

holds, where there is no loss of generality in using the vector of zeros instead of an arbitrary given vector. In doing so, we can cast (3) in terms of a nested composite parametric null hypothesis, within the class of alternatives

$$\Phi(L;\theta)\,X_t \;=\; U_t \qquad\qquad t \;=\; 1,2,\ldots . \tag{5}$$

We take $\Phi(z)$ to have u^{th} diagonal element of form

$$\rho_u(z) \;=\; (1-z)^{\gamma_1^u}\,(1+z)^{\gamma_2^u}\prod_{j=3}^{h^u}(1-2\cos w_j^u z + z^2)^{\gamma_j^u}$$

for a given h^u, given distinct real numbers w_j^u, $j=3,4,..h^u$ on the interval $(0,\pi)$ and given real numbers γ_j^u for $j=1,...h^u$. Thus, a model like (3) will include a wide range of possibilities to be tested for each time series, such as I(d) processes with a single root at zero frequency, if $\rho_u(z) = (1 - z)^d$;[2] quarterly I(d) processes with four roots if $\rho_u(z) = (1 - z^4)^d$; 1/f noise processes if $\rho_u(z) = (1 - z)^{1/2}$, etc.

We specify now $\Phi(z;\theta)$ in a way such that we take each diagonal element of $\Phi(z;\theta)$, $\rho_u(z;\theta)$, to depend on θ but not necessarily involving all elements of θ. To do that, we take

$$\rho_u(z;\theta) \;=\; (1-z)^{\gamma_1^u+\theta_{i_1}^u}(1+z)^{\gamma_2^u+\theta_{i_2}^u}\prod_{j=3}^{h^u}(1-2\cos w_j^u z + z^2)^{\gamma_j^u+\theta_{i_j}^u} \tag{6}$$

where for each combination (u,j), $\theta_{ij}^u = \theta_l$ for some l; and for each l, there is at least one combination (u,j) such that $\theta_{ij}^u = \theta_l$, where θ_l corresponds to the l^{th} element of θ. This is a fairly general specification in the sense that we allow for duplications not only within equations but also across equations. Furthermore, this way of specifying $\Phi(z;\theta)$ permits us to specifically consider situations where θ is the same across all equations, and also the case when θ is taken as strictly different for each equation. This will be illustrated with some examples in Section 4.

We adopt the normalization $\rho_u(0;\theta) = 1$ for all θ and u = 1,2,...N, and we assume that $\rho_u(z;\theta)$ is differentiable in θ on a neighbourhood of $\theta = 0$ for all $|z| = 1$. Also we assume that for any u,v = 1,2,...,N

$$\det(E_{uv}) < \infty \tag{7}$$

[2] Note that this specification includes the unit root case when d = 1.

$$where \quad E_{uv} = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\epsilon_{(u)}(\lambda)\bar{\epsilon}_{(v)}(\lambda)' + \epsilon_{(v)}(\lambda)\bar{\epsilon}_{(u)}(\lambda)'\right)d\lambda$$

$$and \quad \epsilon_{(u)}(\lambda) = \frac{\partial \log \rho_u(e^{i\lambda};\theta)}{\partial\theta} \tag{8}$$

for real λ, and $\bar{\varepsilon}_{(u)}(\lambda)$ as the conjugate vector of $\varepsilon_{(u)}(\lambda)$. Note that the (p x 1) vector $\varepsilon_{(u)}(\lambda)$ is independent of θ given the linearity of log $\rho_u(e^{i\lambda};\theta)$ with respect to θ in (6). In particular, its real part takes the form

$$\delta_{1l}^u \log\left|2\sin\frac{\lambda}{2}\right| + \delta_{2l}^u \log\left(2\cos\frac{\lambda}{2}\right) + \sum_{j=3}^{h^u} \delta_{jl}^u \log\left|2(\cos\lambda - \cos w_j^u)\right|,$$

for l=1,...,p and $|\lambda| < \pi$, where $\delta_{jl}^u = 1$ if $\theta_{ij}^u = \theta_l$ and 0 otherwise. Condition (7) is not satisfied when testing unit roots nested in AR alternatives of form: $\rho_u(z;\theta) = (1 - (1+\theta)z)$, but it is satisfied by fractional alternatives of form: $\rho_u(z;\theta) = (1 - z)^{1+\theta}$, for example.

It should also be noted that under the null hypothesis, defined in (4), the model will be completely specified by (1)-(3), and it can be redefined as

$$\Phi(L)Y_t = W_t(\delta) + U_t \tag{9}$$

where $W_t(\delta) = (W_{1t}(\delta);W_{2t}(\delta);...;W_{Nt}(\delta))'$, with $W_{ut}(\delta) = \rho_u(L)Z_{ut}(\delta)$. (9) is a very general form of a regression model which includes multivariate linear and non-linear models and simultaneous equation systems, and its possible non-linear nature is motivated given that in economics and the physical sciences, multivariate regression models that are essentially of a non-linear nature have frequently been proposed to describe phenomena that may be of a continuous nature but are sampled at equal intervals of time. (See e.g. Robinson (1972), (1977)).

The initial discussion of the tests will assume that U_t in (3) is a white noise vector process, so the only nuisance parameters will be the elements of $Z_t(\delta)$ in (1) and those of the variance-covariance matrix of U_t. Then, we will extend the tests to a quite general form of I(0) autocorrelation in U_t, which will include as specific examples, the type of multiple autoregressive-moving average (ARMA) models.

We will start by presenting the functional forms of the test statistics based on the three general principles when deriving nested parametric hypotheses, namely, the score, Wald and likelihood-ratio principles, and we will do so for the two situations mentioned above, that is, white noise and weak parametric autocorrelation in U_t. As usual, it should be possible to show that the tests based on these three principles will have the same null limit distribution (a χ_p^2 distribution where p is the number of restrictions tested). However, we do not present rigorous proofs of the asymptotic properties, but rather informal statements. It will undoubtedly be possible to extend the asymptotic null and local distribution theory of Robinson (1994) for the scalar case, to our multivariate situation under natural generalizations of his conditions. Once we have obtained the functional forms of the tests, we will rewrite them for two cases of particular interest: First, when θ in (5) is the same across all diagonal elements in $\Phi(z;\theta)$ and then, we will consider the case when θ is strictly different for each element in $\Phi(z;\theta)$. Finally, some simulations based on Monte Carlo experiments will be carried out in order to study the finite-sample behaviour of versions of the tests. Appendices 1 and 2 show the derivations of the test statistics of Sections 2 and 3 respectively.

2. Score test for white noise U_t

In this section I describe a score test for the null hypothesis (4) in a model given by (1), (2) and (5), under the presumption that U_t in (5) is a vector sequence of zero mean uncorrelated in time random variables, with unknown variance-covariance matrix K. One definition for the score test is as follows. Let $L(\eta)$ be an objective function (such as the negative of the log-likelihood) and take $\eta = (\theta',\upsilon')'$, where $\tilde{\eta} = (0',\tilde{v}')'$ are the values that minimizes $L(\eta)$ under the null hypothesis. A score test (see Rao (1973), page 418) is then given by

$$\frac{\partial L(\eta)}{\partial \eta'}\left[E_o\left(\frac{\partial L(\eta)}{\partial \eta}\frac{\partial L(\eta)}{\partial \eta'}\right)\right]^{-1}\frac{\partial L(\eta)}{\partial \eta}\Big|_{\substack{\theta=0\\ v=\tilde{v}}} \tag{10}$$

where the expectation is taken under the null hypothesis prior to substitution of $\tilde{v}$. However, the same asymptotic behaviour will be expected if we replace the inverted matrix appearing in (10) by alternative forms such as the sample average or the Hessian. For convenience in the derivation below, we will make use of the expected information matrix, so the score test will take the form

$$\frac{\partial L(\eta)}{\partial \eta'}\left[E_o\,\frac{\partial L^2(\eta)}{\partial \eta\,\partial \eta'}\right]^{-1}\frac{\partial L(\eta)}{\partial \eta}\Big|_{\substack{\theta=0.\\ v=\tilde{v}}} \tag{11}$$

We now describe the test statistic. We take L in (11), with $\eta = (\theta',\delta',\alpha')'$

and $\alpha = v(K)$, to be the negative of the log-likelihood based on Gaussian U_t. In Appendix 1 it is shown that (11) takes the form

$$\hat{S}^t = T\,\hat{a}^{t\prime}\,(\hat{A}^t)^{-1}\,\hat{a}^t \tag{12}$$

where $\hat{a}^t$ is a (p x 1) vector of form

$$\hat{a}^t = -\sum_{u=1}^{N}\sum_{v=1}^{N}\hat{\sigma}^{uv}\sum_{s=1}^{T-1}\psi_s^{(u)}C_{uv}(s;\hat{\delta}), \tag{13}$$

and $\psi_s^{(u)}$ is obtained by expanding

$$\psi_{(u)}(\lambda) = Re[\epsilon_{(u)}(\lambda)] \quad as \quad \sum_{s=1}^{\infty}\psi_s^{(u)}\cos\lambda s.$$

$$\hat{A}^t = \sum_{u=1}^{N}\sum_{v=1}^{N}\hat{\sigma}^{uv}\hat{\sigma}_{uv}\sum_{s=1}^{T-1}\left(1-\frac{s}{T}\right)\psi_s^{(u)}\psi_s^{(v)\prime}, \tag{14}$$

$\hat{\sigma}^{uv}$ is the $(u,v)^{th}$ element of $\hat{K}^{-1}$; $\hat{\sigma}_{uv}$ is the $(u,v)^{th}$ element of $\hat{K}$; and $C_{uv}(s;\hat{\delta})$ is the $(u,v)^{th}$ element of $C_{\hat{U}}(s)$, where

$$\hat{K} = \frac{1}{T}\sum_{t=1}^{T}\hat{U}_t(\delta)\,\hat{U}_t(\delta)'; \qquad C_{\hat{U}}(s) = \frac{1}{T}\sum_{t=1}^{T-s}\hat{U}_t(\delta)\,\hat{U}_{t+s}(\delta)';$$

$\hat{U}_t(\delta) = \Phi(L)Y_t - W_t(\hat{\delta})$, and $\hat{\delta}$ must be at least a $T^{1/2}$- consistent estimate of the true value δ.

Clearly, as in the univariate tests of Robinson (1994), concise formulas for $\psi_s^{(u)}$ are available in some simple cases; for example, $\psi_s^{(u)} = -s^{-1}$, when $\rho_u(L;\theta) = (1 - L)^{d+\theta}$, for any real d. However, we can also express the test statistic in the frequency domain and, under certain suitable conditions[1], approximate this to obtain an alternative, asymptotically equivalent, form. $\hat{a}^t$ in (13) can be written as

$$-\frac{1}{2}\sum_{u=1}^{N}\sum_{v=1}^{N}\hat{\sigma}^{uv}\int_{-\pi}^{\pi}(\epsilon_{(u)}(\lambda) + \bar{\epsilon}_{(v)}(\lambda))I_{uv}(\lambda;\hat{\delta})d\lambda,$$

where $\varepsilon_{(u)}(\lambda)$ is as in (8) and $\bar{\varepsilon}_{(v)}(\lambda)$ is the conjugate vector of $\varepsilon_{(v)}(\lambda)$; $I_{uv}(\lambda;\hat{\delta})$ is

[1] These conditions are basically a generalization of those of Robinson (1994), requiring technical assumptions on ρ_u (and thus on $\varepsilon_{(u)}(\lambda)$) to justify approximating integrals by sums.

the (u,v)th element in the cross-periodogram of $\hat{U}_t(\delta) = (\hat{U}_{1t}(\delta);...;\hat{U}_{Nt}(\delta))'$:

$$I_{uv}(\lambda,\hat{\delta}) = W_u(\lambda;\hat{\delta})\overline{W_v(\lambda;\hat{\delta})}, \quad W_u(\lambda;\hat{\delta}) = \frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T}\hat{U}_{ut}(\delta)e^{i\lambda t},$$

where the line over $W_v(\lambda;\delta)$ denotes complex conjugate. To see the previous result note that $\hat{a}^t$ in (13) can be decomposed into

$$-\left(\sum_{u=1}^{N}\hat{\sigma}^{uu}\sum_{s=1}^{T-1}\psi_s^{(u)}C_{uu}(s;\hat{\delta}) + \frac{1}{2}\sum_{u=1}^{N}\sum_{\substack{v=1\\v\neq u}}^{N}\hat{\sigma}^{uv}\sum_{s=1}^{T-1}(\psi_s^{(u)}C_{uv}(s;\hat{\delta}) + \psi_s^{(v)}C_{vu}(s;\hat{\delta}))\right),$$

and it can be shown that

$$\sum_{s=1}^{T-1}\psi_s^{(u)}C_{uu}(s;\hat{\delta}) = \frac{1}{2}\int_{-\pi}^{\pi}(\epsilon_{(u)}(\lambda) + \overline{\epsilon}_{(u)}(\lambda))I_{uu}(\lambda;\hat{\delta})d\lambda,$$

$$\textit{and} \qquad \sum_{s=1}^{T-1}(\psi_s^{(u)}C_{uv}(s;\hat{\delta}) + \psi_s^{(v)}C_{vu}(s;\hat{\delta})) =$$

$$\frac{1}{2}\int_{-\pi}^{\pi}(\epsilon_{(u)}(\lambda) + \overline{\epsilon}_{(v)}(\lambda))I_{uv}(\lambda;\hat{\delta})d\lambda + \frac{1}{2}\int_{-\pi}^{\pi}(\epsilon_{(v)}(\lambda) + \overline{\epsilon}_{(u)}(\lambda))I_{vu}(\lambda;\hat{\delta})d\lambda.$$

Also, under suitable conditions, keeping $\hat{\sigma}^{uv}$ and $\hat{\sigma}_{uv}$ fixed, $\hat{A}^t$ in (14) becomes asymptotically

$$\sum_{u=1}^{N}\sum_{v=1}^{N}\hat{\sigma}^{uv}\hat{\sigma}_{uv}\sum_{s=1}^{\infty}\psi_s^{(u)}\psi_s^{(v)'}, \tag{15}$$

and using Parseval's relationship, this quantity can be expressed as

$$\sum_{u=1}^{N}\sum_{v=1}^{N}\hat{\sigma}^{uv}\hat{\sigma}_{uv}\frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\epsilon_{(u)}(\lambda)\overline{\epsilon}_{(v)}(\lambda)' + \epsilon_{(v)}(\lambda)\overline{\epsilon}_{(u)}(\lambda)'\right)d\lambda = \sum_{u}^{N}\sum_{v=1}^{N}\hat{\sigma}^{uv}\hat{\sigma}_{uv}E_{uv}$$

since (15) can also be decomposed into

$$\sum_{u=1}^{N}\hat{\sigma}^{uu}\hat{\sigma}_{uu}\sum_{s=1}^{\infty}\psi_s^{(u)}\psi_s^{(u)'} = \sum_{u=1}^{N}\hat{\sigma}^{uu}\hat{\sigma}_{uu}\frac{1}{2\pi}\int_{-\pi}^{\pi}(\epsilon_{(u)}(\lambda)\overline{\epsilon}_{(u)}(\lambda)')d\lambda,$$

$$\textit{and} \quad \frac{1}{2}\sum_{u=1}^{N}\sum_{\substack{v=1\\v\neq u}}^{N}\hat{\sigma}^{uv}\hat{\sigma}_{uv}\sum_{s=1}^{\infty}(\psi_s^{(u)}\psi_s^{(v)'} + \psi_s^{(v)}\psi_s^{(u)'}) =$$

$$= \sum_{u=1}^{N} \sum_{\substack{v=1 \\ v \neq u}}^{N} \hat{\sigma}^{uv} \hat{\sigma}_{uv} \frac{1}{4\pi} \int_{-\pi}^{\pi} \epsilon_u(\lambda)\overline{\epsilon_v}(\lambda)' + \epsilon_v(\lambda)\overline{\epsilon_u}(\lambda)' d\lambda.$$

Therefore, the score statistic in (12) can be approximated in the frequency domain by the expression

$$\hat{S}^f = T\, \hat{a}^{f'}\, (\hat{A}^f)^{-1}\, \hat{a}^f \tag{16}$$

where

$$\hat{a}^f = \frac{-\pi}{T} \sum_{u=1}^{N} \sum_{v=1}^{N} \hat{\sigma}^{uv} \sum_{r}^{*} (\epsilon_{(u)}(\lambda_r) + \overline{\epsilon}_{(v)}(\lambda_r))\, I_{uv}(\lambda_r;\hat{\delta}), \tag{17}$$

and

$$\hat{A}^f = \frac{1}{2T} \sum_{u=1}^{N} \sum_{v=1}^{N} \hat{\sigma}^{uv} \hat{\sigma}_{uv} \sum_{r}^{*} (\epsilon_{(u)}(\lambda_r)\overline{\epsilon}_{(v)}(\lambda_r)' + \epsilon_{(v)}(\lambda_r)\overline{\epsilon}_{(u)}(\lambda_r)'), \tag{18}$$

$\lambda_r = 2\pi r/T$, and the sums on the asterisk are over λ_r in M where $M = \{\lambda;\ -\pi < \lambda \leq \pi;\ \lambda \notin (\rho_l - \lambda; \rho_l + \lambda),\ l=1,2,...,s\}$, such that ρ_l, $l=1,2,...s$ are the distinct poles on $\varepsilon_{(u)}(\lambda)$ on $(\pi,\pi]$ for $u = 1,2,...,N$. Note that if, for example, $\rho_u(L;\theta)$ is given by $(1-L)^{d+\theta}$, we calculate $\varepsilon_{(u)}(\lambda_r)$ as

$$Re[\epsilon_{(u)}(\lambda_r)] = \psi_{(u)}(\lambda_r) = \log\left|2\sin\frac{\lambda_r}{2}\right|, \quad and \quad Im[\epsilon_{(u)}(\lambda_r)] = \frac{\lambda_r - \pi}{2},$$

with $r = 1,2,...,T-1$, (see e.g., Zygmund (1979), page 5).

We should expect that under some regularity conditions, (basically a natural generalization of those in Robinson (1994)), the test described below will have the same optimal asymptotic properties as Robinson's (1994) univariate tests. These conditions impose a martingale difference assumption on the white noise vector U_t;[2] also W as defined in Appendix 1 must be a positive definite matrix; and $\rho_u(z;\theta)$, $u = 1,2...,N$ must belong to Class H as defined in Robinson (1994), with $\varepsilon_{(u)}(\lambda)$ satisfying the same conditions as $\psi(\lambda)$ in that paper. We believe that under these conditions, (12) and (16) will have a null limit χ_p^2 distribution, and under local alternatives of form H_a: $\theta = \theta_T = \delta T^{-1/2}$, a $\chi_p^2(\nu)$ distribution with a non-centrality parameter ν, which is optimal under

---

[2] That is, $E(U_t \mid B_{t-1}) = 0$ and $E(U_t U_t' \mid B_{t-1}) = K$, where B_t is the σ-field of events generated by U_s, $s \leq t$.

Gaussianity of U_t.

Thus, a large-sample $100\alpha\%$-level test for rejecting H_o (4) against the alternative: H_1: $\theta \neq 0$, will be given by the rule: "Reject H_o if $\hat{S}^t$ (or $\hat{S}^f$) > $\chi^2_{p,\alpha}$", where $P(\chi_p^2 > \chi^2_{p,\alpha}) = \alpha$.

3. Score test for weakly parametrically correlated U_t

The test statistics presented in Section 2 can be robustified to allow weakly parametrically autocorrelated U_t. We can consider the model in (1), (2), and (5), with U_t in (5) as a vector process with N components generated by a parametric model of form

$$U_t = \sum_{j=0}^{\infty} A(j;\tau)\epsilon_{t-j} \qquad t = 1,2,..., \tag{19}$$

where ε_t is a vector white noise process, and K is now the unknown variance-covariance matrix of ε_t. In relation with (19), the corresponding spectral density matrix is

$$f(\lambda;\tau) = \frac{1}{2\pi} k(\lambda;\tau) \; K \; k(\lambda;\tau)^*, \tag{20}$$

where $k(\lambda;\tau) = \sum_{j=0}^{\infty} A(j;\tau)e^{i\lambda j}$, and k^* means the complex conjugate transpose of k.

A number of conditions are required on A and f in Appendix 2 when deriving the test statistic; their practical implications being that though U_t is capable of exhibiting a much stronger degree of autocorrelation than multiple autoregressive moving average ARMA processes, its spectral density matrix must be finite, with eigenvalues bounded and bounded away from zero. Thus, it cannot include fractional processes with positive or negative differencing parameters.

By extending the argument in Section 2 and Appendix 1, we show in Appendix 2 that, under Gaussianity of U_t, an approximate score statistic for testing (4) in (1), (2), (5) and (19) is

$$\tilde{S} = T\,\tilde{b}'\,\tilde{B}^{-1}\,\tilde{b} \tag{21}$$

$$\tilde{b} = -\frac{1}{2T}\sum_{r}^{*}\sum_{u=1}^{N}\sum_{v=1}^{N}\left(\epsilon_{(u)}(\lambda_r) + \overline{\epsilon}_{(v)}(\lambda_r)\right) I_{uv}(\lambda_r;\hat{\delta})\hat{f}^{vu}(\lambda_r;\tilde{\tau}),$$

and $\tilde{B}$ is $\tilde{C} - \tilde{D}'\tilde{E}^{-1}\tilde{D}$, where

$$\tilde{C} = \frac{1}{2T}\sum_{r}^{*}\sum_{u,v=1}^{N}\left(\epsilon_{(u)}(\lambda_r)\overline{\epsilon}_{(v)}(\lambda_r)' + \overline{\epsilon}_{(v)}(\lambda_r)\epsilon_{(u)}(\lambda_r)'\right)\hat{f}_{uv}(\lambda_r;\tilde{\tau})\hat{f}^{vu}(\lambda_r;\tilde{\tau}),$$

$$\tilde{D}' = -\frac{1}{2T}\sum_{r}^{*}\sum_{u,v=1}^{N}\left(\epsilon_{(u)}(\lambda_r) + \overline{\epsilon}_{(v)}(\lambda_r)'\right)\hat{f}^{vu}(\lambda_r;\tilde{\tau})\,\frac{\partial \hat{f}_{uv}(\lambda_r;\tilde{\tau})}{\partial \tau'},$$

and

$$(\tilde{E})_{uv} = \frac{1}{2T}\sum_{r}^{*}\mathrm{tr}\left(\hat{f}^{-1}(\lambda_r;\tilde{\tau})\frac{\partial \hat{f}(\lambda_r;\tilde{\tau})}{\partial \tau_u}\hat{f}^{-1}(\lambda_r;\tilde{\tau})\frac{\partial \hat{f}(\lambda_r;\tilde{\tau})}{\partial \tau_v}\right).$$

$I_{uv}(\lambda;\hat{\delta})$ is the $(u,v)^{th}$ element of the periodogram of $\hat{U}_t$, $I_U(\lambda;\hat{\delta})$, as was given in Section 2; $\hat{f}_{uv}(\lambda_r;\tilde{\tau})$ and $\hat{f}^{uv}(\lambda_r;\tilde{\tau})$ correspond to the $(u,v)^{th}$ elements of $\hat{f}(\lambda_r;\tilde{\tau})$ and $\hat{f}^{-1}(\lambda_r;\tilde{\tau})$ respectively, with

$$\hat{f}(\lambda;\tilde{\tau}) = \frac{1}{2\pi}k(\lambda;\tilde{\tau})\ \hat{K}\ \ k(\lambda;\tilde{\tau})^{*}$$

and

$$\tilde{\tau} = \mathrm{argmin}_{\tau\in T^{*}}\left(\frac{T}{2}\log\det\ \hat{f}(\lambda_r;\tau) + \frac{1}{2}\sum_{r}^{*}\mathrm{tr}[\hat{f}^{-1}(\lambda_r;\tau)I_U(\lambda_r;\hat{\delta})\right], \qquad (22)$$

where T^{*} is a compact subset of q-dimensional Euclidean space.

We can see that the test statistic obtained in (21) becomes (16) when we consider the case of white noise U_t. In such situation, $\hat{f}_{uv}(\lambda_r;\tau) = \hat{\sigma}_{uv}/2\pi$, and $\hat{f}^{uv}(\lambda_r;\tau) = 2\pi\,\hat{\sigma}^{uv}$. Then,

$$\tilde{b} = \frac{-\pi}{T}\sum_{r}^{*}\sum_{u=1}^{N}\sum_{v=1}^{N}\left(\epsilon_{(u)}(\lambda_r) + \overline{\epsilon}_{(v)}(\lambda_r)\right) I_{uv}(\lambda_r;\hat{\delta})\,\hat{\sigma}^{vu} =$$

$$= \frac{-\pi}{T}\sum_{u=1}^{N}\sum_{v=1}^{N}\hat{\sigma}^{uv}\sum_{r}^{*}\left(\epsilon_{(u)}(\lambda_r) + \overline{\epsilon}_{(v)}(\lambda_r)\right) I_{uv}(\lambda_r;\hat{\delta}) = \hat{a}^{f} \text{ in } (17).$$

Similarly,

$$\tilde{C} = \frac{1}{2T}\sum_{r}^{*}\sum_{u=1}^{N}\sum_{v=1}^{N}\left(\epsilon_{(u)}(\lambda_r)\bar{\epsilon}_{(v)}(\lambda_r)' + \epsilon_{(v)}(\lambda_r)\bar{\epsilon}_{(u)}(\lambda_r)'\right)\hat{\sigma}_{uv}\hat{\sigma}^{vu} = \hat{A}^{f} \text{ in (18)},$$

and finally, $\tilde{D}$ and $\tilde{E}$ are now zero matrices, so we have that $\tilde{S}$ in (21) takes the same form as $\hat{S}^f$ in (16).

Extending the conditions in Robinson (1994) to this multivariate context, we should expect that, allowing a martingale difference assumption on ε_t in (19), with $\sum_{j=1}^{\infty} j^{1/2}\|A(j;\tau)\| < \infty$, where $\|A\|$ means any norm for the matrix A, for example the square root of the maximum eigenvalue of A^*A; with W as a positive definite matrix; ρ_u, u = 1,...,N, satisfying the same conditions as in Section 2; and $f_{uv}(\lambda;\tau)$, and $\partial f_{uv}(\lambda;\tau)/\partial\tau$ satisfying a Lipschitz condition in λ of order $\eta > 1/2$, for all u,v = 1,...,N, then, under H_o (4): $\tilde{S} \rightarrow_d \chi_p^{\,2}$ as $T \rightarrow \infty$, and $\tilde{S}$ should also satisfy the same asymptotic efficiency properties as $\hat{S}^t$ and $\hat{S}^f$ in Section 2.

4. Particular cases of the score tests

In the preceding sections we have presented three different versions of the score test statistic: (12), which corresponds to the time domain representation of the test for white noise U_t; (16), which approximates (12) in the frequency domain; and (21) which is the frequency domain version of the test statistic for weakly parametrically autocorrelated U_t. In this section we consider two particular cases of interest for each version of these tests. The first case corresponds to the test statistic when we take θ in (5) as a (p x 1) vector containing exactly the same elements across all diagonal elements in $\Phi(z;\theta)$, while the second case takes this vector θ as strictly different for each diagonal element in $\Phi(z;\theta)$.

We illustrate this with two simple examples in a bivariate model: First we test if one of the series is an $I(d_1)$ process and if the other is $I(d_2)$. Thus, we consider that both series have a root at the same zero frequency, though with different integration orders. In the second example, we consider that the series might differ in the number of roots in its bivariate representation. Thus, we test the same hypothesis, $(I(d_1))$, for the first series and a quarterly $I(d_2)$ process in the second one. Therefore, the model will be specified, under the null hypothesis, in the first of these examples as

$$\begin{pmatrix} (1-L)^{d_1} & 0 \\ 0 & (1-L)^{d_2} \end{pmatrix} \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} U_{1t} \\ U_{2t} \end{pmatrix} \qquad t = 1,2,...., \qquad (E1)$$

and in the second as

$$\begin{pmatrix} (1-L)^{d_1} & 0 \\ 0 & (1-L^4)^{d_2} \end{pmatrix} \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} U_{1t} \\ U_{2t} \end{pmatrix} \qquad t = 1,2,...., \qquad (E2)$$

where $X_t = (X_{1t},X_{2t})' = 0$ for $t \leq 0$, and $U_t = (U_{1t},U_{2t})'$ follows an I(0) process.

4.a Same θ across the equations

We consider the model in (1), (2), and (5), but now we take $\Phi(z;\theta)$ to be of form such that its u^{th} diagonal element is

$$\rho_u(z;\theta) = (1 - z)^{\gamma_1^u + \theta_{i_1}} (1 + z)^{\gamma_2^u + \theta_{i_2}} \prod_{j=3}^{h^u} (1 - 2\cos w_j z + z^2)^{\gamma_j^u + \theta_{i_j}},$$

and for each j, $\theta_{ij} = \theta_l$ for some l, and for each l, there is at least one j such that $\theta_{ij} = \theta_l$. Therefore we take the parameter vector θ to be exactly the same across all equations in (5), and the difference between one equation and another comes now through the coefficients γ_i^u for $i=1,2,...,h^u$ and $u=1,2,...,N$. Thus, in the first example, the model will be specified as

$$\begin{pmatrix} (1-L)^{d_1+\theta} & 0 \\ 0 & (1-L)^{d_2+\theta} \end{pmatrix} \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} U_{1t} \\ U_{2t} \end{pmatrix} \qquad t = 1,2,....,$$

and we will test here the null hypothesis, H_o: $\theta = 0$, against the alternative, H_a: $\theta \neq 0$. Given that in this case θ is a scalar, we can also consider one-sided tests for the same null hypothesis against the alternatives: H_{a1}: $\theta < 0$ or H_{a2}: $\theta > 0$.

In the second example, the model will take the form

$$\begin{pmatrix} (1-L)^{d_1+\theta_1}(1+L)^{\theta_2}(1+L^2)^{\theta_3} & 0 \\ 0 & (1-L)^{d_2+\theta_1}(1+L)^{d_2+\theta_2}(1+L^2)^{d_2+\theta_3} \end{pmatrix} \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} U_{1t} \\ U_{2t} \end{pmatrix},$$

which, under the null hypothesis, H_o: $\theta = (\theta_1,\theta_2,\theta_3)' = 0$, becomes (E2), implying that X_{2t} behaves as a quarterly $I(d_2)$ process, and therefore, with all roots with the same integration order d_2. Clearly we could also have tested a model, allowing different integration orders at zero and at seasonal frequencies.

This specification is a particular case of the general model presented in Section 1 where now

$$\epsilon_{(u)}(\lambda) = \frac{\partial \log \rho_u(e^{i\lambda};\theta)}{\partial\theta} = \epsilon(\lambda) \quad \textit{for} \quad \textit{all} \quad u=1,2,\ldots,N. \tag{23}$$

(23) implies that $\psi_s^{(u)} = \psi_s$ for all u = 1,2,...,N, and then, we can immediately describe the functional forms of the three test statistics. Starting with white noise U_t, substituting (23) in (12) - (14), the time domain version of the test statistic is

$$\hat{S}^{t^1} = T\,\hat{a}^{t^1\prime}\,(\hat{A}^{t^1})^{-1}\,\hat{a}^{t^1} \tag{24}$$

where

$$\hat{a}^{t^1} = -\sum_{s=1}^{T-1}\psi_s\sum_{u=1}^{N}\sum_{v=1}^{N}\hat{\sigma}^{uv}C_{uv}(s;\hat{\delta}) = -\sum_{s=1}^{T-1}\psi_s\, tr\left[\hat{K}^{-1}\, C_{\hat{U}}(s)\right],$$

and

$$\hat{A}^{t^1} = \sum_{s=1}^{T-1}\left(1-\frac{s}{T}\right)\psi_s\psi_s'\sum_{u=1}^{N}\sum_{v=1}^{N}\hat{\sigma}^{uv}\hat{\sigma}_{uv} = N\sum_{s=1}^{T-1}\left(1-\frac{s}{T}\right)\psi_s\,\psi_s'.$$

Expressing now the test statistic in terms of its frequency domain representation

$$\hat{S}^{f^1} = T\,\hat{a}^{f^1\prime}\,(\hat{A}^{f^1})^{-1}\,\hat{a}^{f^1} \tag{25}$$

where

$$\hat{a}^{f^1} = \frac{-\pi}{T}\sum_{\substack{u=1\\v=1}}^{N}\hat{\sigma}^{uv}\sum_r^{*}\left(\epsilon(\lambda_r)+\bar{\epsilon}(\lambda_r)\right)I_{uv}(\lambda_r;\hat{\delta}) = \frac{-2\pi}{T}\sum_r^{*}\psi(\lambda_r)\,tr[\hat{K}^{-1}\,I_{\hat{U}}(\lambda_r,\hat{\delta})],$$

$$\hat{A}^{f^1} = \frac{1}{2T}\sum_r^{*}\left(\epsilon(\lambda_r)\bar{\epsilon}(\lambda_r)'+\epsilon(\lambda_r)\bar{\epsilon}(\lambda_r)'\right)\sum_{\substack{u=1\\v=1}}^{N}\hat{\sigma}^{uv}\hat{\sigma}_{uv} = \frac{1}{T}\sum_r^{*}\epsilon(\lambda_r)\bar{\epsilon}(\lambda_r)'\sum_{\substack{u=1\\v=1}}^{N}\hat{\sigma}^{uv}\hat{\sigma}_{uv}$$

$$= \frac{N}{T}\sum_r^{*}\epsilon(\lambda_r)\,\bar{\epsilon}(\lambda_r)' = \frac{2N}{T}\sum_r^{*}\psi(\lambda_r)\,\psi(\lambda_r)'.$$

Finally, allowing weak parametric autocorrelation in U_t, substituting (23) in (21), we obtain that the test statistic is

$$\tilde{S}^1 = T\,\tilde{b}^{1\prime}\left(\tilde{C}^1 - \tilde{D}^{1\prime}\,\tilde{E}^{-1}\,\tilde{D}^1\right)^{-1}\tilde{b}^1 \tag{26}$$

where

$$\tilde{b}^1 = -\frac{1}{2T}\sum_r^{*}\left(\epsilon(\lambda_r) + \overline{\epsilon}(\lambda_r)\right)\sum_{u=1}^{N}\sum_{v=1}^{N} I_{uv}(\lambda_r;\hat{\delta})\,\hat{f}^{vu}(\lambda_r;\tilde{\tau}) =$$

$$\frac{-1}{T}\sum_r^{*}\psi(\lambda_r)\sum_{u=1}^{N}\sum_{v=1}^{N} I_{uv}(\lambda_r;\hat{\delta})\,\hat{f}^{vu}(\lambda_r;\tilde{\tau}) = \frac{-1}{T}\sum_r^{*}\psi(\lambda_r)\,tr\left[I_U(\lambda_r;\hat{\delta})\hat{f}(\lambda_r;\tilde{\tau})^{-1}\right],$$

and

$$\tilde{C}^1 = \frac{1}{2T}\sum_r^{*}\left(\epsilon(\lambda_r)\overline{\epsilon}(\lambda_r)' + \epsilon(\lambda_r)\overline{\epsilon}(\lambda_r)'\right)\sum_{u=1}^{N}\sum_{v=1}^{N}\hat{f}_{uv}(\lambda_r;\tilde{\tau})\,\hat{f}^{vu}(\lambda_r;\tilde{\tau})$$

$$= \frac{N}{T}\sum_r^{*}\epsilon(\lambda_r)\,\overline{\epsilon}(\lambda_r)' = \frac{2N}{T}\sum_r^{*}\psi(\lambda_r)\,\psi(\lambda_r)', \tag{27}$$

$$\tilde{D}^{1\prime} = -\frac{1}{2T}\sum_r^{*}\left(\epsilon(\lambda_r) + \overline{\epsilon}(\lambda_r)\right)\sum_{u=1}^{N}\sum_{v=1}^{N}\hat{f}^{vu}(\lambda_r;\tilde{\tau})\frac{\partial\hat{f}_{uv}(\lambda_r;\tilde{\tau})}{\partial\tau'} =$$

$$\frac{-1}{T}\sum_r^{*}\psi(\lambda_r)\left[tr\left(\hat{f}(\lambda_r;\tilde{\tau})^{-1}\frac{\partial\hat{f}(\lambda_r;\tilde{\tau})}{\partial\tau_1}\right);\ldots;tr\left(\hat{f}(\lambda_r;\tilde{\tau})^{-1}\frac{\partial\hat{f}(\lambda_r;\tilde{\tau})}{\partial\tau_q}\right)\right], \tag{28}$$

$$\tilde{E}_{uv} = \frac{1}{2T}\sum_r^{*}tr\left[\hat{f}(\lambda_r;\tilde{\tau})^{-1}\frac{\partial\hat{f}(\lambda_r;\tilde{\tau})}{\partial\tau_u}\hat{f}(\lambda_r;\tilde{\tau})^{-1}\frac{\partial\hat{f}(\lambda_r;\tilde{\tau})}{\partial\tau_v}\right]. \tag{29}$$

4.b Different θ's across equations

A second case of interest might be when we take the (p x 1) vector θ appearing in (5) to be equal to $(\theta^{1'};\theta^{2'};\ldots;\theta^{N'})'$, where θ^u is a $(p_u \times 1)$ vector affecting only the u^{th} equation. That is, the vector of parameters involving θ will be strictly different for each equation. We can now write down the u^{th} diagonal element in $\Phi(z;\theta)$ as

$$\rho_u(z;\theta^u) = (1-z)^{\gamma_1^u+\theta_{i_1}^u}\,(1+z)^{\gamma_2^u+\theta_{i_2}^u}\prod_{j=3}^{h^u}(1-2\cos w_j z+z^2)^{\gamma_j^u+\theta_{i_j}^u} \tag{30}$$

where for each j, $\theta_{i_j}^u = \theta_l^u$ for some l, and for each l, there is at least one j such that $\theta_{i_j}^u = \theta_l^u$. Thus, in the first of the examples mentioned above, the model will be of form

$$\begin{pmatrix}(1-L)^{d_1+\theta_1^1} & 0\\ 0 & (1-L)^{d_2+\theta_1^2}\end{pmatrix}\begin{pmatrix}X_{1t}\\ X_{2t}\end{pmatrix} = \begin{pmatrix}U_{1t}\\ U_{2t}\end{pmatrix} \qquad t = 1,2,....,$$

with $\theta = (\theta^1; \theta^2)' = (\theta_1^1; \theta_1^2)'$, and in the second example

$$\begin{pmatrix}(1-L)^{d_1+\theta_1^1}(1+L)^{\theta_2^1}(1+L^2)^{\theta_3^1} & 0\\ 0 & (1-L)^{d_2+\theta_1^2}(1+L)^{d_2+\theta_2^2}(1+L^2)^{d_2+\theta_3^2}\end{pmatrix}\begin{pmatrix}X_{1t}\\ X_{2t}\end{pmatrix} = \begin{pmatrix}U_{1t}\\ U_{2t}\end{pmatrix},$$

with $\theta = (\theta^{1'}; \theta^{2'})' = (\theta_1^1, \theta_2^1, \theta_3^1; \theta_1^2, \theta_2^2, \theta_3^2)'$.

Again this way of specifying the model is a particular case of the general model presented in Section 1. We need to define the (p_u x 1) vectors

$$e_{(u)}(\lambda) = \frac{\partial \log \rho_u(e^{i\lambda};\theta^u)}{\partial \theta^u}; \qquad f_{(u)}(\lambda) = Re[e_{(u)}(\lambda)],$$

for all u = 1,2,...,N, sharing the same properties as $\varepsilon_{(u)}(\lambda)$ and $\psi_{(u)}(\lambda)$ in Sections 1-3. To show that this is a particular case of the general specification given before, we just need to note that

$$\epsilon_{(u)}(\lambda) = P_u\, e_{(u)}(\lambda) \tag{31}$$

where P_u is a (p x p_u) matrix of 1's and 0's of form

$$P_u = \begin{pmatrix} 0 \\ --- \\ I_{p_u} \\ --- \\ 0 \end{pmatrix},$$

and substituting (31) in (12), (16) and (21) we can easily obtain the functional forms for the three test statistics. Starting again with white noise U_t, in the time domain representation, and noting that $\psi_s^{(u)} = P_u f_s^{(u)}$ where $f_s^{(u)}$ comes from expanding $f_{(u)}(\lambda)$ as $\sum_{s=1}^{\infty} f_s^{(u)} \cos\lambda s,$ the test statistic takes the form

$$\hat{S}^{t^2} = T\,\hat{a}^{t^2\prime}\,(\hat{A}^{t^2})^{-1}\,\hat{a}^{t^2} \tag{32}$$

$$where\ \hat{a}^{t^2} = \begin{pmatrix} \hat{a}_1^{t^2} \\ \hat{a}_2^{t^2} \\ . \\ . \\ \hat{a}_N^{t^2} \end{pmatrix}, \quad with\ \hat{a}_u^{t^2} = -\sum_{v=1}^{N} \hat{\sigma}^{uv} \sum_{s=1}^{T-1} C_{uv}(s;\hat{\delta})\, f_s^{(u)};$$

$$and\ \hat{A}^{t^2} = \begin{pmatrix} \hat{a}_{11}^{t} & \ldots & \hat{a}_{1N}^{t} \\ . & \ldots & . \\ \hat{a}_{N1}^{t} & \ldots & \hat{a}_{NN}^{t} \end{pmatrix}, \quad with\ \hat{a}_{uv}^{t} = \hat{\sigma}^{uv}\hat{\sigma}_{uv} \sum_{s=1}^{T-1} \left(1 - \frac{s}{T}\right) f_s^{(u)} f_s^{(v)\prime}. \tag{33}$$

The corresponding test statistic in the frequency domain representation is

$$\hat{S}^{f^2} = T\,\hat{a}^{f^2\prime}\,(\hat{A}^{f^2})^{-1}\,\hat{a}^{f^2} \tag{34}$$

$$where\ \hat{a}^{f^2} = \begin{pmatrix} \hat{a}_1^{f^2} \\ \hat{a}_2^{f^2} \\ . \\ . \\ \hat{a}_N^{f^2} \end{pmatrix}, \quad with\ \hat{a}_u^{f^2} = \frac{-2\pi}{T} \sum_{v=1}^{N} \hat{\sigma}^{uv} \sum_{r}^{*} e_{(u)}(\lambda_r) I_{uv}(\lambda_r;\hat{\delta}),$$

and

$$\hat{A}^{f^2} = \begin{pmatrix} \hat{a}_{11}^{f} & \ldots & \hat{a}_{1N}^{f} \\ . & \ldots & . \\ \hat{a}_{N1}^{f} & \ldots & \hat{a}_{NN}^{f} \end{pmatrix}, \tag{35}$$

with

$$\hat{a}_{uv}^{f} = \frac{1}{T}\hat{\sigma}^{uv}\hat{\sigma}_{uv} \sum_{r}^{*} e_{(u)}(\lambda_r)\bar{e}_{(v)}(\lambda_r)' = \frac{2}{T}\hat{\sigma}_{uv}\hat{\sigma}^{uv} \sum_{r}^{*} f_{(u)}(\lambda_r) f_{(v)}(\lambda_r)'. \tag{36}$$

Finally, the test statistic in the frequency domain for weakly parametrically autocorrelated U_t takes the form

$$\tilde{S}^2 = T\,\tilde{b}^{2\prime}\left(\tilde{C}^2 - \tilde{D}^{2\prime}(\tilde{E})^{-1}\tilde{D}^2\right)^{-1}\tilde{b}^2 \tag{37}$$

$$\tilde{b}^2 = \begin{pmatrix} \tilde{b}_1^2 \\ \tilde{b}_2^2 \\ \cdot \\ \tilde{b}_N^2 \end{pmatrix} \quad with \quad \tilde{b}_u^2 = -\frac{1}{T} Re\left(\sum_r^{*} e_{(u)}(\lambda_r)\sum_{v=1}^{N} I_{uv}(\lambda_r;\hat{\delta})\hat{f}^{vu}(\lambda_r;\tilde{\tau})\right),$$

$$\tilde{C}^2 = \begin{pmatrix} \tilde{c}_{11} & \dots & \tilde{c}_{1N} \\ \cdot & \dots & \cdot \\ \tilde{c}_{N1} & \dots & \tilde{c}_{NN} \end{pmatrix}$$

with

$$\tilde{c}_{uv} = \frac{1}{T} Re\left(\sum_r^{*} e_{(u)}(\lambda_r)\bar{e}_{(v)}(\lambda_r)'\hat{f}_{uv}(\lambda_r;\tilde{\tau})\hat{f}^{vu}(\lambda_r;\tilde{\tau})\right), \tag{38}$$

$$\tilde{D}^{2\prime} = \begin{pmatrix} \tilde{D}_1' \\ \tilde{D}_2' \\ \dots \\ \tilde{D}_N' \end{pmatrix} \tag{39}$$

with

$$\tilde{D}_u' = -\frac{1}{T} Re\left(\sum_r^{*} e_{(u)}(\lambda_r)\sum_{v=1}^{N} \hat{f}^{uv}(\lambda_r;\tilde{\tau})\frac{\partial \hat{f}_{vu}(\lambda_r;\tilde{\tau})}{\partial \tau'}\right), \tag{40}$$

and $\tilde{E}_{uv}$ remains unchanged, i.e. as in (29).

5. Wald tests

Once we have obtained the functional forms of the score test statistics, we can use and extend the derivations of previous sections to obtain representations of the tests based on the Wald and likelihood-ratio principles. In this section we concentrate on Wald tests, and present functional forms of the three cases studied before, i.e., the time domain and the frequency domain versions of the tests for white noise U_t, and the frequency domain representation when U_t is weakly parametrically autocorrelated.

5.a Wald test for white noise U_t

Here we describe a Wald test for the null hypothesis (4) in the model (1), (2), and (5) under the presumption that U_t in (5) is a vector sequence of zero mean uncorrelated random variables, with unknown variance-covariance matrix K. Recalling from Section 2, $\eta = (\theta',\delta',\alpha')'$, $L(\eta)$ is the negative of the log-likelihood based on Gaussian U_t, (with a minimum at $\eta = \bar{\eta}$), and given the asymptotic block diagonality of the second derivative matrix of $L(\eta)$, (see (A13) in Appendix 1), a general form of the Wald test can be written as

$$\bar{\theta}' \, E\left(\frac{\partial^2 L(\bar{\eta})}{\partial\theta \, \partial\theta'} \right) \bar{\theta}, \tag{41}$$

though any other $T^{1/2}$-consistent estimate of η, under (4) could also be adopted in (41).

We start by specifying the test statistic in its time domain representation. Denoting $\dot{\eta}$ any admissible value of η, the negative of the log-likelihood, apart from a constant, can be expressed as

$$L^t(\dot{\eta}) \;=\; \frac{T}{2}\log\det(K(\dot{\alpha})) + \frac{1}{2}\sum_{t=1}^{T} U_t(\dot{\theta},\dot{\delta})' K(\dot{\alpha})^{-1} U_t(\dot{\theta},\dot{\delta}) \tag{42}$$

where $U_t(\theta,\delta) = \Phi(L;\theta)\, Y_t - W_t(\delta)$, and the supscript 't' on $L(\eta)$ indicates the time domain form of the log-likelihood. By the same arguments as those given in Appendix 1 it can be shown that

$$\frac{\partial^2 L^t(\dot{\eta})}{\partial\dot{\theta}\,\partial\dot{\theta}'} \;=\; \sum_{u=1}^{N}\sum_{s=1}^{T-1} \psi_s^{(u)} \sum_{t=1}^{T-s}\left(\sum_{m=1}^{\infty} \psi_m^{(u)'} U_{u,t-m}(\dot{\theta},\dot{\delta})\right) \sum_{v=1}^{N} \dot{\sigma}^{uv} U_{v,t+s}(\dot{\theta},\dot{\delta}) \;+$$

$$\sum_{u=1}^{N}\sum_{s=1}^{T-1} \psi_s^{(u)} \sum_{t=1}^{T-s} U_{u,t}(\dot{\theta},\dot{\delta}) \sum_{v=1}^{N} \dot{\sigma}^{uv} \sum_{m=1}^{\infty} \psi_m^{(v)'} U_{v,t+s-m}(\dot{\theta},\dot{\delta}),$$

where $\dot{\sigma}^{uv} = \{K(\dot{\alpha})^{-1}\}_{uv}$ and $U_{ut}(\dot{\theta},\dot{\delta})$ is the u^{th} element of $U_t(\dot{\theta},\dot{\delta})$. Taking now the expectation in this last expression, evaluated under the null (4) and at $\delta = \dot{\delta}$, it is zero for the first summand, and for the second term becomes

$$(T-s)\sum_{u=1}^{N}\sum_{v=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\psi_s^{(v)\prime}\ \hat{\sigma}^{uv}\ E(\hat{U}_{u,t}(\delta)\hat{U}_{v,t}(\delta))\ =$$

$$=\ T\sum_{u=1}^{N}\sum_{v=1}^{N}\hat{\sigma}^{uv}\hat{\sigma}_{uv}\sum_{s=1}^{T-1}\left(1-\frac{s}{T}\right)\psi_s^{(u)}\psi_s^{(v)\prime}, \tag{43}$$

given the uncorrelatedness in U_t.

Substituting now (43) in (41), we obtain that a Wald test statistic in the time domain context takes the form

$$\hat{W}^t\ =\ T\,\hat{\theta}^{t\prime}\,\hat{A}^t\,\hat{\theta}^t, \tag{44}$$

where $\hat{\theta}^t$ is obtained throughout the minimization of $L^t(\eta)$ in (42), using $T^{1/2}$ - consistent estimates $\hat{\delta}$ and $\hat{\alpha}$, under the null hypothesis (4), and

$$\hat{A}^t\ =\ \sum_{s=1}^{T-1}\left(1-\frac{s}{T}\right)\sum_{u=1}^{N}\sum_{v=1}^{N}\hat{\sigma}^{uv}\hat{\sigma}_{uv}\psi_s^{(u)}\psi_s^{(v)\prime},$$

that is, adopting the same form as in (14).

For the frequency domain version of the test statistic, we can approximate the negative of the log-likelihood function as

$$L^f(\dot{\eta})\ =\ \frac{T}{2}\log\det\left(\frac{1}{2\pi}K(\dot{\alpha})\right)\ +\ \pi\sum_r^{*} tr\left[K(\dot{\alpha})^{-1}I_U(\lambda_r;\dot{\theta};\dot{\delta})\right], \tag{45}$$

where $I_U(\lambda_r,\dot{\theta},\delta)$ is now the cross-periodogram of $U_t(\dot{\theta},\delta)$ evaluated at $\lambda_r=2\pi r/T$.

Starting with the derivation with respect to $\dot{\theta}$,

$$\frac{\partial L^f(\dot{\eta})}{\partial\dot{\theta}}\ =\ \frac{\partial}{\partial\dot{\theta}}\left(\pi\sum_r^{*}\ tr\left[K(\dot{\alpha})^{-1}I_U(\lambda_r;\dot{\theta};\dot{\delta})\right]\right)\ =$$

$$\pi\sum_r^{*}\left(\frac{\partial}{\partial\dot{\theta}}vec'\left(I_U(\lambda_r;\dot{\theta};\dot{\delta})\right)\right)vec(K(\dot{\alpha})^{-1})\ =\ \pi\sum_r^{*}\sum_{u=1}^{N}\sum_{v=1}^{N}\frac{\partial I_{uv}(\lambda_r;\dot{\theta};\dot{\delta})}{\partial\dot{\theta}}\dot{\sigma}^{vu},$$

and using the same arguments as in Appendix 2, under suitable conditions, this last expression becomes asymptotically

$$\pi\sum_r^{*}\sum_{u=1}^{N}\sum_{v=1}^{N}\left(\epsilon_{(u)}(\lambda_r)\ +\ \bar{\epsilon}_{(v)}(\lambda_r)\right)I_{uv}(\lambda_r;\dot{\theta};\dot{\delta})\,\dot{\sigma}^{vu},$$

and thus, $\partial^2 L^f(\eta)/\partial\dot{\theta}\,\partial\dot{\theta}'$ evaluated at $\dot{\theta}=0$ and at $\delta=\hat{\delta}$ becomes asymptotically

$$\pi \sum_{r}^{*} \sum_{u=1}^{N} \sum_{v=1}^{N} (\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r))\,(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r))'\, I_{uv}(\lambda_r;\hat{\delta})\,\hat{\sigma}^{vu},$$

whose expectation for large T will be given by

$$\frac{1}{2} \sum_{r}^{*} \sum_{u=1}^{N} \sum_{v=1}^{N} (\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r))\,(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r))'\ \hat{\sigma}_{uv}\,\hat{\sigma}^{vu}.$$

Therefore, a Wald test statistic in this context will adopt the form

$$\hat{W}^f = T\,\hat{\theta}^{f\prime}\,\hat{A}^f\,\hat{\theta}^f \tag{46}$$

where $\hat{\theta}^f$ is obtained now throughout the mininimization of $L^f(\eta)$ in (45) with $T^{1/2}$ - consistent estimates $\hat{\delta}$ and $\hat{\alpha}$ under the null, and

$$\hat{A}^f = \frac{1}{2T} \sum_{r}^{*} \sum_{u=1}^{N} \sum_{v=1}^{N} \left(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r)\right) \left(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r)\right)'\ \hat{\sigma}^{uv}\,\hat{\sigma}_{uv}$$

$$= \frac{1}{2T} \sum_{r}^{*} \sum_{u=1}^{N} \sum_{v=1}^{N} \left(\epsilon_{(u)}(\lambda_r)\,\bar{\epsilon}_{(v)}(\lambda_r)' + \bar{\epsilon}_{(v)}(\lambda_r)\,\epsilon_{(u)}(\lambda_r)'\right)\ \hat{\sigma}^{uv}\ \hat{\sigma}_{uv},$$

by the same arguments as those given in previous sections.

5.b Wald test for weakly parametrically correlated U_t

Analogously to what we did for the score test, we can now robustify the test statistic in (46), to allow for weak parametric autocorrelation in U_t. We take U_t as in (19) and again here, the same conditions as those given in Section 3 and Appendix 2 will be required on U_t to obtain the test statistic. Recalling η from Section 3, the Wald test in this context will take the form

$$\theta'\,F^{\theta\theta}\,\theta\ \Big|_{\eta=\bar{\eta}}$$

where $\bar{\eta}$ is the value that minimizes $L(\eta)$ in Appendix 2, though again any other $T^{1/2}$-consistent estimate can be adopted, and

$$F^{\theta\theta} = F_{\theta\theta} - F_{\theta\tau}\,F_{\tau\tau}^{-1}\,F_{\tau\theta}$$

$$\textit{where}\quad F = \begin{pmatrix} F_{\theta\theta} & F_{\theta\tau} \\ F_{\tau\theta} & F_{\tau\tau} \end{pmatrix}$$

is the expected information matrix. Now, given the derivations carried out in Appendix 2, a Wald test statistic will adopt the form

$$\tilde{W} \;=\; T\,\tilde{\theta}'\,(\tilde{C} - \tilde{D}'\,(\tilde{E})^{-1}\,\tilde{D}\,)\,\tilde{\theta} \qquad (47)$$

with $\tilde{C}$,$\tilde{D}$ and $\tilde{E}$ as in (21), $\tilde{\tau}$ as in (22) and $\tilde{\theta}$ obtained by minimizing $L(\eta)$ in (B4) in Appendix 2 with $\tau = \tilde{\tau}$.

5.c Particular cases

We can stress the two cases of interest mentioned in Section 4. First, we consider θ is exactly the same parameter vector across all equations in (5). The test statistic for white noise U_t in the time domain representation takes the form

$$\hat{W}^{t^1} \;=\; T\,\hat{\theta}^{t^1\prime}\,(\hat{A}^{t^1})\,\hat{\theta}^{t^1} \qquad (48)$$

$$\textit{with} \quad \hat{A}^{t^1} \;=\; N\sum_{s=1}^{T-1}\left(1-\frac{s}{T}\right)\psi_s\,\psi_s',$$

and $\hat{\theta}^{t1}$ as in (44) but minimizing $L^t(\eta)$ with $\Phi(L;\theta)$ as defined in Section 4.a.

The frequency domain version is

$$\hat{W}^{f^1} \;=\; T\,\hat{\theta}^{f^1\prime}\,(\hat{A}^{f^1})\,\hat{\theta}^{f^1} \qquad (49)$$

$$\textit{with} \;\; \hat{\theta}^{f^1} \; \textit{as in} \; (46), \quad \textit{and} \quad \hat{A}^{f^1} \;=\; \frac{2N}{T}\sum_{r}^{*}\psi(\lambda_r)\,\psi(\lambda_r)',$$

and if U_t is weakly parametrically autocorrelated, the test statistic becomes

$$\tilde{W}^1 \;=\; T\,\tilde{\theta}^{1\prime}\,(\tilde{C}^1 - \tilde{D}^{1\prime}\,(\tilde{E})^{-1}\,\tilde{D}^1)\,\tilde{\theta}^1 \qquad (50)$$

with $\tilde{\theta}^1$ as in (47), and $\tilde{C}^1$, $\tilde{D}^1$ and $\tilde{E}$ as in (27), (28) and (29) respectively.

Finally, we consider the different versions of the test statistics when we take the parameter vector θ to be strictly different for each equation in (5). The time domain representation for white noise U_t is

$$\hat{W}^{t^2} \;=\; T\,\hat{\theta}^{t^2\prime}\,(\hat{A}^{t^2})\,\hat{\theta}^{t^2} \qquad (51)$$

with $\hat{\theta}^{t2}$ as in (44), i.e. minimizing (42) under the null hypothesis (4) and using now the new matrix $\Phi(L;\theta)$ specified in (30) and $\hat{A}^{t2}$ as in (33). The frequency

The frequency domain version of the test statistic is

$$\hat{W}^{f2} = T\, \hat{\theta}^{f2\prime}\, (\hat{A}^{f2})\, \hat{\theta}^{f2} \tag{52}$$

with $\hat{\theta}^{f2}$ as in (46) and $\hat{A}^{f2}$ as in (35) and (36); and finally, if U_t is weakly parametrically autocorrelated, the test statistic becomes

$$\tilde{W}^{2} = T\, \tilde{\theta}^{2\prime}\, (\tilde{C}^{2} - \tilde{D}^{2\prime}\, (\tilde{E})^{-1}\, \tilde{D}^{2})\, \tilde{\theta}^{2} \tag{53}$$

with $\tilde{\theta}^2$ as in (47) and $\tilde{C}^2$, $\tilde{D}^2$ and $\tilde{E}$ as in (38-40).

6. Likelihood ratio tests

We can also compute pseudo likelihood ratio test statistics under the same situations as in previous sections. Starting with the case of white noise U_t, a pseudo log-likelihood ratio test will adopt the form

$$LR = 2\,(L(\hat{\eta}) - L(\bar{\eta}))$$

where $L(\eta)$ is the negative of the log-likelihood; $\hat{\eta} = (0'; \hat{\delta}'; \hat{\alpha}')'$ as in Section 2, and $\bar{\eta} = (\hat{\theta}'; \hat{\delta}; \bar{\alpha}')'$, where $\hat{\theta}$ minimizes $L(\theta'; \hat{\delta}'; \hat{\alpha}')$ and $\bar{\alpha}$ is obtained using $\hat{\theta}$ and $\hat{\delta}$. First we concentrate on the time domain version of the test. From previous sections, we can write

$$L^{t}(\hat{\eta}) = \frac{T}{2}\log\det K(\hat{\alpha}) + \frac{1}{2}\sum_{t=1}^{T} \hat{U}_t(\delta)' K(\hat{\alpha})^{-1} \hat{U}_t(\delta)$$

$$= \frac{T}{2}\log\det K(\hat{\alpha}) + \frac{T}{2} tr\left[K(\hat{\alpha})^{-1} \frac{1}{T}\sum_{t=1}^{T} \hat{U}_t(\delta)\, \hat{U}_t(\delta)'\right]$$

$$= \frac{T}{2}\log\det K(\hat{\alpha}) + \frac{NT}{2}, \tag{54}$$

and similarly,

$$L^{t}(\bar{\eta}) = \frac{T}{2}\log\det K(\bar{\alpha}) + \frac{1}{2}\sum_{t=1}^{T} U_t(\hat{\theta}^t,\hat{\delta})' K(\bar{\alpha})^{-1} U_t(\hat{\theta}^t,\hat{\delta})$$

$$= \frac{T}{2}\log\det K(\bar{\alpha}) + \frac{T}{2} tr\left[K(\bar{\alpha})^{-1} \frac{1}{T}\sum_{t=1}^{T} U_t(\hat{\theta}^t,\hat{\delta})\, U_t(\hat{\theta}^t,\hat{\delta})'\right]$$

$$= \frac{T}{2}\log\det K(\bar{\alpha}) + \frac{NT}{2}. \tag{55}$$

Using (54) and (55), we can write a pseudo log-likelihood ratio test statistic as

$$LR^t = T\log\frac{\det K(\hat{\alpha})}{\det K(\bar{\alpha})} \tag{56}$$

where

$$K(\hat{\alpha}) = \frac{1}{T}\sum_{t=1}^{T}\hat{U}_t(\delta)\,\hat{U}_t(\delta)',$$

$\hat{U}_t(\delta) = U_t(\hat{\delta})$, and $\hat{\delta}$ is as given in Section 2 (i.e., a $T^{1/2}$-consistent estimate of δ under the null hypothesis) , and

$$K(\bar{\alpha}) = \frac{1}{T}\sum_{t=1}^{T}U_t(\hat{\theta}^t,\hat{\delta})\,U_t(\hat{\theta}^t,\hat{\delta})'$$

and $\hat{\theta}^t$ obtained throughout the minimization of $L^t(\eta)$ based on $\hat{\delta}$ and $\hat{\alpha}$.

Similarly, we can derive the test statistic in its frequency domain representation. Again from previous sections we have that

$$L^f(\hat{\eta}) = \frac{T}{2}\log\det\left(\frac{1}{2\pi}K(\hat{\alpha})\right) + \pi\sum_{r}^{*} tr\,[K(\hat{\alpha})^{-1}I_U(\lambda_r;\hat{\delta})]$$

$$= \frac{-NT}{2}\log 2\pi + \frac{T}{2}\log\det K(\hat{\alpha}) + \pi\,tr\left[K(\hat{\alpha})^{-1}\sum_{r}^{*} I_U(\lambda_r;\hat{\delta})\right]$$

$$= C + \frac{T}{2}\log\det(K(\hat{\alpha}), \quad where \quad C = \frac{NT}{2}(1-\log 2\pi)$$

and similarly,

$$L^f(\bar{\eta}) = C + \frac{T}{2}\log\det K(\bar{\alpha}).$$

Thus, a pseudo LR test statistic in this context can be approximated by

$$LR^f = T\log\frac{\det K(\hat{\alpha})}{\det K(\bar{\alpha})} \tag{57}$$

where now

$$K(\hat{\alpha}) \quad = \quad \frac{2\pi}{T}\sum_{r}^{*} I_U(\lambda_r;\hat{\delta})$$

and

$$K(\bar{\alpha}) \quad = \quad \frac{2\pi}{T}\sum_{r}^{*} I_U(\lambda_r;\hat{\theta}^f;\hat{\delta}),$$

and $\hat{\theta}^f$ minimizes the frequency domain version of the log-likelihood based on δ and $\hat{\alpha}$.

Extending the tests for weak parametric autocorrelation in U_t, the test statistic takes the form

$$\overline{LR} \quad = \quad T \log \frac{\det K(\hat{\alpha})}{\det K(\bar{\alpha})} \tag{58}$$

where

$$K(\hat{\alpha}) \quad = \quad \frac{2\pi}{T}\sum_{r}^{*} tr[\hat{f}(\lambda_r;\tilde{\tau})^{-1}\; I_U(\lambda_r;\hat{\delta})]$$

and

$$K(\bar{\alpha}) \quad = \quad \frac{2\pi}{T}\sum_{r}^{*} tr[\hat{f}(\lambda_r;\tilde{\tau})^{-1}\; I_U(\lambda_r;\tilde{\theta}^f;\hat{\delta})],$$

with $\hat{f}$, $\tilde{\tau}$, $\tilde{\theta}^f$ and $\hat{\delta}$ as they were given in all previous pages.

Finally, for the two particular cases considered in Section 4, the test statistics will take the same form as in (56), (57) and (58) with the only difference in the specification of the matrix $\Phi(L;\theta)$ appearing in (5).

7. Finite sample performace

In this final section we examine the finite sample behaviour of some of the test statistics presented in previous sections, by means of Monte Carlo simulations. All calculations were carried out using Fortran and the NAG's library random number generator, on LSE's VAX computer. Given the variety of tests and the number of possibilities covered by them, we concentrate on a bivariate model where the null hypothesis will be two time series following a random walk. We will consider a model of form

$$\begin{pmatrix} (1-L)^{1+\theta_1} & 0 \\ 0 & (1-L)^{1+\theta_2} \end{pmatrix} \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} U_{1t} \\ U_{2t} \end{pmatrix} \qquad t = 1,2,\ldots,T \tag{59}$$

$$\mathbf{X_t = (X_{1t}, X_{2t})' = 0 \quad \text{for all } t \leq 0,} \tag{60}$$

where under the null hypothesis given by:

$$\mathbf{H_o: \theta = (\theta_1, \theta_2)' = 0,} \tag{61}$$

$U_t = (U_{1t}, U_{2t})'$ will be initially, a white noise vector process with mean zero and variance-covariance matrix Σ. First, and without loss of generality, we assume that $\Sigma = I_2$, but we also present results, for a given positive definite matrix Σ, in order to check if the test statistics are robust for a different specification of Σ. We look first at rejection frequencies of the score test statistic given in (32), for fractional alternatives, where $(\theta_i)_{i=1,2}$ in (59) takes values: -0.8; -0.6; -0.4; -0.2; 0; 0.2; 0.4; 0.6 and 0.8. Then, we generate Gaussian series for different sample sizes (50, 100 and 200 observations) taking 5000 replications of each case, and present results for four different nominal sizes: 10%, 5%, 2.5% and 1%. The reason for focusing on the test statistic given in (32), (i.e., the time domain version), rather than in its frequency domain representation (i.e., (34)), is that the latter form of the test statistic is much more expensive computationally in terms of CPU time. We know that in finite samples, the results of the two test statistics can vary substantially, though asymptotically the difference will be negligible.

In Table 1 we present rejection frequencies of the test statistic $\hat{S}^{t2}$ in (32) when $\Sigma = I_2$, for three different sample sizes (T = 50, 100 and 200) and a nominal size of 10%. Tables 2-4 are similar to Table 1 but with nominal sizes of 5%, 2.5% and 1%, respectively. Looking across these tables, we see that the sizes of the tests are too small in all cases, however they tend to improve as we increase the number of observations. For example, we observe in Table 1 (α = 10%), that when the sample size is 50, the size is 3.3%, but increases to 5.3% when T = 100, and to 7.2% when T = 200. Similarly in Table 2 (α = 5%), the

(Tables 1 - 4 about here)

sizes are 1.2% for T = 50, 2.0% for T = 100, and 3.2% for T = 200. The same behaviour is observed in Tables 3 and 4, with all sizes smaller than nominal ones but increasing with the number of observations. If we concentrate now on small departures from the null (61), we observe that these rejection frequencies increase strongly, especially when the sample size is large (e.g. T = 200). This increase is more marked when θ_1 and θ_2 take the same value, though it is also noticeable when θ_1 and θ_2 are different. In Table 1c (T = 200, α = 10%) we see that the lowest rejection probability, apart from that of the true model ($\theta_1 = \theta_2 = 0$), is 0.827 which is obtained when $\theta_1 = 0$ and $\theta_2 = -0.2$, and becomes 0.993 when $\theta_1 = \theta_2 = -0.2$. Similarly in Table 2c, (when T = 200 and α = 5%), the values for the same alternatives are 0.671 and 0.997; in Table 3c (α = 2.5%) are 0.495 and 0.941, and in Table 4c (α = 1%) 0.279 and 0.848.

Another remarkable feature of these results is that when the sample size is small (e.g. T = 50), it seems that there is a bias toward positive values of θ_1 and θ_2. This bias is especially clear when the nominal size is also small. We can see through Tables 2a, 3a and 4a that if θ_1 and θ_2 are both greater than or equal to 0, rejection frequencies are always greater than those obtained when the values of θ_1 and θ_2 were less than or equal to 0. Taking nominal sizes of 2.5% and 1%, this bias also appears for a sample size of 100 observations (Tables 3b and 4b); however, increasing the sample size to 200 observations, the bias tends to disappear. A particularly poor result is obtained in Table 4a (T =50; α = 1%), when θ_1 (or θ_2) is equal to 0 and θ_2 (or θ_1) is negative. In such situations, the rejection probabilities never exceed 0.100. Again these results improve considerably when we increase the sample size to 100 or 200 observations (Tables 4b and 4c). Finally we observe that in all cases, rejection frequencies increase with absolute value of θ and with sample size T, and when T = 200, the rejection probability of 1 is obtained in most of the cases when $|\theta_i|_{i=1,2} \geq 0.4$ for α = 10% and 5%, and when $|\theta_i|_{i=1,2} \geq 0.6$ for α = 2.5% and 1%.

Tables 5-8 report rejection frequencies of the same statistic as above, but now we take Σ as a positive definite matrix of form: [(1,1)'; (1,2)']. In doing so, we can see if the test statistic is robust to a different specification of the variance-covariance matrix of the differenced residuals. Table 5 is the counterpart of Table 1 for the new variance-covariance matrix Σ. Similarly, Tables 6-8 corresponds to Tables 2-4 above. We observe now that sizes are slightly greater than before, but again too small with respect to nominal ones though increasing with the sample size T. In Table 5 (α = 10%), we see that

(Tables 5 - 8 about here)

sizes are now 3.9% for T = 50; 6.1% for T = 100; and 7.5% for T = 200. Across Tables 6-8 we see that in five cases (Tables 6c, 7b, 7c, 8a and 8c), sizes are the same as when $\Sigma = I_2$, while in the other four cases (Tables 6a, 6b, 7a and 8b) they are slightly greater, but not exceeding in 0.02% those results obtained across Tables 2-4. A bias for positive values of θ_1 and θ_2 is again observed when nominal sizes and sample sizes are small; however, the pathological cases observed in Table 4a have now disappeared (Table 8a). All rejection frequencies increase with sample size T, but in a few cases, we now observe a lack of monotonicity of these rejections with respect to $(\theta_i)_{i=1,2}$, when the sample size is small and $(\theta_i)_{i=1,2}$ takes low values. Comparing these results in Tables 5-8 with those obtained in Tables 1-4, we see that in most of the cases, rejection frequencies are now slightly greater, but in general, results are similar across all tables, suggesting that the test statistic is not affected much by the different specifications of the variance-covariance matrix Σ.

In Tables 9 and 10 we present empirical sizes of the test in the frequency domain representation. Table 9 reports sizes of the test statistic $\hat{S}^{f2}$ in (34), assuming first, in Table 9a, that $\Sigma = I_2$, while in Table 9b we take Σ = [(1,1)'; (1,2)']. As in all previous tables, we see that sizes are very small when T = 50, however they improve considerably when we increase the sample size. Comparing empirical sizes in Table 9a with those in Tables 1-4, we see that they are very similar. When T = 50 the sizes are now slightly smaller than in the time domain versions of the tests, but when T = 100 or 200, they are slightly greater. We should mention here that results obtained in Table 9 (and also in Table 10) have been obtained using 1000 replications of each case, (unlike the 5000 replications used in Tables 1-8). Therefore the difference may be largely due to the different number of replications used. When $\Sigma \neq I_2$ (Table 9b) the

(Tables 9 and 10 about here)

same conclusions hold, with empirical sizes smaller than nominal ones but increasing with T, and observing few differences with respect to empirical sizes obtained in the time domain representation of the tests across Tables 5-8. Comparing results in Table 9b with those obtained in Table 9a, we again observe few differences, with the highest one occurring when T = 50 and α = 10%; in this case, the empirical sizes are 2.8% in Table 9a and 3.6% in Table 9b, while in the remaining cases, the differences are not greater than 0.03% between both tables.

Finally, Table 10 reports sizes for the test statistic $\tilde{S}^2$ in (37), i.e., the frequency domain representation of the test when U_t is weakly parametrically autocorrelated. In Table 10a we assume that U_t follows a VAR(1) representation, and we choose the parameterization

$$\begin{pmatrix} U_{1t} \\ U_{2t} \end{pmatrix} = \begin{pmatrix} 0.5 & 0.2 \\ 0.3 & 0.5 \end{pmatrix} \begin{pmatrix} U_{1t-1} \\ U_{2t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}, \tag{62}$$

where ε_t is normally distributed with mean zero and variance-covariance matrix I_2. In Table 10b we consider a VMA(1) structure on U_t using the same parameters as in the VAR(1) case. That is,

$$\begin{pmatrix} U_{1t} \\ U_{2t} \end{pmatrix} = \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} + \begin{pmatrix} 0.5 & 0.2 \\ 0.3 & 0.5 \end{pmatrix} \begin{pmatrix} \epsilon_{1t-1} \\ \epsilon_{2t-1} \end{pmatrix}, \tag{63}$$

and again ε_t normally distributed with mean 0 and variance I_2.

In both tables we see that sizes are now too large for all nominal sizes, especially when T = 50, however, as we increase the number of observations, these empirical sizes reduce and then tend to approximate to nominal ones. Thus, for the VAR(1) case (Table 10a), we see that if the number of observations is 200, the sizes are 10.4% for $\alpha = 10\%$; 6.0% for $\alpha = 5\%$; 3.1% for $\alpha = 2.5\%$; and 1.2% for $\alpha = 1\%$. When the VMA(1) structure is considered (Table 10b), empirical sizes are now slightly greater than in the VAR(1) case, but again we observe a considerable improvement when we increase the number of observations. Similar results were obtained when we used different parameters in (62) and (63) and a different variance-covariance matrix for the residuals ε_t.

As a conclusion, we can summarize the results obtained across these tables by saying that the score test statistics obtained in sections 2 - 5 seem to be adequate to test the null hypothesis of a random walk in this bivariate context. Though sizes are smaller than nominal ones in most of the cases, the performance of these tests seems quite good even for small departures of the null hypothesis (61), especially as we increase the number of observations.

References

Brillinger, D. R. (1981). Time series. Data analysis and theory. The University of California, Berkeley, McGraw-Hill, Inc.

Johansen, S. (1988). Statistical analysis of cointegration vectors. J. Econom. Dynamics Control **12,** 231-254.

Rao, C. R. (1973). Linear Statistical Inference and its applications. Wiley, New York.

Robinson, P. M. (1972). Non-linear regression for multiple time series. J. Appl. Probab. **9,** 758-768.

Robinson, P. M. (1977). The construction and estimation of continuous time models and discrete approximations in econometrics. J. Econometrics **6,** 173-197.

Robinson, P. M. (1994). Efficient tests of nonstationary hypotheses. J. Amer. Statist. Assoc. **89,** 1420-1437.

Zygmund, A. (1979). Trygonometric Series, Vol 1. Cambridge University Press, Cambridge.

Appendix 1: Derivation of the score statistic $\hat{S}^t$

The negative of the log-likelihood under (1), (2), (5), and Gaussianity of U_t, can be expressed, apart from a constant as

$$L(\theta,\dot{\delta},\dot{\alpha}) = \frac{T}{2}\log\det K(\dot{\alpha}) + \frac{1}{2}\sum_{t=1}^{T} U_t(\theta,\dot{\delta})' K(\dot{\alpha})^{-1} U_t(\theta,\dot{\delta})$$

$$= \frac{T}{2}\log\det K(\dot{\alpha}) + \frac{1}{2}\sum_{t=1}^{T} X_t(\dot{\delta})'\Phi(L;\theta)\, K(\dot{\alpha})^{-1}\,\Phi(L;\theta) X_t(\dot{\delta}), \qquad (A1)$$

for any admissible $\dot{\alpha}$ and $\dot{\delta}$, where $U_t(\theta,\dot{\delta}) = \Phi(L;\theta)X_t(\dot{\delta})$ and $X_t(\dot{\delta}) = Y_t - Z_t(\dot{\delta})$.

Starting with the first derivatives in (11),

$$\frac{\partial L(\theta,\dot{\delta},\dot{\alpha})}{\partial\theta} = \frac{\partial}{\partial\theta}\left[\frac{1}{2}\sum_{t=1}^{T} X_t(\dot{\delta})'\Phi(L;\theta) K(\dot{\alpha})^{-1}\Phi(L;\theta) X_t(\dot{\delta})\right]$$

$$= \sum_{t=1}^{T}\left[\frac{\partial\rho_1(L;\theta)}{\partial\theta}X_{1t}(\dot{\delta}); \dots ; \frac{\partial\rho_N(L;\theta)}{\partial\theta}X_{Nt}(\dot{\delta})\right] K(\dot{\alpha})^{-1} U_t(\theta,\dot{\delta})$$

$$= \sum_{t=1}^{T}\left[\frac{\partial\log\rho_1(L;\theta)}{\partial\theta}U_{1t}(\theta,\dot{\delta}); \dots ; \frac{\partial\log\rho_N(L;\theta)}{\partial\theta}U_{Nt}(\theta,\dot{\delta})\right] K(\dot{\alpha})^{-1} U_t(\theta,\dot{\delta})$$

where $U_t(\theta,\dot{\delta}) = (U_{1t}(\theta,\dot{\delta});..;U_{Nt}(\theta,\dot{\delta}))'$ and $X_t(\dot{\delta}) = (X_{1t}(\dot{\delta});..;X_{Nt}(\dot{\delta}))'$, and evaluating now this last expression at $\theta = 0$ we obtain

$$\sum_{t=1}^{T}\left[\epsilon_{(1)}(L)\, U_{1t}(\dot{\delta}); \dots ; \epsilon_{(N)}(L)\, U_{Nt}(\dot{\delta})\right] K(\dot{\alpha})^{-1} U_t(\dot{\delta}) \qquad (A2)$$

where $\epsilon_{(u)}(L) = \dfrac{\partial\log\rho_u(L;\theta)}{\partial\theta}$ *can be expanded as* $\sum_{s=1}^{\infty}\psi_s^{(u)} L^s$,

in view of (8) and below, and the expression in (A2) becomes

$$\sum_{t=1}^{T}\left[\left(\sum_{s=1}^{\infty}\psi_s^{(1)} U_{1,t-s}(\dot{\delta})\right); \dots ;\left(\sum_{s=1}^{\infty}\psi_s^{(N)} U_{N,t-s}(\dot{\delta})\right)\right] K(\dot{\alpha})^{-1} U_t(\dot{\delta})$$

$$= \sum_{t=1}^{T}\left[\left(\sum_{s=1}^{\infty}\psi_s^{(1)}U_{1,t-s}(\dot{\delta})\right);\ldots;\left(\sum_{s=1}^{\infty}\psi_s^{(N)}U_{N,t-s}(\dot{\delta})\right)\right]\begin{bmatrix}\sum_{v=1}^{N}\dot{\sigma}^{1v}U_{vt}(\dot{\delta})\\ --\\ \sum_{v=1}^{N}\dot{\sigma}^{Nv}U_{vt}(\dot{\delta})\end{bmatrix} \qquad (A3)$$

$$=\sum_{u=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\sum_{t=1}^{T-s}U_{ut}(\dot{\delta})\sum_{v=1}^{N}\dot{\sigma}^{uv}U_{v,t+s}(\dot{\delta}) = T\sum_{\substack{u=1\\v=1}}^{N}\dot{\sigma}^{uv}\sum_{s=1}^{T-1}\psi_s^{(u)}C_{uv}(s,\dot{\delta}), \qquad (A4)$$

where $\dot{\sigma}^{uv}$ is the $(u,v)^{th}$ element of $K(\dot{\alpha})^{-1}$ and

$$C_{uv}(s,\dot{\delta}) = \frac{1}{T}\sum_{t=1}^{T-s}U_{u,t}(\dot{\delta})\,U_{v,t+s}(\dot{\delta}).$$

Calling $L_o = L(\dot{\eta})_{\theta=0}$, the first derivative with respect to δ is

$$\frac{\partial L_o}{\partial\dot{\delta}} = \frac{\partial}{\partial\dot{\delta}}\left[\frac{1}{2}\sum_{t=1}^{T}\left(Y_t-Z_t(\dot{\delta})\right)'\Phi(L)K(\dot{\alpha})^{-1}\Phi(L)\left(Y_t-Z_t(\dot{\delta})\right)\right]$$

$$= \frac{\partial}{\partial\dot{\delta}}[\frac{1}{2}\sum_{t=1}^{T}W_t(\dot{\delta})'K(\dot{\alpha})^{-1}W_t(\dot{\delta}) - \frac{1}{2}\sum_{t=1}^{T}W_t(\dot{\delta})'K(\dot{\alpha})^{-1}\Phi(L)Y_t -$$

$$\frac{1}{2}\sum_{t=1}^{T}Y_t'\,\Phi(L)\,K(\dot{\alpha})^{-1}\,W_t(\dot{\delta})]$$

$$= \frac{\partial}{\partial\dot{\delta}}\left[\frac{1}{2}\sum_{t=1}^{T}W_t(\dot{\delta})'K(\dot{\alpha})^{-1}W_t(\dot{\delta}) - \sum_{t=1}^{T}W_t(\dot{\delta})'K(\dot{\alpha})^{-1}\Phi(L)Y_t\right]$$

$$= \sum_{t=1}^{T}\frac{\partial W_t(\dot{\delta})}{\partial\dot{\delta}}K(\dot{\alpha})^{-1}W_t(\dot{\delta}) - \sum_{t=1}^{T}\frac{\partial W_t(\dot{\delta})}{\partial\dot{\delta}}K(\dot{\alpha})^{-1}\Phi(L)Y_t$$

$$= -\sum_{t=1}^{T}\frac{\partial W_t(\dot{\delta})}{\partial\dot{\delta}}K(\dot{\alpha})^{-1}U_t(\dot{\delta}). \qquad (A5)$$

From (A1) we have that L_o can be expressed as

$$\frac{T}{2}\log\det K(\dot\alpha) \ + \ \frac{1}{2}tr\left[K(\dot\alpha)^{-1}S(\dot\delta)\right],$$

where $S(\dot\delta) = \sum_{t=1}^{T} U_t(\dot\delta)\,U_t(\dot\delta)'$, and differentiating L_o with respect to $\dot\alpha$ leads to

$$\frac{T}{2}tr\,[K(\dot\alpha)^{-1}(dK(\dot\alpha))] \ - \ \frac{1}{2}tr\,[K(\dot\alpha)^{-1}(dK(\dot\alpha))\,K(\dot\alpha)^{-1}S(\dot\delta)] \qquad (A6)$$

$$= \ -\frac{1}{2}tr\ \left[(dK(\dot\alpha)\,K(\dot\alpha)^{-1}\ (S(\dot\delta) \ - \ TK(\dot\alpha))\ K(\dot\alpha)^{-1}\right]$$

$$= \ -\frac{1}{2}(vec\,(dK(\dot\alpha)))'\ \left(K(\dot\alpha)^{-1}\ \otimes\ K(\dot\alpha)^{-1}\right)\ vec\,(S(\dot\delta) - TK(\dot\alpha)\,)$$

$$= -\frac{1}{2}dv(K(\dot\alpha))'\,D_m'\left(K(\dot\alpha)^{-1}\ \otimes\ K(\dot\alpha)^{-1}\right)vec\,(S(\dot\delta) - TK(\dot\alpha)), \qquad (A7)$$

where D_m is the duplication matrix, and using the well known result that tr[ABCD] = (vec A)'(D' ⊗ B)(vec C). Then, from (A7) we easily observe that

$$\frac{\partial L_o}{\partial\dot\alpha} \ = \ -\frac{1}{2}D_m'\left(K(\dot\alpha)^{-1}\ \otimes\ K(\dot\alpha)^{-1}\right)\ vec\ \left(S(\dot\delta) - \ TK(\dot\alpha)\right). \qquad (A8)$$

Next we look at the second derivative matrices appearing in (11), and first concentrate on the (pxp) matrix $\partial^2 L_o/\partial\theta\,\partial\theta'$. From the left-side in the equality in (A4)

$$\frac{\partial L_o}{\partial\theta} \ = \ \sum_{u=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\sum_{t=1}^{T-s} U_{ut}(\dot\delta)\sum_{v=1}^{N}\dot\sigma^{uv}\,U_{v,t+s}(\dot\delta),$$

and then we have that

$$\frac{\partial^2 L_o}{\partial\theta\partial\theta'} \ = \ \sum_{u=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\sum_{t=1}^{T-s}\frac{\partial\log\rho_u(L;0)}{\partial\theta'}U_{ut}(\dot\delta)\sum_{v=1}^{N}\dot\sigma^{uv}U_{v,t+s}(\dot\delta) \ +$$

$$\sum_{u=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\sum_{t=1}^{T-s}U_{ut}(\dot\delta)\sum_{v=1}^{N}\dot\sigma^{uv}\frac{\partial\log\rho_v(L;0)}{\partial\theta'}U_{v,t+s}(\dot\delta)$$

$$= \sum_{u=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\sum_{t=1}^{T-s}\left(\sum_{m=1}^{\infty}\psi_m^{(u)'}U_{u,t-m}(\dot\delta)\right)\sum_{v=1}^{N}\dot\sigma^{uv}U_{v,t+s}(\dot\delta) \ +$$

$$\sum_{u=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\sum_{t=1}^{T-s}U_{ut}(\dot{\delta})\sum_{v=1}^{N}\dot{\sigma}^{uv}\left(\sum_{m=1}^{\infty}\psi_m^{(v)'}U_{v,t+s-m}(\dot{\delta})\right).$$

In order to form (11), we need to take the expectation of this last expression. (Note that it is evaluated at $\theta = 0$, i.e. under H_o (4)). It is zero for the first summand given the uncorrelatedness in U_t and since it involves terms of the form $U_{u,t-m}$ and $U_{v,t+s}$, for m,s > 0. The expectation of the second summand is

$$\sum_{u=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\sum_{v=1}^{N}\dot{\sigma}^{uv}\psi_s^{(v)'}\sum_{t=1}^{T-s}E\left(U_{ut}(\dot{\delta})\,U_{vt}(\dot{\delta})\right) =$$

$$\sum_{u=1}^{N}\sum_{v=1}^{N}\sum_{s=1}^{T-1}(T-s)\,\psi_s^{(u)}\,\psi_s^{(v)'}\,\dot{\sigma}^{uv}\,\dot{\sigma}_{uv} = T\sum_{s=1}^{T-1}\left(1-\frac{s}{T}\right)\sum_{u=1}^{N}\sum_{v=1}^{N}\dot{\sigma}^{uv}\,\dot{\sigma}_{uv}\,\psi_s^{(u)}\,\psi_s^{(v)'}.$$

Again from the first equality in (A4), we have that $\partial L_o/\partial\theta$ can be rewritten as

$$\sum_{u=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\sum_{t=1}^{T-s}\left(\rho_u(L)\,Y_{ut} - W_{ut}(\dot{\delta})\right)\sum_{v=1}^{N}\dot{\sigma}^{uv}\left(\rho_v(L)\,Y_{v,t+s} - W_{v,t+s}(\dot{\delta})\right),$$

and from this expression, we observe that

$$\frac{\partial^2 L_o}{\partial\theta\,\partial\dot{\delta}'} = \frac{\partial}{\partial\dot{\delta}'}[\sum_{u=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\sum_{t=1}^{T-s}(W_{ut}(\dot{\delta})\sum_{v=1}^{N}\dot{\sigma}^{uv}W_{v,t+s}(\dot{\delta}) - \rho_u(L)\,Y_{ut}\ x$$

$$x\ \sum_{v=1}^{N}\dot{\sigma}^{uv}\ W_{v,t+s}(\dot{\delta}) - W_{ut}(\dot{\delta})\sum_{v=1}^{N}\dot{\sigma}^{uv}\ \rho_v(L)\,Y_{v,t+s})] =$$

$$\sum_{u=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\sum_{t=1}^{T-s}(\frac{\partial W_{ut}(\dot{\delta})}{\partial\dot{\delta}'}\sum_{v=1}^{N}\dot{\sigma}^{uv}W_{v,t+s}(\dot{\delta}) + W_{ut}(\dot{\delta})\sum_{v=1}^{N}\dot{\sigma}^{uv}\frac{\partial W_{v,t+s}(\dot{\delta})}{\partial\dot{\delta}'}$$

$$-\ \rho_u(L)\,Y_{ut}\sum_{v=1}^{N}\dot{\sigma}^{uv}\frac{\partial W_{v,t+s}(\dot{\delta})}{\partial\dot{\delta}'} - \frac{\partial W_{ut}(\dot{\delta})}{\partial\dot{\delta}'}\sum_{v=1}^{N}\dot{\sigma}^{uv}\ \rho_v(L)\,Y_{v,t+s})$$

$$= -\left(\sum_{u=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\sum_{t=1}^{T-s}\left[\frac{\partial W_{ut}(\dot{\delta})}{\partial\dot{\delta}'}\sum_{v=1}^{N}\dot{\sigma}^{uv}U_{v,t+s}(\dot{\delta}) + U_{ut}(\dot{\delta})\sum_{v=1}^{N}\dot{\sigma}^{uv}\frac{\partial W_{v,t+s}(\dot{\delta})}{\partial\dot{\delta}'}\right]\right)$$

$$= -\sum_{u=1}^{N}\sum_{s=1}^{T-1}\psi_s^{(u)}\sum_{v=1}^{N}\dot{\sigma}^{uv}\sum_{t=1}^{T-s}\left[\frac{\partial W_{ut}(\dot{\delta})}{\partial\dot{\delta}'}U_{v,t+s}(\dot{\delta}) + U_{ut}(\dot{\delta})\frac{\partial W_{v,t+s}(\dot{\delta})}{\partial\dot{\delta}'}\right]. \qquad (A9)$$

For the derivation of $\partial^2 L_o/\partial\theta\partial\dot{\alpha}'$, we have that calling $P_t(\dot{\delta}) = [P_{1t}(\dot{\delta});...;P_{Nt}(\dot{\delta})]$ the (p x N) matrix appearing in (A3), then

$$\frac{\partial L_o}{\partial\theta} = \sum_{t=1}^{T} P_t(\dot{\delta})K(\dot{\alpha})^{-1}U_t(\dot{\delta}),$$ and differentiating this expression with respect to $\dot{\alpha}$:

$$-\sum_{t=1}^{T} P_t(\dot{\delta})\dot{K}^{-1}(d\dot{K})\dot{K}^{-1}U_t(\dot{\delta}) = -\sum_{t=1}^{T}\left(U_t(\dot{\delta})\otimes P_t(\dot{\delta})\right)vec(\dot{K}^{-1}(d\dot{K})\dot{K}^{-1})$$

$$= -\sum_{t=1}^{T}\left(U_t(\dot{\delta})\otimes P_t(\dot{\delta})\right)(\dot{K}^{-1}\otimes\dot{K}^{-1})D_m\, dv(\dot{K})$$

where $\dot{K} = K(\dot{\alpha})$, and therefore,

$$\frac{\partial^2 L_o}{\partial\theta\partial\dot{\alpha}'} = -\sum_{t=1}^{T}(U_t(\dot{\delta})\otimes P_t(\dot{\delta}))(K(\dot{\alpha})^{-1}\otimes K(\dot{\alpha})^{-1})D_m. \qquad (A10)$$

Finally in order to complete the Hessian in (11) we still have to calculate some second derivatives with respect to δ and $\dot{\alpha}$. From (A5)

$$\frac{\partial^2 L_o}{\partial\dot{\delta}\partial\dot{\delta}'} = \frac{\partial}{\partial\dot{\delta}'}\left(-\sum_{t=1}^{T}\frac{\partial W_t(\dot{\delta})}{\partial\dot{\delta}}K(\dot{\alpha})^{-1}(\Phi(L)Y_t - W_t(\dot{\delta}))\right) =$$

$$\sum_{t=1}^{T}\left(\frac{\partial W_t(\dot{\delta})}{\partial\dot{\delta}}K(\dot{\alpha})^{-1}\frac{\partial W_t(\dot{\delta})}{\partial\dot{\delta}'} - \left(U_t(\dot{\delta})'K(\dot{\alpha})^{-1}\otimes I_k\right)\frac{\partial vec}{\partial\dot{\delta}'}\left(\frac{\partial W_t(\dot{\delta})}{\partial\dot{\delta}}\right)\right). \qquad (A11)$$

Next we consider $\partial^2 L_o/\partial\delta\partial\dot{\alpha}'$, and since

$$\frac{\partial L_o}{\partial\dot{\delta}} = -\sum_{t=1}^{T}\frac{\partial W_t(\dot{\delta})}{\partial\dot{\delta}}K(\dot{\alpha})^{-1}U_t(\dot{\delta}),$$ differentiating this expression w.r.t. $\dot{\alpha}$,

$$\sum_{t=1}^{T} \frac{\partial W_t(\dot{\delta})}{\partial \dot{\delta}} K(\dot{\alpha})^{-1} (dK(\dot{\alpha}))\, K(\dot{\alpha})^{-1}\, U_t(\dot{\delta})$$

$$= \sum_{t=1}^{T} \left(U_t(\dot{\delta})' \otimes \frac{\partial W_t(\dot{\delta})}{\partial \dot{\delta}} \right) vec\,[K(\dot{\alpha})^{-1} (dK(\dot{\alpha}))\, K(\dot{\alpha})^{-1}]$$

$$= \sum_{t=1}^{T} \left(U_t(\dot{\delta})' \otimes \frac{\partial W_t(\dot{\delta})}{\partial \dot{\delta}} \right) (K(\dot{\alpha})^{-1} \otimes K(\dot{\alpha})^{-1})\, D_m\; dv(K(\dot{\alpha})),$$

and therefore,

$$\frac{\partial^2 L_o}{\partial \dot{\delta}\, \partial \dot{\alpha}'} = \sum_{t=1}^{T} \left(U_t(\dot{\delta})' \otimes \frac{\partial W_t(\dot{\delta})}{\partial \dot{\delta}} \right) (K(\dot{\alpha})^{-1} \otimes K(\dot{\alpha})^{-1})\, D_m. \qquad (A12)$$

The final term in (11) that we should look at is $\partial^2 L_o / \partial \dot{\alpha}\, \partial \dot{\alpha}'$. Differentiating (A6) with respect to $\dot{\alpha}$, and recalling again $\dot{K} = K(\dot{\alpha})$, we have

$$-\frac{T}{2} tr[\dot{K}^{-1}(d\dot{K})\dot{K}^{-1}(d\dot{K})] + \frac{1}{2} tr[\dot{K}^{-1}(d\dot{K})\dot{K}^{-1}(d\dot{K})\dot{K}^{-1} S(\dot{\delta})]$$

$$+ \frac{1}{2} tr[\dot{K}^{-1}(d\dot{K})\dot{K}^{-1}(d\dot{K})\dot{K}^{-1} S(\dot{\delta})]$$

$$= -\frac{T}{2} tr[(d\dot{K})\dot{K}^{-1}(d\dot{K})\dot{K}^{-1}] + tr[(d\dot{K})\dot{K}^{-1}(d\dot{K})\dot{K}^{-1} S(\dot{\delta})\dot{K}^{-1}]$$

$$= -\frac{T}{2} vec(d\dot{K})'(\dot{K}^{-1} \otimes \dot{K}^{-1})\, vec(d\dot{K}) + vec(d\dot{K})'(\dot{K}^{-1} S(\dot{\delta})\dot{K}^{-1} \otimes \dot{K}^{-1})\, vec(d\dot{K})$$

$$= -\frac{T}{2} dv(\dot{K})' D_m' (\dot{K}^{-1} \otimes \dot{K}^{-1}) D_m\, dv(\dot{K}) + dv(\dot{K})' D_m' (\dot{K}^{-1} S(\dot{\delta})\dot{K}^{-1} \otimes \dot{K}^{-1}) D_m\, dv(\dot{K}),$$

obtaining as a final expression for $\partial^2 L_o / \partial \dot{\alpha}\, \partial \dot{\alpha}'$

$$-\frac{T}{2} D_m' (K(\dot{\alpha})^{-1} \otimes K(\dot{\alpha})^{-1}) D_m + D_m' (K(\dot{\alpha})^{-1} S(\dot{\delta}) K(\dot{\alpha})^{-1} \otimes K(\dot{\alpha})^{-1}) D_m.$$

We can get now consistent and efficient estimates of δ and α by equating

(A5) and (A8) to zero; however, for practical purposes and in order to simplify the computations, we can take any $T^{1/2}$-consistent estimates of δ and α. We will assume that $\hat{\delta}$ is a consistent estimate of δ and we will take $\hat{K} = K(\hat{\alpha}) = T^{-1} S(\hat{\delta})$. It follows then from previous pages that

$$\frac{\partial L(0,\hat{\delta},\hat{\alpha})}{\partial \theta} = \sum_{u=1}^{N} \sum_{s=1}^{T-1} \psi_s^{(u)} \sum_{t=1}^{T-s} \hat{U}_{ut}(\delta) \sum_{v=1}^{N} \hat{\sigma}^{uv} \hat{U}_{v,t+s}(\delta) =$$

$$T \sum_{u=1}^{N} \sum_{s=1}^{T-1} \sum_{v=1}^{N} \psi_s^{(u)} \hat{\sigma}^{uv} \frac{1}{T} \sum_{t=1}^{T-s} \hat{U}_{ut}(\delta) \hat{U}_{v,t+s}(\delta) = T \sum_{u=1}^{N} \sum_{s=1}^{T-1} \sum_{v=1}^{N} \psi_s^{(u)} \hat{\sigma}^{uv} C_{uv}(s;\hat{\delta}).$$

$$\textit{Also,}\quad E\left(\frac{\partial^2 L(0,\hat{\delta},\hat{\alpha})}{\partial \theta \partial \theta'}\right) = T \sum_{u=1}^{N} \sum_{v=1}^{N} \hat{\sigma}^{uv} \hat{\sigma}_{uv} \sum_{s=1}^{T-1} \left(1 - \frac{s}{T}\right) \psi_s^{(u)} \psi_s^{(v)\prime} = T \hat{A}^{t},$$

and the asymptotic expectation matrix in (11) multiplied by 1/T will take the form

$$\begin{pmatrix} \bar{A} & 0 & 0 \\ 0 & W & 0 \\ 0 & 0 & \frac{1}{2} D_m'(K^{-1} \otimes K^{-1}) D_m \end{pmatrix} \qquad (A13)$$

$$\text{where } \bar{A} = \sum_{u=1}^{N} \sum_{v=1}^{N} \sigma^{uv} \sigma_{uv} \sum_{s=1}^{\infty} \psi_s^{(u)} \psi_s^{(v)\prime};$$

$$W = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \left(\frac{\partial W_t(\delta)}{\partial \delta} K^{-1} \frac{\partial W_t(\delta)}{\partial \delta'} \right)$$

is a positive definite matrix by assumption. (Note that the block diagonality in (A13) follows from expressions (A9), (A10) and (A12), given that $\hat{\sigma}^{uv}$ consistently estimates σ^{uv} and $\hat{U}_t(\delta)$ has zero expectation).

Appendix 2: Derivation of the score statistic $\tilde{S}$

For the derivation of the score test statistic in this context of weak parametric autocorrelation in U_t, we assume that k and K in (20) are parameterized separately, so $\dot{\tau}$ is taken to specify k and $\dot{\alpha}$ to specify K. Thus, the spectral density matrix of $U_t(\theta;\delta)$ for any admissible δ and $\dot{\tau}$ is

$$f(\lambda;\dot{\alpha};\dot{\tau}) = \frac{1}{2\pi} k(\lambda;\dot{\tau})\,K(\dot{\alpha})\,k(\lambda;\dot{\tau})^* \qquad (B1)$$

$$\text{where} \quad k(\lambda;\dot{\tau}) = \sum_{j=0}^{\infty} A(j;\dot{\tau})e^{i\lambda j}.$$

It is also assumed that $A(0;\dot{\tau}) = I_N$ (the N-rowed identity matrix) for any $\dot{\tau}$ in Euclidean space R^q, and that $f(\lambda;\dot{\alpha};\dot{\tau})$ is a finite, positive matrix, with eigenvalues bounded and bounded away from zero at any frequency on a neighborhood N^* of τ and M^* of α. Also, we assume that each element of $\hat{f}(\lambda;\dot{\tau})$, $\hat{f}_{uv}(\lambda;\dot{\tau})$, as defined below (B4), must be continuous in $(\lambda,\dot{\tau})$ for $\dot{\tau} \in N^*$ and have first and second derivatives with respect to $\dot{\tau}$ continuous in $(\lambda,\dot{\tau})$ for $\dot{\tau} \in N^*$.

Taking now $\eta = (\theta';\alpha';\delta';\tau')'$, the negative of the log-likelihood based on Gaussianity of U_t can be expressed as

$$l(\dot{\eta}) = \frac{1}{2}\log\det J(\dot{\alpha};\dot{\tau}) + \frac{1}{2} U(\theta,\dot{\delta})' J^{-1}(\dot{\alpha};\dot{\tau})\,U(\theta,\dot{\delta}), \qquad (B2)$$

where $U(\theta,\delta) = (U_1(\theta,\delta);\ U_2(\theta,\delta);\ldots;U_T(\theta,\delta))'$, and $J(\dot{\alpha},\dot{\tau})$ is a (NT x NT) matrix with $J_{s-t}(\dot{\alpha},\dot{\tau}) = \int_{-\pi}^{\pi} e^{i(s-t)\lambda} f(\lambda;\dot{\alpha};\dot{\tau})d\lambda$ in the (t,s) block of N^2 elements, for any admissible $\dot{\alpha}$, δ and $\dot{\tau}$. However, given the computational difficulty of this expression, especially when N and T are large, under suitable conditions, (B2) can be approximated by

$$L(\theta;\dot{\alpha};\dot{\delta};\dot{\tau}) = \frac{T}{2}\log\det f(\lambda_r;\dot{\alpha};\dot{\tau}) + \frac{1}{2}\sum_r^{*} \mathrm{tr}\,[f^{-1}(\lambda_r;\dot{\alpha};\dot{\tau})\,I_U(\lambda_r;\theta;\dot{\delta})], \qquad (B3)$$

where $I_U(\lambda_r;\theta;\delta)$ is the periodogram of $U_t(\theta;\delta)$ evaluated at frequencies $\lambda_r = 2\pi r/T$ and the sum on * is as described in Section 2.

Calling now $\hat{\delta}$ any $T^{1/2}$-consistent estimate of δ and $\hat{\alpha}$ as defined in Appendix 1, we can concentrate both out and consider

$$\hat{L}(\theta;\dot{\tau}) = L(\theta;\hat{\alpha};\hat{\delta};\dot{\tau}) = \frac{T}{2}\log\det \hat{f}(\lambda_r;\dot{\tau}) + \frac{1}{2}\sum_r^{*} \mathrm{tr}\,[\hat{f}^{-1}(\lambda_r;\dot{\tau})\,\hat{I}_U(\lambda_r;\theta)], \qquad (B4)$$

where

$$\hat{f}(\lambda_r;\dot{\tau}) = \frac{1}{2\pi} k(\lambda_r;\dot{\tau}) K(\hat{\alpha}) k(\lambda_r;\dot{\tau})^*,$$

and

$$\hat{I}_U(\lambda_r;\theta) = \hat{W}(\lambda_r;\theta)\,\overline{\hat{W}(\lambda_r;\theta)}, \quad \textit{with} \quad \hat{W}(\lambda_r;\theta) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} U_t(\theta;\hat{\delta})\, e^{i\lambda_r t}.$$

Then we can express a score test statistic as:

$$\frac{\partial \hat{L}(\theta;\dot{\tau})}{\partial \theta'} \left[E\left(\frac{\partial^2 \hat{L}(\theta;\dot{\tau})}{\partial\theta\partial\theta'}\right) - E\left(\frac{\partial^2 \hat{L}(\theta;\dot{\tau})}{\partial\theta\partial\dot{\tau}'}\right)\left(E\left(\frac{\partial^2 \hat{L}(\theta;\dot{\tau})}{\partial\dot{\tau}\partial\dot{\tau}'}\right)\right)^{-1} E\left(\frac{\partial^2 \hat{L}(\theta;\dot{\tau})}{\partial\dot{\tau}\partial\theta'}\right)\right]^{-1} \frac{\partial \hat{L}(\theta;\dot{\tau})}{\partial\theta} \Big|_{\theta=0,\dot{\tau}=\tilde{\tau}}, \quad \textbf{\textit{(B5)}}$$

where the expectations are taken under the null hypothesis (4) prior to substitution of $\tilde{\tau}$, where $\tilde{\tau}$ can be any consistent estimate of τ under (4).

We start with $\partial\hat{L}(\theta;\tau)/\partial\theta$, and from (B4), we see that it is

$$\frac{\partial}{\partial\theta}\left(\frac{1}{2}\sum_r^* \mathrm{tr}[\hat{f}^{-1}(\lambda_r;\dot{\tau})\, \hat{I}_U(\lambda_r;\theta)]\right)$$

$$= \frac{1}{2}\sum_r^* \left(\frac{\partial}{\partial\theta} \mathit{vec}'(\hat{I}_U(\lambda_r;\theta))\right) \mathit{vec}(\hat{f}^{-1}(\lambda_r;\dot{\tau})')$$

$$= \frac{1}{2}\sum_r^* \sum_{u=1}^{N} \sum_{v=1}^{N} \frac{\partial \hat{I}_{uv}(\lambda_r;\theta)}{\partial\theta} \hat{f}^{vu}(\lambda_r;\dot{\tau}), \qquad \textbf{\textit{(B6)}}$$

where $\hat{I}_{uv}(\lambda_r;\theta)$ is the $(u,v)^{th}$ element of $\hat{I}_U(\lambda_r;\theta)$, and $\hat{f}^{uv}(\lambda_r;\tau)$ is the $(u,v)^{th}$ element of $\hat{f}^{-1}(\lambda_r;\tau)$. We first concentrate on

$$\frac{\partial \hat{I}_{uv}(\lambda_r;\theta)}{\partial\theta}\Big|_{\theta=0} = \frac{\partial}{\partial\theta}\left(\frac{1}{2\pi T}\sum_{t=1}^{T}\sum_{s=1}^{T} U_{u,t}(\theta;\hat{\delta})\, U_{v,s}(\theta;\hat{\delta})\, e^{i(t-s)\lambda_r}\right)\Big|_{\theta=0}$$

$$= \frac{1}{2\pi T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\frac{\partial \log \rho_u(L;\theta)}{\partial\theta} U_{u,t}(\theta;\hat{\delta})\right) U_{v,s}(\theta;\hat{\delta})\, e^{i(t-s)\lambda_r}\Big|_{\theta=0} \quad +$$

$$\frac{1}{2\pi T}\sum_{t=1}^{T}\sum_{s=1}^{T} U_{u,t}(\theta;\hat{\delta})\left(\frac{\partial \log \rho_v(L;\theta)}{\partial\theta} U_{v,s}(\theta;\hat{\delta})\right) e^{i(t-s)\lambda_r}\Big|_{\theta=0}$$

$$= \frac{1}{2\pi T}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{m=1}^{\infty} \psi_m^{(u)} U_{u,t-m}(\hat{\delta})\, U_{v,s}(\hat{\delta})\, e^{i(t-s)\lambda_r} + \frac{1}{2\pi T}\sum_{t=1}^{T}\sum_{s=1}^{T} U_{u,t}(\hat{\delta}) \sum_{m=1}^{\infty} \psi_m^{(v)} U_{v,s}$$

$$= \sum_{m=1}^{T-1} \psi_m^{(u)} e^{i\lambda_r m} \frac{1}{2\pi T} \sum_{t=1}^{T-m} \sum_{s=1}^{T} U_{ut}(\hat{\delta}) U_{vs}(\hat{\delta}) e^{i\lambda_r(t-s)} + \sum_{m=1}^{T-1} \psi_m^{(v)} e^{-i\lambda_r m} \frac{1}{2\pi T} \sum_{t=1}^{T} \sum_{s=1}^{T-m} U_{ut}(\hat{\delta}) U_{vs}$$

and, under suitable conditions, (with m = 1,2,...,M < T-1, for sufficiently large M), this expression becomes asymptotically

$$\left(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r)\right) I_{uv}(\lambda_r;\hat{\delta}). \tag{B7}$$

Substituting now (B7) in (B6) we obtain that $\partial \hat{L}(\theta;\dot{\tau}) / \partial\theta \,|_{\theta=0}$ is asymptotically

$$\frac{1}{2}\sum_{r}^{*} \sum_{u=1}^{N} \sum_{v=1}^{N} \left(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r)\right) I_{uv}(\lambda_r;\hat{\delta}) \hat{f}^{vu}(\lambda_r;\dot{\tau}). \tag{B8}$$

We next examine the second derivative matrices appearing in (B5), and we start with $\partial^2 \hat{L}(\theta;\tau)/\partial\theta\,\partial\theta'$.

$$\frac{\partial^2 \hat{L}(\theta;\dot{\tau})}{\partial\theta\,\partial\theta'} = \frac{1}{2}\sum_{r}^{*} \sum_{u=1}^{N} \sum_{v=1}^{N} \left(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r)\right) \frac{\partial \hat{I}_{uv}(\lambda_r;\theta)}{\partial\theta'} \hat{f}^{vu}(\lambda_r;\dot{\tau})$$

and using again (B7), this last expression evaluated at $\theta = 0$, becomes for large T

$$\frac{1}{2}\sum_{r}^{*} \sum_{u=1}^{N} \sum_{v=1}^{N} \left(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r)\right) \left(\epsilon_{(u)}(\lambda_r)' + \bar{\epsilon}_{(v)}(\lambda_r)'\right) I_{uv}(\lambda_r;\hat{\delta}) \hat{f}^{vu}(\lambda_r;\dot{\tau}),$$

whose asymptotic expectation is

$$\frac{1}{2}\sum_{r}^{*} \sum_{u=1}^{N} \sum_{v=1}^{N} \left(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r)\right) \left(\epsilon_{(u)}(\lambda_r)' + \bar{\epsilon}_{(v)}(\lambda_r)'\right) \hat{f}_{uv}(\lambda_r;\dot{\tau}) \hat{f}^{vu}(\lambda_r;\dot{\tau}),$$

given that, heuristically, if $f(\lambda;\tau)$ is continuous in λ, $E(I_{uv}(\lambda)) \to_{T\to\infty} f_{uv}(\lambda;\tau)$, for fixed λ. (See Brillinger (1981)). We can write this last expression as

$$\frac{1}{2}\sum_{r}^{*} \sum_{\substack{u=1\\v=1}}^{N} \left(\epsilon_{(u)}(\lambda_r)\epsilon_{(u)}(\lambda_r)' + \epsilon_{(u)}(\lambda_r)\bar{\epsilon}_{(v)}(\lambda_r)' + \bar{\epsilon}_{(v)}(\lambda_r)\epsilon_{(u)}(\lambda_r)' + \bar{\epsilon}_{(v)}(\lambda_r)\bar{\epsilon}_{(v)}(\lambda_r)'\right) x$$

$$x\ \hat{f}_{uv}(\lambda_r;\dot{\tau}) \hat{f}^{vu}(\lambda_r;\dot{\tau}) = \frac{1}{2}\sum_{r}^{*} \sum_{u=1}^{N} \epsilon_{(u)}(\lambda_r)\epsilon_{(u)}(\lambda_r)' \sum_{v=1}^{N} \hat{f}_{uv}(\lambda_r;\dot{\tau}) \hat{f}^{vu}(\lambda_r;\dot{\tau})$$

$$+ \frac{1}{2}\sum_{r}^{*} \sum_{v=1}^{N} \bar{\epsilon}_{(v)}(\lambda_r)\, \bar{\epsilon}_{(v)}(\lambda_r)' \sum_{u=1}^{N} \hat{f}_{uv}(\lambda_r;\dot{\tau}) \hat{f}^{vu}(\lambda_r;\dot{\tau})$$

$$+\ \frac{1}{2}\sum_{r}^{*}\sum_{u=1}^{N}\sum_{v=1}^{N}\left(\epsilon_{(u)}(\lambda_r)\bar{\epsilon}_{(v)}(\lambda_r)' + \bar{\epsilon}_{(v)}(\lambda_r)\epsilon_{(u)}(\lambda_r)'\right)\hat{f}_{uv}(\lambda_r;\dot{\tau})\hat{f}^{vu}(\lambda_r;\dot{\tau}), \qquad \textbf{(B9)}$$

which first two summands will be approximately zero noting that

$$\sum_{v=1}^{N}\hat{f}_{uv}(\lambda_r;\dot{\tau})\hat{f}^{vu}(\lambda_r;\dot{\tau}) = \sum_{u=1}^{N}\hat{f}_{uv}(\lambda_r;\dot{\tau})\hat{f}^{vu}(\lambda_r;\dot{\tau}) = 1 \quad \textit{and}$$

$$\sum_{r}^{*}\epsilon_{(u)}(\lambda_r)\epsilon_{(u)}(\lambda_r)' = \sum_{r}^{*}\bar{\epsilon}_{(v)}(\lambda_r)\,\bar{\epsilon}_{(v)}(\lambda_r)' \simeq 0, \qquad \textbf{(B10)}$$

for all u,v = 1,2,..,N. To see this last result, note that approximating the sum by an integral

$$\int_{-\pi}^{\pi}\epsilon_{(u)}(\lambda)\epsilon_{(u)}(\lambda)'d\lambda = \int_{-\pi}^{\pi}\sum_{s=1}^{\infty}\psi_s^{(u)}e^{i\lambda s}\sum_{m=1}^{\infty}\psi_m^{(u)'}e^{i\lambda m}d\lambda$$

$$= \int_{-\pi}^{\pi}\sum_{s=1}^{\infty}\psi_s^{(u)}(\cos\lambda s + i\sin\lambda s)\sum_{m=1}^{\infty}\psi_m^{(u)'}(\cos\lambda m + i\sin\lambda m)d\lambda$$

$$= \sum_{s=1}^{\infty}\sum_{m=1}^{\infty}\psi_s^{(u)}\psi_m^{(u)'}\left(\int_{-\pi}^{\pi}\cos\lambda s\cos\lambda m\,d\lambda - \int_{-\pi}^{\pi}\sin\lambda s\sin\lambda m\,d\lambda\right) = 0,$$

and identically for the second term in (B10).

Now we look at the (p x q) matrix $\partial^2\hat{L}(\theta;\tau)/\partial\theta\,\partial\tau'$ in (B5) which, evaluated at $\theta = 0$, is

$$\frac{\partial}{\partial\dot{\tau}'}\left(\frac{1}{2}\sum_{r}^{*}\sum_{u=1}^{N}\sum_{v=1}^{N}(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r))I_{uv}(\lambda_r;\hat{\delta})\hat{f}^{vu}(\lambda_r;\dot{\tau})\right)$$

$$= \frac{1}{2}\sum_{r}^{*}\sum_{u=1}^{N}\sum_{v=1}^{N}(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r))I_{uv}(\lambda_r;\hat{\delta})\frac{\partial\hat{f}^{vu}(\lambda_r;\dot{\tau})}{\partial\dot{\tau}'},$$

and whose expectation for large T is

$$\frac{1}{2}\sum_{r}^{*}\sum_{u=1}^{N}\sum_{v=1}^{N}(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r))\hat{f}_{uv}(\lambda_r;\dot{\tau})\frac{\partial\hat{f}^{vu}(\lambda_r;\dot{\tau})}{\partial\dot{\tau}'}. \qquad \textbf{(B11)}$$

This last expression can also be shown in terms of the derivatives of $\hat{f}$ with respect to $\dot{\tau}$ (instead of the derivatives of its inverse, $\hat{f}^{-1}$). (B11) can be expressed as

$$\frac{1}{2}\sum_{r}^{*}\sum_{u=1}^{N} \epsilon_{(u)}(\lambda_r) \sum_{v=1}^{N} \hat{f}_{uv}(\lambda_r;\dot{\tau}) \frac{\partial \hat{f}^{vu}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}'} \quad + \qquad (B12)$$

$$\frac{1}{2}\sum_{r}^{*}\sum_{v=1}^{N} \bar{\epsilon}_{(v)}(\lambda_r) \sum_{u=1}^{N} \hat{f}_{uv}(\lambda_r;\dot{\tau}) \frac{\partial \hat{f}^{vu}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}'}, \qquad (B13)$$

where $$\frac{\partial \hat{f}^{vu}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}'} = \left(\frac{\partial \hat{f}^{vu}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}_1}; \dots ; \frac{\partial \hat{f}^{vu}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}_q} \right).$$

Now using the relationship

$$\frac{\partial \hat{f}^{-1}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}_i} = - \hat{f}^{-1}(\lambda_r;\dot{\tau}) \frac{\partial \hat{f}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}_i} \hat{f}^{-1}(\lambda_r;\dot{\tau}),$$

we have that

$$\hat{f}(\lambda_r;\dot{\tau}) \frac{\partial \hat{f}^{-1}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}_i} = - \frac{\partial \hat{f}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}_i} \hat{f}^{-1}(\lambda_r;\dot{\tau})$$

and

$$\frac{\partial \hat{f}^{-1}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}_i} \hat{f}(\lambda_r;\dot{\tau}) = - \hat{f}^{-1}(\lambda_r;\dot{\tau}) \frac{\partial \hat{f}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}_i},$$

implying this two equalities that

$$\sum_{v=1}^{N} \hat{f}_{uv}(\lambda_r;\dot{\tau}) \frac{\partial \hat{f}^{vu}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}_i} = - \sum_{v=1}^{N} \frac{\partial \hat{f}_{uv}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}_i} \hat{f}^{vu}(\lambda_r;\dot{\tau}) \qquad (B14)$$

and

$$\sum_{u=1}^{N} \frac{\partial \hat{f}^{vu}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}_i} \hat{f}_{uv}(\lambda_r;\dot{\tau}) = - \sum_{u=1}^{N} \hat{f}^{vu}(\lambda_r;\dot{\tau}) \frac{\partial \hat{f}_{uv}(\lambda_r;\dot{\tau})}{\partial \dot{\tau}_i} \qquad (B15)$$

respectively. Substituting now (B14) in (B12) and (B15) in (B13), (B11) becomes

$$-\frac{1}{2}\sum_r^* \sum_{u=1}^N \epsilon_{(u)}(\lambda_r) \sum_{v=1}^N \hat{f}^{vu}(\lambda_r;\dot{t}) \frac{\partial \hat{f}_{uv}(\lambda_r;\dot{t})}{\partial \dot{t}'} - \frac{1}{2}\sum_r^* \sum_{v=1}^N \bar{\epsilon}_{(v)}(\lambda_r) \sum_{u=1}^N \hat{f}^{vu}(\lambda_r;\dot{t})$$

$$= -\frac{1}{2}\sum_r^* \sum_{u=1}^N \sum_{v=1}^N \left(\epsilon_{(u)}(\lambda_r) + \bar{\epsilon}_{(v)}(\lambda_r)\right) \hat{f}^{vu}(\lambda_r;\dot{t}) \frac{\partial \hat{f}_{uv}(\lambda_r;\dot{t})}{\partial \dot{t}'}. \qquad \textit{(B16)}$$

Finally, we look at the (qxq) matrix $\partial\, ^2\hat{L}(\theta;t)/\partial\, t\, \partial\, t'$. The u^{th} element of $\partial\hat{L}(\theta;t)/\partial\, t$ is

$$\frac{\partial \hat{L}(\theta;\dot{t})}{\partial \dot{t}_u} = \frac{1}{2}\sum_r^* \mathrm{tr}\left(\hat{f}^{-1}(\lambda_r;\dot{t}) \frac{\partial \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_u}\right) + \frac{1}{2}\sum_r^* \mathrm{tr}\left(\frac{\partial \hat{f}^{-1}(\lambda_r;\dot{t})}{\partial \dot{t}_u} \hat{I}_U(\lambda_r;\theta)\right) =$$

$$= \frac{1}{2}\sum_r^* \mathrm{tr}\left(\hat{f}^{-1}(\lambda_r;\dot{t}) \frac{\partial \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_u} + \frac{\partial \hat{f}^{-1}(\lambda_r;\dot{t})}{\partial \dot{t}_u} \hat{I}_U(\lambda_r;\theta)\right)$$

$$= \frac{1}{2}\sum_r^* tr\left(\hat{f}^{-1}(\lambda_r;\dot{t}) \frac{\partial \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_u} - \hat{f}(\lambda_r;\dot{t})^{-1} \frac{\partial \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_u} \hat{f}^{-1}(\lambda_r;\dot{t}) \hat{I}_U(\lambda_r;\theta)\right).$$

Then, $\dfrac{\partial \hat{L}(\theta;\dot{t})}{\partial \dot{t}_u \partial \dot{t}_v}$, ***evaluated at*** $\theta = 0$, ***becomes:***

$$\frac{1}{2}\sum_r^* tr\Big(\frac{\partial \hat{f}^{-1}(\lambda_r;\dot{t})}{\partial \dot{t}_v} \frac{\partial \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_u} + \hat{f}^{-1}(\lambda_r;\dot{t}) \frac{\partial^2 \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_u \partial \dot{t}_v} - \frac{\partial \hat{f}^{-1}(\lambda_r;\dot{t})}{\partial \dot{t}_v} \frac{\partial \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_u} \hat{f}^{-1}(\lambda_r;\dot{t}) I_U(\lambda_r;\hat{\delta})$$

$$- \hat{f}^{-1}(\lambda_r;\dot{t}) \frac{\partial^2 \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_u \partial \dot{t}_v} \hat{f}^{-1}(\lambda_r;\dot{t}) I_U(\lambda_r;\hat{\delta}) - \hat{f}^{-1}(\lambda_r;\dot{t}) \frac{\partial \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_u} \frac{\partial \hat{f}^{-1}(\lambda_r;\dot{t})}{\partial \dot{t}_v} I_U(\lambda_r;\hat{\delta})\Big)$$

$$\frac{1}{2}\sum_r^* tr\Big(-\hat{f}^{-1}(\lambda_r;\dot{t}) \frac{\partial \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_v} \hat{f}^{-1}(\lambda_r;\dot{t}) \frac{\partial \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_u} + \hat{f}^{-1}(\lambda_r;\dot{t}) \frac{\partial^2 \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_u \partial \dot{t}_v} + \hat{f}^{-1}(\lambda_r;\dot{t})$$

$$x\ \frac{\partial \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_v} \frac{\partial \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_v} \hat{f}^{-1}(\lambda_r;\dot{t}) I_U(\lambda_r;\hat{\delta}) - \hat{f}^{-1}(\lambda_r;\dot{t}) \frac{\partial^2 \hat{f}(\lambda_r;\dot{t})}{\partial \dot{t}_u \partial \dot{t}_v} \hat{f}^{-1}(\lambda_r;\dot{t})\ I_U(\lambda_r;\hat{\delta}) +$$

$$\hat{f}^{-1}(\lambda_r;\dot{\tau})\frac{\partial\hat{f}(\lambda_r;\dot{\tau})}{\partial\dot{\tau}_u}\hat{f}^{-1}(\lambda_r;\dot{\tau})\frac{\partial\hat{f}(\lambda_r;\dot{\tau})}{\partial\dot{\tau}_v}\hat{f}^{-1}(\lambda_r;\dot{\tau})I_U(\lambda_r;\hat{\delta})\ ),$$

and whose asymptotic expectation is

$$\frac{1}{2}\sum_{r}^{*}\operatorname{tr}\left(\hat{f}^{-1}(\lambda_r;\dot{\tau})\frac{\partial\hat{f}(\lambda_r;\dot{\tau})}{\partial\dot{\tau}_u}\hat{f}^{-1}(\lambda_r;\dot{\tau})\frac{\partial\hat{f}(\lambda_r;\dot{\tau})}{\partial\dot{\tau}_v}\right). \tag{B17}$$

Substituting now (B8), (B9), (B11) and (B17) in (B5), evaluated at $\tau = \tilde{\tau}$, we form (21).

TABLE 1

Rejection frequencies of $\hat{S}^{t2}$ in (32) with $\Sigma = I_2$

True model: $\theta_1 = \theta_2 = 0$. **No. of replications: 5000**

$\alpha = 10\ \%$

Table 1a): T = 50

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	.999	.998	.977	.939	.952	.988	.998	1.000
-0.6	.999	.998	.984	.904	.772	.829	.953	.994	.999
-0.4	.996	.982	.910	.677	.428	.576	.871	.982	.997
-0.2	.979	.902	.674	.322	.128	.316	.751	.957	.996
0	.933	.768	.430	.126	**.033**	.206	.660	.939	.992
0.2	.954	.823	.563	.308	.203	.336	.725	.944	.994
0.4	.985	.952	.863	.746	.661	.724	.894	.980	.997
0.6	.999	.994	.978	.953	.930	.944	.979	.995	.999
0.8	1.000	.999	.999	.995	.992	.992	.998	.999	1.000

Table 1b): T = 100

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	.999	.999	1.000	1.000	1.000
-0.4	1.000	1.000	1.000	.989	.933	.979	.999	1.000	1.000
-0.2	1.000	1.000	.989	.805	.417	.751	.987	1.000	1.000
0	1.000	.998	.935	.411	**.053**	.516	.964	.999	1.000
0.2	1.000	.999	.971	.756	.518	.767	.985	.999	1.000
0.4	1.000	1.000	.999	.987	.967	.984	.998	1.000	1.000
0.6	1.000	1.000	1.000	.999	.999	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 1c): T = 200

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.2	1.000	1.000	1.000	.993	.834	.984	1.000	1.000	1.000
0	1.000	1.000	1.000	.827	**.072**	.849	.999	1.000	1.000
0.2	1.000	1.000	1.000	.988	.858	.984	1.000	1.000	1.000
0.4	1.000	1.000	1.000	1.000	.999	.999	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE 2

Rejection frequencies of $\hat{S}^{t2}$ in (32) with $\Sigma = I_2$

True model: $\theta_1 = \theta_2 = 0$. **No. of replications: 5000**

α = **5 %**

Table 2a): T = 50

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	.999	.998	.981	.889	.744	.811	.947	.993	.999
-0.6	.997	.986	.918	.708	.462	.601	.879	.984	.998
-0.4	.980	.919	.726	.400	.181	.358	.778	.962	.996
-0.2	.888	.700	.394	.132	.039	.209	.660	.935	.990
0	.741	.457	.170	.037	**.012**	.147	.585	.910	.987
0.2	.805	.591	.350	.205	.137	.256	.646	.923	.989
0.4	.944	.873	.769	.659	.582	.648	.846	.968	.996
0.6	.993	.977	.959	.927	.904	.920	.967	.992	.998
0.8	1.000	.999	.996	.991	.987	.988	.997	.999	1.000

Table 2b): T = 100

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	.998	.989	.996	1.000	1.000	1.000
-0.4	1.000	1.000	.998	.958	.803	.925	.998	1.000	1.000
-0.2	1.000	.998	.958	.618	.221	.632	.976	1.000	1.000
0	1.000	.987	.798	.206	**.020**	.426	.945	.999	1.000
0.2	1.000	.996	.922	.637	.425	.693	.976	.999	1.000
0.4	1.000	1.000	.997	.975	.949	.975	.997	1.000	1.000
0.6	1.000	1.000	1.000	.999	.999	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 2c): T = 200

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	1.000	1.000	.999	1.000	1.000	1.000	1.000
-0.2	1.000	1.000	1.000	.997	.680	.966	1.000	1.000	1.000
0	1.000	1.000	.999	.671	**.032**	.793	.999	1.000	1.000
0.2	1.000	1.000	.999	.969	.808	.972	1.000	1.000	1.000
0.4	1.000	1.000	1.000	.999	.999	.999	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE 3

Rejection frequencies of $\hat{S}^{t2}$ in (32) with $\Sigma = I_2$

True model: $\theta_1 = \theta_2 = 0$. **No. of replications: 5000**

α = 2.5 %

Table 3a): T = 50

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	.997	.982	.905	.672	.416	.563	.869	.982	.998
-0.6	.985	.926	.750	.428	.187	.371	.784	.965	.995
-0.4	.905	.745	.471	.170	.054	.232	.682	.940	.992
-0.2	.666	.412	.174	.044	.010	.143	.585	.906	.986
0	.404	.182	.053	.011	**.004**	.105	.521	.880	.981
0.2	.548	.358	.228	.137	.100	.198	.575	.895	.984
0.4	.863	.776	.681	.578	.516	.581	.798	.951	.993
0.6	.976	.960	.935	.900	.872	.891	.956	.988	.998
0.8	.999	.997	.993	.986	.979	.984	.994	.999	1.000

Table 3b): T = 100

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	.999	.998	.999	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	.995	.955	.983	.999	1.000	1.000
-0.4	1.000	1.000	.992	.880	.602	.834	.993	1.000	1.000
-0.2	1.000	.992	.880	.413	.092	.526	.963	.999	1.000
0	.997	.953	.591	.089	**.010**	.354	.924	.998	1.000
0.2	.999	.980	.836	.527	.349	.621	.963	.999	1.000
0.4	1.000	.999	.992	.961	.928	.963	.996	1.000	1.000
0.6	1.000	1.000	1.000	.999	.999	.999	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3c): T = 200

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	1.000	1.000	.996	.999	1.000	1.000	1.000
-0.2	1.000	1.000	1.000	.941	.495	.934	1.000	1.000	1.000
0	1.000	1.000	.996	.495	**.014**	.742	.999	1.000	1.000
0.2	1.000	1.000	.999	.941	.754	.957	1.000	1.000	1.000
0.4	1.000	1.000	1.000	.999	.999	.999	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE 4

Rejection frequencies of $\hat{S}^{t2}$ in (32) with $\Sigma = I_2$

True model: $\theta_1 = \theta_2 = 0$. **No. of replications: 5000**

$\alpha = 1\ \%$

Table 4a): T = 50

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	.962	.874	.651	.289	.096	.293	.740	.956	.994
-0.6	.875	.704	.406	.134	.035	.204	.661	.933	.990
-0.4	.646	.403	.164	.039	.010	.141	.583	.905	.986
-0.2	.288	.136	.040	.006	.001	.096	.500	.869	.979
0	.098	.034	.010	.003	**.001**	.070	.438	.832	.973
0.2	.281	.205	.136	.091	.066	.141	.500	.853	.976
0.4	.734	.663	.573	.494	.442	.500	.739	.932	.989
0.6	.950	.928	.898	.862	.830	.850	.931	.983	.997
0.8	.995	.991	.985	.977	.970	.974	.988	.997	.999

Table 4b): T = 100

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	.995	.979	.991	1.000	1.000	1.000
-0.6	1.000	1.000	.997	.956	.805	.925	.998	1.000	1.000
-0.4	1.000	.998	.962	.688	.307	.685	.982	1.000	1.000
-0.2	.997	.958	.692	.199	.026	.416	.941	.998	1.000
0	.976	.801	.295	.025	**.004**	.281	.892	.997	1.000
0.2	.990	.923	.685	.405	.279	.530	.944	.999	1.000
0.4	.999	.997	.980	.937	.897	.945	.994	.999	1.000
0.6	1.000	1.000	1.000	.999	.998	.999	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4c): T = 200

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	1.000	.999	.975	.997	1.000	1.000	1.000
-0.2	1.000	1.000	.999	.848	.275	.886	.999	1.000	1.000
0	1.000	1.000	.971	.279	**.005**	.673	.999	1.000	1.000
0.2	1.000	1.000	.998	.892	.686	.932	1.000	1.000	1.000
0.4	1.000	1.000	1.000	.999	.998	.999	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE 5

Rejection frequencies of $\hat{S}^{t2}$ in (32) with $\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$

True model: $\theta_1 = \theta_2 = 0$. **No. of replications: 5000**

α = 10 %

Table 5a): T = 50

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	.999	.998	.997	.998	.999	.999	1.000	1.000
-0.6	1.000	.998	.989	.969	.982	.997	.999	1.000	1.000
-0.4	.998	.987	.912	.770	.841	.975	.998	1.000	1.000
-0.2	.998	.968	.768	.346	.319	.816	.983	.999	1.000
0	.998	.985	.834	.323	**.039**	.431	.921	.994	.999
0.2	.999	.997	.974	.813	.442	.340	.824	.987	.998
0.4	1.000	.999	.998	.981	.917	.824	.895	.985	.999
0.6	1.000	1.000	1.000	.999	.995	.987	.987	.996	.999
0.8	1.000	1.000	1.000	.999	.999	.999	.999	.999	1.000

Table 5b): T = 100

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1 000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	.999	.996	.998	1.000	1.000	1.000	1.000
-0.2	1.000	1.000	.996	.813	.764	.997	1.000	1.000	1.000
0	1.000	1.000	.998	.768	**.061**	.821	.998	1.000	1.000
0.2	1.000	1.000	1.000	.995	.825	.765	.992	1.000	1.000
0.4	1.000	1.000	1.000	1.000	.999	.994	.998	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 5c): T = 200

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.2	1.000	1.000	1.000	.993	.991	1.000	1.000	1.000	1.000
0	1.000	1.000	1.000	.990	**.075**	.989	1.000	1.000	1.000
0.2	1.000	1.000	1.000	.988	.989	.983	1.000	1.000	1.000
0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE 6

Rejection frequencies of $\hat{S}^{t2}$ in (32) with $\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$

True model: $\theta_1 = \theta_2 = 0$. **No. of replications: 5000**

$\alpha = 5\ \%$

Table 6a): T = 50

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	.999	.998	.990	.979	.986	.997	.999	1.000	1.000
-0.6	.997	.987	.932	.866	.921	.986	.998	.999	1.000
-0.4	.988	.930	.743	.503	.625	.920	.992	.999	1.000
-0.2	.975	.861	.500	.152	.144	.688	.964	.996	.999
0	.986	.913	.618	.149	**.014**	.327	.882	.987	.999
0.2	.996	.984	.916	.685	.340	.256	.766	.978	.997
0.4	.999	.998	.992	.960	.879	.764	.846	.976	.997
0.6	1.000	1.000	.999	.998	.991	.979	.977	.992	.999
0.8	1.000	1.000	.999	.999	.999	.999	.999	.999	1.000

Table 6b): T = 100

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	.999	.982	.993	1.000	1.000	1.000	1.000
-0.2	1.000	.999	.983	.630	.571	.989	1.000	1.000	1.000
0	1.000	1.000	.992	.594	**.021**	.750	.998	1.000	1.000
0.2	1.000	1.000	1.000	.985	.754	.689	.988	1.000	1.000
0.4	1.000	1.000	1.000	1.000	.999	.989	.997	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 6c): T = 200

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.2	1.000	1.000	1.000	.979	.970	1.000	1.000	1.000	1.000
0	1.000	1.000	1.000	.968	**.032**	.979	1.000	1.000	1.000
0.2	1.000	1.000	1.000	1.000	.980	.972	1.000	1.000	1.000
0.4	1.000	1.000	1.000	.999	1.000	.999	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE 7

Rejection frequencies of $\hat{S}^{t2}$ in (32) with $\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$

True model: $\theta_1 = \theta_2 = 0$. **No. of replications: 5000**

$\alpha = 2.5$ %

Table 7a): T = 50

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	.997	.987	.942	.889	.921	.981	.996	.999	.999
-0.6	.984	.928	.786	.634	.753	.939	.992	.999	.999
-0.4	.937	.786	.487	.243	.366	.817	.977	.996	.999
-0.2	.876	.627	.247	.051	.005	.553	.934	.991	.999
0	.918	.737	.374	.061	**.005**	.251	.835	.982	.998
0.2	.980	.937	.814	.556	.261	.196	.709	.968	.995
0.4	.996	.992	.973	.927	.826	.702	.798	.963	.996
0.6	.999	.999	.998	.993	.983	.969	.968	.988	.997
0.8	.999	.999	.999	.999	.999	.998	.997	.998	1.000

Table 7b): T = 100

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	.999	1.000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	.993	.939	.975	1.000	1.000	1.000	1.000
-0.2	1.000	.999	.938	.431	.380	.972	1.000	1.000	1.000
0	1.000	1.000	.974	.392	**.010**	.668	.996	1.000	1.000
0.2	1.000	1.000	.999	.971	.689	.618	.983	1.000	1.000
0.4	1.000	1.000	1.000	1.000	.997	.985	.997	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	.999	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 7c): T = 200

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.2	1.000	1.000	1.000	.942	.922	1.000	1.000	1.000	1.000
0	1.000	1.000	1.000	.920	**.014**	.966	1.000	1.000	1.000
0.2	1.000	1.000	1.000	1.000	.968	.957	1.000	1.000	1.000
0.4	1.000	1.000	1.000	1.000	1.000	.999	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE 8

Rejection frequencies of $\hat{S}^{t2}$ in (32) with $\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$

True model: $\theta_1 = \theta_2 = 0$. **No. of replications: 5000**

$\alpha = 1$ %

Table 8a): T = 50

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	.966	.890	.722	.552	.651	.883	.975	.993	.998
-0.6	.885	.720	.452	.263	.402	.797	.965	.992	.998
-0.4	.716	.446	.184	.006	.137	.643	.938	.991	.998
-0.2	.553	.260	.062	.007	.016	.404	.880	.984	.998
0	.638	.400	.140	.021	**.001**	.182	.769	.973	.995
0.2	.879	.791	.634	.415	.190	.140	.632	.947	.994
0.4	.971	.955	.930	.870	.763	.640	.735	.947	.994
0.6	.993	.993	.989	.982	.971	.949	.952	.982	.996
0.8	.999	.999	.998	.999	.997	.996	.996	.998	.999

Table 8b): T = 100

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	.999	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	.999	.994	.997	1.000	1.000	1.000	1.000
-0.4	1.000	.999	.963	.792	.907	.998	1.000	1.000	1.000
-0.2	1.000	.995	.798	.213	.205	.938	.999	1.000	1.000
0	1.000	.998	.898	.198	**.003**	.579	.992	1.000	1.000
0.2	1.000	1.000	.997	.934	.600	.533	.976	1.000	1.000
0.4	1.000	1.000	1.000	1.000	.993	.974	.994	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	.999	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 8c): T = 200

θ_1 / θ_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.2	1.000	1.000	1.000	.853	.810	1.000	1.000	1.000	1.000
0	1.000	1.000	1.000	.819	**.005**	.942	1.000	1.000	1.000
0.2	1.000	1.000	1.000	1.000	.948	.929	1.000	1.000	1.000
0.4	1.000	1.000	1.000	1.000	1.000	.999	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE 9

Table 9a: Empirical sizes of $\hat{S}^{f2}$ in (34) with $\Sigma = I_2$

True model: $\theta_1 = \theta_2 = 0$. **No. of replications: 1000**

T \ α	10%	5%	2.5%	1%
50	0.028	0.012	0.001	0.000
100	0.058	0.019	0.010	0.006
200	0.074	0.038	0.020	0.008

Table 9b: Empirical sizes of $\hat{S}^{f2}$ in (34) with $\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$

True model: $\theta_1 = \theta_2 = 0$. **No. of replications: 1000**

T \ α	10%	5%	2.5%	1%
50	0.036	0.012	0.002	0.000
100	0.057	0.021	0.008	0.005
200	0.076	0.035	0.017	0.006

TABLE 10

Table 10a: Empirical sizes of $\tilde{S}^2$ in (37) with a VAR(1) structure on U_t

True model: $\theta_1 = \theta_2 = 0$. **No. of replications: 1000**

T \ α	10%	5%	2.5%	1%
50	0.134	0.074	0.040	0.017
100	0.123	0.069	0.035	0.014
200	0.104	0.060	0.031	0.012

Table 10b: Empirical sizes of $\tilde{S}^2$ in (37) with a VMA(1) structure on U_t

True model: $\theta_1 = \theta_2 = 0$. **No. of replications: 1000**

T \ α	10%	5%	2.5%	1%
50	0.207	0.154	0.127	0.097
100	0.137	0.090	0.054	0.045
200	0.131	0.062	0.038	0.023

EUI WORKING PAPERS

EUI Working Papers are published and distributed by the
European University Institute, Florence

Copies can be obtained free of charge
– depending on the availability of stocks – from:

The Publications Officer
European University Institute
Badia Fiesolana
I-50016 San Domenico di Fiesole (FI)
Italy

Please use order form overleaf

Publications of the European University Institute

To The Publications Officer
European University Institute
Badia Fiesolana
I-50016 San Domenico di Fiesole (FI) – Italy
Telefax No: +39/55/4685 636
e-mail: publish@datacomm.iue.it
http://www.iue.it

From Name . . . .
Address. . . . .
. . . .
. . . .
. . . .
. . . .

❐ Please send me a complete list of EUI Working Papers
❐ Please send me a complete list of EUI book publications
❐ Please send me the EUI brochure Academic Year 1998/99

Please send me the following EUI Working Paper(s):

No, Author . . . .
Title: . . . .
No, Author . . . .
Title: . . . .
No, Author . . . .
Title: . . . .
No, Author . . . .
Title: . . . .

Date . . . .

Signature . . . .

Working Papers of the Department of Economics Published since 1997

ECO No. 97/1
Jonathan SIMON
The Expected Value of Lotto when not all Numbers are Equal

ECO No. 97/2
Bernhard WINKLER
Of Sticks and Carrots: Incentives and the Maastricht Road to EMU

ECO No. 97/3
James DOW/Rohit RAHI
Informed Trading, Investment, and Welfare

ECO No. 97/4
Sandrine LABORY
Signalling Aspects of Managers' Incentives

ECO No. 97/5
Humberto LÓPEZ/Eva ORTEGA/Angel UBIDE
Dating and Forecasting the Spanish Business Cycle

ECO No. 97/6
Yadira GONZÁLEZ de LARA
Changes in Information and Optimal Debt Contracts: The Sea Loan

ECO No. 97/7
Sandrine LABORY
Organisational Dimensions of Innovation

ECO No. 97/8
Sandrine LABORY
Firm Structure and Market Structure: A Case Study of the Car Industry

ECO No. 97/9
Elena BARDASI/Chiara MONFARDINI
The Choice of the Working Sector in Italy: A Trivariate Probit Analysis

ECO No. 97/10
Bernhard WINKLER
Coordinating European Monetary Union

ECO No. 97/11
Alessandra PELLONI/Robert WALDMANN
Stability Properties in a Growth Model

ECO No. 97/12
Alessandra PELLONI/Robert WALDMANN
Can Waste Improve Welfare?

ECO No. 97/13
Christian DUSTMANN/Arthur van SOEST
Public and Private Sector Wages of Male Workers in Germany

ECO No. 97/14
Søren JOHANSEN
Mathematical and Statistical Modelling of Cointegration

ECO No. 97/15
Tom ENGSTED/Søren JOHANSEN
Granger's Representation Theorem and Multicointegration

ECO No. 97/16
Søren JOHANSEN/
Ernst SCHAUMBURG
Likelihood Analysis of Seasonal Cointegration

ECO No. 97/17
Maozu LU/Grayham E. MIZON
Mutual Encompassing and Model Equivalence

ECO No. 97/18
Dimitrios SIDERIS
Multilateral Versus Bilateral Testing for Long Run Purchasing Power Parity: A Cointegration Analysis for the Greek Drachma

ECO No. 97/19
Bruno VERSAEVEL
Production and Organizational Capabilities

ECO No. 97/20
Chiara MONFARDINI
An Application of Cox's Non-Nested Test to Trinomial Logit and Probit Models

ECO No. 97/21
James DOW/Rohit RAHI
Should Speculators be Taxed?

*out of print

ECO No. 97/22
Kitty STEWART
Are Intergovernmental Transfers in Russia Equalizing?

ECO No. 97/23
Paolo VITALE
Speculative Noise Trading and Manipulation in the Foreign Exchange Market

ECO No. 97/24
Günther REHME
Economic Growth, (Re-)Distributive Policies, Capital Mobility and Tax Competition in Open Economies

ECO No. 97/25
Susana GARCIA CERVERO
A Historical Approach to American Skill Differentials

ECO No. 97/26
Susana GARCIA CERVERO
Growth, Technology and Inequality: An Industrial Approach

ECO No. 97/27
Bauke VISSER
Organizational Structure and Performance

ECO No. 97/28
Pompeo DELLA POSTA
Central Bank Independence and Public Debt Convergence in an Open Economy Dynamic Game

ECO No. 97/29
Matthias BRUECKNER
Voting and Decisions in the ECB

ECO No. 97/30
Massimiliano MARCELLINO
Temporal Disaggregation, Missing Observations, Outliers, and Forecasting: A Unifying Non-Model Based Procedure

ECO No. 97/31
Marion KOHLER
Bloc Formation in International Monetary Policy Coordination

ECO No. 97/32
Marion KOHLER
Trade Blocs and Currency Blocs: A Package Deal?

ECO No. 97/33
Lavan MAHADEVA
The Comparative Static Effects of Many Changes

ECO No. 97/34
Lavan MAHADEVA
Endogenous Growth with a Declining Rate of Interest

ECO No. 97/35
Spyros VASSILAKIS
Managing Design Complexity to Improve on Cost, Quality, Variety, and Time-to-Market Performance Variables

ECO No. 97/36
Spyros SKOURAS
Analysing Technical Analysis

***

ECO No. 98/1
Bauke VISSER
Binary Decision Structures and the Required Detail of Information

ECO No. 98/2
Michael ARTIS/Massimiliano MARCELLINO
Fiscal Solvency and Fiscal Forecasting in Europe

ECO No. 98/3
Giampiero M. GALLO/Barbara PACINI
Early News is Good News: The Effects of Market Opening on Market Volatility

ECO No. 98/4
Michael J. ARTIS/Zenon G. KONTOLEMIS
Inflation Targeting and the European Central Bank

ECO No. 98/5
Alexandre KOLEV
The Distribution of Enterprise Benefits in Russia and their Impact on Individuals' Well-Being

ECO No. 98/6
Kitty STEWART
Financing Education at the Local Level: A Study of the Russian Region of Novgorod

*out of print

ECO No. 98/7
Anna PETTINI/Louis PHLIPS
A Redistributive Approach to Price and Quality Discrimination

ECO No. 98/8
Aldo RUSTICHINI/Andrea ICHINO/Daniele CHECCHI
More Equal but Less Mobile? Education Financing and Intergenerational Mobility in Italy and in the US

ECO No. 98/9
Andrea ICHINO/Pietro ICHINO
Discrimination or Individual Effort? Regional Productivity Differentials in a Large Italian Firm

ECO No. 98/10
Andrea ICHINO/Rudolf WINTER-EBMER
The Long-Run Educational Cost of World War II. An Example of Local Average Treatment Effect Estimation

ECO No. 98/11
Luca FLABBI/Andrea ICHINO
Productivity, Seniority and Wages. New Evidence from Personnel Data

ECO No. 98/12
Jian-Ming ZHOU
Is Nominal Public but *De Facto* Private Land Ownership Appropriate? A Comparative Study among Cambodia, Laos, Vietnam; Japan, Taiwan Province of China, South Korea; China, Myanmar; and North Korea

ECO No. 98/13
Anna PETTINI
Consumers' Tastes and the Optimal Price Gap

ECO No. 98/14
Christian DUSTMANN/Najma RAJAH/Arthur VAN SOEST
School Quality, Exam Performance, and Career Choice

ECO No. 98/15
Ramon MARIMON/Juan Pablo NICOLINI/Pedro TELES
Electronic Money: Sustaining Low Inflation?

ECO No. 98/16
Michael ARTIS/Marion KOHLER/Jacques MÉLITZ
Trade and the Number of OCA's in the World

ECO No. 98/17
Nuala O'DONNELL
Why Did Earnings Inequality Increase in Ireland: 1987-1994?

ECO No. 98/18
Luis A. GIL-ALAÑA
Fractional Integration in the Purchasing Power Parity

ECO No. 98/19
Luis A. GIL-ALAÑA
Multivariate Tests of Fractionally Integrated Hypotheses

*out of print