Economics Letters 125 (2014) 1–4

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

On testing for nonlinearity in multivariate time series☆

Zacharias Psaradakis [a], Marián Vávra [b,*]

[a] Department of Economics, Mathematics and Statistics, Birkbeck, University of London, United Kingdom
[b] Research Department, National Bank of Slovakia, Slovakia

ARTICLE INFO

Article history:
Received 26 May 2014
Accepted 28 July 2014
Available online 8 August 2014

JEL classification:
C12
C15
C32

Keywords:
Multivariate time series
Nonlinearity tests
Principal components

ABSTRACT

This paper considers a multivariate extension of the test for neglected nonlinearity proposed by Tsay (1986) that uses principal components to overcome the problem of dimensionality that is common with tests of this type. Monte Carlo experiments reveal that the modified multivariate test provides a significant dimensional reduction without suffering from any systematic level distortion or power loss, and is more powerful than univariate nonlinearity tests.

1. Introduction

Recent years have witnessed a growing interest in tests for neglected nonlinearity in time series models (see, e.g., Tong, 1990; Teräsvirta et al., 2010). Such tests have become an essential first step in model-building exercises since, due to the difficulties associated with the statistical analysis of nonlinear models, it is often desirable to establish the adequacy or otherwise of a linear data representation before exploring more complicated nonlinear structures.

Although much of the relevant literature has focused on univariate models, there are situations in which relationships between two or more time series may have a nonlinear structure. In such cases it is reasonable to expect that more powerful inference procedures may be obtained by considering tests for neglected nonlinearity in multivariate instead of univariate models. A test of this type was considered by Harvill and Ray (1999), who developed a multivariate generalization of the nonlinearity test proposed by Tsay (1986) and Luukkonen et al. (1988). A practical difficulty with the application of such a test to real-world data is the large number of terms required to construct the relevant artificial test regression. The inclusion of these terms may induce substantial collinearity and necessitates the use of relatively long time series for the effective implementation of the test.

The present paper offers a way of overcoming these difficulties by introducing a multivariate test for neglected nonlinearity which achieves a reduction in the dimension of the set of relevant test variables through the use of principal components. The modified multivariate test is straightforward to construct and provides a significant dimensional reduction without suffering from any systematic level distortion or power loss relative to the original test. This makes the modified test quite attractive for applications in which relatively long stretches of data may not be available, as is often the case, for example, in macroeconometrics. What is more, as Harvill and Ray (1999) also observed, multivariate tests are generally considerably more powerful than univariate tests applied to the components of a nonlinear multiple time series, suggesting that there are clear advantages to testing the component series jointly rather than individually.

The tests to be considered are described in Section 2. A simulation study of the properties of the tests is presented in Section 3. Section 4 summarizes and concludes.

2. Tests for neglected nonlinearity

Consider the vector autoregressive (VAR) model for a k-variate time series $\{\boldsymbol{x}_t\}$ given by

$$\boldsymbol{x}_t = \boldsymbol{\mu} + \sum_{j=1}^{p} \boldsymbol{A}_j \boldsymbol{x}_{t-j} + \boldsymbol{u}_t, \quad t = 0, \pm 1, \pm 2, \ldots, \tag{1}$$

☆ The authors would like to thank Timo Teräsvirta and participants in the Research Seminar at the National Bank of Slovakia for helpful comments and suggestions.

* Corresponding author.
E-mail addresses: z.psaradakis@bbk.ac.uk (Z. Psaradakis), marian.vavra@nbs.sk (M. Vávra).

http://dx.doi.org/10.1016/j.econlet.2014.07.031
0165-1765/

where $p \geqslant 1$ is a fixed integer, $\boldsymbol{\mu}$ is a $k \times 1$ vector of real constants, $\boldsymbol{A}_j$ ($j = 1, \ldots, p$) are $k \times k$ matrices of real constants, and $\{\boldsymbol{u}_t\}$ is a sequence of independent, identically distributed k-dimensional real random vectors with $\mathbb{E}(\boldsymbol{u}_t) = \boldsymbol{0}$, $\det \mathbb{E}(\boldsymbol{u}_t \boldsymbol{u}_t') \neq 0$, and $\mathbb{E}(\|\boldsymbol{u}_t\|^4) < \infty$. It is also assumed that $\det(\boldsymbol{I}_k - \sum_{j=1}^{p} \boldsymbol{A}_j z^j) \neq 0$ for all complex z such that $|z| \leqslant 1$, where $\boldsymbol{I}_k$ denotes the identity matrix of order k. Under these assumptions the VAR equations (1) have a unique causal, stationary and ergodic solution. The assumptions are also sufficient for the least-squares estimator of the parameters of the model to be consistent and asymptotically normal (e.g., Lütkepohl, 2005, Sec. 3.2.2). We are interested in testing the hypothesis that there is no neglected nonlinearity in (1).

Given a sample $(\boldsymbol{x}_{-p+1}, \ldots, \boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$, the test for neglected nonlinearity considered by Harvill and Ray (1999) may be implemented as a test for the hypothesis $\boldsymbol{B}_2 = \boldsymbol{0}$ in the auxiliary multivariate regression

$$\hat{\boldsymbol{u}}_t = \boldsymbol{b}_0 + \boldsymbol{B}_1 \boldsymbol{v}_t + \boldsymbol{B}_2 \boldsymbol{w}_t + \boldsymbol{\eta}_t, \quad t = 1, 2, \ldots, T, \tag{2}$$

where $\hat{\boldsymbol{u}}_t$ is the $k \times 1$ vector of least-squares residuals from (1), $\boldsymbol{v}_t$ is the $kp \times 1$ vector defined as $\boldsymbol{v}_t = (\boldsymbol{x}_{t-1}', \ldots, \boldsymbol{x}_{t-p}')'$, $\boldsymbol{w}_t$ is the $\frac{1}{2}kp(kp+1) \times 1$ vector defined as $\boldsymbol{w}_t = \operatorname{vech}(\boldsymbol{v}_t \boldsymbol{v}_t')$, $(\boldsymbol{b}_0, \boldsymbol{B}_1, \boldsymbol{B}_2)$ are artificial parameters, and $\boldsymbol{\eta}_t$ is an artificial error term. Putting $m = \frac{1}{2}kp(kp+1)$, the linearity hypothesis is rejected for large values of the likelihood-ratio statistic

$$\Lambda_{\mathrm{HR}} = (T - \tau)(\ln \det \boldsymbol{S}_0 - \ln \det \boldsymbol{S}_1), \tag{3}$$

where $\boldsymbol{S}_1$ and $\boldsymbol{S}_0$ are the least-squares residual sum of squares matrices from (2) with $\boldsymbol{B}_2$ unrestricted and $\boldsymbol{B}_2 = \boldsymbol{0}$, respectively, and $\tau = kp + \frac{1}{2}(k+m+3)$ is Bartlett's correction factor (see Anderson, 2003, Sec. 8.5.2). When $\{\boldsymbol{x}_t\}$ satisfies (1), Λ_{HR} has an approximate χ^2_{km} distribution for large T.[1]

An obvious difficulty with the application of a nonlinearity test based on (3) in practice is the large dimension m of the squares and cross-products vector $\boldsymbol{w}_t$. As a result, relatively long time series are required for the implementation of the test procedure. In addition, the components of $\boldsymbol{w}_t$ are likely to be highly collinear, something which can have adverse effects on the finite-sample performance of the test.

We argue that the dimensionality and collinearity problems may be effectively alleviated by the use of principal components. Specifically, we suggest replacing $\boldsymbol{w}_t$ in (2) by the n-dimensional vector $\boldsymbol{y}_t = (Y_{1t}, \ldots, Y_{nt})'$, $1 \leqslant n \leqslant m$, consisting of the first n sample principal components of $\boldsymbol{w}_t$. Letting $\lambda_1 \geqslant \cdots \geqslant \lambda_m$ denote the eigenvalues of the sample correlation matrix of $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_T)$, the ith principal component is computed as $Y_{it} = \boldsymbol{\xi}_i' \boldsymbol{w}_t^*$ ($i = 1, \ldots, m$), where $\boldsymbol{\xi}_i$ is the normalized eigenvector associated with λ_i and $\boldsymbol{w}_t^*$ is the standardized version of $\boldsymbol{w}_t$. A test for nonlinearity may then be implemented as a test for the hypothesis $\boldsymbol{C}_2 = \boldsymbol{0}$ in the auxiliary multivariate regression

$$\hat{\boldsymbol{u}}_t = \boldsymbol{c}_0 + \boldsymbol{C}_1 \boldsymbol{v}_t + \boldsymbol{C}_2 \boldsymbol{y}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, 2, \ldots, T, \tag{4}$$

where $(\boldsymbol{c}_0, \boldsymbol{C}_1, \boldsymbol{C}_2)$ are artificial parameters and $\boldsymbol{\varepsilon}_t$ is an artificial error term. Linearity is thus rejected for large values of the likelihood-ratio statistic

$$\Lambda_{\mathrm{PC}} = (T - \bar{\tau})(\ln \det \boldsymbol{S}_0 - \ln \det \boldsymbol{S}_2), \tag{5}$$

where $\bar{\tau} = kp + \frac{1}{2}(k+n+3)$ and $\boldsymbol{S}_2$ is the least-squares residual sum of squares matrix from (4). For large T, Λ_{PC} may be approximately treated as χ^2_{kn} under the null hypothesis that $\{\boldsymbol{x}_t\}$ satisfies the linear model (1).

In addition to the dimensional reduction achieved by transforming into principal components, the collinearity problem associated with the use of $\boldsymbol{w}_t$ is effectively eliminated since sample principal components are uncorrelated. A decision, however, needs to be made in the implementation of the test based on Λ_{PC} on the number of principal components to be used. Among the various methods available in the literature, the following rules for selecting n are popular in applied work and are used in the sequel[2]:

R1: n is the smallest integer such that $m^{-1} \sum_{i=1}^{n} \lambda_i \geqslant 0.95$ (proportion-of-variance rule);

R2: n is the smallest integer such that $\lambda_{n+1} \leqslant \tilde{\lambda}$ for some prespecified $\tilde{\lambda} > 0$; following a recommendation of Jolliffe (1972), we set $\tilde{\lambda} = 0.7$ (average-root rule);

R3: n is the smallest integer such that $\lambda_{n+1} \leqslant m^{-1} \sum_{i=n+1}^{m} i^{-1}$ (broken-stick rule).

It is finally worth remarking that the test procedures based on criteria like those in (3) and (5) may be easily modified to allow for a VARMA or VARMAX structure under the null hypothesis of linearity (cf. Harvill and Ray, 1999). Furthermore, the finite-order VAR model used in the construction of the tests may be viewed as only an approximation to a potentially infinite-order VAR structure for $\{\boldsymbol{x}_t\}$. Asymptotic justification of inference procedures in this case requires that the order of the VAR model fitted to the data increases, at some appropriate rate, simultaneously with the sample size (cf. Lütkepohl, 2005, Ch. 15).

3. Monte Carlo simulations

To assess the finite-sample properties of the tests based on the statistics in (3) and (5), we carry out some Monte Carlo experiments. We consider bivariate time series $\{\boldsymbol{x}_t\}$ satisfying the following models:

M1:
$$\boldsymbol{x}_t = \begin{pmatrix} 0.4 & 0 \\ 0 & 0.4 \end{pmatrix} \boldsymbol{x}_{t-1} + \begin{pmatrix} 0.3 & 0 \\ 0 & 0.3 \end{pmatrix} \boldsymbol{x}_{t-2} + \boldsymbol{u}_t$$

M2:
$$\boldsymbol{x}_t = \begin{pmatrix} 0.4 & 0.3 \\ 0.3 & 0.4 \end{pmatrix} \boldsymbol{x}_{t-1} + \boldsymbol{u}_t$$

M3:
$$\boldsymbol{x}_t = \begin{pmatrix} 0.4 & -0.3 \\ -0.3 & 0.4 \end{pmatrix} \boldsymbol{x}_{t-1} + \boldsymbol{u}_t$$

N1:
$$\boldsymbol{x}_t = \begin{pmatrix} 0.4 & -0.3 \\ -0.3 & 0.4 \end{pmatrix} \boldsymbol{x}_{t-1} + \begin{pmatrix} 0.10 & -0.05 \\ -0.05 & 0.10 \end{pmatrix} \times (\boldsymbol{x}_{t-1} \circ \boldsymbol{u}_{t-1}) + \boldsymbol{u}_t$$

N2:
$$\boldsymbol{x}_t = \begin{pmatrix} 0.4 & -0.3 \\ -0.3 & 0.4 \end{pmatrix} \boldsymbol{x}_{t-1} + \begin{pmatrix} -0.05 & 0.10 \\ 0.10 & -0.05 \end{pmatrix} \times (\boldsymbol{x}_{t-1} \circ \boldsymbol{u}_{t-1}) + \boldsymbol{u}_t$$

N3:
$$\boldsymbol{x}_t = \begin{pmatrix} 0.4 & -0.3 \\ -0.3 & 0.4 \end{pmatrix} \boldsymbol{x}_{t-1} + \begin{pmatrix} 0.0 & 0.1 \\ 0.1 & 0.0 \end{pmatrix} (\boldsymbol{x}_{t-1} \circ \boldsymbol{x}_{t-1}) + \boldsymbol{u}_t$$

N4:
$$\boldsymbol{x}_t = \begin{pmatrix} 0.4 & -0.3 \\ -0.3 & 0.4 \end{pmatrix} \boldsymbol{x}_{t-1} + \begin{pmatrix} -0.05 & 0.10 \\ 0.10 & -0.05 \end{pmatrix} \times (\boldsymbol{x}_{t-1} \circ \boldsymbol{x}_{t-1}) + \boldsymbol{u}_t$$

[1] Note that Harvill and Ray (1999) use a test criterion based on an F-approximation to Wilks' lambda statistic ($\det \boldsymbol{S}_1 / \det \boldsymbol{S}_0$) instead of Λ_{HR}.

[2] For a detailed discussion of these rules the reader is referred to Jolliffe (2002, Ch. 6).

Table 1
Rejection frequencies of tests (linear time series).

	Test	$T = 200$			$T = 500$		
		M1	M2	M3	M1	M2	M3
$\sigma_{12} = 0$	Λ_{HR}	0.044	0.038	0.042	0.048	0.051	0.042
	$\Lambda_{\mathrm{PC}}(\mathrm{R1})$	0.041	0.042	0.037	0.045	0.055	0.047
	$\Lambda_{\mathrm{PC}}(\mathrm{R2})$	0.042	0.042	0.037	0.052	0.056	0.042
	$\Lambda_{\mathrm{PC}}(\mathrm{R3})$	0.046	0.046	0.040	0.047	0.059	0.044
	$\Lambda_{\mathrm{HR}}(1)$	0.039	0.049	0.045	0.046	0.058	0.054
	$\Lambda_{\mathrm{HR}}(2)$	0.047	0.051	0.046	0.046	0.055	0.062
$\sigma_{12} = 0.5$	Λ_{HR}	0.061	0.041	0.052	0.062	0.042	0.045
	$\Lambda_{\mathrm{PC}}(\mathrm{R1})$	0.060	0.040	0.054	0.058	0.038	0.050
	$\Lambda_{\mathrm{PC}}(\mathrm{R2})$	0.056	0.044	0.052	0.061	0.038	0.052
	$\Lambda_{\mathrm{PC}}(\mathrm{R3})$	0.064	0.042	0.055	0.056	0.042	0.052
	$\Lambda_{\mathrm{HR}}(1)$	0.042	0.044	0.044	0.040	0.059	0.061
	$\Lambda_{\mathrm{HR}}(2)$	0.040	0.058	0.044	0.042	0.048	0.062

Table 2
Rejection frequencies of tests (nonlinear time series).

	Test	$T = 200$							$T = 500$						
		N1	N2	N3	N4	N5	N6	N7	N1	N2	N3	N4	N5	N6	N7
$\sigma_{12} = 0$	Λ_{HR}	0.35	0.39	0.76	0.87	0.46	1.00	1.00	0.86	0.89	0.99	1.00	0.88	1.00	1.00
	$\Lambda_{\mathrm{PC}}(\mathrm{R1})$	0.36	0.37	0.77	0.87	0.46	1.00	1.00	0.82	0.84	0.99	1.00	0.88	1.00	1.00
	$\Lambda_{\mathrm{PC}}(\mathrm{R2})$	0.32	0.36	0.77	0.87	0.46	1.00	1.00	0.79	0.80	1.00	1.00	0.87	1.00	1.00
	$\Lambda_{\mathrm{PC}}(\mathrm{R3})$	0.37	0.38	0.77	0.88	0.46	1.00	1.00	0.86	0.89	1.00	1.00	0.88	1.00	1.00
	$\Lambda_{\mathrm{HR}}(1)$	0.33	0.11	0.04	0.17	0.10	1.00	1.00	0.69	0.25	0.05	0.32	0.20	1.00	0.99
	$\Lambda_{\mathrm{HR}}(2)$	0.23	0.08	0.05	0.18	0.09	0.30	0.19	0.60	0.19	0.04	0.35	0.21	0.52	0.93
$\sigma_{12} = 0.5$	Λ_{HR}	0.60	0.57	0.80	0.95	0.22	1.00	1.00	0.97	0.97	1.00	1.00	0.54	1.00	1.00
	$\Lambda_{\mathrm{PC}}(\mathrm{R1})$	0.54	0.53	0.79	0.92	0.20	1.00	1.00	0.94	0.95	0.99	0.99	0.52	1.00	1.00
	$\Lambda_{\mathrm{PC}}(\mathrm{R2})$	0.49	0.48	0.81	0.95	0.21	1.00	1.00	0.90	0.92	0.99	1.00	0.55	1.00	1.00
	$\Lambda_{\mathrm{PC}}(\mathrm{R3})$	0.60	0.58	0.80	0.96	0.20	1.00	1.00	0.98	0.98	1.00	1.00	0.54	1.00	1.00
	$\Lambda_{\mathrm{HR}}(1)$	0.36	0.14	0.07	0.14	0.07	0.99	1.00	0.67	0.26	0.09	0.31	0.14	1.00	1.00
	$\Lambda_{\mathrm{HR}}(2)$	0.23	0.08	0.07	0.16	0.08	0.96	0.34	0.58	0.18	0.08	0.35	0.14	1.00	1.00

N5:

$$\boldsymbol{x}_t = \begin{pmatrix} 0.4 & -0.3 \\ -0.3 & 0.4 \end{pmatrix} \boldsymbol{x}_{t-1} + \begin{pmatrix} 0.1 & 0.0 \\ 0.1 & 0.0 \end{pmatrix} \times (\boldsymbol{x}_{t-1} \circ \boldsymbol{P}\boldsymbol{x}_{t-1}) + \boldsymbol{u}_t$$

N6:

$$\boldsymbol{x}_t = \begin{pmatrix} 0.7 & 0.0 \\ 0.3 & 0.7 \end{pmatrix} \boldsymbol{x}_{t-1} \mathbb{I}(\boldsymbol{e}_1' \boldsymbol{x}_{t-1} \leqslant 0) + \begin{pmatrix} -0.7 & 0.0 \\ -0.3 & -0.7 \end{pmatrix} \boldsymbol{x}_{t-1} \mathbb{I}(\boldsymbol{e}_1' \boldsymbol{x}_{t-1} > 0) + \boldsymbol{u}_t$$

N7:

$$\boldsymbol{x}_t = \begin{pmatrix} 0.7 & 0.0 \\ 0.3 & 0.7 \end{pmatrix} \boldsymbol{x}_{t-1} \{1 - G(\boldsymbol{e}_1' \boldsymbol{x}_{t-1})\} + \begin{pmatrix} -0.7 & 0.0 \\ -0.3 & -0.7 \end{pmatrix} \boldsymbol{x}_{t-1} G(\boldsymbol{e}_1' \boldsymbol{x}_{t-1}) + \boldsymbol{u}_t.$$

Here, $\{\boldsymbol{u}_t\}$ are independent, Gaussian, two-dimensional random vectors having zero mean and covariance matrix $\boldsymbol{\Sigma} = (\sigma_{ij})$, with $\sigma_{11} = \sigma_{22} = 1$ and $\sigma_{12} \in \{0, 0.5\}$, $\boldsymbol{P}$ is a permutation matrix defined as $\boldsymbol{P} = (\boldsymbol{e}_2, \boldsymbol{e}_1)$, where $\boldsymbol{e}_i$ ($i = 1, 2$) is the ith column of $\boldsymbol{I}_2$, $\mathbb{I}(\cdot)$ is the indicator function, $G(\cdot)$ is the logistic distribution function given by $G(z) = (1 + e^{-z})^{-1}$, and $\circ$ signifies element-wise multiplication. M1–M3 are linear VAR models, N1–N2 are vector bilinear models, N3–N5 are nonlinear VAR models, N6 is a threshold VAR model, and N7 is a smooth-transition VAR model.[3]

In the experiments, 1000 independent artificial time series of length $T + 100$, with $T \in \{200, 500\}$, are generated from each of the models above. The first 100 data points of each series are then discarded in order to eliminate start-up effects, and the remaining T data points are used to implement the multivariate nonlinearity tests based on the statistics Λ_{HR} and Λ_{PC}. In addition, we also carry out a univariate variant of the test based on Λ_{HR} applied to the two components of $\boldsymbol{x}_t$; these tests are denoted in the sequel by $\Lambda_{\mathrm{HR}}(1)$ and $\Lambda_{\mathrm{HR}}(2)$. In either case, the order of the autoregressive models fitted to the data is determined by means of the Bayesian criterion of Schwarz (1978), with the maximum order allowed set equal to the integer part of $8(T/100)^{1/4}$. For the test based on Λ_{PC}, the number of principal components used is selected according to the rules R1, R2 and R3 stated in Section 2; the resulting tests are denoted by $\Lambda_{\mathrm{PC}}(\mathrm{R1})$, $\Lambda_{\mathrm{PC}}(\mathrm{R2})$ and $\Lambda_{\mathrm{PC}}(\mathrm{R3})$, respectively.[4]

The Monte Carlo rejection frequencies of tests of nominal level 0.05 are reported in Tables 1 and 2, while Table 3 shows the number of original nonlinear terms (m) and principal components (n) used in the test regressions (averaged across Monte Carlo replications). The results reveal the following:

(i) For most design points, both the multivariate and univariate tests have empirical levels which do not differ significantly from the nominal 0.05 level. A small level distortion is observed in some cases, but even then the distortion is not of a magnitude that makes the tests unattractive for applications.
(ii) The multivariate tests are highly successful in detecting neglected nonlinearity in the larger of the two sample sizes considered, the only exception being the case of time series satisfying N5 with $\sigma_{12} = \frac{1}{2}$. The tests are also very successful

[3] Note that the chosen parameter values ensure stationarity. The maximum-over-sum procedure discussed in Embrechts et al. (1997, Sec. 6.2.6) indicates that the nonlinear time series have finite fourth moments.

[4] In addition to principal components obtained from the classical Pearson correlation matrix of $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_T)$, we also considered principal components based on the Spearman rank correlation matrix. Since no significant differences in the properties of the resulting tests were observed, we only report results based on the Pearson correlation matrix.

Table 3
Average number of nonlinear terms and principal components.

	Test	$T = 200$										$T = 500$									
		M1	M2	M3	N1	N2	N3	N4	N5	N6	N7	M1	M2	M3	N1	N2	N3	N4	N5	N6	N7
$\sigma_{12} = 0$	Λ_{HR}	11.1	5.7	5.6	11.1	11.0	4.2	4.4	5.9	4.5	3.9	10.6	8.1	8.1	12.9	12.6	4.6	5.4	8.2	4.7	4.3
	$\Lambda_{\mathrm{PC}}(\mathrm{R1})$	7.1	3.7	3.7	8.0	8.1	2.8	2.9	3.9	3.0	2.7	6.9	5.5	5.5	9.6	9.4	3.2	3.7	5.5	3.2	2.0
	$\Lambda_{\mathrm{PC}}(\mathrm{R2})$	5.9	3.0	3.0	5.9	5.9	2.8	2.6	3.1	2.6	3.0	6.0	3.6	3.6	6.5	6.5	3.0	2.8	3.7	2.5	1.9
	$\Lambda_{\mathrm{PC}}(\mathrm{R3})$	8.2	4.1	4.0	9.4	9.3	3.3	3.2	4.3	3.2	3.4	7.6	6.0	6.0	11.5	11.3	3.7	3.9	6.0	3.4	2.6
$\sigma_{12} = 0.5$	Λ_{HR}	11.3	4.5	5.3	11.4	11.5	4.5	4.9	5.0	4.5	3.8	10.6	5.3	7.2	13.8	14.2	5.0	6.1	7.2	4.4	4.1
	$\Lambda_{\mathrm{PC}}(\mathrm{R1})$	6.4	1.7	3.5	8.4	8.5	2.9	3.2	3.3	1.9	2.7	6.2	2.0	4.9	10.3	10.5	3.4	4.2	4.8	1.8	1.9
	$\Lambda_{\mathrm{PC}}(\mathrm{R2})$	4.5	1.7	3.5	7.1	7.2	3.1	3.5	3.4	1.8	3.0	3.9	2.0	4.5	8.5	8.8	3.4	4.2	4.5	1.7	1.8
	$\Lambda_{\mathrm{PC}}(\mathrm{R3})$	7.5	2.2	4.1	10.1	10.1	3.4	4.0	3.8	2.6	3.4	7.3	2.5	5.6	12.4	12.8	3.9	5.2	5.5	2.5	2.5

for shorter time series, particularly those satisfying N3, N4, N6 and N7. Except for the case of N5, tests tend to have higher power when $\sigma_{12} \neq 0$.

(iii) There appear to be no systematic and significant differences in the rejection frequencies of the multivariate test based on Λ_{HR} and those based on Λ_{PC}. Furthermore, tests based on Λ_{PC} are insensitive with respect to the rule used to select the number of principal components. Rule R3 appears to have a small advantage over R1 and R2 in terms of test power in some cases, although even then the differences across the three rules are extremely small. While all three rules achieve a dimensional reduction relative to the original test regression, R2 generally results in a smaller number of principal components being selected.

(iv) The multivariate tests based on Λ_{HR} and Λ_{PC} have a significant advantage over the univariate $\Lambda_{\mathrm{HR}}(1)$ and $\Lambda_{\mathrm{HR}}(2)$ tests. The differences in power are particularly prominent for time series satisfying N2, N3, N4 and N5; in these cases, univariate tests applied to individual time series have very little power to detect nonlinearity.

4. Conclusion

This paper has considered a test for neglected nonlinearity in multivariate time series which achieves a reduction in the dimension of the set of relevant test variables through the use of principal components. Monte Carlo experiments have revealed that principal components provide a simple and effective way of overcoming dimensionality and multicollinearity difficulties, with the resulting tests displaying no systematic level distortion or power loss relative to the original test considered by Harvill and Ray (1999). The simulation experiments have also revealed that, in a multivariate context, there are considerable advantages to testing for joint nonlinearity since univariate tests applied to individual components of a multiple time series may lack power to detect a nonlinear structure.

References

Anderson, T.W., 2003. An Introduction to Multivariate Statistical Analysis, third ed. Wiley, New York.
Embrechts, P., Kluppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer-Verlag, Berlin.
Harvill, J.L., Ray, B.K., 1999. A note on tests for nonlinearity in a vector time series. Biometrika 86, 728–734.
Jolliffe, I.T., 1972. Discarding variables in a principal component analysis I: artificial data. Appl. Stat. 21, 160–173.
Jolliffe, I.T., 2002. Principal Component Analysis, second ed. Springer, New York.
Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer-Verlag, Berlin.
Luukkonen, R., Saikkonen, P., Teräsvirta, T., 1988. Testing linearity against smooth transition autoregressive models. Biometrika 75, 491–499.
Schwarz, G., 1978. Estimating the dimension of a model. Ann. Statist. 6, 461–464.
Teräsvirta, T., Tjøstheim, D., Granger, C.W.J., 2010. Modelling Nonlinear Economic Time Series. Oxford University Press, Oxford.
Tong, H., 1990. Non-linear Time Series: A Dynamical System Approach. Oxford University Press, Oxford.
Tsay, R.S., 1986. Nonlinearity tests for time series. Biometrika 73, 461–466.