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

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).