ECONOMETRIC REVIEWS
http://dx.doi.org/10.1080/07474938.2016.1261015

Taylor & Francis
Taylor & Francis Group

Portmanteau tests for linearity of stationary time series

Zacharias Psaradakis[a] and Marián Vávra[b]

[a]Department of Economics, Mathematics and Statistics, Birkbeck, University of London, London, UK; [b]Advisor to the Governor and Research Department, National Bank of Slovakia, Bratislava, Slovakia

ABSTRACT
This article considers the problem of testing for linearity of stationary time series. Portmanteau tests are discussed which are based on generalized correlations of residuals from a linear model (that is, autocorrelations and cross-correlations of different powers of the residuals). The finite-sample properties of the tests are assessed by means of Monte Carlo experiments. The tests are applied to 100 time series of stock returns.

KEYWORDS
Autocorrelation; cross-correlation; nonlinearity; portmanteau test; stock returns

JEL CLASSIFICATION
C12; C22; C52

1. Introduction

The problem of testing for neglected nonlinearity in time series models has attracted a great deal of interest in recent years. A multitude of statistical procedures designed to test the null hypothesis of linearity against nonlinear alternatives are available in the literature, including general portmanteau tests without a specific alternative as well as tests with fully specified parametric alternatives; Tong (1990) and Teräsvirta et al. (2010) provide useful overviews. Linearity 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.

The present article contributes to this literature by considering portmanteau tests for linearity of stationary time series based on "generalized correlations" of residuals from a finite-parameter linear model, that is to say autocorrelations and cross-correlations of different powers of the residuals.[1] Such tests are similar in spirit to the popular test proposed by McLeod and Li (1983), which is based on the empirical autocorrelations of squared residuals. The McLeod–Li test is known to respond well to autoregressive conditional heteroskedasticity (ARCH) but tends to lack power against many other interesting types of nonlinearity that do not have apparent ARCH structures.

In addition to tests based on the empirical autocorrelations of the second or higher power of residuals, we also investigate tests that involve empirical cross-correlations between residuals and their squares (or, more generally, cross-correlations between different powers of the residuals). Lawrance and Lewis (1985, 1987) put forward the idea of using such cross-correlations to identify nonlinear dependence and examined analytically the cross-correlation functions for certain types of nonlinear models. Their analysis, however, focused only on visual inspection of individual cross-correlations and they did not consider the effects of parameter estimation.

In what follows, we tackle these problems by developing portmanteau tests based on the generalized correlations of residuals from linear models. The proposed tests are easy to implement and the relevant test statistics have standard asymptotic null distributions under general regularity conditions.

CONTACT Zacharias Psaradakis z.psaradakis@bbk.ac.uk Department of Economics, Mathematics and Statistics, Birkbeck, University of London, Malet Street, London WC1E 7HX, UK.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lecr.

[1]It is worth noting that our use of the term "generalized correlations" differs from the concept of "generalized autocorrelations" introduced recently in Proietti and Luati (2015).

Furthermore, tests based on cross-correlations are shown to be more powerful against many types of nonlinearity compared to the familiar test based on squared-residual autocorrelations.

The article is organized as follows. In Section 2, we discuss residual-based generalized correlations and the associated portmanteau tests for linearity, and present some relevant asymptotic results. Section 3 examines the finite-sample properties of the proposed tests by means of Monte Carlo experiments. Section 4 presents an application to time series of stock returns. Section 5 summarizes and concludes.

2. Generalized correlations and portmanteau statistics

Consider a second-order stationary, short-range dependent, real-valued stochastic process $\{X_t\}$ with mean μ satisfying

$$X_t - \mu = \Psi(L)\varepsilon_t, \qquad t \in \mathbb{Z}, \tag{1}$$

where

$$\Psi(z) = 1 + \sum_{j=1}^{\infty} \psi_j(\boldsymbol{\delta}) z^j, \qquad z \in \mathbb{C},$$

$\{\psi_j(\boldsymbol{\delta})\}$ is an absolutely summable sequence of weights, assumed to be known functions of a finite-dimensional (row) vector $\boldsymbol{\delta}$ of unknown parameters, $\{\varepsilon_t\}$ is strictly stationary white noise, and L denotes the lag operator. A leading example of a parametric model which gives rise to a process that is representable as in (1) is the autoregressive moving average (ARMA) model. In this case, the transfer function $\Psi(z)$ is of the form

$$\Psi(z) = B(z)/A(z), \qquad z \in \mathbb{C}, \tag{2}$$

where, for some fixed $p, q \in \mathbb{N} \cup \{0\}$ such that $p + q > 0$, $A(z) = 1 - \sum_{i=1}^{p} \alpha_i z^i$, with $A(z) \neq 0$ for all $|z| \leqslant 1$, $B(z) = 1 + \sum_{i=1}^{q} \beta_i z^i$, and $\boldsymbol{\delta} = (\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q)$.

A stochastic process $\{X_t\}$ is typically characterized as *linear* if it admits the moving-average (MA) representation (1) with $\{\varepsilon_t\}$ being independent and identically distributed (i.i.d.) random variables. This is the notion of linearity considered by McLeod and Li (1983), Lawrance and Lewis (1985, 1987), Bickel and Bühlmann (1996), Berg et al. (2010), and Giannerini et al. (2015), among many others, and is the one adopted in this article.[2] It is worth noting, however, that this is not the only characterization of linearity found in the literature. Hannan (1973), for instance, considers a second-order stationary process to be linear if its best one-step-ahead linear predictor is the best predictor (both in the mean-square sense), which is equivalent to $\{\varepsilon_t\}$ in (1) being a square-integrable martingale-difference sequence relative to its natural filtration. This alternative characterization of linearity does not lend itself to the type of statistical tests considered in the sequel.[3]

The focus of attention here are the generalized correlations of the noise $\{\varepsilon_t\}$ in (1). For $r, s \in \mathbb{N}$ such that $E(|\varepsilon_0|^{r+s}) < \infty$, we define the generalized correlations of $\{\varepsilon_t\}$ at lag k as

$$\rho_{rs}(k) = \{\gamma_{rr}(0)\gamma_{ss}(0)\}^{-1/2}\gamma_{rs}(k), \qquad k \in \mathbb{Z}, \tag{3}$$

where $\gamma_{rs}(k) = \mathrm{Cov}(\varepsilon_0^r, \varepsilon_k^s)$. Thus, (3) gives the autocorrelations of $\{\varepsilon_t\}$ for $r = s = 1$, the autocorrelations of $\{\varepsilon_t^2\}$ for $r = s = 2$, and cross-correlations of the type considered by Lawrance and Lewis (1985, 1987) for $(r, s) \in \{(1, 2), (2, 1)\}$. Linearity of $\{X_t\}$ implies that $\rho_{rs}(k) = 0$ for all $k \neq 0$.

When an estimator $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\boldsymbol{\delta}})$ of $\boldsymbol{\theta} = (\mu, \boldsymbol{\delta})$ is available, one may use residuals $\{\hat{\varepsilon}_t; t = 1, 2, \ldots, T\}$ (to be defined in a precise manner later) in place of the unobservable noise $\{\varepsilon_t\}$. For $r, s \in \mathbb{N}$, we define

[2] For example, a causal ARMA process satisfying (1)–(2) is considered to be linear if $\{\varepsilon_t\}$ are i.i.d. but nonlinear if $\{\varepsilon_t\}$ form an uncorrelated but not independent sequence (e.g., an infinite-order ARCH sequence with $\varepsilon_t = \eta_t \left(a_0 + \sum_{j=1}^{\infty} a_j \varepsilon_{t-j}^2\right)^{1/2}$ and $\mathrm{E}(\eta_0^2) \sum_{j=1}^{\infty} a_j < 1$, $\{\eta_t\}$ being i.i.d. zero-mean random variables).

[3] A test for linearity of the best predictor is discussed in Terdik and Máth (1998).

the empirical generalized correlations of the residuals at lag k as

$$\hat{\rho}_{rs}(k) = \left\{\hat{\gamma}_{rr}(0)\hat{\gamma}_{ss}(0)\right\}^{-1/2} \hat{\gamma}_{rs}(k), \qquad k = 0, \pm 1, \ldots, \pm(T-1), \tag{4}$$

where $\hat{\gamma}_{rs}(k) = T^{-1}\sum_{t=1}^{T-k} f_r(\hat{\varepsilon}_t)f_s(\hat{\varepsilon}_{t+k})$ for $k \geqslant 0$, $\hat{\gamma}_{rs}(k) = \hat{\gamma}_{sr}(-k)$ for $k < 0$, and $f_b(\xi_t) = \xi_t^b - T^{-1}(\xi_1^b + \cdots + \xi_T^b)$ for any collection of random variables $\{\xi_t\}$ and $b \in \mathbb{N}$. Tests for linearity of $\{X_t\}$ may then be based on test statistics of the form

$$\mathring{Q}_{rs}(m) = T\sum_{k=1}^{m} \hat{\rho}_{rs}^2(k), \tag{5}$$

for some $r, s, m \in \mathbb{N}$ such that $r + s > 2$ and $m < T$. Asymptotically equivalent statistics of the form

$$Q_{rs}(m) = T(T+2)\sum_{k=1}^{m}(T-k)^{-1}\hat{\rho}_{rs}^2(k), \tag{6}$$

may also be considered (cf. McLeod and Li, 1983), which are similar in spirit to the modification of the Box–Pierce statistic $\mathring{Q}_{11}(m)$ proposed by Ljung and Box (1978).

In order to develop asymptotic distribution theory for residual-based generalized correlations and associated portmanteau tests, the following assumptions are made (in the sequel, limits in stochastic-order symbols are taken by letting $T \to \infty$):

A1: $\{\varepsilon_t\}$ are i.i.d. with $E(\varepsilon_0) = 0$ and $0 < E(\varepsilon_0^2) < \infty$.

A2: $\Psi(z)$ is holomorphic in an open neighbourhood of the closed disc $|z| \leqslant 1$, does not vanish at any $|z| \leqslant 1$, and is differentiable with respect to $\boldsymbol{\delta}$.

A3: $\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = O_p(1)$.

A4: $\partial\tilde{\gamma}_{rs}(k)/\partial\boldsymbol{\theta} = O_p(T^{-1/2})$ for $k \in \{0, 1, \ldots, T-1\}$ and $r, s \in \mathbb{N}$ such that $r+s > 2$ and $E[|\varepsilon_0|^{2(r+s)}] < \infty$, where $\tilde{\gamma}_{rs}(k) = T^{-1}\sum_{t=1}^{T-k} f_r(\varepsilon_t)f_s(\varepsilon_{t+k})$.

Assumption A1 amounts to linearity of $\{X_t\}$ in our setting. Under A2, $1/\Psi(z)$ has the convergent power series expansion $1/\Psi(z) = \phi_0(\boldsymbol{\delta}) - \sum_{j=1}^{\infty} \phi_j(\boldsymbol{\delta})z^j$ for $|z| \leqslant 1$, with $\phi_0(\boldsymbol{\delta}) = 1$ and

$$\phi_j(\boldsymbol{\delta}) = \psi_j(\boldsymbol{\delta}) - \sum_{i=1}^{j-1} \phi_{j-i}(\boldsymbol{\delta})\psi_i(\boldsymbol{\delta}), \qquad j \in \mathbb{N},$$

and, consequently, $\{X_t\}$ admits the autoregressive (AR) representation

$$X_t - \mu = \sum_{j=1}^{\infty} \phi_j(\boldsymbol{\delta})(X_{t-j} - \mu) + \varepsilon_t, \qquad t \in \mathbb{Z}. \tag{7}$$

Hence, given an estimator $\hat{\boldsymbol{\theta}}$ based on a finite stretch $(X_0, X_1, \ldots, X_T)$ of $\{X_t\}$, residuals may be defined as (cf. Kreiss, 1991)

$$\hat{\varepsilon}_t = X_t - \hat{\mu} - \sum_{j=1}^{t} \phi_j(\hat{\boldsymbol{\delta}})(X_{t-j} - \hat{\mu}), \qquad t = 1, 2, \ldots, T.$$

Estimators of $\boldsymbol{\theta}$ satisfying assumption A3 may be obtained by quasi-maximum likelihood or instrumental-variables methods under suitable regularity conditions (see, e.g., Hannan, 1973; Dunsmuir, 1979; Hosoya and Taniguchi, 1982; Kuersteiner, 2001). In the ARMA case specified by (2), assumptions A2–A4 hold true, under an i.i.d. assumption about $\{\varepsilon_t\}$, as long as the polynomials $A(z)$ and $B(z)$ have no zeros in common and $A(z)B(z) \neq 0$ for all $|z| \leqslant 1$.

We have the following result for the asymptotic distribution of a finite set of empirical generalized correlations of the residuals defined by (4) under the assumption that $\{X_t\}$ is linear.

Theorem 1. *Suppose that* $\{X_t\}$ *satisfies (1) and assumptions* A1–A4 *hold. Then, for any fixed* $m \in \mathbb{N}$ *and* $r, s \in \mathbb{N}$ *such that* $r+s > 2$ *and* $E[|\varepsilon_0|^{2(r+s)}] < \infty$*, the asymptotic distribution of* $\sqrt{T}\left(\hat{\rho}_{rs}(1), \dots, \hat{\rho}_{rs}(m)\right)$*, as* $T \to \infty$*, is Gaussian with zero mean vector and identity covariance matrix.*

Proof. For a fixed $m < T$, a Taylor expansion of $\hat{\gamma}_{rs}(k)$ about $\boldsymbol{\theta}$ leads to

$$\hat{\gamma}_{rs}(k) = \tilde{\gamma}_{rs}(k) + \frac{\partial \tilde{\gamma}_{rs}(k)}{\partial \boldsymbol{\theta}}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})' + O_p(T^{-1}) = \tilde{\gamma}_{rs}(k) + O_p(T^{-1}), \quad k = 0, 1, \dots, m.$$

Hence, the distribution of $\sqrt{T}(\hat{\gamma}_{rs}(1) - \gamma_{rs}(1), \dots, \hat{\gamma}_{rs}(m) - \gamma_{rs}(m))$ is asymptotically the same as the distribution of $\sqrt{T}(\tilde{\gamma}_{rs}(1) - \gamma_{rs}(1), \dots, \tilde{\gamma}_{rs}(m) - \gamma_{rs}(m))$. Furthermore, putting $\dot{f}_b(\varepsilon_t) = \varepsilon_t^b - E(\varepsilon_0^b)$, $b \in \mathbb{N}$, and noting that $T^{-1}\sum_{t=1}^{T}\dot{f}_b(\varepsilon_t) = O_p(T^{-1/2})$ for $b \in \{r, s\}$, it is not difficult to show that $\tilde{\gamma}_{rs}(k) - T^{-1}\sum_{t=1}^{T}\dot{f}_r(\varepsilon_t)\dot{f}_s(\varepsilon_{t+k}) = o_p(T^{-1/2})$ for $0 \leqslant k \leqslant m$. Therefore, recalling that $\gamma_{rs}(k) = 0$ for all $k \neq 0$ under assumption A1, by an application of the central limit theorem for strictly stationary, finitely dependent sequences (e.g., Anderson, 1971, Theorem 7.7.6) to the normalized partial sum $T^{-1/2}\sum_{t=1}^{T}(\dot{f}_r(\varepsilon_t)\dot{f}_s(\varepsilon_{t+1}), \dots, \dot{f}_r(\varepsilon_t)\dot{f}_s(\varepsilon_{t+m}))$, we may conclude that, as $T \to \infty$, the distribution of $\sqrt{T}\{\gamma_{rr}(0)\gamma_{ss}(0)\}^{-1/2}(\hat{\gamma}_{rs}(1), \dots, \hat{\gamma}_{rs}(m))$ converges weakly to the standard normal distribution on $\mathbb{R}^m$. The assertion of the theorem follows from this result and the fact that $\hat{\gamma}_{bb}(0) = \tilde{\gamma}_{bb}(0) + O_p(T^{-1}) = \gamma_{bb}(0) + o_p(1)$ for $b \in \{r, s\}$.

Theorem 1 generalizes the central limit theorem of McLeod and Li (1983), which is restricted to the case where $r = s = 2$ and the transfer function $\Psi(z)$ is rational. It is readily seen that, under the conditions of Theorem 1, the asymptotic distribution of the test statistics defined in (5) and (6) is chi-square with m degrees of freedom. The implementation of tests based on statistics such as $\mathring{Q}_{rs}(m)$ and $Q_{rs}(m)$ is straightforward and computationally inexpensive.[4]

3. Monte Carlo simulations

This section presents simulation results regarding the properties of portmanteau tests for linearity. In addition to the finite-sample size and power properties of the tests, we also examine the effects of non-Gaussian noise, measurement errors, correlation order, and multiple testing.

3.1. Simulation design

The following data-generating processes (DGPs) are used in the simulations:
M1: $X_t = 0.8X_{t-1} + \varepsilon_t$
M2: $X_t = 0.6X_{t-1} - 0.5X_{t-2} + \varepsilon_t$
M3: $X_t = 0.8\varepsilon_{t-1} + \varepsilon_t$
M4: $X_t = 0.8X_{t-1} + 0.15X_{t-2} + 0.3\varepsilon_{t-1} + \varepsilon_t$
M5: $X_t = 0.6X_{t-1} + 0.4\varepsilon_{t-1} + \varepsilon_t$
M6: $X_t = 0.8X_{t-1}I(X_{t-1} \leqslant -1) - 0.8X_{t-1}I(X_{t-1} > -1) + \varepsilon_t$
M7: $X_t = -0.5X_{t-1}I(X_{t-1} \leqslant 1) + 0.4X_{t-1}I(X_{t-1} > 1) + \varepsilon_t$
M8: $X_t = -0.5X_{t-1}\{1 - G(X_{t-1})\} + 0.4X_{t-1}G(X_{t-1}) + \varepsilon_t$
M9: $X_t = 0.8X_{t-1}\{1 - G(X_{t-1})\} - 0.8X_{t-1}G(X_{t-1}) + \varepsilon_t$
M10: $X_t = 0.8|X_{t-1}|^{1/2} + \varepsilon_t$
M11: $X_t = Y_t^2 + \varepsilon_t, \quad Y_t = 0.6Y_{t-1} + \nu_t$
M12: $X_t = \sigma_t\varepsilon_t, \quad \sigma_t^2 = 0.1 + 0.6X_{t-1}^2$
M13: $X_t = \sigma_t\varepsilon_t, \quad \sigma_t^2 = 0.01 + 0.12X_{t-1}^2 + 0.85\sigma_{t-1}^2$

[4]For example, the full set of Monte Carlo experiments reported in Section 3 took approximately 3 hours to carry out in MATLAB running under Windows 7 (64-bit) on a laptop with Intel Core i7 2.60 GHz processor and 8 GB of RAM.

M14: $X_t = \sigma_t \varepsilon_t, \quad \ln \sigma_t^2 = 0.01 + 0.3\{|\varepsilon_{t-1}| - E(|\varepsilon_{t-1}|)\} - 0.8\varepsilon_{t-1} + 0.9 \ln \sigma_{t-1}^2$
M15: $X_t = 0.4X_{t-1} - 0.3X_{t-2} + (0.8 + 0.5X_{t-1})\varepsilon_{t-1} + \varepsilon_t$
M16: $X_t = 0.5 - (0.4 - 0.4\varepsilon_{t-1})X_{t-1} + \varepsilon_t$
M17: $X_t = 0.8\varepsilon_{t-2}^2 + \varepsilon_t$
M18: $X_t = -0.3\varepsilon_{t-1} + (0.2 + 0.4\varepsilon_{t-1} - 0.25\varepsilon_{t-2})\varepsilon_{t-2} + \varepsilon_t$
Unless stated otherwise, $\{\varepsilon_t\}$ and $\{\nu_t\}$ are i.i.d. standard normal random variables independent of each other, $G(x) = 1/(1+e^{-x})$ is the logistic distribution function, and $I(A)$ denotes the indicator of event A. The DGPs cover a variety of linear and nonlinear processes used in the literature, namely ARMA [M1–M5], threshold AR (TAR) [M6, M7], smooth-transition AR (STAR) [M8, M9], fractional AR (FAR) [M10], square AR (SQAR) [M11], ARCH [M12], generalized ARCH (GARCH) [M13], exponential GARCH (EGARCH) [M14], bilinear (BL) [M15, M16], and nonlinear MA (NLMA) [M17, M18].[5]

In the experiments, 5,000 independent artificial time series $\{X_t\}$ of length $100 + T$, with $T \in \{200, 500, 1000\}$, are generated according to M1–M18, but only the last T data points of each series are used to carry out portmanteau tests for linearity. As preliminary analysis indicated that, for relatively short time series, tests based on the statistics $Q_{rs}(m)$ defined in (6) control the Type I error probability somewhat more successfully (albeit marginally) than tests based on the statistics $\mathring{Q}_{rs}(m)$ defined in (5), we shall henceforth focus on the former.

Unless indicated otherwise, the tests are applied to least-squares residuals from an AR model for $\{X_t\}$ the order of which is determined by the Bayesian information criterion (BIC). The BIC is defined according to Method 1 of Ng and Perron (2005) with the maximum allowable order set equal to $\lfloor 8(T/100)^{1/4} \rfloor$, where $\lfloor x \rfloor$ denotes the largest integer not exceeding x.[6] Employing an AR model with data-dependent order as the null specification is not only computationally convenient but also theoretically attractive. Even when the DGP is not a finite-order AR process, an AR model the order of which increases simultaneously with the sample size may be viewed as a finite-parameter approximation to a linear process that admits the infinite-order AR representation (7). If the order of the AR approximation grows at a suitable rate, the approximation error becomes small as T increases, and estimates of the parameters in (7) obtained from the approximating autoregression are consistent and asymptotically normal (see Berk, 1974; Bhansali, 1978; Lewis and Reinsel, 1985).

3.2. *Empirical size and power*

The Monte Carlo rejection frequencies of tests, of nominal level 0.05, based on the statistics $Q_{12}(m)$, $Q_{21}(m)$ and $Q_{22}(m)$, with $m \in \{1, 2, \ldots, \lfloor \sqrt{T} \rfloor\}$, are shown in Figs. 1–3.[7] Under linear DGPs (M1–M5), all three portmanteau tests have empirical levels which do not differ significantly from the nominal level regardless of the sample size T and the number of generalized correlations m used to construct the test statistic. It is noteworthy that the tests work well in the case of linear DGPs which do not have a finite-order AR structure (M3–M5), suggesting that AR approximations provide a useful way of modelling dynamics under the null hypothesis in this context.

For all but two of the nonlinear DGPs (M6–M18), at least one of the two cross-correlation tests Q_{12} and Q_{21} has higher rejection frequencies than the Q_{22} test, especially when $T \leqslant 500$. The test based on Q_{22} has a clear advantage in the case of time series generated according to M12 and M13, which is not perhaps surprising since Q_{22} is asymptotically equivalent to a Lagrange multiplier statistic for testing linearity against ARCH (see Luukkonen et al., 1988). The power of all the tests generally improves as T increases.

[5] The DGPs are taken from Lee et al. (1993) [M11, M15, M18], Barnett et al. (1997) [M4], Hong and Lee (2003) [M2, M14], Hong and White (2005) [M10], and Giannerini et al. (2015) [M1, M3, M5, M6, M7, M12, M13, M16, M17]; M8 and M9 are smooth-transition variants of M7 and M6, respectively.
[6] Very similar results are obtained using Akaike's information criterion instead of the BIC.
[7] Simulation results for tests of nominal level 0.01 and 0.10 are not reported, due to space constraints, but are available upon request.

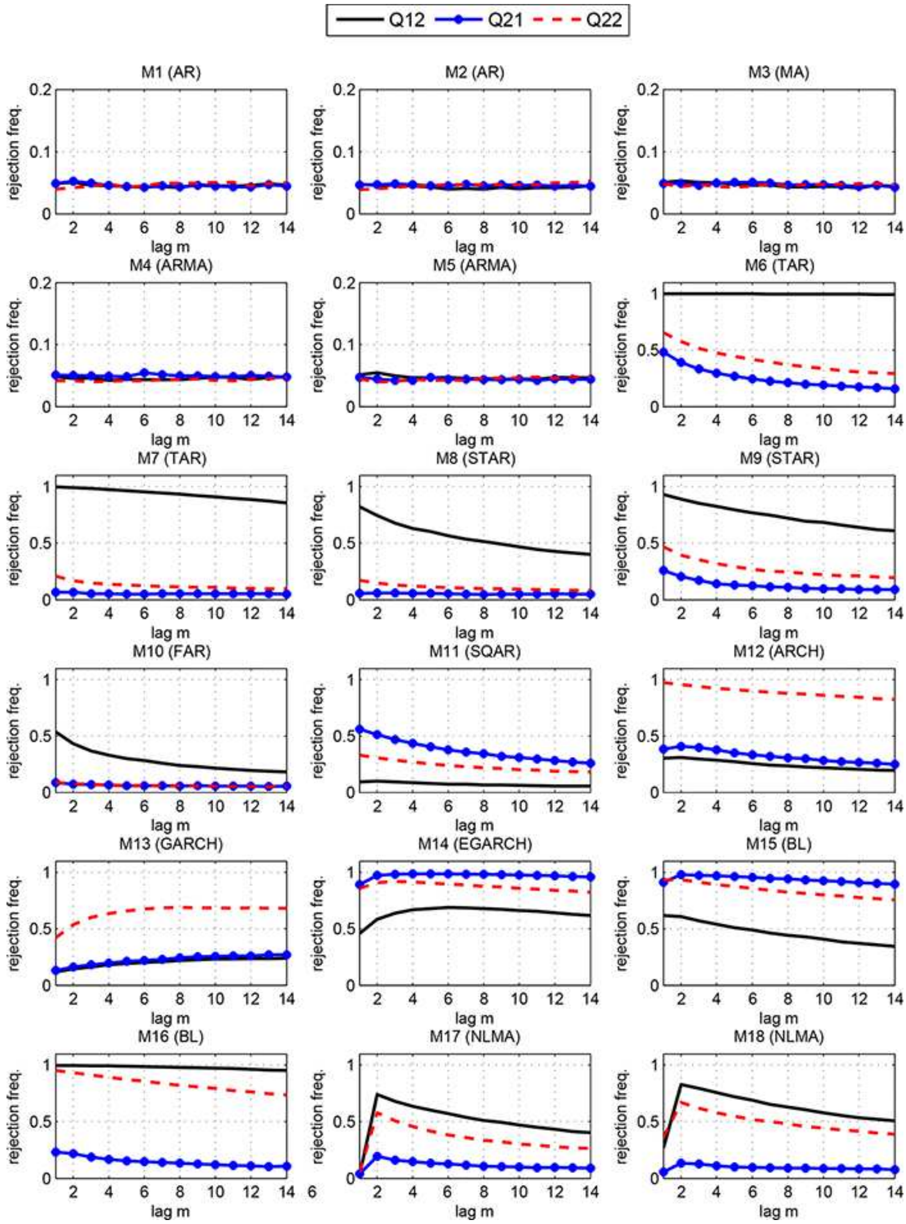


Figure 1. Rejection frequencies of Q_{rs} tests: $T = 200$.

3.3. Non-Gaussian noise

To investigate the sensitivity of the simulation results with respect to non-Gaussianity of the noise in the DGP, we consider artificial time series (of length $T = 500$) generated according to M1–M18 with ε_t having either Student's t distribution with d degrees of freedom or a gamma distribution with shape parameter d and scale parameter 1. (The distributions are recentred and/or rescaled so as to have zero mean and unit variance). We take $d \in \{10, 11, \ldots, 19, 20\}$, a range of values which is sufficiently representative of some of the distributional characteristics (e.g., mild asymmetry and leptokurtosis) of many economic and financial time series. Following the suggestion in Tong (1990, p. 324) that, when

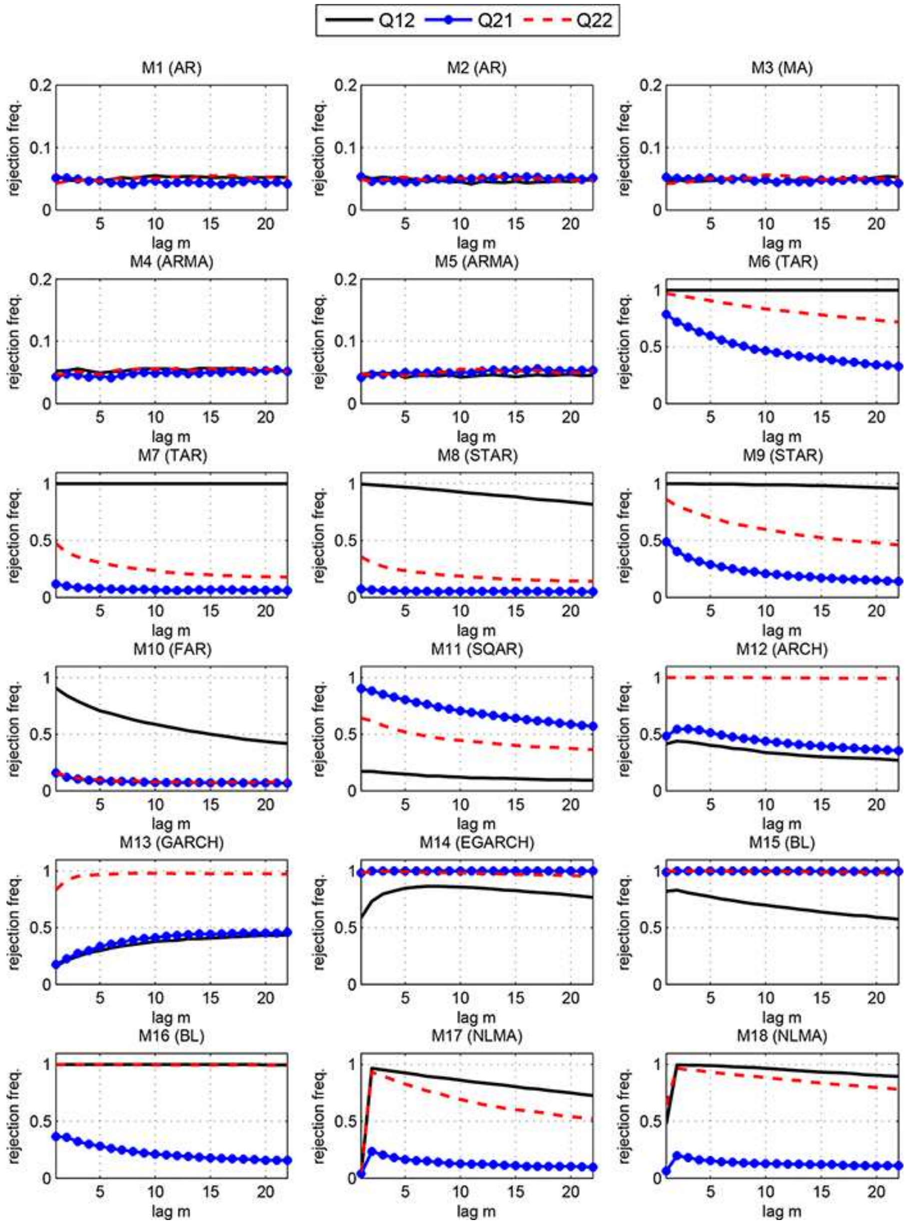


Figure 2. Rejection frequencies of Q_{rs} tests: $T = 500$.

constructing tests for uncorrelatedness, autocorrelations at low lags should be watched more closely than autocorrelations at high lags, we set $m = \lfloor \ln T \rfloor$ (see also Tsay, 2010, p. 33).

For the sake of expositional simplicity and space conservation, the rejection frequencies of tests (of nominal level 0.05) are averaged over the linear (M1–M5) and nonlinear (M6–M18) DGPs, and are shown in Fig. 4 (straight lines indicate the average rejection frequencies of tests under Gaussian noise). The results indicate that the level and power properties of the tests are generally insensitive with respect to the value of the parameter d. In the case of gamma distributed noise, Q_{22} (Q_{21}) has marginally lower (higher) average power compared to the Gaussian case.

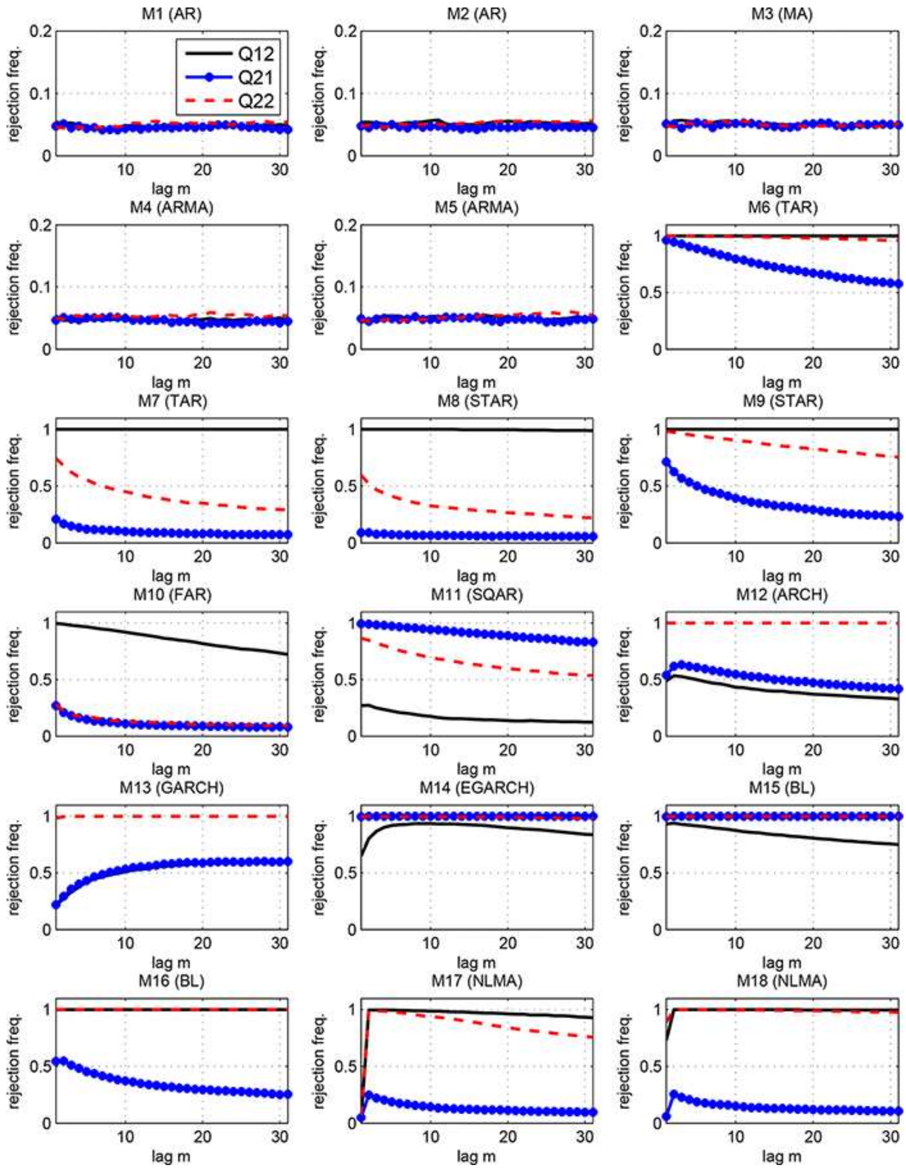
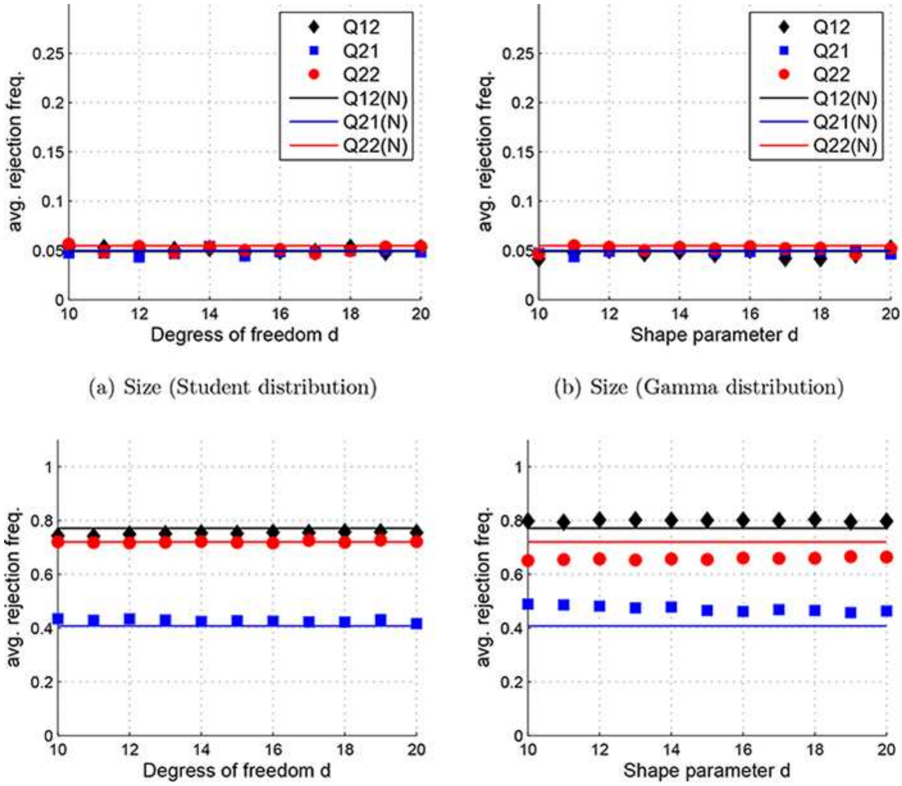


Figure 3. Rejection frequencies of Q_{rs} tests: $T = 1{,}000$.

3.4. Measurement errors

Economic and financial time series are often contaminated by measurement errors due to, *inter alia*, sampling, self-reporting, or imperfect data sources. To investigate the potential effect of such measurement errors on tests for nonlinearity, we consider contaminated series (of length $T = 500$) generated according to $X_t^* = X_t + \sigma_\eta \eta_t$, where X_t comes from M1–M18 and $\{\eta_t\}$ are i.i.d. random variables, independent of $\{\varepsilon_t\}$ and $\{\nu_t\}$, having either Student's t distribution with 10 degrees of freedom or a gamma distribution with shape parameter 10 and scale parameter 1 (recentered and/or rescaled to have zero mean and unit variance). The variance of the measurement error is allowed to be proportional to the sample variance $\hat{\sigma}_x^2$ of $(X_1, \ldots, X_T)$, that is $\sigma_\eta^2 = \omega^2\,\hat{\sigma}_x^2$, where $\omega^2 \in \{0.005, 0.010, \ldots, 0.060, 0.065\}$. The range of values for the noise-to-signal ratio ω is calibrated according to Koreisha and Fang (1999) and allows for up to 25% contamination by measurement errors.



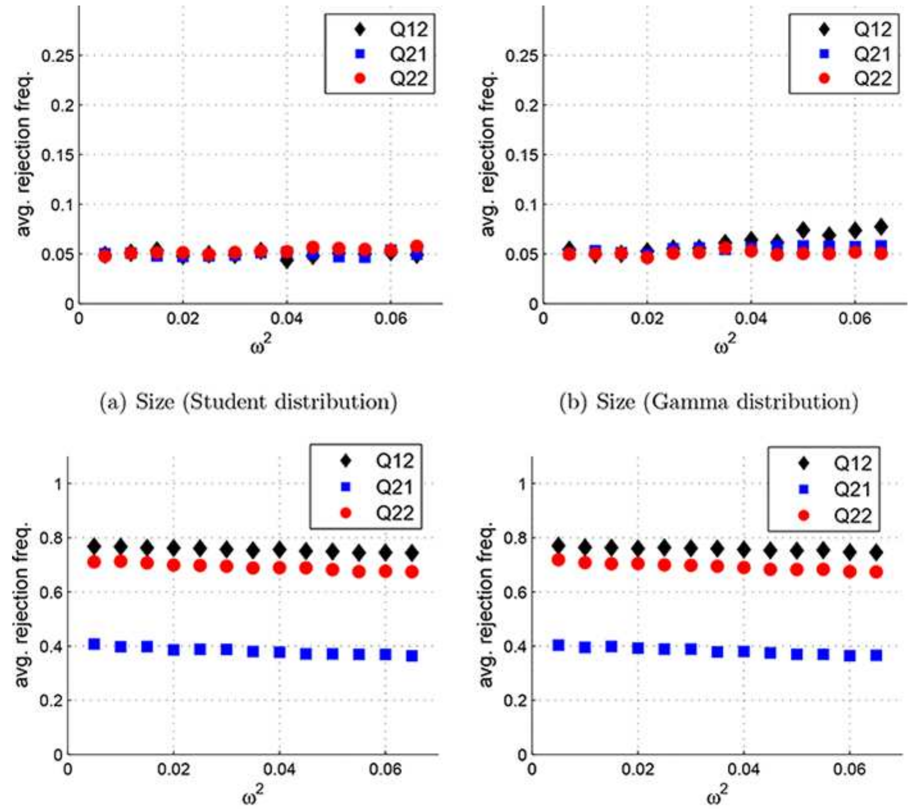
(a) Size (Student distribution)
(b) Size (Gamma distribution)
(c) Power (Student distribution)
(d) Power (Gamma distribution)

Figure 4. Rejection frequencies of Q_{rs} tests under non-Gaussian noise.

Tests for linearity based on $Q_{rs}(m)$, with $r, s \in \{1, 2\}$ and $m = \lfloor \ln T \rfloor$, are implemented as described in Section 3.1 using $\{X_t^*\}$ in place of $\{X_t\}$. The rejection frequencies of tests (of nominal level 0.05), averaged across the linear (M1–M5) and nonlinear (M6–M18) DGPs, are shown in Fig. 5. The tests exhibit no substantial size distortion regardless of the contamination rate and the distribution of the noise. Some power loss is observed as the contamination rate increases, but the reduction in power is not of the magnitude that makes the tests unattractive for applications.

3.5. Higher-order correlations

Although the discussion in much of the article focuses on tests with $r, s \in \{1, 2\}$, the use of higher values for (r, s) is, of course, possible. To examine whether power gains may be made by using higher-order generalized correlations, we compute the empirical power of tests based on $Q_{rs}(m)$ with $r, s \in \{1, 2, \ldots, 6\}$ and $m = \lfloor \ln T \rfloor$. The rejection frequencies of tests (of nominal level 0.05) for $T = 500$, averaged across the nonlinear DGPs (M6–M18), are reported in Table 1. The results indicate that there are generally no power improvements associated with the use of higher-order generalized correlations; for instance, tests based on Q_{12} and Q_{32} have almost the same (average) rejection frequencies. Furthermore, it is worth bearing in mind that the asymptotic justification of portmanteau tests associated with high values of (r, s) requires finiteness of a fairly large number of moments (cf. Theorem 1). This requirement may be at odds with the characteristics of many economic and financial time series (e.g., equity returns, exchange rate returns, interest rates), for which it is often argued that they only possess unconditional moments of relatively low order (see, e.g., Koedijk et al., 1990; Jansen and de Vries, 1991; de Lima, 1997).



(a) Size (Student distribution) (b) Size (Gamma distribution)

(c) Power (Student distribution) (d) Power (Gamma distribution)

Figure 5. Rejection frequencies of Q_{rs} tests under contamination.

Table 1. Rejection frequencies of Q_{rs} tests.

w	Q_{1w}	Q_{w1}	Q_{2w}	Q_{w2}	Q_{ww}
2	0.77	0.40	0.72	0.72	0.72
3	0.19	0.24	0.42	0.76	0.38
4	0.65	0.42	0.60	0.66	0.53
5	0.24	0.32	0.36	0.66	0.32
6	0.53	0.42	0.48	0.60	0.34

3.6. Multiple testing

In practice, linearity is often tested using several tests (e.g., $Q_{rs}(m)$, $r, s \in \{1, 2\}$) jointly and/or several values of m. However, unless adjustments for multiple testing are made, there is an increased risk of overstating the significance of nonlinearity when many tests are carried out using the same set of data (see Psaradakis, 2000). This is due to the fact that, if the linearity hypothesis is rejected when at least one of the tests leads to a rejection, the overall Type I error probability associated with the multiple testing procedure (i.e., the probability of at least one erroneous rejection) can be well in excess of the nominal level of each individual test.

A simple Bonferroni-type adjustment for multiple testing based on Simes (1986) is considered here, which may be implemented as follows. Let $P_{(1)} \leqslant P_{(2)} \leqslant \cdots \leqslant P_{(N)}$ denote the ordered (asymptotic) P-values associated with the set of portmanteau test statistics under consideration. Multiplicity-adjusted P-values are then calculated as $\widetilde{P}_{(i)} = \min\{NP_{(i)}/i, 1\}$, $i \in \{1, 2, \ldots, N\}$, and the null hypothesis of linearity is rejected at overall level $\alpha \in (0, 1)$ if $\min_{1 \leqslant i \leqslant N} \widetilde{P}_{(i)} \leqslant \alpha$. Simes' procedure is generally less

Table 2. Rejection frequencies under multiple testing.

	Case A ($N = 3$)		Case B ($N = 18$)	
	Unadjusted	Adjusted	Unadjusted	Adjusted
M1	0.141	0.048	0.286	0.042
M2	0.136	0.044	0.287	0.039
M3	0.136	0.048	0.296	0.041
M4	0.137	0.054	0.284	0.043
M5	0.141	0.055	0.287	0.042
M6	1.000	1.000	1.000	1.000
M7	1.000	1.000	1.000	1.000
M8	0.964	0.919	0.997	0.966
M9	0.997	0.992	1.000	0.997
M10	0.723	0.542	0.931	0.681
M11	0.799	0.703	0.941	0.788
M12	1.000	0.999	1.000	0.999
M13	0.980	0.965	0.984	0.944
M14	1.000	1.000	1.000	1.000
M15	1.000	1.000	1.000	1.000
M16	1.000	1.000	1.000	1.000
M17	0.946	0.894	0.989	0.920
M18	0.998	0.991	1.000	0.995

conservative than the classical Bonferroni or Šidák procedures, especially when several highly correlated test statistics are involved. It also yields the same critical values as the multiple testing procedure of Benjamini and Hochberg (1995), which controls the so-called false discovery rate (i.e., the expected proportion of erroneous rejections among all rejections) at level α.

In Table 2 we report Monte Carlo estimates of the probability that at least one of the tests under consideration rejects the null hypothesis of linearity at the 0.05 level (when $T = 500$). Case A refers to the situation when linearity is tested using the statistics $Q_{12}(m)$, $Q_{21}(m)$, and $Q_{22}(m)$ with $m = \lfloor \ln T \rfloor = 6$ ($N = 3$); in Case B linearity is tested using $Q_{12}(m)$, $Q_{21}(m)$, $Q_{22}(m)$ and six different values of m, namely $m \in \{1, \dots, 6\}$ ($N = 18$). The advantage of adjusting for multiplicity in testing is immediately evident. Using unadjusted P-values, the probability that one or more of the tests will erroneously reject the null hypothesis under M1–M5 ranges from 0.14 to 0.29. By contrast, the multiple testing procedures generally have an overall Type I error probability that is quite close to the nominal 0.05 level (in spite of the fact that they do not account for dependence among the individual test statistics). Moreover, the protection against an excessive overall Type I error probability is not achieved at the cost of a systematic loss of the ability of the tests to reject correctly the linearity hypothesis under M6–M18.

4. Empirical application

In this section, portmanteau tests for linearity are applied to a set of weekly stock returns, spanning the period 1993–2007 (781 observations), for 100 companies from the Standard & Poor's 500 Composite index. The selected series are part of the data set analyzed by Kapetanios (2009) and are such that the hypothesis of strict stationarity cannot be rejected for any of them (at 5% significance level). The presence of nonlinearity in asset returns has important implications for, *inter alia*, pricing, risk management, and forecasting.

As in Section 3, we test for neglected nonlinearity in an AR model for each time series, the order of which is determined by minimizing the BIC over the range $\{0, 1, \dots, \lfloor 8(T/100)^{1/4} \rfloor\}$. The asymptotic P-values for tests based on $Q_{12}(m)$, $Q_{21}(m)$, and $Q_{22}(m)$, with $m = \lfloor \ln T \rfloor$, are reported in Table 3. In order to guard against the danger of overstating the significance of nonlinearity because of the use of three different tests, we also report the P-values of the individual test statistics adjusted for multiplicity using the methods of Simes (1986) and Benjamini and Hochberg (1995). In the notation of Section 3.6, the Benjamini–Hochberg adjusted P-values are computed as $\check{P}_{(i)} = \min\{NP_{(i)}/i, \check{P}_{(i+1)}\}$ for $i \in \{1, \dots, N-1\}$ and $\check{P}_{(N)} = P_{(N)}$.

Table 3. Unadjusted and multiplicity-adjusted P-values.

	Unadjusted P-values			Simes P-values			Benjamini-Hochberg P-values			AR order
Company	Q_{12}	Q_{21}	Q_{22}	Q_{12}	Q_{21}	Q_{22}	Q_{12}	Q_{21}	Q_{22}	
Alcoa Inc	0.20	0.00	0.00	0.20	0.01	0.00	0.20	0.01	0.00	1
Apple Inc.	0.01	0.70	0.47	0.02	0.70	0.70	0.02	0.70	0.70	2
Adobe Systems Inc	0.21	0.01	0.00	0.21	0.01	0.00	0.21	0.01	0.00	1
Analog Devices Inc	0.07	0.00	0.00	0.07	0.00	0.00	0.07	0.00	0.00	1
Archer-Daniels-Midland	0.56	0.99	0.04	0.85	0.99	0.11	0.85	0.99	0.11	1
Autodesk Inc	0.10	0.19	0.01	0.15	0.19	0.02	0.15	0.19	0.02	1
American Electric Power	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
AES Corp	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
AFLAC Inc	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
Allergan Inc	0.04	0.09	0.00	0.06	0.09	0.00	0.06	0.09	0.00	1
American Intl Group Inc	0.03	0.00	0.00	0.03	0.00	0.00	0.03	0.00	0.00	1
Aon plc	0.02	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00	1
Apache Corporation	0.37	0.02	0.00	0.37	0.03	0.00	0.37	0.03	0.00	1
Anadarko Petroleum	0.73	0.05	0.00	0.73	0.08	0.00	0.73	0.08	0.00	1
Avon Products	0.19	0.00	0.00	0.19	0.00	0.00	0.19	0.00	0.00	1
Avery Dennison Corp	0.00	0.12	0.00	0.01	0.12	0.00	0.01	0.12	0.00	1
American Express Co	0.24	0.00	0.00	0.24	0.00	0.00	0.24	0.00	0.00	1
Bank of America Corp	0.21	0.00	0.00	0.21	0.00	0.00	0.21	0.00	0.00	1
Baxter International Inc.	0.06	0.00	0.28	0.09	0.00	0.28	0.09	0.00	0.28	1
BBT Corporation	0.58	0.01	0.00	0.58	0.02	0.00	0.58	0.02	0.00	1
Best Buy Co. Inc.	0.75	0.00	0.00	0.75	0.00	0.00	0.75	0.00	0.00	1
Bard (C.R.) Inc.	0.71	0.03	0.01	0.71	0.05	0.02	0.71	0.05	0.02	1
Becton Dickinson	0.24	0.07	0.00	0.24	0.11	0.00	0.24	0.11	0.00	1
Franklin Resources	0.38	0.00	0.00	0.38	0.00	0.00	0.38	0.00	0.00	1
Brown-Forman Corp	0.29	0.00	0.04	0.29	0.00	0.06	0.29	0.00	0.06	1
Baker Hughes Inc	0.24	0.00	0.00	0.24	0.00	0.00	0.24	0.00	0.00	1
The Bank of NY Mellon	0.23	0.00	0.00	0.23	0.00	0.00	0.23	0.00	0.00	1
Ball Corp	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
Boston Scientific	0.95	0.01	0.01	0.95	0.02	0.04	0.95	0.02	0.02	1
Cardinal Health Inc.	0.83	0.05	0.02	0.83	0.08	0.05	0.83	0.08	0.05	3
Caterpillar Inc.	0.08	0.05	0.01	0.08	0.07	0.04	0.08	0.07	0.04	1
Chubb Corp.	0.01	0.04	0.00	0.02	0.04	0.00	0.02	0.04	0.00	1
Coca-Cola Enterprises	0.26	0.31	0.06	0.39	0.31	0.18	0.31	0.31	0.18	1
Carnival Corp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
CIGNA Corp.	0.24	0.00	0.70	0.36	0.00	0.70	0.36	0.00	0.70	1
Cincinnati Financial	0.00	0.63	0.00	0.00	0.63	0.00	0.00	0.63	0.00	1
Clorox Co.	0.18	0.00	0.00	0.18	0.00	0.00	0.18	0.00	0.00	1
Comerica Inc.	0.39	0.00	0.01	0.39	0.00	0.02	0.39	0.00	0.02	1
CMS Energy	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2
CenterPoint Energy	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
Cabot Oil and Gas	1.00	0.02	0.00	1.00	0.03	0.00	1.00	0.03	0.00	1
ConocoPhillips	0.18	0.38	0.49	0.53	0.57	0.49	0.49	0.49	0.49	1
Campbell Soup	0.86	0.00	0.00	0.86	0.00	0.00	0.86	0.00	0.00	1
CSX Corp.	0.08	0.04	0.00	0.08	0.06	0.01	0.08	0.06	0.01	1
CenturyLink Inc	0.10	0.00	0.05	0.10	0.01	0.08	0.10	0.01	0.08	1
Cablevision Corp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
Chevron Corp.	0.04	0.33	0.01	0.06	0.33	0.03	0.06	0.33	0.03	1
Dominion Resources	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
Deere and Co.	0.86	0.04	0.00	0.86	0.07	0.00	0.86	0.07	0.00	2
D. R. Horton	0.39	0.02	0.01	0.39	0.03	0.04	0.39	0.03	0.03	1
Danaher Corp.	0.01	0.05	0.00	0.02	0.05	0.00	0.02	0.05	0.00	3
Walt Disney Co.	0.62	0.05	0.01	0.62	0.07	0.04	0.62	0.07	0.04	1
Dow Chemical	0.19	0.09	0.00	0.19	0.14	0.00	0.19	0.14	0.00	1
Duke Energy	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
Ecolab Inc.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
Equifax Inc.	0.13	0.23	0.12	0.20	0.23	0.37	0.20	0.23	0.20	1
Edison Int'l	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	2
EMC Corp.	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	1
Emerson Electric	0.31	0.03	0.00	0.31	0.04	0.00	0.31	0.04	0.00	1
Equity Residential	0.45	0.22	0.00	0.45	0.34	0.00	0.45	0.34	0.00	1

(*Continued*)

Table 3. Continued.

Company	Unadjusted *P*-values			Simes *P*-values			Benjamini-Hochberg *P*-values			AR order
	Q_{12}	Q_{21}	Q_{22}	Q_{12}	Q_{21}	Q_{22}	Q_{12}	Q_{21}	Q_{22}	
EQT Corporation	0.20	0.00	0.00	0.20	0.00	0.00	0.20	0.00	0.00	1
Eaton Corp.	0.47	0.22	0.40	0.47	0.67	0.59	0.47	0.47	0.47	1
Entergy Corp.	0.01	0.11	0.00	0.01	0.11	0.00	0.01	0.11	0.00	1
Exelon Corp.	0.45	0.43	0.02	0.45	0.64	0.05	0.45	0.45	0.05	1
Ford Motor	0.01	0.15	0.00	0.01	0.15	0.00	0.01	0.15	0.00	1
Fastenal Co	0.28	0.06	0.03	0.28	0.09	0.08	0.28	0.09	0.08	1
Family Dollar Stores	0.86	0.05	0.00	0.86	0.08	0.00	0.86	0.08	0.00	1
FedEx Corporation	0.04	0.17	0.00	0.06	0.17	0.00	0.06	0.17	0.00	1
Fiserv Inc	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3
Fifth Third Bancorp	0.04	0.00	0.00	0.04	0.01	0.00	0.04	0.01	0.00	1
Fluor Corp.	0.00	0.03	0.12	0.00	0.04	0.12	0.00	0.04	0.12	1
Forest Laboratories	0.03	0.08	0.64	0.09	0.12	0.64	0.09	0.12	0.64	1
Frontier Commun.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
Gannett Co.	0.64	0.01	0.00	0.64	0.01	0.00	0.64	0.01	0.00	1
General Dynamics	0.50	0.01	0.00	0.50	0.02	0.00	0.50	0.02	0.00	1
General Electric	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
General Mills	0.98	0.87	0.98	0.98	1.00	1.00	0.98	0.98	0.98	1
Genuine Parts	0.02	0.02	0.00	0.02	0.03	0.00	0.02	0.02	0.00	1
Gap (The)	0.81	0.00	0.00	0.81	0.00	0.00	0.81	0.00	0.00	1
Grainger Inc.	0.37	0.05	0.00	0.37	0.07	0.00	0.37	0.07	0.00	2
Halliburton Co.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1
Harman Int'l Ind.	0.63	0.35	0.54	0.63	1.00	0.81	0.63	0.63	0.63	2
Hasbro Inc.	0.16	0.14	0.34	0.24	0.42	0.34	0.24	0.24	0.34	1
Huntington Bancshares	0.12	0.01	0.00	0.12	0.02	0.00	0.12	0.02	0.00	1
Health Care REIT	0.98	0.11	0.00	0.98	0.16	0.00	0.98	0.16	0.00	1
Home Depot	0.35	0.07	0.00	0.35	0.11	0.00	0.35	0.11	0.00	1
Hess Corporation	0.47	0.28	1.00	0.71	0.85	1.00	0.71	0.71	1.00	1
Harley-Davidson	0.07	0.00	0.74	0.11	0.00	0.74	0.11	0.00	0.74	1
Honeywell Int'l Inc.	0.32	0.13	0.68	0.48	0.40	0.68	0.48	0.40	0.68	1
Hewlett-Packard	0.16	0.01	0.00	0.16	0.01	0.00	0.16	0.01	0.00	1
Block H and R	0.13	0.02	0.00	0.13	0.02	0.00	0.13	0.02	0.00	1
Hormel Foods Corp.	0.46	0.09	0.02	0.46	0.14	0.06	0.46	0.14	0.06	1
The Hershey Company	1.00	0.20	0.61	1.00	0.61	0.91	1.00	0.61	0.91	1
Intel Corp.	0.92	0.25	0.00	0.92	0.38	0.00	0.92	0.38	0.00	1
International Paper	0.17	0.26	0.00	0.26	0.26	0.00	0.26	0.26	0.00	1
Interpublic Group	0.56	0.00	0.00	0.56	0.00	0.00	0.56	0.00	0.00	1
Ingersoll-Rand PLC	0.47	0.05	0.00	0.47	0.07	0.00	0.47	0.07	0.00	1
Johnson Controls	0.03	0.00	0.00	0.03	0.00	0.00	0.03	0.00	0.00	1
Jacobs Eng. Group	0.13	0.01	0.20	0.19	0.02	0.20	0.19	0.02	0.20	1
Johnson and Johnson	0.13	0.00	0.00	0.13	0.00	0.00	0.13	0.00	0.00	1

Using unadjusted test *P*-values, evidence against linearity is found in 82 stock returns (at 5% significance level) on the basis of the Q_{22} test. This arguably is not a very surprising finding since conditional heteroskedasticity is a characteristic feature of many asset returns. Linearity is also rejected by at least one of the cross-correlation Q_{12}/Q_{21} tests in 75 cases. Using multiplicity-adjusted *P*-values, evidence against linearity is found by at least one of the three tests in 85% of stock returns (at 5% significance level). We conclude, therefore, that the vast majority of the stock returns considered in our analysis exhibit nonlinear features which cannot be captured by a linear model with i.i.d. noise.

5. Summary

This article considered portmanteau tests for linearity of stationary time series based on generalized correlations of residuals. The proposed tests are easy to implement, have a chi-square large-sample null distribution, and good size and power properties in finite samples. The simulation results indicated that the cross-correlation tests Q_{12} and Q_{21} are useful in identifying various types of nonlinearity and are generally more powerful than the popular Q_{22} test based on squared-residual autocorrelations. An application to time series of stock returns illustrated the practical use of the tests.

Acknowledgments

The authors are grateful to two anonymous referees for many helpful comments and suggestions in improving an earlier version of the paper.

References

Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. New York: Wiley.
Barnett, W. A., Gallant, A. R., Hinich, M. J., Jungeilges, J. A., Kaplan, D. T., Jensen, M. J. (1997). A single-blind controlled competition among tests for nonlinearity and chaos. *Journal of Econometrics* 82:157–192.
Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B* 57:289–300.
Berg, A., Paparoditis, E., Politis, D. N. (2010). A bootstrap test for time series linearity. *Journal of Statistical Planning and Inference* 140:3841–3857.
Berk, K. N. (1974). Consistent autoregressive spectral estimates. *Annals of Statistics* 2:489–502.
Bhansali, R. J. (1978). Linear prediction by autoregressive model fitting in the time domain. *Annals of Statistics* 6:224–231.
Bickel, P. J., Bühlmann, P. (1996). What is a linear process? *Proceedings of the National Academy of Sciences of the USA* 93:12128–12131.
de Lima, P. J. F. (1997). On the robustness of nonlinearity tests to moment condition failure. *Journal of Econometrics* 76:251–280.
Dunsmuir, W. (1979). A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise. *Annals of Statistics* 7:490–506.
Giannerini, S., Maasoumi, E., Bee Dagum, E. (2015). Entropy testing for nonlinear serial dependence in time series. *Biometrika* 102:661–675.
Hannan, E. J. (1973). The asymptotic theory of linear time-series models. *Journal of Applied Probability* 10:130–145.
Hong, Y., Lee, T.-H. (2003). Diagnostic checking for the adequacy of nonlinear time series models. *Econometric Theory* 19:1065–1121.
Hong, Y., White, H. (2005). Asymptotic distribution theory for nonparametric entropy measures of serial dependence. *Econometrica* 73:837–901.
Hosoya, Y., Taniguchi, M. (1982). A central limit theorem for stationary processes and the parameter estimation of linear processes. *Annals of Statistics* 10:132–153.
Jansen, D. W., de Vries, C. G. (1991). On the frequency of large stock returns: Putting booms and busts into perspective. *Review of Economics and Statistics* 73:18–24.
Kapetanios, G. (2009). Testing for strict stationarity in financial variables. *Journal of Banking and Finance* 33:2346–2362.
Koedijk, K., Schafgans, M., de Vries, C. (1990). The tail index of exchange rate returns. *Journal of International Economics* 29:93–108.
Koreisha, S. G., Fang, Y. (1999). The impact of measurement errors on ARMA prediction. *Journal of Forecasting* 18:95–109.
Kreiss, J.-P. (1991). Estimation of the distribution function of noise in stationary processes. *Metrika* 38:285–297.
Kuersteiner, G. M. (2001). Optimal instrumental variables estimation for ARMA models. *Journal of Econometrics* 104:359–405.
Lawrance, A. J., Lewis, P. A. W. (1985). Modelling and residual analysis of nonlinear autoregressive time series in exponential variables. *Journal of the Royal Statistical Society B* 47:165–202.
Lawrance, A. J., Lewis, P. A. W. (1987). Higher-order residual analysis for nonlinear time series with autoregressive correlation structures. *International Statistical Review* 55:21–35.
Lee, T.-H., White, H., Granger, C. W. J. (1993). Testing for neglected nonlinearity in time series models: A comparison of neural network methods and alternative tests. *Journal of Econometrics* 56:269–290.
Lewis, R., Reinsel, G. C. (1985). Prediction of multivariate time series by autoregressive model fitting. *Journal of the Multivariate Analysis* 16:393–411.
Ljung, G. M., Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika* 65:297–303.
Luukkonen, R., Saikkonen, P., Teräsvirta, T. (1988). Testing linearity in univariate time series models. *Scandinavian Journal of Statistics* 15:161–175.
McLeod, A. I., Li, W. K. (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis* 4:269–273.
Ng, S., Perron, P. (2005). A note on the selection of time series models. *Oxford Bulletin of Economics and Statistics* 67:115–134.

Proietti, T., Luati, A. (2015). The generalised autocovariance function. *Journal of Econometrics* 186:245–257.
Psaradakis, Z. (2000). *P*-value adjustments for multiple tests for nonlinearity. *Studies in Nonlinear Dynamics & Econometrics* 4:95–100.
Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika* 73:751–754.
Teräsvirta, T., Tjøstheim, D., Granger, C. W. J. (2010). *Modelling Nonlinear Economic Time Series*. Oxford: Oxford University Press.
Terdik, G., Máth, J. (1998). A new test of linearity of time series based on the bispectrum. *Journal of Time Series Analysis* 19:737–753.
Tong, H. (1990). *Non-linear Time Series: A Dynamical System Approach*. Oxford: Oxford University Press.
Tsay, R. S. (2010). *Analysis of Financial Time Series*. 3rd ed. New York: Wiley.