Econometric Theory, **38**, 2022, 913–941.
doi:10.1017/S0266466620000341

ROBUST TESTS FOR WHITE NOISE AND CROSS-CORRELATION

VIOLETTA DALLA
National and Kapodistrian University of Athens

LIUDAS GIRAITIS
Queen Mary University of London

PETER C. B. PHILLIPS
Yale University,
University of Auckland,
University of Southampton,
Singapore Management University

Commonly used tests to assess evidence for the absence of autocorrelation in a univariate time series or serial cross-correlation between time series rely on procedures whose validity holds for i.i.d. data. When the series are not i.i.d., the size of correlogram and cumulative Ljung–Box tests can be significantly distorted. This paper adapts standard correlogram and portmanteau tests to accommodate hidden dependence and nonstationarities involving heteroskedasticity, thereby uncoupling these tests from limiting assumptions that reduce their applicability in empirical work. To enhance the Ljung–Box test for non-i.i.d. data, a new cumulative test is introduced. Asymptotic size of these tests is unaffected by hidden dependence and heteroskedasticity in the series. Related extensions are provided for testing cross-correlation at various lags in bivariate time series. Tests for the i.i.d. property of a time series are also developed. An extensive Monte Carlo study confirms good performance in both size and power for the new tests. Applications to real data reveal that standard tests frequently produce spurious evidence of serial correlation.

1. INTRODUCTION

Temporal dependence is one of the primary characteristics of economic and financial data that are measured sequentially over time. In studying such data, estimation of and inference on the serial correlation $\rho_k = \mathrm{corr}(x_t, x_{t-k})$ is a common first step in the analysis of time series data $\{x_t\}$ or regression residuals. For a sample $x_1, \ldots, x_n$, estimation of ρ_k by the sample serial correlation $\widehat{\rho}_k$ for various lags $k = 1, 2, \ldots$ and testing whether it is significant dates back to the early years of the twentieth century, primarily to Yule (1926) who introduced the terminology serial

Dalla acknowledges financial support from ELKE-EKPA. Phillips acknowledges support from the Kelly Fund at the University of Auckland, a KLC Fellowship at Singapore Management University, and the NSF under Grant No. SES 18-50860. Address correspondence to Liudas Giraitis, Queen Mary University of London, London, UK; e-mail: l.giraitis@qmul.ac.uk.

© The Author(s), 2020. Published by Cambridge University Press.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

correlation. Yule highlighted the need to understand the degree of time persistence in the data prior to applying correlation/regression analysis and characterized this phenomenon as the "time correlation problem" in his earlier Royal Society address (Yule, 1921). To aid analysis, Yule introduced the sample serial correlation $\widehat{\rho}_k$ along with the standard confidence band $\pm z_{\alpha/2}/\sqrt{n}$ for testing its significance, $H_0 : \rho_k = 0$ under the simplifying assumption that the data are identically and independently distributed (i.i.d.), bringing the problem into the existing framework of the Pearson correlation coefficient.

Bartlett (1946) provided a major step forward in a more general analysis by deriving an asymptotic formula, now known as Bartlett's formula, for $\text{cov}(\widehat{\rho}_j, \widehat{\rho}_k)$ for a stationary linear process $\{x_t\}$ driven by i.i.d. errors. The joint asymptotic distribution of $\widehat{\rho} = (\widehat{\rho}_1, \ldots, \widehat{\rho}_m)'$ was given by Anderson and Walker (1964) and was found to be normal with variance–covariance matrix $n^{-1}W$ where the elements of W are given by Bartlett's formula. An important aspect of this formula is that the asymptotic variance matrix depends only on the autocorrelations ρ_k themselves and not fourth moments, as is the case for sample autocovariances.[1] Hannan and Hyde (1972) relaxed the i.i.d. assumption on the errors and showed that asymptotic normality remains valid under some additional regularity assumptions on the noise.

Besides testing for significant serial correlation at one lag k, it is common to test the cumulative hypothesis $H_0 : \rho_1 = \cdots = \rho_m = 0$ using the portmanteau statistics of Box and Pierce (1970) and Ljung and Box (1978). The Box–Pierce statistic is based on the observation that the matrix W in the asymptotic distribution of $\widehat{\rho}$ reduces to the identity matrix under the i.i.d. assumption[2] on x_t, while the Ljung–Box statistic entails a slightly better performance in finite samples. They find that for i.i.d. and normally distributed data, the cumulative statistics have a χ^2_m limit distribution when applied to raw data and a χ^2_{m-p-q} limit distribution when used for residuals of fitted ARMA(p, q) models. They indicated that the normality assumption is not essential for these results.

Concern that these standard tests of $H_0 : \rho_k = 0$ and $H_0 : \rho_1 = \cdots = \rho_m = 0$ are not suitable under heteroskedasticity or nonindependence of uncorrelated noise x_t was highlighted by Granger and Andersen (1978) and by Taylor (1984). The first paper warned against the use of standard tests in bilinear data and the second raised concerns for testing in models where the $\{x_t\}$ are heteroskedastic. Taylor (1984) provided a modified standard error for $\widehat{\rho}_k$, resulting in a robust confidence band and a robust t-statistic $\widetilde{t}_k$, given in (4) below, as well as a robust cumulative statistic, given in (12).

Since then, various authors have modified the statistic $t_k = \sqrt{n}\widehat{\rho}_k$ and/or its cumulative portmanteau versions in similar ways to Taylor (1984) so that they are applicable for testing uncorrelated non-i.i.d. noise in which the covariance matrix W is diagonal but not the identity–among others, see Diebold (1986), Lo and

[1] The asymptotic variance of $\widehat{\rho}_h$, for instance, is $n^{-1}\sum_{j=1}^{\infty}(\rho_{h+j} + \rho_{h-j} - 2\rho_h\rho_j)^2$, a simple derivation of which is given in Phillips and Solo (1992).

[2] In this case, $\text{var}(\widehat{\rho}_h) = n^{-1}\sum_{j=1}^{\infty}(\rho_{h+j} + \rho_{h-j} - 2\rho_h\rho_j)^2 = n^{-1}$ as the only nonzero element in the sum occurs when $j = h$.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

MacKinlay (1989), Robinson (1991), Francq and Zakoïan (2009), and Kokoszka and Politis (2011).

However, the matrix W is not always diagonal. Practical settings involving economic and financial uncorrelated data x_t for which a nondiagonal W is relevant appear in Cumby and Huizinga (1992), Guo and Phillips (2001), Lobato, Nankervis, and Savin (2002), and Francq, Roy, and Zakoïan (2005). These papers typically assume that $\{x_t\}$ is stationary and has a martingale difference structure or is strongly mixing. Guo and Phillips (2001) estimated the covariance matrix W by its empirical counterpart, while the other papers used nonparametric procedures. Taking a different approach Romano and Thombs (1996) and Horowitz, Lobato, and Savin (2006) used bootstrap methods to obtain suitable critical values for standard test procedures under the null assumption that the uncorrelated data $\{x_t\}$ are stationary and strongly mixing.

Similar issues arise in testing cross-correlation in bivariate time series $\{x_t, y_t\}$, where interest and early attempts date back over a century, for instance to Hooker (1901). Unsurprisingly, the original work on estimation and testing for zero cross-correlation $\rho_{xy,k} = \text{corr}(x_t, y_{t-k})$ relied again on sample cross-correlations based on the theory of the Pearson correlation coefficient. Later studies such as Haugh (1976) and Haugh and Box (1977) examined testing for zero cross-correlation at an individual lag and using cumulative versions of the statistics. For the same reason as in the univariate case, tests for the absence of cross-correlation may be invalidated when the time series are not i.i.d. Corrected versions of the statistics have been examined (e.g., Kyriazidou, 1998, that extend to the bivariate case the univariate results of Cumby and Huizinga, 1992). However, the assumptions used are restrictive, imposing additional technical conditions and excluding unconditional heteroskedasticity.

Test statistics based on the correlogram either with the standard confidence band $\pm z_{\alpha/2}/\sqrt{n}$ suggested by Yule (1926) or that based on Bartlett's (1946) formula or the cumulative statistics of Box and Pierce (1970) and Ljung and Box (1978) are all still in extensive use today and are present in most statistical packages. Despite the literature addressing the complications of departures from i.i.d. noise, problems with finite sample performance and complexity of implementation seem to have prevented replacement of these methods with procedures that extend applicability to more general settings. Similarly, cross-correlogram tests are usually implemented with standard confidence bands and the cumulative Haugh–Box test for cross-correlation is rarely reported in applications.

The goal of the present paper is to develop a formulation and method of implementation that will enable testing with both univariate and bivariate time series that is robust to multiple forms of heteroskedastic and dependence departures from i.i.d. noise. Our approach is based on extending the original robust test by Taylor (1984) for the absence of correlation at an individual lag and a corresponding cumulative portmanteau test, together with analogous testing procedures for the bivariate case. Taylor's test for correlation at a specific lag and our cumulative test are both easy to apply and demonstrate good size control for a large class of

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

uncorrelated data covering martingale difference noise of unspecified form with time varying unconditional variance.

The rest of the paper is organized as follows. Section 2 outlines the model, introduces the tests and presents limit theory for the case of univariate time series testing. Section 3 develops corresponding tests for zero cross-correlation in the bivariate case, a problem that has attracted much less attention in the literature in spite of its links to Granger causality testing. Section 4 considers direct testing of the hypothesis that a time series is i.i.d. Various tests of this hypothesis have been used in the literature and often relate to testing the random walk hypothesis using variate differences. Standard tests based on the squared time series have been proposed by McLeod and Li (1983) and Campbell, Lo and MacKinlay (1997, Chapter 2) provide a brief summary. We suggest testing procedures (both cumulative and individual lag tests) that combine the tests on correlation of the data in levels with absolutes or squared values. Wong and Ling (2005) and Zhu (2013) used similar methods to test for absence of autocorrelation in residuals and absolute (squared) residuals in a general linear model, in particular, of a fitted ARMA-GARCH model. Different from our tests, those tests do not require demeaning of residuals and their asymptotic distributions depend on the parameters of a fitted model. An extensive Monte Carlo study was conducted and the results are presented in Section 5 with recommended guidelines for practical implementation of the various tests. Section 6 reports an empirical application of the test procedures to financial data. The paper is accompanied by an Online Supplement consisting of two documents. The first (I) contains proofs of all the results in the paper. The second (II) reports the findings of the full Monte Carlo study.

An R package and an EViews add-in (named *testcorr*) have been created by the authors and are now available to implement all the testing procedures developed in the paper. The R package is available on CRAN, https://cran.r-project.org/package=testcorr and the EViews add-in at https://www.eviews.com/Addins/addins.shtml.

2. TESTS FOR ZERO CORRELATION

While stationarity is commonly assumed, it is not necessary for testing absence of correlation in a time series. Indeed, for a series $\{x_t\}$ of uncorrelated random variables the condition that the autocorrelation function (ACF) $\rho_k = \text{corr}(x_t, x_{t-k}) = 0$ at lag $k = 1, 2, \ldots$. is well defined for all t with or without an assumption of stationarity. Its empirical version, the sample autocorrelation $\widehat{\rho}_k$ at the lag $k \neq 0$, based on observed data $x_1, \ldots, x_n$,

$$\widehat{\rho}_k = \frac{\sum_{t=k+1}^{n}(x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}, \quad \bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t \tag{1}$$

remains a valid tool for testing for zero correlation at lag k. Such testing does not require assumptions of independence or stationarity of $\{x_t\}$, thereby enabling a more general approach to testing for white noise uncorrelatedness in data.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

There is, of course, a large literature on estimation of the ACF ρ_k by $\widehat{\rho}_k$ for stationary times series $\{x_t\}$. The asymptotic distribution of the sample autocorrelations $(\widehat{\rho}_1, \ldots, \widehat{\rho}_m)'$ for a stationary linear process was given by Anderson and Walker (1964) and Hannan and Hyde (1972) and has the form

$$\sqrt{n}(\widehat{\rho}_1 - \rho_1, \ldots, \widehat{\rho}_m - \rho_m) \to_D \mathcal{N}(0, W), \tag{2}$$

where W is a matrix whose elements are given by Bartlett's formula. If the $\{x_t\}$ are i.i.d. random variables, the matrix W reverts to the identity matrix I_m and (2) reduces to the standard asymptotic result

$$\sqrt{n}(\widehat{\rho}_1, \ldots, \widehat{\rho}_m) \to_D \mathcal{N}(0, I_m) \tag{3}$$

used for testing $H_0 : \rho_k = 0$ at lag k, just as in Yule (1926), with the confidence band $\pm z_{\alpha/2}/\sqrt{n}$ for zero correlation at significance level α. Methods based on this procedure are still heavily used and came into prominence following the influential work of Box and Jenkins (1970).

When $\{x_t\}$ is uncorrelated but not i.i.d., the standard method for testing zero serial correlation based on the asymptotic distribution in (3) generally fails. This was first noted by Granger and Andersen (1978) and Taylor (1984). Taylor (1984) suggested correcting the standard error of $\widehat{\rho}_k$, leading to the robust t-statistic

$$\widetilde{t}_k = \frac{\sum_{t=k+1}^{n} e_{tk}}{(\sum_{t=k+1}^{n} e_{tk}^2)^{1/2}}, \quad e_{tk} = (x_t - \bar{x})(x_{t-k} - \bar{x}), \tag{4}$$

so that in testing H_0 the sample autocorrelation $\widehat{\rho}_k$ should be corrected for its variance

$$\widetilde{t}_k = \widehat{\rho}_k \widehat{c}_k \to_D \mathcal{N}(0, 1), \quad \widehat{c}_k = \frac{\widetilde{t}_k}{\widehat{\rho}_k}. \tag{5}$$

This correction leads to a $\pm z_{\alpha/2}/\widehat{c}_k$ confidence band for zero correlation at lag k.

Taylor showed that $\widehat{c}_k^{-2}$ is an unbiased estimate of the variance of $\widehat{\rho}_k$ when x_t has symmetric density but he did not prove the asymptotic normality given in (5). He also suggested a corrected cumulative test, as given in (12) below.

The t-statistic $\widetilde{t}_k$ takes the form of a self-normalizing sum. Our aim is to establish asymptotic normality for $\widetilde{t}_k$ as well as corresponding cumulative tests in a general setting where the observed data sample $x_1, \ldots, x_n$ is a series of uncorrelated random variables that may be dependent and nonstationary. We seek an approach that does not require verification of additional technical assumptions and allows the applied researcher to be somewhat agnostic about the structure and generating mechanism of the uncorrelated data x_t.

In this paper, we assume that x_t satisfies

$$x_t = \mu + h_t \varepsilon_t, \tag{6}$$

where $\{\varepsilon_t\}$ is a stationary ergodic martingale difference (m.d.) sequence with respect to some σ-field filtration $\mathcal{F}_t$ that includes the natural filtration with $\mathbb{E}[\varepsilon_t | \mathcal{F}_{t-1}] = 0$, $\mathbb{E}\varepsilon_t^4 < \infty$, $\mu = \mathbb{E}x_t$ and a deterministic scaling factor $h_t \equiv h_{tn}$ is a sequence of real numbers for which, as $n \to \infty$,

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

$$\max_{1\le t\le n} h_t^4 = o\left(\sum_{t=1}^n h_t^4\right), \qquad \sum_{t=2}^n (h_t - h_{t-1})^4 = o\left(\sum_{t=1}^n h_t^4\right). \tag{7}$$

Then, ε_t allows for conditional heteroskedasticity via $\mathbb{E}[\varepsilon_t^2|\mathcal{F}_{t-1}]$ and h_t introduces unconditional heterogeneity over t. For any $k \neq 0$, we have

$$\operatorname{corr}(x_t, x_{t-k}) = \operatorname{corr}(\varepsilon_t, \varepsilon_{t-k}) = 0 \quad \text{for all } t.$$

Since m.d. variables ε_t are uncorrelated, the variables x_t are also uncorrelated. In simulations involving h_t, we use the examples $h_t = t/n$ and $h_t = 1 + aI(\tau_1 < t/n \le \tau_2)$, where $0 \le \tau_1 < \tau_2 \le 1$, $a \neq 0$ and I is the indicator function. In general, it is easy to see that the (weak trend) scaling factor $h_t = v(t/n)$ satisfies (7) if $v(x)$, $0 \le x \le 1$, is a piecewise bounded differentiable function with a bounded derivative. Our main example for an ergodic m.d. sequence $\{\varepsilon_t\}$ is a stationary GARCH noise. More generally, if $\{\eta_t\}$ is an i.i.d. noise and f is a measurable function then $\varepsilon_t = \eta_t f(\eta_{t-1}, \eta_{t-2}, \dots)$ is a stationary ergodic m.d. sequence (e.g. Stout (1974, Thm. 3.5.8)), an example being $\varepsilon_t = \eta_t \eta_{t-1}$.

The statistics $t_k = \sqrt{n}\,\widehat{\rho}_k, \widetilde{t}_k$ and their cumulative versions have been examined by various authors either for raw data or for residuals from some estimated model. In either case, it is common to assume that $\{x_t\}$ is as in (6) but stronger assumptions on $\{h_t, \varepsilon_t\}$ are imposed. Most authors assume that $h_t = 1$ and ε_t is an m.d. noise of a specific type. A few authors allow h_t to be deterministic and to vary with t, but require $\{\varepsilon_t\}$ to be i.i.d. so that the x_t remain independent. An exception is Lo and MacKinlay (1989) where h_t is permitted to be time varying and mixing conditions are imposed on $\{\varepsilon_t\}$ but with restrictive moment conditions that exclude, for example, GARCH data with skewed innovations.

The following result establishes the asymptotic distribution of the t-statistic $\widetilde{t}_k$ given in (4).

THEOREM 2.1. *Suppose that $x_1, \dots, x_n$ is a sample from a sequence given in (6) and the h_t satisfy (7). Then for any fixed integer $k \neq 0$, as $n \to \infty$,*

$$\widetilde{t}_k \to_D \mathcal{N}(0,1). \tag{8}$$

Moreover, for $m = 1, 2, \dots$,

$$(\widetilde{t}_1, \dots, \widetilde{t}_m) \to_D \mathcal{N}(0, R), \tag{9}$$

where $R = (r_{jk})$ is an $m \times m$ matrix with elements

$$r_{jk} = \operatorname{corr}(\varepsilon_1 \varepsilon_{1-j}, \varepsilon_1 \varepsilon_{1-k}), \quad j,k = 1, \dots, m.$$

By virtue of (8), $\widehat{c}_k \widehat{\rho}_k \to_D \mathcal{N}(0,1)$ where

$$\widehat{c}_k = c_k(1 + o_p(1)), \quad c_k := \frac{\sum_{t=1}^n h_t^2}{(\sum_{t=1}^n h_t^4)^{1/2}} \frac{\mathbb{E}\varepsilon_1^2}{(\mathbb{E}\varepsilon_1^2 \varepsilon_{1-k}^2)^{1/2}} \to \infty, \tag{10}$$

which leads to a $\pm z_{\alpha/2}/\widehat{c}_k$ confidence band for zero correlation at lag k.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

Box and Pierce (1970) and Ljung and Box (1978) suggested the well-known cumulative statistics

$$BP_m = n\sum_{k=1}^{m}\widehat{\rho}_k^2, \qquad LB_m = (n+2)n\sum_{k=1}^{m}\frac{\widehat{\rho}_k^2}{n-k} \tag{11}$$

for portmanteau testing of the joint null hypothesis $H_0 : \rho_1 = \cdots = \rho_m = 0$. These tests are based on the property (3) of the sample ACF $\widehat{\rho}_k$'s showing that under H_0 the tests are asymptotically χ^2_m distributed.

When $\{x_t\}$ is uncorrelated but not i.i.d. and (3) fails these cumulative tests produce size distortions in testing. In turn, Taylor (1984) suggested the corrected-for-variance cumulative statistic

$$\sum_{k=1}^{m}\widetilde{t}_k^2 \tag{12}$$

for testing $H_0 : \rho_1 = \cdots = \rho_m = 0$. This formulation corresponds to the diagonal matrix $R = I$ in (9) which holds only when the variables $\omega_j = \varepsilon_1\varepsilon_{1-j}$ are uncorrelated. Setting $\widetilde{t} = (\widetilde{t}_1, \ldots, \widetilde{t}_m)'$, result (9) of Theorem 2.1 implies that

$$\widetilde{t}' R^{-1}\widetilde{t} \to_D \chi^2_m, \qquad \widetilde{t}'\widehat{R}^{-1}\widetilde{t} \to_D \chi^2_m, \tag{13}$$

for any consistent estimate $\widehat{R} \to_p R$ of R. The matrix R is positive definite and R^{-1} exists if $\varepsilon_1 \neq 0$ *a.s.*, see Lemma 3.1 below.

As discussed in the Introduction, various authors have examined statistics (4) and (12). However, as noted by Guo and Phillips (2001) (see also Cumby and Huizinga, (1992), Lobato *et al.*, (2002), and Francq *et al.*, (2005)) who arrived at (13) under different assumptions on the data generating process $\{x_t\}$, there are sequences $\{x_t\}$ that are uncorrelated but not independent for which the matrix R is not diagonal and therefore the cumulative statistic (12) is invalid. Guo and Phillips (2001) use a similar estimate $\widehat{R}$ of R to our estimate given in (14) below. Cumby and Huizinga (1992) and Lobato *et al.* (2002) use kernel nonparametric methods and Francq *et al.* (2005) introduce an autoregressive approximation method to estimate W and R. These authors all assume stationarity of $\{x_t\}$, thereby excluding unconditional heterogeneity.

We will estimate R by $\widehat{R} = (\widehat{r}_{jk})$ where $\widehat{r}_{jk}$ are sample cross-correlations of the variables $(e_{tj}, t = 1, \ldots, n)$ and $(e_{tk}, t = 1, \ldots, n)$:

$$\widehat{r}_{jk} = \frac{\sum_{t=\max(j,k)+1}^{n} e_{tj}e_{tk}}{\left(\sum_{t=\max(j,k)+1}^{n} e_{tj}^2\right)^{1/2}\left(\sum_{t=\max(j,k)+1}^{n} e_{tk}^2\right)^{1/2}}. \tag{14}$$

To improve the finite sample performance of the cumulative test, we use the thresholded version $\widehat{R}^* = (\widehat{r}^*_{jk})$ of $\widehat{R}$ where

$$\widehat{r}^*_{jk} = \widehat{r}_{jk} I(|\tau_{jk}| > \lambda) \quad \text{with } \lambda = 2.576, \tag{15}$$

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

where τ_{jk} is a self-normalized t-type statistic

$$\tau_{jk} = \frac{\sum_{t=\max(j,k)+1}^{n} e_{tj} e_{tk}}{\left(\sum_{t=\max(j,k)+1}^{n} e_{tj}^2 e_{tk}^2\right)^{1/2}}. \tag{16}$$

Notice that $\widehat{R}^*$ is the sample analogue of the variance–covariance matrix of $\widetilde{t}$ for which we threshold the off-diagonal elements by checking at the 1% significance level whether they are significant. This is a simpler approach to that undertaken by Cumby and Huizinga (1992), Lobato *et al*. (2002), or Francq *et al*. (2005).

The next theorem establishes the asymptotic properties for the cumulative tests

$$Q_m = \widetilde{t}' \widehat{R}^{-1} \widetilde{t}, \qquad \widetilde{Q}_m = \widetilde{t}' \widehat{R}^{*-1} \widetilde{t} \tag{17}$$

for the joint hypothesis $H_0 : \rho_1 = \cdots = \rho_m = 0$. In the empirical applications and Monte Carlo study described later in the paper we use the cumulative statistic $\widetilde{Q}_m$ with the suggested threshold setting $\lambda = 2.576$.

THEOREM 2.2. *Under the assumptions of Theorem 2.1, for any* $m = 1, 2, \ldots$ *and with any threshold* $\lambda > 0$, *as* $n \to \infty$,

$$\widehat{R} \to_p R, \qquad \widehat{R}^* \to_p R, \tag{18}$$

$$Q_m \to_D \chi_m^2, \qquad \widetilde{Q}_m \to_D \chi_m^2. \tag{19}$$

The assumptions of Theorems 2.1 and 2.2 are minimal and less restrictive than those used so far in the literature. They allow for both nonstationarity (unconditional heteroskedasticity) and dependent martingale difference type uncorrelated noise including GARCH type conditional heteroskedasticity.

The test Q_m without thresholding based on property $\widehat{R} \to_p R$ suffers some size distortion in finite samples for moderate and large m and moderate sample sizes n. We will show that thresholding has no impact on consistency asymptotically, i.e. for any $\lambda > 0$, $\widehat{R}^* \to_p R$, but it will improve finite sample performance in many cases where the off-diagonal elements are either very small or zero, in particular when R is sparse (see Online Supplement II). The following simple arguments show consistency $\widehat{r}_{jk}^* = \widehat{r}_{jk} I(\tau_{jk} \geq \lambda) \to_p r_{jk}$. Indeed, $\widehat{r}_{jk} \to_p r_{jk}$, and $I(\tau_{jk} \geq \lambda) \to_p 1$ if $r_{jk} \neq 0$. The objective of thresholding is to improve estimation of r_{jk} when taking zero values. Here λ plays the role of a critical level in testing the null hypothesis that r_{jk} (or its numerator) is zero. We recommend using it as a preselected tuning parameter λ which does not depend on n. If $\tau_{jk} \to_D \mathcal{N}(0,1)$, thresholding with $\lambda = 2.576$ can be interpreted as testing for $r_{jk} = 0$ at a 1% significance level. Use of thresholding has negligible impact on power, which is shown in the Online Supplement II.

As will be apparent in the next section, the assumptions of Theorems 3.1 and 3.2 in the bivariate case render the above methods of analysis valid for a univariate series in a straightforward way. In particular, since any measurable function

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

$y_t = f(x_t, x_{t-1}, \ldots, x_{t-k})$ of a stationary ergodic process $\{x_t\}$ is also a stationary ergodic process (e.g., Stout (1974, Cor. 3.5.2)), if $\mathbb{E}|y_t| < \infty$ it follows that

$$\mathbb{E}\left|n^{-1}\sum_{t=1}^{n} y_t - \mathbb{E}y_1\right| \to 0, \quad n \to \infty.$$

Thus the bivariate series $\{x_t, y_t\}$ with $y_t = x_t$ satisfies the conditions of Theorems 3.1 and 3.2 in the following section, which imply the results of Theorems 2.1 and 2.2, and Corollary 3.1 implies (10) above.

3. TESTS FOR ZERO CROSS-CORRELATION

For bivariate time series $\{x_t, y_t\}$ we observe data $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ and are interested in testing possible cross-correlation between these time series at various lags. Denote by $\widehat{\rho}_{xy,k}$ the kth sample cross-correlation estimate of the kth cross-correlation $\rho_{xy,k} = \mathrm{corr}(x_t, y_{t-k})$ for $k = 0, 1, \ldots$

$$\widehat{\rho}_{xy,k} = \frac{\sum_{t=k+1}^{n}(x_t - \bar{x})(y_{t-k} - \bar{y})}{\left(\sum_{t=1}^{n}(x_t - \bar{x})^2 \sum_{t=1}^{n}(y_t - \bar{y})^2\right)^{1/2}}, \quad \bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t,\ \bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t.$$

The asymptotic theory for such sample cross-correlations was given in Hannan (1970). Haugh and Box (1977) developed a test for cross-correlation under the assumption of independent series $\{x_t\}$ and $\{y_t\}$. But there is little literature on testing cross-correlation using the sample statistic $\widehat{\rho}_{xy,k}$ when the series are not independent or when they are heteroskedastic. Although in regression analysis the issue of heteroskedasticity has been addressed, graphs of the sample cross-correlations typically display confidence bands based on $\pm z_{\alpha/2}/\sqrt{n}$, corresponding to the t-statistic

$$t_{xy,k} = \sqrt{n}\widehat{\rho}_{xy,k} \tag{20}$$

for testing $H_0 : \rho_{xy,k} = 0$ under independence conditions. Further, bivariate analogues of cumulative standard statistics are rarely analyzed and often involve additional tests for the significance of the univariate autocorrelations of $\{x_t\}$ and $\{y_t\}$, as in Tsay (2010) for example. In what follows, we examine the Haugh (1976) and Haugh and Box (1977) test for cross-correlation

$$HB_{xy,m} = n^2 \sum_{k=0}^{m} \frac{\widehat{\rho}_{xy,k}^2}{n-k} \tag{21}$$

for testing $H_0 : \rho_{xy,0} = \rho_{xy,1} = \cdots = \rho_{xy,m} = 0$ which assumes independence of the time series $\{x_t\}$ and $\{y_t\}$.

Similar arguments to those of the univariate case in Section 2 show that standard normal and χ_m^2 approximations for the statistics in (20) and (21) are not always valid for bivariate times series $\{x_t, y_t\}$ which are not uncorrelated and stationary, independent noises. To provide a more general framework, we assume in this paper that $\{x_t, y_t\}$ satisfy

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

$$x_t = \mu_x + h_t\varepsilon_t, \quad y_t = \mu_y + g_t\eta_t, \tag{22}$$

where $\{\varepsilon_t\}, \{\eta_t\}$ are stationary sequences, $\mathbb{E}\varepsilon_t = \mathbb{E}\eta_t = 0$, $\mathbb{E}\varepsilon_t^4 < \infty$, $\mathbb{E}\eta_t^4 < \infty$ and h_t, g_t are real numbers that satisfy conditions made explicit in Assumption A and Theorem 3.1 below.

Notice that $\rho_{xy,k} = \operatorname{corr}(x_t, y_{t-k}) = \operatorname{corr}(\varepsilon_t, \eta_{t-k})$. However, when $\rho_{xy,k} = 0$, the standard normal approximation for $t_{xy,k}$ may not hold because of the presence of heteroskedasticity in $\{x_t\}$ and $\{y_t\}$ or dependence between the two uncorrelated series. For that reason, in order to develop a robust test of $H_0 : \rho_{xy,k} = 0$, $k \geq 0$, we define the following robust t-statistic

$$\widetilde{t}_{xy,k} = \frac{\sum_{t=k+1}^{n} e_{xy,tk}}{\left(\sum_{t=k+1}^{n} e_{xy,tk}^2\right)^{1/2}}, \quad e_{xy,tk} = (x_t - \bar{x})(y_{t-k} - \bar{y}). \tag{23}$$

Our objective is to establish the asymptotic normality

$$\widetilde{t}_{xy,k} = \widehat{\rho}_{xy,k}\widehat{c}_{xy,k} \to_D \mathcal{N}(0,1), \quad \widehat{c}_{xy,k} = \frac{\widetilde{t}_{xy,k}}{\widehat{\rho}_{xy,k}} \tag{24}$$

for $k \geq 0$ such that $\operatorname{corr}(\varepsilon_t, \eta_{t-k}) = 0$. This correction leads to a $\pm z_{\alpha/2}/\widehat{c}_{xy,k}$ confidence band for zero cross-correlation $\rho_{xy,k}$ at lag k.

As we see below, (24) requires the first series $\{x_t\}$ to be an uncorrelated noise or, more specifically, an m.d. sequence. In what follows, we first allow the uncorrelated series $\{x_t\}$ and $\{y_t\}$ to be mutually dependent, and subsequently examine the case where they are mutually independent. Test consistency is considered last.

Clearly, for (24) to hold, we need additional assumptions on $\{\varepsilon_t, \eta_t\}$. The conditions below are formulated in terms of the product series $\omega_{tk} := \varepsilon_t\eta_{t-k}$, $k \geq 0$.

Assumption A. *For $j, k \geq 0$, $\{\omega_{tk}\}$, $\{\omega_{tj}\omega_{tk}\}$ are stationary sequences, $\mathbb{E}\omega_{tk}^2 < \infty$, and*

$$\mathbb{E}\left|\left(n^{-1}\sum_{t=1}^{n}\omega_{tj}\omega_{tk}\right) - \mathbb{E}[\omega_{1j}\omega_{1k}]\right| \to 0, \qquad n \to \infty. \tag{25}$$

The weights h_t, g_t satisfy the following conditions: setting $q_n = \sum_{t=1}^{n} h_t^2 g_t^2$,

$$\begin{aligned} &\max_{t=1,\ldots,n} h_t^4 = o(q_n), \quad \left(\sum_{t=1}^{n} h_t^4\right)^{1/2}\left(\sum_{t=2}^{n}(g_t - g_{t-1})^4\right)^{1/2} = o(q_n), \\ &\max_{t=1,\ldots,n} g_t^4 = o(q_n), \quad \left(\sum_{t=1}^{n} g_t^4\right)^{1/2}\left(\sum_{t=2}^{n}(h_t - h_{t-1})^4\right)^{1/2} = o(q_n). \end{aligned} \tag{26}$$

These conditions lead to a transparent asymptotic theory. It can be shown that robust testing procedures remain valid for a large class of nonsmooth scaling factors h_t and g_t that meet these conditions.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

I: The case of mutually dependent series $\{x_t\}$ and $\{y_t\}$.

First, we assume that x_t and the lagged variables y_{t-k} are uncorrelated but not mutually independent. For example, suppose that in (22) $\{\varepsilon_t\}$ is an m.d. sequence with respect to some σ-field $\mathcal{F}_t$, and η_{t-k} is $\mathcal{F}_{t-1}$ measurable. Then $\varepsilon_t\eta_{t-k}$ is also an m.d. sequence so that $\mathbb{E}[\varepsilon_t\eta_{t-k}|\mathcal{F}_{t-1}] = 0$, and thus $\operatorname{corr}(x_t, y_{t-k}) = 0$.

Bivariate series x_t and y_{t-k} can be uncorrelated at some lags, say $k = m_0, \ldots, m$ where $0 \le m_0 \le m$, and correlated at other lags, say $k = 0$. The next theorem establishes the multivariate limit distribution of the vector $(\widetilde{t}_{xy,m_0}, \ldots, \widetilde{t}_{xy,m})$ when $\operatorname{corr}(x_t, y_{t-k}) = 0$ for $k = m_0, \ldots, m$. Subsequently, we use this vector to test the hypothesis $\operatorname{corr}(x_t, y_{t-k}) = 0$, $k = m_0, \ldots, m$. To show asymptotic normality for the statistic $\widetilde{t}_{xy,k}$ based on the centered variables $x_t - \bar{x}$ and $y_t - \bar{y}$, we make the following assumption.

Assumption B. The autocovariance functions $\operatorname{cov}(\varepsilon_t, \varepsilon_{t-k}) = \gamma_{\varepsilon,k}$, $\operatorname{cov}(\eta_t, \eta_{t-k}) = \gamma_{\eta,k}$ of the stationary sequences $\{\varepsilon_t\}$ and $\{\eta_t\}$ satisfy the following covariance summability conditions

$$\sum\nolimits_k |\gamma_{\varepsilon,k}| < \infty, \quad \sum\nolimits_k |\gamma_{\eta,k}| < \infty. \tag{27}$$

These conditions are clearly satisfied by white noise/m.d. sequences $\{\varepsilon_t\}$ or $\{\eta_t\}$.

THEOREM 3.1. *Suppose that $\{x_t, y_t\}$ in (22) satisfy Assumptions A, B, and (26). If $\{\varepsilon_t\eta_{t-k}\}$, $k = m_0, \ldots, m$ $(0 \le m_0 \le m)$ are m.d. sequences with respect to the same σ-field $\mathcal{F}_t$, with $\mathbb{E}[\varepsilon_t\eta_{t-k}|\mathcal{F}_{t-1}] = 0$, then, as $n \to \infty$,*

$$(\widetilde{t}_{xy,m_0}, \ldots, \widetilde{t}_{xy,m}) \to_D \mathcal{N}(0, R_{xy}), \tag{28}$$

where $R_{xy} = (r_{xy,jk})_{j,k=m_0,\ldots,m}$ is a matrix with elements

$$r_{xy,jk} = \operatorname{corr}(\varepsilon_1\eta_{1-j}, \varepsilon_1\eta_{1-k}).$$

In particular, $\widetilde{t}_{xy,k} \to_D \mathcal{N}(0,1)$ for $k = m_0, \ldots, m$.

Simulations confirm that the test for zero cross-correlation at individual lag $k = 0, 1, 2, \ldots$ based on $\widetilde{t}_{xy,k}$ is well sized in finite samples for numerous lags when $\{x_t\}$ is serially uncorrelated and $\{y_t\}$ is serially uncorrelated or a series of dependent random variables.

Theorem 3.1 implies that $\widehat{c}_{xy,k}\widehat{\rho}_{xy,k} \to_D \mathcal{N}(0,1)$ for $k = m_0, \ldots, m$. The corollary below shows that the *"standard error"* $\widehat{c}^{-1}_{xy,k} = \widehat{\rho}_{xy,k}/\widetilde{t}_{xy,k}$ has a deterministic asymptotic form. This implies that the robust nonrejection region $\pm z_{\alpha/2}/\widehat{c}_{xy,k}$ of the null hypothesis of a lag k zero correlation at the α significance level can be interpreted as a $1-\alpha$ confidence interval for zero correlation. For this result, we employ some additional assumptions on the scaling factors $\{h_t, g_t\}$ and the autocovariograms of the stationary sequences $\{\varepsilon_t^2\}$ and $\{\eta_t^2\}$:

$$\max\nolimits_{1\le t\le n} h_t^2 = o\left(\sum\nolimits_{t=1}^n h_t^2\right), \max\nolimits_{1\le t\le n} g_t^2 = o\left(\sum\nolimits_{t=1}^n g_t^2\right),$$
$$\operatorname{cov}(\varepsilon_k^2, \varepsilon_0^2) \to 0, \operatorname{cov}(\eta_k^2, \eta_0^2) \to 0 \text{ as } k \to \infty. \tag{29}$$

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

COROLLARY 3.1. *Let (29) and the assumptions of Theorem 3.1 hold. Then,*

$$\widehat{c}_{xy,k} = \left(\frac{\sum_{t=1}^{n} h_t^2 \sum_{t=1}^{n} g_t^2}{\sum_{t=1}^{n} h_t^2 g_t^2} \frac{\mathbb{E}\varepsilon_1^2 \mathbb{E}\eta_1^2}{\mathbb{E}[\varepsilon_1^2 \eta_{1-k}^2]} \right)^{1/2} (1 + o_p(1)). \tag{30}$$

We next consider cumulative tests. First, note that the matrix R_{xy} in (28) is positive definite and R_{xy}^{-1} exists if $\varepsilon_1 \neq 0$ *a.s.*; see Lemma 3.1 below. The convergence (28) implies

$$\widetilde{t}_{xy}' R_{xy}^{-1} \widetilde{t}_{xy} \to_D \chi^2_{m-m_0+1} \quad \text{where } \widetilde{t}_{xy} = (\widetilde{t}_{xy,m_0}, \ldots, \widetilde{t}_{xy,m})'. \tag{31}$$

For testing the cumulative hypothesis $H_0 : \rho_{xy,m_0} = \cdots = \rho_{xy,m} = 0$, we suggest the following standardized statistics

$$Q_{xy,m} = \widetilde{t}_{xy}' \widehat{R}_{xy}^{-1} \widetilde{t}_{xy}, \qquad \widetilde{Q}_{xy,m} = \widetilde{t}_{xy}' \widehat{R}_{xy}^{*-1} \widetilde{t}_{xy}, \tag{32}$$

where $\widehat{R}_{xy} = (\widehat{r}_{xy,jk})_{j,k=m_0,\ldots,m}$ and $\widehat{R}_{xy}^{*}$ are consistent estimates of the matrix R_{xy}. We define

$$\widehat{r}_{xy,jk} = \frac{\sum_{t=\max(j,k)+1}^{n} e_{xy,tj} e_{xy,tk}}{\left(\sum_{t=\max(j,k)+1}^{n} e_{xy,tj}^2 \right)^{1/2} \left(\sum_{t=\max(j,k)+1}^{n} e_{xy,tk}^2 \right)^{1/2}}, \quad j,k = m_0, \ldots, m. \tag{33}$$

To improve the finite sample performance of the cumulative test, we suggest the thresholded version $\widehat{R}_{xy}^{*} = (\widehat{r}_{xy,jk}^{*})_{j,k=m_0,\ldots,m}$ of $\widehat{R}_{xy}$, where

$$\widehat{r}_{xy,jk}^{*} = \widehat{r}_{xy,jk} I(|\tau_{xy,jk}| > \lambda) \quad \text{with } \lambda = 2.576 \tag{34}$$

and $\tau_{xy,jk}$ is a t-statistic constructed as

$$\tau_{xy,jk} = \frac{\sum_{t=\max(j,k)+1}^{n} e_{xy,tj} e_{xy,tk}}{\left(\sum_{t=\max(j,k)+1}^{n} e_{xy,tj}^2 e_{xy,tk}^2 \right)^{1/2}}. \tag{35}$$

Just as in the univariate case, $\widehat{R}_{xy}^{*}$ is the sample analogue of the variance–covariance matrix of $\widetilde{t}_{xy}$ where we threshold its off-diagonal elements by checking at the 1% level whether they are significant. In our simulations and applications we set $\lambda = 2.576$, but in theory other threshold values $\lambda > 0$ can be used.

The limit theory of these statistics for testing cross-correlation between two serially uncorrelated time series at cumulative lags is given in the following result.

THEOREM 3.2. *Under the assumptions of Theorem 3.1, for any $\lambda > 0$, as $n \to \infty$,*

$$\widehat{R}_{xy} \to_p R_{xy}, \quad \widehat{R}_{xy}^{*} \to_p R_{xy}, \tag{36}$$

$$Q_{xy,m} \to_D \chi^2_{m-m_0+1}, \quad \widetilde{Q}_{xy,m} \to_D \chi^2_{m-m_0+1}. \tag{37}$$

Simulations show that if both $\{x_t\}$ and $\{y_t\}$ are serially uncorrelated the robust cumulative test $\widetilde{Q}_{xy,m}$ based on the thresholded estimate $\widehat{R}_{xy}^{*}$ with parameter

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

$\lambda = 2.576$ in most cases corrects adequately for size, whereas for large m the cumulative test based on $Q_{xy,m}$ suffers some size distortion. We note that in simulations for large values of m the statistic $\widetilde{Q}_{xy,m}$ may sometimes be negative since the matrix $\widehat{R}^*_{xy}$ is not necessarily positive definite.

Different from the univariate case, Assumption A is employed to avoid assuming ergodicity of the stationary sequence $\{\varepsilon_t^2 \eta_{t-j}\eta_{t-k}\}$, which implies (25), see Stout (1974, Cor. 3.5.2). In general, ergodicity of the separate sequences $\{\varepsilon_t\}$ and $\{\eta_t\}$ does not necessarily imply ergodicity of $\{\varepsilon_t^2 \eta_{t-j}\eta_{t-k}\}$ and thus (25). Property (25) alone is sufficient for the proofs to hold.

Theorem 3.1 assumes $\{\varepsilon_t \eta_{t-k}\}$ to be an m.d. sequence. This is a weak assumption and allows for various types of dependence between the sequences $\{\varepsilon_t\}$ and $\{\eta_t\}$. For example, if $\{\varepsilon_t\}$ is an m.d. sequence with $\mathbb{E}[\varepsilon_t|\mathcal{F}_{t-1}] = 0$ and $\eta_t = g(\varepsilon_t, \varepsilon_{t-1}, \ldots)$ is a measurable function of ε_s, $s \le t$, then $\{\varepsilon_t \eta_{t-k}\}$ is an m.d. sequence with respect to $\mathcal{F}_t$ for $k \ge 1$, so that

$$\mathbb{E}[\varepsilon_t \eta_{t-k}|\mathcal{F}_{t-1}] = \eta_{t-k}\mathbb{E}[\varepsilon_t|\mathcal{F}_{t-1}] = 0.$$

Simulations show that both the modified tests $\widetilde{t}_{xy,k}$ and $\widetilde{Q}_{xy,m}$ for zero cross-correlation (at individual and cumulative lags) manifest good size control when $\{x_t\}$ and $\{y_t\}$ are series of uncorrelated variables with time varying variances, and when they are mutually uncorrelated but not necessarily independent. The standard tests $t_{xy,k}$ and $HB_{xy,m}$ require the series $\{x_t\}$ and $\{y_t\}$ to be mutually independent stationary uncorrelated noises. The test $\widetilde{t}_{xy,k}$ at individual lags also achieves good size control when the lagged series is correlated, whereas the standard test $t_{xy,k}$ requires independence of $\{x_t\}$ and $\{y_t\}$ for good performance.

The next result verifies the existence of R_{xy}^{-1}. For $z = (z_1, \ldots, z_m)$, $m \ge 1$ define $\mathrm{Cov}(z) = (\mathrm{cov}(z_j, z_k))$ and $\mathrm{Corr}(z) = (\mathrm{corr}(z_j, z_k))$.

LEMMA 3.1. *Let* $\{\eta_t\}$ *be a stationary sequence with* $\mathbb{E}\eta_t = 0$, $0 < \mathbb{E}\eta_t^2 < \infty$, *autocovariance sequence* $\gamma_\eta(h) = \mathbb{E}(\eta_t \eta_{t-h})$, *and spectral density* $f_\eta(x)$, $x \in [-\pi, \pi]$. *Then:*

(i) $\mathrm{Cov}(\eta)$ *and* $\mathrm{Corr}(\eta)$ *are positive definite for any* $\eta = (\eta_1, \ldots, \eta_m)$; *and*
(ii) *if* $z_j = \varepsilon_1 \eta_{1-j}$, $\mathbb{E}z_j = 0$, $\mathbb{E}z_j^2 < \infty$ *for* $j \ge 1$ *and* $\varepsilon_1 \ne 0$ *a.s., then* $\mathrm{Cov}(z)$ *and* $\mathrm{Corr}(z)$ *are positive definite for any* $z = (z_1, \ldots, z_m)$.

II: The case of mutually independent series $\{x_t\}$ and $\{y_t\}$.

If $\{x_t\}$ and $\{y_t\}$ are mutually independent, then $\mathrm{corr}(x_t, y_{t-k}) = 0$ and $\mathrm{corr}(y_t, x_{t-k}) = 0$ for $k \ge 0$. Properties (28) and (37) of the test statistic $\widetilde{t}_{xy,k}$ for an individual lag and the corresponding cumulative test statistic are preserved if the series $\{x_t\}$ is an uncorrelated m.d. sequence while $\{y_t\}$ may be a sequence of dependent variables.

THEOREM 3.3. *Assume that* $\{x_t, y_t\}$ *in (22) satisfy Assumptions A, B, and (26) and* $\{x_t\}$ *and* $\{y_t\}$ *are mutually independent. Suppose* $\{\varepsilon_t\}$ *in (26) is an m.d. sequence. Then, for any* $0 \le m_0 \le m$, *as* $n \to \infty$,

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

$$\widetilde{\tau}_{xy} = (\widetilde{\tau}_{xy,m_0}, \ldots, \widetilde{\tau}_{xy,m})' \to_D \mathcal{N}(0, R_{xy}),$$
$$\widetilde{\tau}_{yx} = (\widetilde{\tau}_{yx,m_0}, \ldots, \widetilde{\tau}_{yx,m})' \to_D \mathcal{N}(0, R_{yx}), \tag{38}$$

where $R_{xy} = R_{yx} = R_y = (r_{y,jk})$ *and* $r_{y,jk} = \text{corr}(y_j, y_k)$, $j,k = m_0, \ldots, m$.

In particular, $\widetilde{\tau}_{xy,k} \to_D \mathcal{N}(0,1)$ *for any lag* $k = \ldots -1, 0, 1, \ldots$. *Moreover,*

$$Q_{xy,m} \to_D \chi^2_{m-m_0+1}, \quad \widetilde{Q}_{xy,m} \to_D \chi^2_{m-m_0+1} \tag{39}$$

$$\widetilde{\tau}'_{yx}\widehat{R}^{-1}_{xy}\widetilde{\tau}_{yx} \to_D \chi^2_{m-m_0+1}, \quad \widetilde{\tau}'_{yx}\widehat{R}^{*-1}_{xy}\widetilde{\tau}_{xy} \to_D \chi^2_{m-m_0+1,} \tag{40}$$

where $\widehat{R}_{xy}$ *and* $\widehat{R}^*_{xy}$ *are defined as above in (33) and (34) with* $0 \le m_0 \le m$.

Consistent estimates $\widehat{R}_{xy}$ and $\widehat{R}^*_{xy}$ of $R_{xy} = R_y$ require the first variable $\{x_t\}$ to be an uncorrelated m.d. sequence. Notice that $\widehat{R}_{yx} \to_p I \neq R_{yx} = R_y$. Hence, applying the test $\widetilde{Q}_{yx,m}$ instead of $\widetilde{\tau}'_{yx}\widehat{R}^{*-1}_{xy}\widetilde{\tau}_{xy}$ in (40) when the lead series $\{y_t\}$ is dependent may lead to significant size distortions.

Remark 3.1. Testing for cross-correlations, the robust cumulative test $\widetilde{Q}_{xy,m}$ is well sized in simulations when both $\{x_t\}$ and $\{y_t\}$ are serially uncorrelated but suffers size distortions after a few lags when one of the series is serially correlated. Clearly, thresholding does not improve the estimate of $R_{xy,m}$ sufficiently in finite samples when $R_{xy,m}$ is not sparse as in Theorem 3.3 and the robust test $\widetilde{Q}_{xy,m}$ produces correct size only for a few low lags (see Figure 3(c)). Size can be distorted for all lags when the lead series is dependent (see Theorem 3.3). Hence, in applied work we recommend using the robust cumulative test $\widetilde{Q}_{xy,k}$ in cases when both series are serially uncorrelated.

Finally, if the time series $\{x_t\}$ and $\{y_t\}$ are mutually independent but neither $\{x_t\}$ nor $\{y_t\}$ is white noise, the standard normal approximation for $\widetilde{\tau}_{xy,k}$ does not generally hold. In such cases, even if

$$\widetilde{\tau}_{xy,k} \to_D \mathcal{N}(0, \sigma^2_{xy}),$$

the variance $\sigma^2_{xy} = \sum_{j=-\infty}^{\infty} \text{corr}(\varepsilon_0, \varepsilon_j)\text{corr}(\eta_0, \eta_j)$ is not unity, as shown in the following result.

PROPOSITION 3.1. *Assume that* $\{x_t, y_t\}$ *in (22) satisfy Assumptions A, B, and (26). Suppose that sequences* $\{x_t\}$ *and* $\{y_t\}$ *are mutually independent. Then, for any* $k \ge 0$,

$$\widetilde{\tau}_{xy,k} = s_{nk} + o_p(1), \quad n \to \infty, \tag{41}$$

where $\mathbb{E}s_{nk} = 0$, $\text{var}(s_{nk}) \to \sigma^2_{xy} = \sum_{j=-\infty}^{\infty} \text{corr}(\varepsilon_0, \varepsilon_j)\text{corr}(\eta_0, \eta_j)$. *For the definition of* s_{nk}, *see (29) in the proof in Online Supplement I.*

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

III: Test consistency.

To conclude this section, we show that the test for correlation at lag k based on $\widetilde{t}_{xy,k}$, is consistent when $\text{corr}(x_t, y_{t-k}) \neq 0$. To do so, we make some further assumptions on the components $\omega_{tk} = \varepsilon_t \eta_{t-k}$ and (h_t, g_t) of $\{x_t, y_t\}$ in (22).

Assumption C. $\{\omega_{tk}\}$ is a stationary sequence whose covariances satisfy the summability condition

$$\sum_{j=-\infty}^{\infty} |\text{cov}(\omega_{1k}, \omega_{jk})| < \infty, \tag{42}$$

and (h_t, g_t) are such that

$$\sum_{t=k+1}^{n} |h_t(g_t - g_{t-k})| = o(\sum_{t=1}^{n} h_t g_t), \quad q_n^{-1/2} \sum_{t=1}^{n} h_t g_t \to \infty, \tag{43}$$

where $q_n = \sum_{t=1}^{n} h_t^2 g_t^2$, as defined earlier.

Condition (42) is a standard weak dependence condition on the covariances of the sequence $\{\omega_{tk}\}$, and condition (43) is satisfied by many weight sequences (h_t, g_t) that induce unconditional heterogeneity in the series over time, such as linear time trends.

THEOREM 3.4. *Let $\{x_t, y_t\}$ be as in (22) and Assumptions A, B, C, and (26) hold. Suppose that for some $k \geq 0$, $\text{corr}(x_t, y_{t-k}) \neq 0$. Then, as $n \to \infty$,*

$$\widetilde{t}_{xy,k} = \frac{\left(\sum_{t=1}^{n} h_t g_t\right)}{\left(\sum_{t=1}^{n} h_t^2 g_t^2\right)^{1/2}} \frac{\text{cov}(\varepsilon_1, \eta_{1-k})}{(\mathbb{E}[\varepsilon_1^2 \eta_{1-k}^2])^{1/2}} (1 + o_p(1)). \tag{44}$$

Consistency of the tests follow directly from (44) in view of (43).

Remark 3.2. Our results cover as a special case the test for zero Pearson correlation r between two variables x and y, which is the cross-correlation $\rho_{xy,0}$ at lag $k = 0$ between x and y. When samples of x and y are selected randomly and do not have the same variance, the test for $r = 0$ based on the standard test statistic $\widehat{t}_{xy,0}$ is susceptible to size distortions whereas the robust test $\widetilde{t}_{xy,0}$ remains well-sized.

4. TESTS FOR THE I.I.D. PROPERTY

In this section, we examine a simple test for the i.i.d. property of a time series $\{x_t\}$ based on a sample $x_1, \ldots, x_n$. Campbell *et al.* (1997, Chapter 2) provide a brief exposition of this approach for use in financial econometrics. We assume that

$$x_t = \mu + \varepsilon_t, \tag{45}$$

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

where $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with $\mathbb{E}\varepsilon_t = 0$, $\mathbb{E}\varepsilon_t^2 < \infty$. Denote $\rho_{x,k} = \mathrm{corr}(x_t, x_{t-k})$ and define

$$\rho_{|x|,k} = \mathrm{corr}\big(|x_t - \mu|, |x_{t-k} - \mu|\big), \quad \rho_{x^2,k} = \mathrm{corr}\big((x_t - \mu)^2, (x_{t-k} - \mu)^2\big).$$

Clearly, if $\{x_t\}$ is i.i.d., then $\rho_{x,k} = \rho_{|x|,k} = \rho_{x^2,k} = 0$ for $k \neq 0$. With this approach, the problem of testing the i.i.d. property of the time series $\{x_t\}$ is reduced to testing for the absence of correlation in $\{x_t\}$ and $\{|x_t - \mathbb{E}x_t|\}$, or alternatively in $\{x_t\}$ and $\{(x_t - \mathbb{E}x_t)^2\}$. Other tests involving nonparametric density estimation (e.g., Gretton and Györfi, 2010) are available but are considerably more complex in their implementation. The present approach has the benefit of simplicity and makes use of the test machinery developed earlier.

Our test statistics combine the levels of the data $\{x_t\}$ and either absolute $\{|x_t - \bar{x}|\}$ or squared $\{(x_t - \bar{x})^2\}$ deviations from the sample mean $\bar{x}$. We denote by $\widehat{\rho}_{x,k}$, $\widehat{\rho}_{|x|,k}$ and $\widehat{\rho}_{x^2,k}$ the sample correlation (1) computed using the data x_t, $|x_t - \bar{x}|$ and $(x_t - \bar{x})^2$, respectively. Define

$$\tau_{x,k} = \frac{n}{(n-k)^{1/2}}\widehat{\rho}_{x,k}, \quad \tau_{|x|,k} = \frac{n}{(n-k)^{1/2}}\widehat{\rho}_{|x|,k}, \quad \tau_{x^2,k} = \frac{n}{(n-k)^{1/2}}\widehat{\rho}_{x^2,k}.$$

Denote $r_{\varepsilon,|\varepsilon|} = \mathrm{corr}(\varepsilon_1, |\varepsilon_1|)$, $r_{\varepsilon,\varepsilon^2} = \mathrm{corr}(\varepsilon_1, \varepsilon_1^2)$, and set

$$V_{x,|x|} = \begin{pmatrix} 1 & r^2_{\varepsilon,|\varepsilon|} \\ r^2_{\varepsilon,|\varepsilon|} & 1 \end{pmatrix}, \quad V_{x,x^2} = \begin{pmatrix} 1 & r^2_{\varepsilon,\varepsilon^2} \\ r^2_{\varepsilon,\varepsilon^2} & 1 \end{pmatrix}.$$

The next theorem establishes the joint distribution for the statistics $(\tau_{x,k}, \tau_{|x|,k})$ and $(\tau_{x,k}, \tau_{x^2,k})$.

THEOREM 4.1. *Let $x_1, \ldots, x_n$ be a sample from an i.i.d. sequence (45). If $\tau_{x^2,k}$ is employed, assume in addition that $\mathbb{E}\varepsilon_1^4 < \infty$. Then for $m \geq 1$,*

$$(\tau_{x,1}, \tau_{|x|,1}, \ldots, \tau_{x,m}, \tau_{|x|,m}) \to_D \mathcal{N}(0, V_{x,|x|,2m}), \tag{46}$$

$$(\tau_{x,1}, \tau_{x^2,1}, \ldots, \tau_{x,m}, \tau_{x^2,m}) \to_D \mathcal{N}(0, V_{x,x^2,2m}), \tag{47}$$

where $V_{x,|x|,2m} = \mathrm{diag}(V_{x,|x|}, \ldots, V_{x,|x|})$ and $V_{x,x^2,2m} = \mathrm{diag}(V_{x,x^2}, \ldots, V_{x,x^2})$ are $2m \times 2m$ block-diagonal matrices. In particular, for $k \geq 1$,

$$(\tau_{x,k}, \tau_{|x|,k}) \to_D \mathcal{N}(0, V_{x,|x|}), \quad (\tau_{x,k}, \tau_{x^2,k}) \to_D \mathcal{N}(0, V_{x,x^2}). \tag{48}$$

Observe that $V_{x,|x|,2m} = I$, $V_{x,x^2,2m} = I$ are identity matrices if the $\{\varepsilon_t\}$ have a symmetric distribution, since then $r_{\varepsilon,|\varepsilon|} = r_{\varepsilon,\varepsilon^2} = 0$. In general, the nondiagonal elements $r^2_{\varepsilon,|\varepsilon|}$ and $r^2_{\varepsilon,\varepsilon^2}$ in the matrices $V_{x,|x|,2m}$ and $V_{x,x^2,2m}$ are likely to be small. Hence, the standard normal limit $\mathcal{N}(0, I)$ may be a good approximation in (46) and (47) in finite samples. This suggests the following approximation

$$\sum_{k=1}^{m}(\tau^2_{x,k} + \tau^2_{|x|,k}) \sim \chi^2_{2m}, \quad \sum_{k=1}^{m}(\tau^2_{x,k} + \tau^2_{x^2,k}) \sim \chi^2_{2m}, \tag{49}$$

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

which is easy to use in applied work. Good performance of the latter statistics is confirmed by simulations in the Monte Carlo study.

To verify the i.i.d. property of $\{x_t\}$, we test at individual lags and cumulatively for the absence of correlation in levels $\{x_t\}$ and absolute values $\{|x_t - \mathbb{E}x_t|\}$ via the following null hypotheses

H_0: $\rho_{x,k} = 0$, $\rho_{|x|,k} = 0$ at individual lag $k \geq 1$,

H_0: $\rho_{x,k} = 0$, $\rho_{|x|,k} = 0$ for $k = 1, \ldots, m$ for $m \geq 1$ (cumulative)

using respectively the statistics

$$J_{x,|x|,k} = \frac{n^2}{n-k}(\widehat{\rho}^2_{x,k} + \widehat{\rho}^2_{|x|,k}), \qquad C_{x,|x|,m} = \sum_{k=1}^{m} J_{x,|x|,k}. \tag{50}$$

Alternatively, we may test for the absence of correlation in levels $\{x_t\}$ and squares $\{(x_t - \mathbb{E}x_t)^2\}$ via the hypotheses

H_0: $\rho_{x,k} = 0$, $\rho_{x^2,k} = 0$ at individual lag $k \geq 1$,

H_0: $\rho_{x,k} = 0$, $\rho_{x^2,k} = 0$ for $k = 1, \ldots, m$ (cumulative),

using respectively the statistics

$$J_{x,x^2,k} = \frac{n^2}{n-k}(\widehat{\rho}^2_{x,k} + \widehat{\rho}^2_{x^2,k}), \qquad C_{x,x^2,m} = \sum_{k=1}^{m} J_{x,x^2,k}. \tag{51}$$

If an i.i.d. time series $\{x_t\}$ has a symmetric distribution, then Theorem 4.1 shows that, as $n \to \infty$,

$$J_{x,|x|,k}, J_{x,x^2,k} \to_D \chi^2_2, \qquad C_{x,|x|,m}, C_{x,x^2,m} \to_D \chi^2_{2m}. \tag{52}$$

Simulations confirm that the χ^2_{2m} distribution provides a good approximation also for i.i.d. time series x_t with nonsymmetric distributions that do not exhibit severe skewness. Simulations show that these tests have good power in the presence of dependence, conditional heteroskedasticity or nonstationarity when $\mathbb{E}x_t$ or $\text{var}(x_t)$ varies with time.

Related to these results, we recall that some standard tests are already in the literature as mentioned in the Introduction. Notably, convergence for the cumulative test statistic

$$\sum_{k=1}^{m} \frac{n^2}{n-k} \widehat{\rho}^2_{x^2,k} \to_D \chi^2_m$$

based on squares x_t^2 of an i.i.d. sequence was established by McLeod and Li (1983). The present tests involve both levels and absolute values or both levels and squares. Similar tests for residuals of a fitted ARMA-GARCH model were used by Wong and Ling (2005) and Zhu (2013). Different from our testing procedures, residual based tests do not require demeaning and their asymptotic distribution may depend on the parameters of a fitted model.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

5. MONTE CARLO EXPERIMENTS

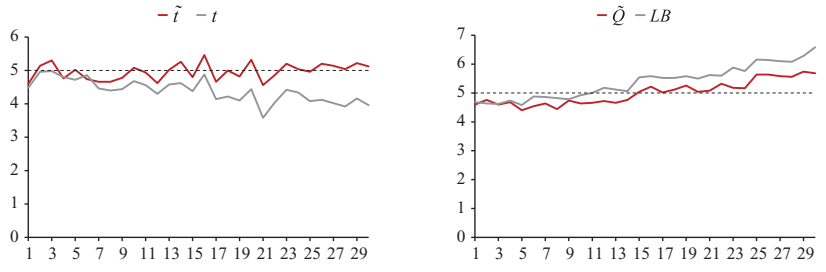
This section presents Monte Carlo findings on the finite sample performance of the standard and robust tests for zero correlation given in Sections 2 and 3 and the tests for the i.i.d. property given in Section 4. We use a variety of models for $\{x_t\}$ and $\{x_t, y_t\}$, sample sizes $n = 100, 300, 1,000$ and the experiments each involve $5,000$ replications. We evaluate the rejection frequency (in %) of the test statistics at significance level $\alpha = 5\%$ using the asymptotic critical values and power is not size corrected. The standard test t_k and the robust $\widetilde{t}_k$ are based on 1.96 critical values, respectively. Figures 1–5 report rejection frequencies (size and power) for a subset of models for $n = 300$ and $k, m = 0, 1, 2, \ldots, 30$ lags. The full findings for $n = 300$ are given in the Online Supplement II of this paper to which readers are referred for complete details and results for $n = 100, 1,000$ are available upon request.

Figures 1 and 2 report size and power of the tests for zero correlation based on the statistics $\widetilde{t}_k$, t_k, $\widetilde{Q}_m$ with $\lambda = 2.576$ and LB_m. Figure 1 shows that the standard statistic t_k and the cumulative LB_m statistic have distorted size when the data are non-i.i.d., see e.g. models (b)–(d). The accumulation of the t-test size distortion in the LB_m test is evident at all lags. On the other hand, the robust statistics $\widetilde{t}_k$ and $\widetilde{Q}_m$ achieve the nominal size of $\alpha = 5\%$ for all models in Figure 1. For the i.i.d. data in model (a), the standard and robust methods give similar results as expected. Figure 2 displays test power. The results show some loss in power for the robust test statistics compared to the standard tests. All tests show spurious power when the data x_t is uncorrelated but has time varying mean, see models (c)–(d).

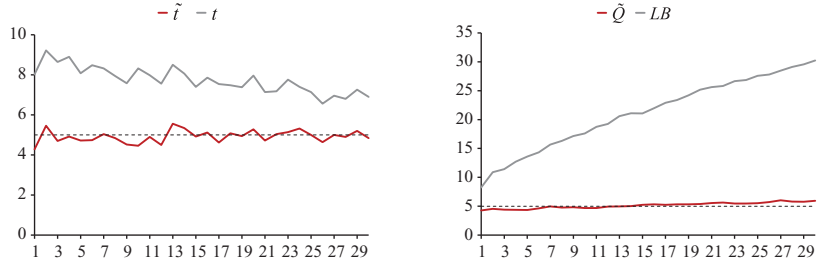
Figures 3 and 4 report test results for zero cross-correlation for bivariate time series $\{x_t\}$ and $\{y_t\}$ based on the statistics $\widetilde{t}_{xy,k}$, $t_{xy,k}$, $\widetilde{Q}_m$ with $\lambda = 2.576$ and HB_m. When series of independent variables $\{x_t\}$ and $\{y_t\}$ are heteroskedastic but jointly independent, as in Figure 3 model (a), the robust statistics produce correct size whereas the standard tests are all oversized. When the time series $\{x_t\}$ and $\{y_t\}$ are uncorrelated but not mutually independent, as in model (b), the robust statistics give correct size whereas the standard ones over-reject. When $\{x_t\}$ and $\{y_t\}$ are mutually independent, but one of the two has either autocorrelation or time varying mean, as in models (c) and (d), then both the standard and robust tests at an individual lag give the correct size, although the cumulative tests tend to become oversized. The latter outcome was unexpected, as the theory of Section 3 for model (c) would suggest the cumulative tests would be well-sized. It seems that $\widehat{R}^*_{xy,m}$ does not estimate well the nonsparse autocorrelation matrix R_y of y_t for this moderate sample size, especially for bigger lags m. When the time series are cross-correlated, as in Figure 4 models (a) and (b), we observe similar power across the standard and robust statistics, with some loss in power at bigger lags for the robust cumulative statistic. In Figure 4 models (c) and (d), the time series $\{x_t\}$ and $\{y_t\}$ are jointly independent but we observe spurious power, when both of them have autocorrelation, like in (c) and as suggested by theory, or when both of them have time varying mean, as in (d).

Figure 5 reports the test results for the i.i.d. property using the statistics $J_{x,|x|,k}, J_{x,x^2,k}$ and $C_{x,|x|,m}, C_{x,x^2,m}$. The size of the tests at individual lags and the

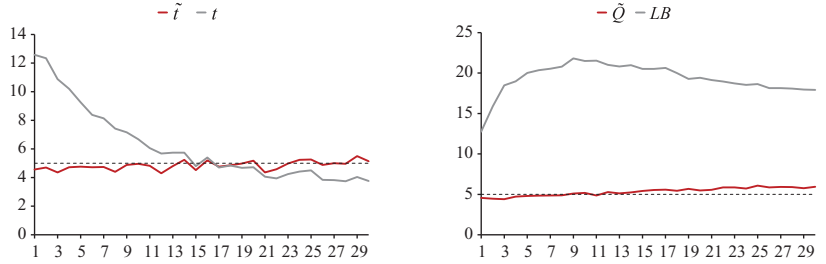
https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press



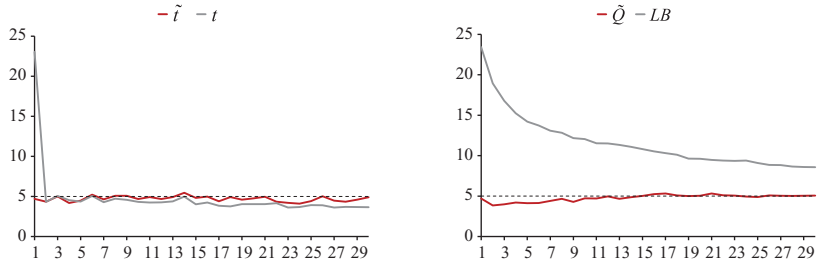
(a) $x_t = \varepsilon_t$



(b) $x_t = (1 + I(t/n > 0.5))\varepsilon_t$



(c) $x_t = r_t$, $r_t = \sigma_t \varepsilon_t$, $\sigma_t^2 = 1 + 0.2r_{t-1}^2 + 0.7\sigma_{t-1}^2$



(d) $x_t = \varepsilon_t \varepsilon_{t-1}$

FIGURE 1. Size. Rejection frequencies (in %) at $\alpha = 5\%$ of robust (red line) $\widetilde{t}_k$ and standard (grey line) t_k tests (left) and robust (red line) $\widetilde{Q}_m$ and standard (grey line) LB_m cumulative tests (right) at lags $k, m = 1, 2, \ldots, 30$. $\varepsilon_t \sim$ i.i.d. $\mathcal{N}(0, 1)$.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

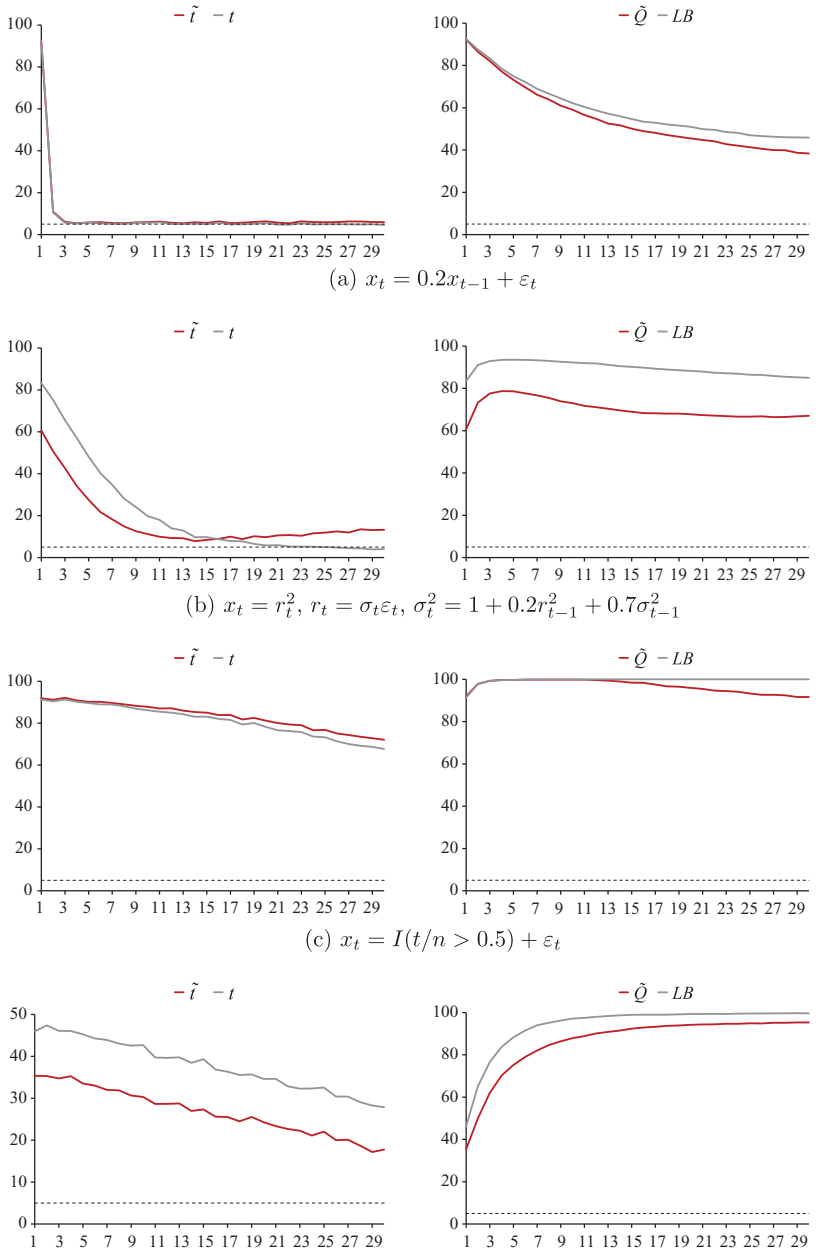


FIGURE 2. Power/spurious power. Rejection frequencies (in %) at $\alpha = 5\%$ of robust (red line) $\widetilde{t}_k$ and standard (grey line) t_k tests (left) and robust (red line) $\widetilde{Q}_m$ and standard (grey line) LB_m cumulative tests (right) at lags $k, m = 1, 2, \ldots, 30$. $\varepsilon_t \sim$ i.i.d. $\mathcal{N}(0, 1)$.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

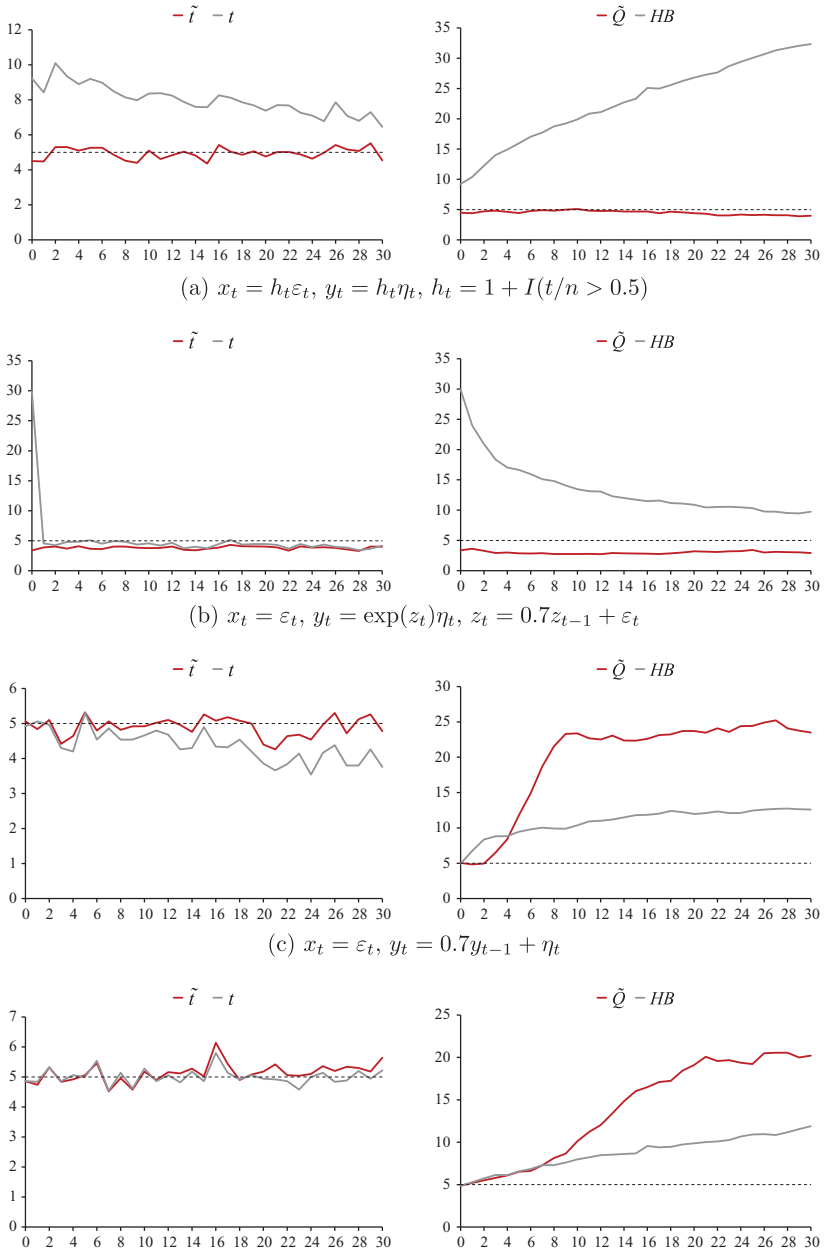
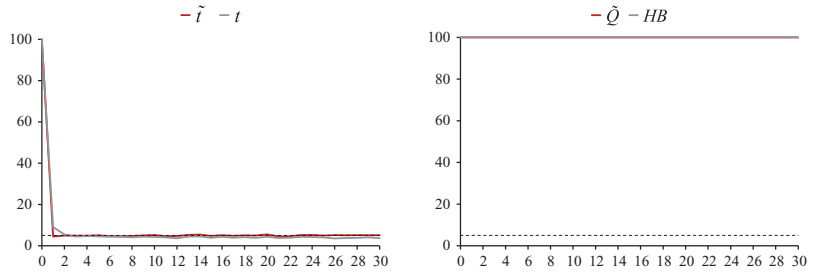
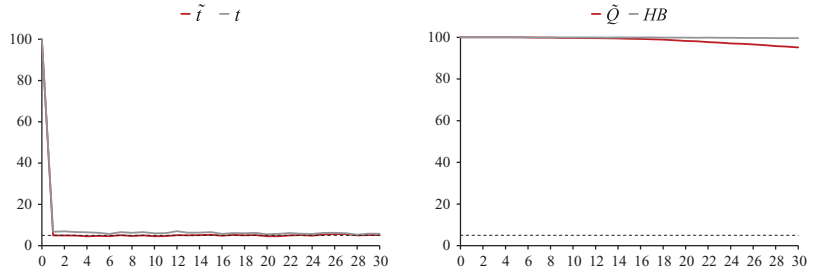


FIGURE 3. Size. Rejection frequencies (in %) at $\alpha = 5\%$ of robust (red line) $\widetilde{\tau}_{xy,k}$ and standard (grey line) $t_{xy,k}$ tests (left) and robust (red line) $\widetilde{Q}_{xy,m}$ and standard (grey line) $HB_{xy,m}$ cumulative tests (right) at lags $k, m = 0, 1, \ldots, 30$. $\varepsilon_t, \eta_t \sim$ i.i.d. $\mathcal{N}(0,1)$, $\{\varepsilon_t\}, \{\eta_t\}$ independent.

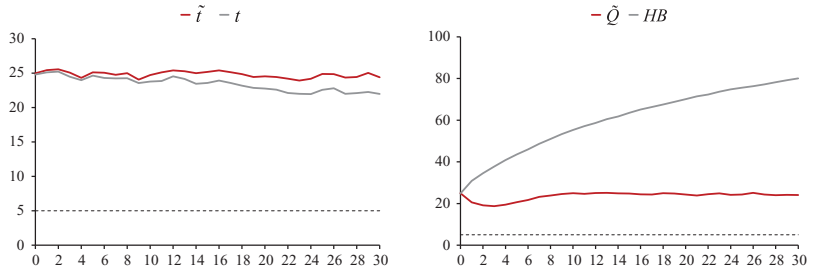
https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press



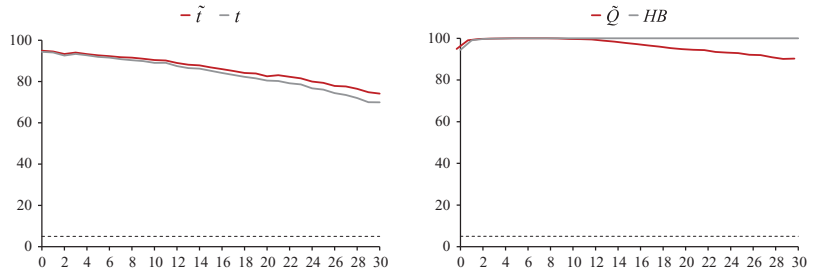
(a) $x_t = \sigma_{x,t}\varepsilon_t$, $y_t = \sigma_{y,t}\varepsilon_t$, $\sigma^2_{x,t} = 1 + 0.2x^2_{t-1}$, $\sigma^2_{y,t} = 1 + 0.2y^2_{t-1} + 0.7\sigma^2_{y,t-1}$



(b) $x_t = h_t\varepsilon_t$, $y_t = x_t + x_{t-1} + x_{t-2} + h_t\eta_t$, $h_t = 1 + I(t/n > 0.5)$



(c) $x_t = 0.7x_{t-1} + \varepsilon_t$, $y_t = 0.7y_{t-1} + \eta_t$



(d) $x_t = m_t + \varepsilon_t$, $y_t = m_t + \eta_t$, $m_t = I(t/n > 0.5)$

FIGURE 4. Power/spurious power. Rejection frequencies (in %) at $\alpha = 5\%$ of robust (red line) $\widetilde{t}_{xy,k}$ and standard (grey line) $t_{xy,k}$ tests (left) and robust (red line) $\widetilde{Q}_{xy,m}$ and standard (grey line) $HB_{xy,m}$ cumulative tests (right) at lags $k, m = 0, 1, \ldots, 30$. $\varepsilon_t, \eta_t \sim$ i.i.d. $\mathcal{N}(0,1)$, $\{\varepsilon_t\}, \{\eta_t\}$ independent.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

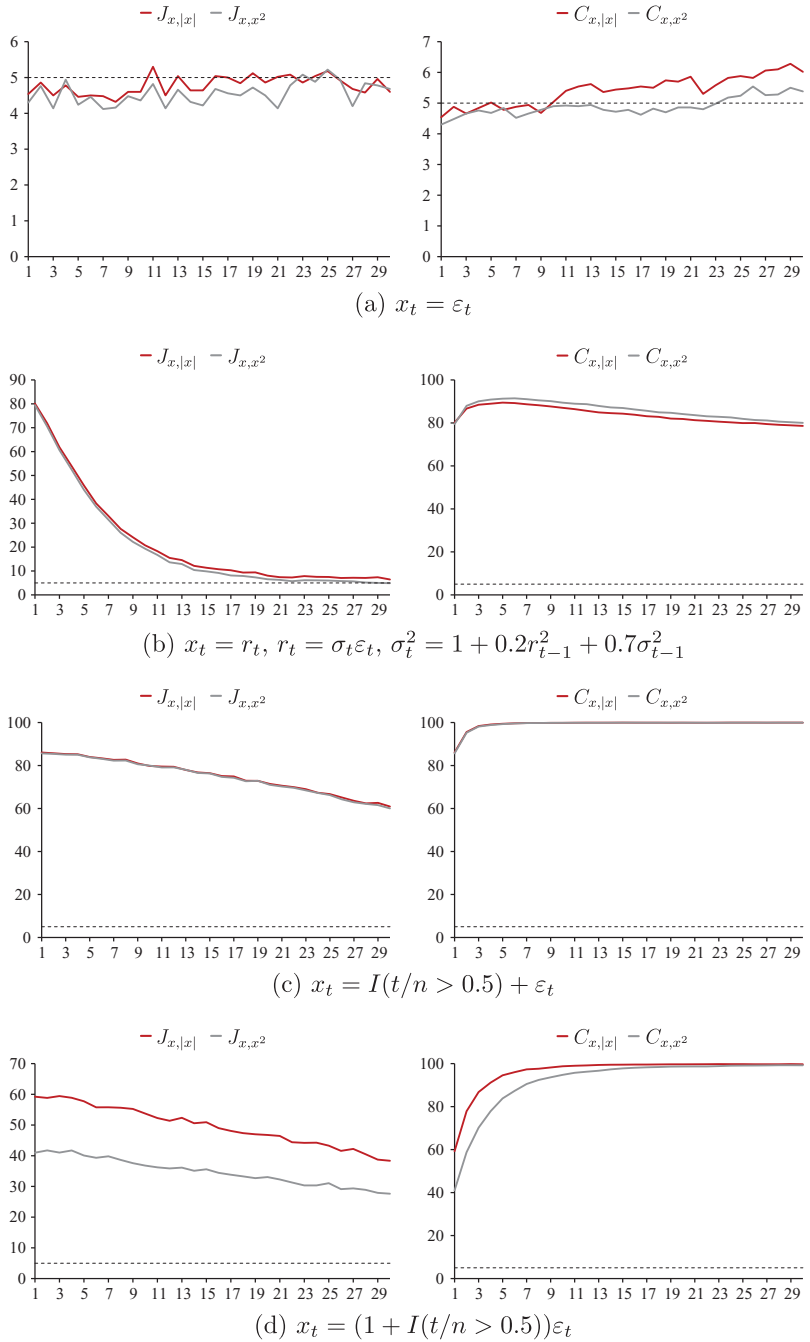


FIGURE 5. Size and power. Rejection frequencies (in %) at $\alpha = 5\%$ of $J_{x,|x|,k}$ (red line) and $J_{x,x^2,k}$ (grey line) tests (left) and $C_{x,|x|,m}$ (red line) and $C_{x,x^2,m}$ (grey line) cumulative tests (right) at lags $k, m = 1, 2, \ldots, 30$. $\varepsilon_t \sim$ i.i.d. $\mathcal{N}(0, 1)$.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

size of the cumulative tests are satisfactory in model (a), and we observe good power in discriminating the non-i.i.d. models (b)–(d). In particular, the statistics based on the absolute values have overall similar or better power properties than those based on the squares.

Some general conclusions from the simulation study are as follows. First, we find that in testing for correlation or for the i.i.d. property, tests at individual lag k perform well at all lags. The cumulative tests with $\lambda = 2.576$ perform overall well at most lags, with tiny distortions at the bigger lags. This conclusion is based on the Monte Carlo size properties of our tests for the models considered in the Online Supplement II. Second, we note that one needs to test for a constant mean (e.g., using the Bai and Perron (1998) test for multiple structural breaks) prior to applying the univariate (both standard and robust) tests for absence of correlation. Third, for bivariate tests it is useful to check for constant mean as well as for serial correlation in each time series prior to applying the tests. Fourth, the findings indicate that our tests for the i.i.d. property based on the χ^2_{2m} approximation perform well unless the distributions are extremely skewed.

Remark 5.1. The theory and Monte Carlo study suggest the following:

(i) In testing for autocorrelation the series needs to have constant mean.
(ii) In testing for cross-correlation each of the series needs to have constant mean and to be serially uncorrelated when applying the portmanteau type statistics or at least one when applying the t-type tests.
(iii) The values $\lambda = 1.96, 2.576$ are good candidates for the threshold in the robust portmanteau statistics, with $\lambda = 2.576$ performing better at relatively large lags.

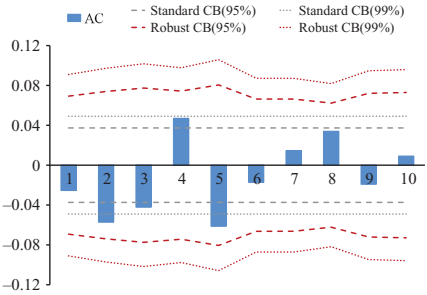
6. EMPIRICAL APPLICATION

We report an application of these methods to financial market data covering the period 2008–2018. Both the standard and robust tests for absence of correlation are used. We analyze univariate time series of daily returns $\{x_t\}$ of the FTSE100 index and the bivariate series $\{x_t, y_t\}$ of the daily returns of the FTSE100 index and gold price.[3] Graphical inspection of the data suggests that the means of both the FTSE100 and gold returns are likely constant.

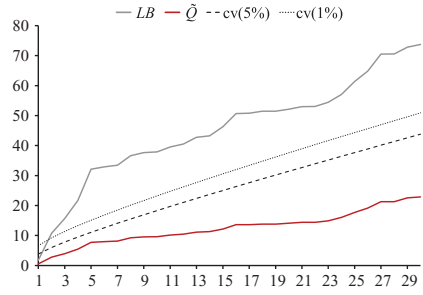
The results of the univariate analyses are shown in Figures 6 and 7. Panel (a) of Figure 6 contains the correlogram $\widehat{\rho}_k$ for lags $k = 1, 2, \ldots, 10$ along with 95% and 99% confidence bands (CBs) for insignificant correlation, the standard CBs are based on $\pm z_{\alpha/2}/\sqrt{n}$ and the robust CBs on $\pm z_{\alpha/2}(\widehat{\rho}_k/\widetilde{t}_k)$ at significance levels $\alpha = 5\%, 1\%$. Panel (b) reports the values of the standard LB_m and robust $\widetilde{Q}_m$ cumulative statistics with threshold $\lambda = 2.576$ for lags $m = 1, 2, \ldots, 30$ along with

[3]The data are sourced from YahooFinance for the FTSE100 index and from the Federal Reserve Bank of St. Louis for the London Bullion Market Association gold price. Both prices are in British pounds. The FTSE100 index is measured at the market closing at 4:30 GMT. The gold price is at 3:00 GMT. Returns are calculated as first differences of log-prices. Missing observations are deleted.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

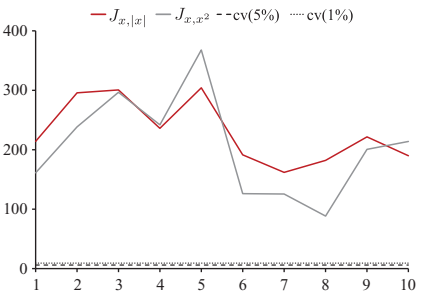


(a) Sample autocorrelation $\widehat{\rho}_k$, standard (grey) and robust (red) CB for non-significance of correlation at lags $k = 1, 2, ..., 10$.

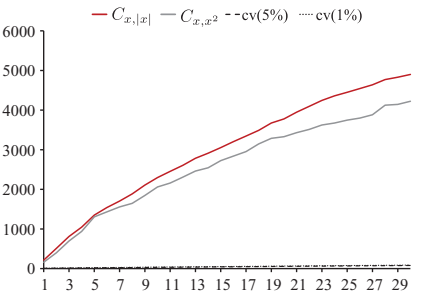


(b) Standard (grey) LB_m and robust (red) $\widetilde{Q}_m$ cumulative tests at lags $m = 1, 2, ..., 30$.

FIGURE 6. FTSE100 daily returns, 2008–2018: Correlogram.



(a) $J_{x,|x|,k}$ (red), $J_{x,x^2,k}$ (grey) tests at lags $k = 1, 2, ..., 10$.



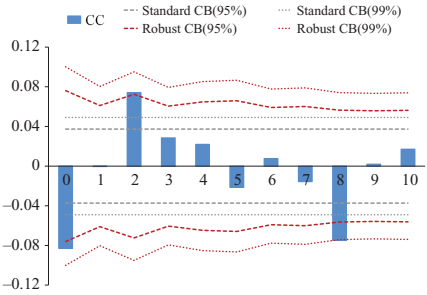
(b) $C_{x,|x|,m}$ (red) and $C_{x,x^2,m}$ (grey) tests at lags $m = 1, 2, ..., 30$.

FIGURE 7. FTSE100 daily returns, 2008–2018: i.i.d. test.

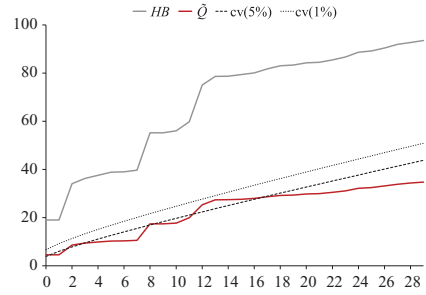
their asymptotic critical values (cv(α)) at significance $\alpha = 5\%, 1\%$ levels. The CBs of the standard test for correlation at individual lags show evidence of serial correlation at lags $k = 2, 5$ at the 1% level and at lags $k = 3, 4$ at the 5% level. The latter is magnified in the LB_m cumulative test. However, in agreement with the robust CBs, the evidence of serial correlation is insignificant at all individual lags at the 5% level. As such, the cumulative hypothesis of no correlation is not to be rejected by the cumulative test $\widetilde{Q}_m$. To test for the i.i.d. property of x_t, we evaluated the statistics $J_{x,|x|,k}$, $J_{x,x^2,k}$, $C_{x,|x|,m}$ and $C_{x,x^2,m}$ at lags $k, m = 1, 2, \ldots, 30$, which are shown in Figure 7. Evidently, this hypothesis is strongly rejected at significance levels $\alpha = 5\%, 1\%$. We therefore conclude that the daily returns of the FTSE100 index during 2008–2018 are uncorrelated, but strong evidence affirms that the series is not i.i.d.

The results of bivariate testing are shown in Figure 8. Just as the FTSE100 returns were found to be uncorrelated, similar analysis (not reported here) confirms uncorrelatedness of the gold returns. Panels (a) and (c) contain the sample

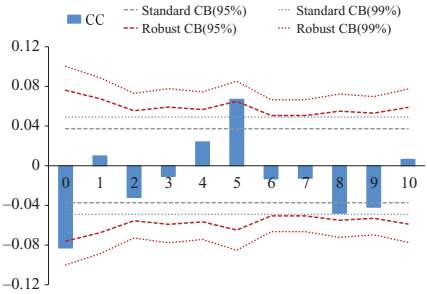
https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press



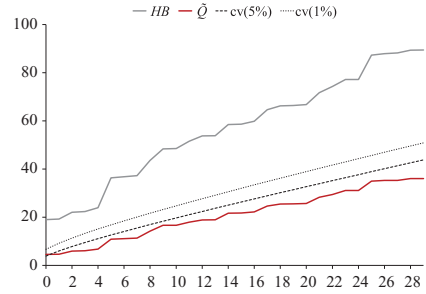
(a) Sample cross-correlation $\widehat{\rho}_{xy,k}$, standard (grey) and robust (red) CB for non-significance of cross-correlation at lags $k = 0, 1, ..., 10$.



(b) Standard (grey) $HB_{xy,m}$ and robust (red) $\widetilde{Q}_{xy,m}$ cumulative tests at lags $m = 0, 1, ..., 30$.



(c) Sample cross-correlation $\widehat{\rho}_{yx,k}$, standard (grey) and robust (red) CB for non-significance of cross-correlation at lags $k = 0, 1, ..., 10$.



(d) Standard (grey) $HB_{yx,m}$ and robust (red) $\widetilde{Q}_{yx,m}$ cumulative tests at lags $m = 0, 1, ..., 30$.

FIGURE 8. FTSE100 (x) and gold (y) daily returns, 2008–2018: Cross-correlogram.

cross-correlograms $\widehat{\rho}_{xy,k}$ and $\widehat{\rho}_{yx,k}$ for lags $k = 0, 1, \ldots, 10$ along with the standard 95% and 99% confidence bands, based on $\pm z_{\alpha/2}/\sqrt{n}$ and the robust ones based on $\pm z_{\alpha/2}(\widehat{\rho}_{xy,k}/\widetilde{t}_{xy,k})$ and $\pm z_{\alpha/2}(\widehat{\rho}_{yx,k}/\widetilde{t}_{yx,k})$ at significance levels $\alpha = 5\%, 1\%$. In panels (b) and (d), we report the standard $HB_{xy,m}, HB_{yx,m}$ and robust $\widetilde{Q}_{xy,m}, \widetilde{Q}_{yx,m}$ cumulative statistics with threshold $\lambda = 2.576$ for lags $m = 0, 1, \ldots, 30$ and their asymptotic critical values at the $\alpha = 5\%, 1\%$ levels. Standard inference suggests evidence of significant contemporaneous cross-correlation $\rho_{xy,0}$ at 1%, as well as cross-correlation $\rho_{xy,k}$ for lags $k = 2, 8$ at 1% and $\rho_{yx,k}$ for leads $k = 5$ at 1% and for $k = 8, 9$ at 5% between FTSE100 and gold returns, x_t and y_t. The cumulative hypothesis of zero cross-correlation is rejected at 1% for all m by standard tests $HB_{xy,m}$, $HB_{yx,m}$. However, the robust CBs do not produce as much evidence of significant cross-correlation. We do find evidence from the robust tests of contemporaneous cross-correlation $\rho_{xy,0}$ for $k = 0$ at 5%, as well as cross-correlation $\rho_{xy,k}$ for lags $k = 2$ at 5% and $k = 8$ at 1% and $\rho_{yx,k}$ for lead $k = 5$ at 5% between FTSE100 and gold returns, x_t and y_t. Subsequently, using modified test statistics the cumulative hypothesis of zero cross-correlation is rejected when

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

$m = 0$ at 5%. Furthermore, when the FTSE100 return x_t is leading, the hypothesis of zero cross-correlation is rejected at 5% for some m and, when it is lagging, the hypothesis of zero cross-correlation is not rejected at 5% for all m. Overall, these results indicate evidence that the daily returns of the FTSE100 index and gold during 2008–2018 have contemporaneous cross-correlation at $k = 0$ but not at different leads and lags for $k \geq 1$.

7. CONCLUSIONS

The procedures developed in this paper belong to a class of econometric tests that robustify existing procedures to take account of realistic features of economic and financial data. Tests for zero autocorrelation and zero cross-correlation are among the fundamental starting points in analyzing time series and they are methods that have remained in common use since influential work by Box and Jenkins (1970) and others. The validity of standard procedures of testing is fragile to latent dependencies and nonstationarities that are well known to be present in much economic and financial data. The methods and limit theory in the present paper correct for such fragilities and in doing so complement and generalize earlier work to accommodate such dependencies. The Monte Carlo experiments corroborate the validity of the proposed methods and provide guidelines for practitioners in implementing the new procedures. The empirical application to financial return data demonstrates the utility of these methods in taking account of latent dependencies and thereby avoiding potentially spurious inferences about autocorrelations and cross-correlations in such data. In subsequent work, we plan to adapt the test procedures developed in this paper to models that involve an evolving mean function and a stochastic heterogeneity factor h_t. We plan to show that robust testing procedures remain valid for scaling factors h_t that do not satisfy the specific smoothness condition (7). Such results naturally involve a more complex correlation matrix R which depends on $\{h_t\}$.

SUPPLEMENTARY MATERIAL

Dalla, V., Giraitis, L. and Phillips, P.C.B. (2020): Supplements I and II to *Robust Tests for White Noise and Cross-Correlation*, Econometric Theory Supplementary Material. To view, please visit: https://doi.org/10.1017/S0266466620000341

REFERENCES

Anderson, T.W. & A.M. Walker (1964) On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process. *The Annals of Mathematical Statistics* 35, 1296–1303.

Bai, J. & P. Perron (1998) Estimating and testing linear models with multiple structural changes. *Econometrica* 66, 47–78.

Bartlett, M.S. (1946) On the theoretical specification and sampling properties of autocorrelated time-series. *Supplement to the Journal of the Royal Statistical Society* 9, 27–41.

Box, G.E.P. & G.M. Jenkins (1970) *Time Series Analysis: Forecasting and Control*. Holden-Day.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

Box, G.E.P. & D.A. Pierce (1970) Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association* 65, 1509–1526.
Campbell, J.Y., A.W. Lo, & A.C. MacKinlay (1997) *The Econometrics of Financial Markets*. Princeton University Press.
Cumby, R.E. & J. Huizinga (1992) Testing the autocorrelation structure of disturbances in ordinary least squares and instrumental variables regressions. *Econometrica* 60, 185–196.
Diebold, F.X. (1986) Testing for serial correlation in the presence of ARCH. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, pp. 323–328.
Francq, C., R. Roy, & J.-M. Zakoïan (2005) Diagnostic checking in ARMA models with uncorrelated errors. *Journal of the American Statistical Association* 100, 532–544.
Francq, C. & J.-M. Zakoïan (2009) Bartlett's formula for a general class of nonlinear processes. *Journal of Time Series Analysis* 30, 449–465.
Granger, C.W.J. & A.P. Andersen (1978) *An Introduction to Bilinear Time Series Models*. Vandenhoeck and Ruprecht.
Gretton, A. & L. Györfi (2010) Consistent nonparametric tests of independence. *Journal of Machine Learning Research* 11, 1391–1423.
Guo, B. & P.C.B. Phillips (2001) Testing for autocorrelation and unit roots in the presence of conditional heteroskedasticity of unknown form. UC Santa Cruz Economics Working Paper 540, pp. 1–55.
Hannan, E.J. (1970) *Multiple Time Series.* Wiley.
Hannan, E.J. & C.C. Hyde (1972) On limit theorems for quadratic functions of discrete time series. *The Annals of Mathematical Statistics* 43, 2058–2066.
Haugh, L.D. (1976) Checking the independence of two covariance-stationary time series: A univariate residual cross-correlation approach. *Journal of the American Statistical Association* 71, 378–385.
Haugh, L.D. & G.E.P. Box (1977) Identification of dynamic regression (distributed lag) models connecting two time series. *Journal of the American Statistical Association* 72, 121–130.
Hooker, R.H. (1901) Correlation of the marriage-rate with trade. *Journal of the Royal Statistical Society* 64, 485–492.
Horowitz, J.L., I.N. Lobato, & N.I. Savin (2006) Bootstrapping the Box-Pierce *Q* test: A robust test of uncorrelatedness. *Journal of Econometrics* 133, 841–862.
Kokoszka, P.S. & D.N. Politis (2011) Nonlinearity of ARCH and stochastic volatility models and Bartlett's formula. *Probability and Mathematical Statistics* 31, 47–59.
Kyriazidou, E. (1998) Testing for serial correlation in multivariate regression models. *Journal of Econometrics* 86, 193–220.
Ljung, G.M. & G.E.P. Box (1978) On a measure of lack of fit in time series models. *Biometrika* 65, 297–303.
Lo, A.W. & A.G. MacKinlay (1989) The size and power of the variance ratio test in finite samples. A Monte Carlo investigation. *Journal of Econometrics* 40, 203–238.
Lobato, I.N., J.C. Nankervis, & N.E. Savin (2002) Testing for zero autocorrelation in the presence of statistical dependence. *Econometric Theory* 18, 730–743.
McLeod, A.I. & W.K. Li (1983) Diagnostic checking ARMA time series models using squared residual autocorrelations. *Journal of Time Series Analysis* 4, 269–273.
Phillips, P.C. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.
Robinson, P.M. (1991) Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics* 47, 67–84.
Romano, J.P. & L.R. Thombs (1996) Inference for autocorrelations under weak assumptions. *Journal of the American Statistical Association* 91, 590–600.
Stout, W. (1974) *Almost Sure Convergence*. Academic Press.
Taylor, S.J. (1984) Estimating the variances of autocorrelations calculated from financial time series. *Journal of the Royal Statistical Society, Series C* 33, 300–308.
Tsay, R.S. (2010) *Analysis of Financial Time Series, 3rd edition*. Wiley.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press

Wong, H. & S. Ling (2005) Mixed portmanteau tests for time-series models. *Journal of Time Series Analysis* 26, 569—579.

Yule, G.U. (1921) On the time-correlation problem, with especial reference to the variate-difference correlation method. *Journal of the Royal Statistical Society* 84, 497–537.

Yule, G.U. (1926) Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series. *Journal of the Royal Statistical Society* 89, 1–63.

Zhu, K. (2013) A mixed portmanteau test for ARMA-GARCH models by the quasi-maximum exponential likelihood estimation approach. *Journal of Time Series Analysis* 34, 230–237.

https://doi.org/10.1017/S0266466620000341 Published online by Cambridge University Press