Econometric Theory, **18**, 2002, 730–743. Printed in the United States of America.
DOI: 10.1017.S0266466602183083

TESTING FOR ZERO AUTOCORRELATION IN THE PRESENCE OF STATISTICAL DEPENDENCE

I.N. LOBATO
Instituto Tecnológico Autónomo de México

JOHN C. NANKERVIS
University of Surrey

N.E. SAVIN
University of Iowa

The problem addressed in this paper is to test the null hypothesis that a time series process is uncorrelated up to lag K in the presence of statistical dependence. We propose an extension of the Box–Pierce Q-test that is asymptotically distributed as chi-square when the null is true for a very general class of dependent processes that includes non-martingale difference sequences. The test is based on a consistent estimator of the asymptotic covariance matrix of the sample autocorrelations under the null. The finite sample performance of this extension is investigated in a Monte Carlo study.

1. INTRODUCTION

Box and Pierce (1970) proposed using the Q_K statistic to test the null hypothesis that the first K autocorrelations of a covariance stationary time series are zero. The Q_K statistic is the sample size times the sum of the squares of the first K sample autocorrelations. Assuming the observations are independent and identically distributed (i.i.d.), the asymptotic covariance matrix of the vector of sample autocorrelations is the inverse of the sample size times the identity matrix. Hence, under the null, Q_K is asymptotically distributed as chi-square with K degrees of freedom, *provided* that the observations are i.i.d. The test can be extended to settings with statistical dependence by using the true asymptotic covariance matrix of the sample autocorrelations, or a consistent estimator, in place of the identity matrix.

We are grateful for the helpful comments by the editor Peter Phillips and four anonymous referees. In addition, we thank David Bates, Kung-Sik Chan, Tom George, Olan Henry, Joel Horowitz, Paul Weller, and seminar participants at Monash University for their suggestions. John Nankervis gratefully acknowledges financial support from the ESRC through Research grant R000222581. Address correspondence to: N.E. Savin, Department of Economics, 108 Pappajohn Business Building, University of Iowa, Iowa City, IA 52242-1000, USA; e-mail: gene-savin@uiowa.edu.

 © 2002 Cambridge University Press 0266-4666/02 $9.50

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

This extension has been carried out for time series generated by a martingale difference sequence (MDS). For certain MDS processes, the asymptotic covariance matrix of the sample autocorrelations is diagonal. This case has been repeatedly addressed in the literature; see, for example, Taylor (1984), Diebold (1986), and Lo and MacKinlay (1989). Guo and Phillips (1998) have extended the test for the case where the asymptotic covariance matrix of the sample autocorrelations may be nondiagonal. This more general MDS setting is of special importance for financial time series. For example, the asymptotic covariance matrix of the sample autocorrelations is nondiagonal for a GARCH (1,1) model with asymmetric errors.

In this paper, we present an extension of the Box–Pierce test to the case where the time series may be generated by a non-MDS process. Our test statistic is in the spirit of a Lagrange multiplier or score statistic because it is based on a consistent estimator of the asymptotic covariance matrix of the sample autocorrelations under the null. Uncorrelated time series can be generated by non-MDS models, including certain bilinear models (Granger and Teräsvirta, 1993) and all-pass models (Breidt, Davis, and Trindade, 1999). An all-pass model is an autoregressive moving average (ARMA) model in which all the roots of the AR polynomial are reciprocals of the roots of the MA polynomial. Both bilinear models and all-pass models have been used in finance applications; for example, see Bera and Higgins (1997) and Terdik (1999) for bilinear examples and Breidt et al. (1999) for all-pass examples.

The performance of the standard Box–Pierce Q_K-test and extensions of the Q_K-test are compared in Monte Carlo experiments. The examples used in the experiments include MDS processes with diagonal and nondiagonal asymptotic covariance matrices of the sample autocorrelations and non-MDS processes. The powers of the various extensions of the Q_K-test are also investigated.

This paper is organized as follows. The notation and assumptions are introduced in Section 2, and our extension of the Q_K-test is developed in Section 3. The results of the Monte Carlo experiments on the probability of making a type I error are reported in Section 4 for MDS and non-MDS examples. The powers of the tests are reported in Section 5. The concluding remarks are in Section 6.

2. PRELIMINARIES

Let $y_1, \ldots, y_n$, denote a real-valued covariance stationary time series that has mean μ and a finite and square integrable spectrum. Define the lag-j autocovariance by $\gamma(j) = E(y_t - \mu)(y_{t-j} - \mu)$ and the lag-j autocorrelation by $\rho(j) = \gamma(j)/\gamma(0)$. Define the sample mean by $\hat{\mu} = (1/n)\sum_{t=1}^{n} y_t$ and the lag-j autocovariance by

$$\hat{\gamma}(j) = \sum_{t=j+1}^{n}(y_t - \hat{\mu})(y_{t-j} - \hat{\mu})/n.$$

The analog estimator of $\rho(j)$ is $r(j) = \hat{\gamma}(j)/\hat{\gamma}(0)$.

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

Denote the vector of sample autocovariances as $\hat{\gamma} = (\hat{\gamma}(0), \hat{\gamma}(1), \ldots, \hat{\gamma}(K))'$ and the vector of population autocovariances as $\gamma = (\gamma(0), \gamma(1), \ldots, \gamma(K))'$. The vector of sample autocorrelations is $r = (r(1), \ldots, r(K))'$, and the vector of population autocorrelations is $\rho = (\rho(1), \ldots, \rho(K))'$. The vector $w_t = (w_{1t}, \ldots, w_{Kt})'$ has as its kth component $w_{kt} = (y_t - \mu)(y_{t-k} - \mu)$ for $k = 1, \ldots K$, and $\hat{w}_t = (\hat{w}_{1t}, \ldots, \hat{w}_{Kt})'$ has as its kth component $\hat{w}_{kt} = (y_t - \hat{\mu})(y_{t-k} - \hat{\mu})$ for $k = 1, \ldots, K$.

We assume that y_t is a weak dependent process for which the vector of sample autocovariances $\hat{\gamma}$ satisfies the following central limit theorem (CLT):

$$\sqrt{n}(\hat{\gamma} - \gamma) \Rightarrow N(0, C),$$

where C (which is assumed to be finite and positive definite) is 2π times the spectral density matrix at zero frequency of the vector w_t and has as its ijth element

$$c_{ij} = \sum_{d=-\infty}^{d=\infty} \{E(y_t - \mu)(y_{t-i} - \mu)(y_{t+d} - \mu)(y_{t+d-j} - \mu) - E(y_t - \mu)(y_{t-i} - \mu)E(y_{t+d} - \mu)(y_{t+d-j} - \mu)\}; \quad i, j = 0, 1, \ldots, K.$$

A straightforward application of the delta method leads, in turn, to a CLT for the sample autocorrelations:

$$\sqrt{n}(r - \rho) \Rightarrow N(0, V),$$

where the ijth element of V is given by

$$v_{ij} = \gamma(0)^{-2}[c_{ij} - \rho(i)c_{0j} - \rho(j)c_{0i} + \rho(i)\rho(j)c_{00}]. \tag{1}$$

Remarks.

(a) The preceding CLT for the sample autocovariances has been derived under a variety of weak dependence assumptions. For instance, Hannan and Heyde (1972) establish it for linear processes with MDS innovations, whereas Romano and Thombs (1996) derive it for processes satisfying a strong mixing condition. For a comparison of both approaches, see Phillips and Solo (1992) and Davidson (2000). Another interesting characterization of weak dependence is given by the idea of near epoch dependence (NED) on a mixing set. This concept has received much attention lately because it has been found to be particularly suitable for handling dependence in nonlinear models. In fact, Davidson (2000) has proved that many nonlinear time series models, such as the bilinear model employed in this paper, satisfy the NED condition.

(b) Notice that for simplicity we have assumed covariance stationarity, but this assumption could be relaxed because the three previously mentioned characterizations of weak dependence allow for some forms of heterogeneity. In this case, the null hypothesis would be that the lag-j autocorrelations are zero for all t and $j = 1, \ldots, K$.

We close this section by presenting a specific assumption under which the preceding CLT for $\hat{\gamma}$ for holds; see De Jong and Davidson (2000).

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

ASSUMPTION 1. Let y_t be a covariance stationary process that satisfies $E|y_t|^s < \infty$ for some $s > 4$ and all t and is L_2-NED of size $-\frac{1}{2}$ on a process U_t where U_t is an α-mixing sequence of size $-s/(s-4)$.

3. TEST STATISTICS

Assuming V is known, H_K: $\rho = 0$ can be tested using a test statistic of the form $nr'V^{-1}r$, which asymptotically follows a chi-square with K degrees of freedom when H_K is true. A feasible test can be obtained either by replacing V by a known matrix or by estimating V. The Box–Pierce Q_K statistic replaces V with the identity matrix.

The spirit of the Lagrange multiplier (LM), or score test, is to estimate V by exploiting the restrictions imposed by the null. Under H_K, the matrix V simplifies to $\tilde{V} = \{\gamma(0)^{-2}\tilde{C}\}$ where $\tilde{C}$ has as its ijth element

$$\tilde{c}_{ij} = \sum_{d=-\infty}^{d=\infty} E(y_t - \mu)(y_{t-i} - \mu)(y_{t+d} - \mu)(y_{t+d-j} - \mu); \qquad i,j = 1,\ldots,K. \tag{2}$$

The extension of the Q_K test we propose is based on the statistic $\tilde{Q}_K = nr'\hat{\tilde{V}}^{-1}r$ where $\hat{\tilde{V}}$ is a consistent estimator of $\tilde{V}$ under H_K. As will be shown subsequently, $\tilde{Q}_K$ is asymptotically chi-squared distributed with K degrees of freedom when H_K is true.

An important case in economics and finance is where the time series is a MDS. For a MDS process, the only possible nonzero elements of $\tilde{C}$ are terms of the form $E(y_t - \mu)^2(y_{t-i} - \mu)(y_{t-j} - \mu)$. In (2) these occur at $d = 0$. Guo and Phillips (1998, Theorem 5) have developed a test of H_K, the GP_K-test, for the MDS case. In our context, the GP_K-test is a special case of the $\tilde{Q}_K$-test where $\tilde{c}_{ij}$ is replaced by the sample analog of $E(y_t - \mu)^2(y_{t-i} - \mu)(y_{t-j} - \mu)$.

The $\tilde{Q}_K$-test can be specialized further by assuming that $\tilde{c}_{ij} = 0$ except when $d = 0$ and $i = j$, in which case $\tilde{V}$ is diagonal and is denoted by V^*. In this diagonal case,

$$c^*_{jj} = E(y_t - \mu)^2(y_{t-j} - \mu)^2, \qquad j = 1,\ldots,K, \tag{3}$$

and the diagonal elements of V^* are

$$v^*_{jj} = c^*_{jj}/\gamma(0)^2.$$

Following Lobato, Nankervis, and Savin (2001), this modified Box–Pierce statistic is denoted by Q^*_K. It is constructed by replacing the diagonal matrix V^* by a consistent estimator $\hat{V}^*$:

$$Q^*_K = nr'[\hat{V}^*]^{-1}r = n\sum_{j=1}^{K}[r(j)]^2/\hat{v}^*_{jj}, \tag{4}$$

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

where $\hat{v}_{jj}^*$ is a consistent estimator of v_{jj}^*:

$$\hat{v}_{jj}^* = \left[n^{-1} \sum_{t=j+1}^{n} (y_t - \hat{\mu})^2 (y_{t-j} - \hat{\mu})^2 / \hat{\gamma}(0)^2 \right].$$

For the hypothesis that a single autocorrelation coefficient is zero, the Q_1^*-test is equivalent to the GP_1-test.

A consistent estimator of $\tilde{V}$ is required to implement the general $\tilde{Q}_K$-test. A consistent estimator can be obtained using $\hat{\gamma}(0)$ to estimate $\gamma(0)$ and a nonparametric estimator of the matrix $\tilde{C}$. Because the matrix $\tilde{C}$ is the spectral density matrix at zero frequency of the K-dimensional vector process w_t under H_K, a nonparametric consistent time domain estimator of $\tilde{C}$ is given by

$$\hat{\tilde{C}} = \sum_j k\left(\frac{j}{\ell}\right) g(j) = \frac{1}{n} \sum_j \sum_t k\left(\frac{j}{\ell}\right) \hat{w}_t \hat{w}_{t-j}', \tag{5}$$

where $g(j) = (1/n) \sum_t \hat{w}_t \hat{w}_{t-j}'$ with $\hat{w}_t$ defined as previously, $\ell > 0$ is the bandwidth parameter, and $k(\bullet)$ is the kernel or lag window.

We assume that the kernel and the bandwidth satisfy the following assumptions.

ASSUMPTION 2. The kernel $k(\bullet)$ belongs to K where K is the class of functions $K = \{k(\bullet):\mathbb{R}; \to [-1,1]\}$ that is symmetric around zero, continuous at zero at all but a finite number of points, and satisfies

$$k(0) = 1, \int_{-\infty}^{\infty} |k(x)| dx < \infty, \int_{-\infty}^{\infty} |\psi(\xi)| d\xi < \infty,$$

where $\psi(\xi) = (2\pi)^{-1} \int_{-\infty}^{\infty} k(x) e^{i\xi x} dx$.

ASSUMPTION 3. The bandwidth sequence satisfies

$$\lim_{n\to\infty} \left(\frac{1}{\ell} + \frac{\ell}{n} \right) = 0.$$

The following lemma (whose proof is a straightforward application of Theorem 2.1 in Davidson and De Jong, 2000) establishes the consistency of the nonparametric estimator of $\tilde{C}$.

LEMMA 1. *Under Assumptions 1–3 and the null hypothesis* H_K, $\hat{\tilde{C}} \xrightarrow{p} \tilde{C}$.

Finally, the next lemma establishes the asymptotic properties of the $\tilde{Q}_K$-test, namely, its null asymptotic distribution and its consistency.

LEMMA 2. *Under Assumptions 1–3 and under the null hypothesis* H_K, *the test statistic* $\tilde{Q}_K$ *converges in distribution to a chi-square distribution with* K *degrees of freedom, and under* $\rho \neq 0$, $\tilde{Q}_K$ *diverges.*

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

Remark. The first part of Lemma 2 is obvious because $\hat{\tilde{V}}$ is consistent for $\tilde{V}$, that is, the covariance matrix of the sample autocorrelations r under H_K. In the second part, using the ergodic theorem, the estimator of $\tilde{V}$ converges in probability to a positive definite matrix, and at least one element of r is $O_P(1)$. Therefore, $\tilde{Q}_K$ diverges.

4. MONTE CARLO EXPERIMENTS UNDER THE NULL

This section reports estimates of the probability of making a type I error for the Q_K-, Q_K^*-, GP_K-, and $\tilde{Q}_K$-tests based on critical values from the chi-square distribution with K degrees of freedom. The estimated rejection probabilities are computed for two MDS examples and two non-MDS examples. Three different hypotheses are considered: $H_K: \rho(1) = \cdots = \rho(K) = 0$, $K = 1, 5, 10$. Each hypothesis is tested at nominal levels 0.01, 0.05, and 0.10. The estimates (empirical rejection probabilities) are calculated using 25,000 replications for sample sizes $n = 200$, 1,000, and 5,000. In the tables, an asterisk denotes that the empirical rejection probability is significantly different at the 0.01 level from the nominal rejection probability, where the significance is evaluated using a 0.01 level two-sided asymptotic test.

The $\tilde{Q}_K$-test requires a consistent estimator of the asymptotic covariance matrix of the sample autocorrelations. In this section, we explore the finite sample performance of two automatic data-based covariance matrix estimation procedures that have been used in the literature. The first procedure (AUTO) employs AR(1) prewhitening/recoloring on each series and selects the bandwidth using formula (2.2) of Newey and West (1994) with weights equal to one and lag truncation equal to $2(n/100)^{2/9}$. The second procedure uses an autoregressive estimator of the covariance matrix where the order of each equation in the VAR is selected automatically. Phillips and Lee (1994) employ an ARMA model in the scalar case and Den Haan and Levin (1997) a VAR in the vector case. We report results for the vector procedure (VAR) with the AIC (Akaike, 1973) and the SC (Schwarz, 1978) criteria. We set the maximum lag length as 3, 10, and 15 for sample sizes $n = 200$, 1,000, and 5,000, respectively, and, for any equation in the VAR, the same lag length is used for each element of the vector process.

4.1. MDS Examples

Monte Carlo experiments are conducted using two examples based on a GARCH (1,1) model, one with normal errors and one where the errors are centered chi-square with 3 degrees of freedom. The GARCH (1,1) model is $y_t = z_t \cdot \sigma_t$ where z_t is an i.i.d. sequence and $\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2$, where α and β are constants such that $\alpha + \beta < 1$. This condition is needed so that y_t is covariance stationary. He and Teräsvirta (1999) show that the unconditional fourth moment

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

of y_t exists for GARCH (1,1) models if and only if $\beta^2 + 2\alpha\beta\nu_2 + \alpha\nu_4 < 1$ where $\nu_i = E|z_t|^i$. We set $\omega = 0.001$, $\alpha = 0.05$, and $\beta = 0.90$. With this parameter setting, the He and Teräsvirta (1999) condition for the existence of the fourth moment of y_t is satisfied for both models we consider. Estimates from stock return data suggest that $\alpha + \beta$ is close to 1 with β also close to 1; for example, see Bera and Higgins (1997).

Example 1 (Gaussian GARCH).

$\{z_t\}$ is a sequence of i.i.d. $N(0, 1)$ random variables. For this process, $\gamma(0) = E(y_t - \mu)^2 = 0.02$, $E(y_t - \mu)^3/\gamma(0)^{3/2} = 0$, $E(y_t - \mu)^4/\gamma(0)^2 = 3.16$, and V is diagonal.

Example 2 (Chi-square (3) GARCH).

$\{z_t\}$ is a sequence of demeaned and standardized i.i.d. chi-square random variables with 3 degrees of freedom. In this case (the skewness is an estimate), $\gamma(0) = E(y_t - \mu)^2 = 0.02$, $E(y_t - \mu)^3/\gamma(0)^{3/2} = 1.72$, and $E(y_t - \mu)^4/\gamma(0)^2 = 8.27$, and V is no longer diagonal.

The empirical rejection probabilities for the GARCH (1,1) examples are reported in Table 1 for $n = 200$ and 1,000. The main facts that emerge from these results are the following: (i) At $n = 200$, the GP_K-test and the three versions of the $\tilde{Q}_K$-test tend to perform poorly, especially for H_{10}. The exceptions are the VAR (AIC) version of the $\tilde{Q}_K$-test for Gaussian GARCH and the VAR (SC) version for the chi-square (3) GARCH. (ii) At $n = 1{,}000$, the GP_K-test tends to work satisfactorily for both Gaussian and chi-square (3) GARCH and similarly for the VAR versions of the $\tilde{Q}_K$-test. AUTO tends to under-reject, especially for H_{10}. The results for the GP_K-test and the VAR (SC) version of the $\tilde{Q}_K$-test tend to be similar because the SC criterion tends to choose zero lags, and GP_K is equivalent to the VAR implementation of $\tilde{Q}_K$ when the lag lengths are all zero.

The results are not reported for $n = 5{,}000$ because at this sample size the asymptotic approximation to the finite sample distribution is accurate for the GP_K-test and the three versions of the $\tilde{Q}_K$-test. Notice that the distortions in the rejection probabilities of the Q_K-test are larger for $n = 1{,}000$ than for $n = 200$. In particular, the Q_K-test suffers from substantial overrejection for all hypotheses, especially for H_{10}.

4.2. Non-MDS Examples

This section investigates the finite sample performance of the tests for uncorrelated non-MDS processes: a bilinear model and an all-pass model. The bilinear model is described in Granger and Teräsvirta (1993) and the all-pass model in Breidt et al. (1999). Both these models can produce time series having similar properties to those of series produced by GARCH models, namely, where the autocorrelation function of the level of the series is flat but the autocorrelation function of the squares (absolute values) of the series declines

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

TABLE 1. Rejection probabilities (percentages) of tests: MDS examples

	H_1			H_5			H_{10}		
	1	5	10	1	5	10	1	5	10
GARCH (1,1) with normal errors, $n = 200$									
Q	1.4*	6.0*	11.3*	1.7*	7.0*	12.5*	1.9*	6.9*	12.4*
Q^*	1.0	4.9	10.1	1.0	4.9	9.8	1.1	5.0	9.3*
GP	1.0	4.9	10.1	0.6*	4.1*	8.9*	0.4*	3.0*	7.4*
AUTO	0.7*	4.8	10.1	0.5*	3.6*	8.1*	0.2*	1.9*	5.3*
VAR (AIC)	0.6*	4.4*	9.7	0.9	4.9	10.1	1.0	4.4*	9.4*
VAR (SC)	0.7*	4.4*	9.7	0.7*	4.4*	9.2*	0.4*	3.3*	7.7*
GARCH (1,1) with normal errors, $n = 1{,}000$									
Q	1.6*	6.5*	12.2*	2.1*	7.9*	14.4*	2.3*	8.5*	15.4
Q^*	1.0	4.9	9.9	1.0	4.8	9.9	1.1	4.8	9.7
GP	1.0	4.9	9.9	0.9	4.7	9.8	0.8	4.3*	9.2*
AUTO	1.0	4.9	9.7	0.8	4.4*	9.5	0.6*	3.6*	8.1*
VAR (AIC)	0.9	4.7	9.7	0.9	5.0	10.2	0.9	4.6	9.8
VAR (SC)	1.0	4.9	9.9	1.0	4.8	10.0	0.8	4.4*	9.4*
GARCH (1,1) with chi-square (3) errors, $n = 200$									
Q	1.9*	7.4*	13.6*	2.9*	9.2*	15.7*	3.4*	9.8	16.4*
Q^*	1.1	5.4	11.0*	1.2	5.7*	11.2*	1.5*	6.1*	11.7*
GP	1.1	5.4	11.0*	0.6*	4.5*	9.9	0.5*	3.7*	8.3*
AUTO	0.8	4.8	10.0	0.7*	3.6*	8.0*	0.5*	2.2*	5.2*
VAR (AIC)	0.6*	4.4*	10.1	1.0	4.7	10.0	1.8*	5.8*	10.9*
VAR (SC)	0.7*	4.5*	10.1	0.9	4.9	10.4	1.1	4.8	9.6
GARCH (1,1) with chi-square (3) errors, $n = 1{,}000$									
Q	2.8*	9.5	15.8*	5.7*	14.8*	22.6*	7.0*	17.0*	25.9*
Q^*	1.1	5.1	9.9	1.4	5.9*	11.2*	1.3*	5.7	10.9*
GP	1.1	5.1	9.9	1.2	5.6*	10.7	0.9	4.7	9.9
AUTO	0.8	4.8	9.8	0.8	4.4*	9.2*	0.5*	3.2*	7.5*
VAR (AIC)	0.8	4.7	9.7	0.9	4.8	10.1	0.8	4.3*	9.3*
VAR (SC)	0.9	4.8	9.8	1.2	5.5*	11.0*	1.2	5.4	11.0*

Notes: The number of replications is 25,000. An asterisk denotes that the empirical rejection probability is significantly different at the 0.01 level from the nominal rejection probability, where the significance is evaluated using a 0.01 level two-sided asymptotic test.

slowly. Neither the Q_K^*-test nor the GP_K-test is asymptotically valid for these two examples.

Example 3 (Bilinear model).

Let $y_t = z_t + bz_{t-1}y_{t-2}$ where $\{z_t\}$ is a sequence of i.i.d. $N(0, \sigma^2)$ random variables. The y_t process is uncorrelated but not independent and is covariance stationary provided that $b^2\sigma^2 < 1$. The fourth moment of this process exists if $3b^4\sigma^4 < 1$. We set $b = 0.50$ and $\sigma^2 = 1.0$. For this process, the first four

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

moments are $\mu = 0$, $\gamma(0) = E(y_t - \mu)^2 = \sigma^2/(1 - b^2\sigma^2) = 1.33$, $E(y_t - \mu)^3/\gamma(0)^{3/2} = 0$, and $E(y_t - \mu)^4/\gamma(0)^2 = 3(1 - b^4\sigma^4)/(1 - 3b^4\sigma^4) = 3.46$. See Granger and Andersen (1978) for more details with respect to this example.

Example 4 (All-Pass ARMA(1,1) model).

Let $(y_t - \mu) - \phi(y_{t-1} - \mu) = (z_t - \phi^{-1}z_{t-1})$ where $\{z_t\}$ is a sequence of i.i.d. random variables and where $\mu = 0$ and $\phi = 0.8$. The y_t process is uncorrelated but not independent if z_t is nonnormal. The fourth moment of y_t is finite provided z_t has a finite fourth moment. In our example, z_t is Student's t with 10 degrees of freedom. For this process, the first four moments are $\mu = 0$, $\gamma(0) = E(y_t - \mu)^2 = 1.95$, $E(y_t - \mu)^3/\gamma(0)^{3/2} = 0$, and $E(y_t - \mu)^4/\gamma(0)^2 = 3.44$. A special feature of this example (where the process is linear and uncorrelated) is that V is the identity matrix; see Bartlett (1946).

Table 2 reports the empirical rejection probabilities for the non-MDS examples for samples of $n = 1{,}000$ and 5,000 because the asymptotic approximation is poor at $n = 200$. For the bilinear example, the empirical rejection probabilities are given in the first and second sections of Table 2. The Q_K-test overrejects by a large margin for the three hypotheses. The Q_K^*- and GP_K-tests perform similarly for the three hypotheses; both tests substantially overreject. The AUTO version of the $\tilde{Q}_K$-test tends to be unsatisfactory. The VAR versions tend to overreject when $n = 1{,}000$, but they are satisfactory for all three hypotheses when $n = 5{,}000$.

For the all-pass ARMA (1,1) examples, sections three and four of Table 2 show that the performance of the Q_K-test is excellent. This is explained by the fact that the V matrix is the identity. The Q_K^*- and GP_K-tests perform similarly; both underreject the null. All three versions of the $\tilde{Q}_K$-test also tend to underreject. The VAR (AIC) version, however, is noticeably better than the other two versions. And it only marginally underrejects when $n = 5{,}000$.

Computing. The random number generator used in the experiments was the very long period generator RANLUX with luxury level $p = 3$; see Hamilton and James (1997). Calculations were performed on a Silicon Graphics R10000 system and on a Sun Enterprise 3000 server using double precision Fortran 77. The program used for VARHAC was a version of the program by den Haan and Levin (http://weber.ucsd.edu/~wdenhaan/varhac.html) modified to run substantially faster. To mitigate the effect of occasional extreme estimates the program was also modified, using the procedure of Andrews and Monahan (1992), by setting the minimum singular value of the recoloring matrix to be 0.005.

5. POWERS

This section reports the empirical powers of the tests in a small Monte Carlo study where the time series are generated by MA processes with uncorrelated errors.

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

TABLE 2. Rejection probabilities (percentages) of tests: Non-MDS examples

	H_1			H_5			H_{10}		
	1	5	10	1	5	10	1	5	10
				Bilinear, $n = 1{,}000$					
Q	5.6*	14.9*	22.4*	6.4*	16.3*	25.0*	4.5*	12.8*	20.6*
Q^*	2.0*	7.9*	14.3*	2.1*	7.8*	14.1*	1.7*	6.7*	12.5*
GP	2.0*	7.9*	14.3*	1.8*	7.4*	13.6*	1.2	5.8*	11.6*
AUTO	1.1	5.8*	11.5*	1.0	5.0	10.3	0.6*	4.0*	8.7*
VAR (AIC)	0.9	4.9	10.0	1.2	5.5*	11.1*	1.2	5.4	10.8*
VAR (SC)	1.2*	5.9*	11.4*	1.5*	6.6*	12.3*	1.2	5.9*	11.4*
				Bilinear, $n = 5{,}000$					
Q	6.1*	15.1*	23.0*	6.6*	17.0*	26.2*	4.7*	13.8*	22.4*
Q^*	2.1*	8.1*	14.2*	1.9*	7.9*	14.2*	1.7*	6.5*	13.3*
GP	2.1*	8.1*	14.2*	1.8*	7.7*	14.1*	1.5*	6.5*	12.8*
AUTO	1.3*	5.9*	11.2*	1.2	5.7*	10.8*	1.0	4.8	10.0
VAR (AIC)	1.0	4.9	10.0	1.1	5.2	10.4	1.1	5.3	10.4
VAR (SC)	1.0	4.7	9.7	1.1	5.3	10.4	1.2	5.9*	11.2*
				All-pass ARMA (1,1), $n = 1{,}000$					
Q	0.9	4.8	9.8	0.9	4.7	9.7	0.9	4.8	9.6
Q^*	0.7*	3.9*	8.5*	0.7*	3.8*	8.4*	0.8	4.2*	8.7*
GP	0.7*	3.9*	8.5*	0.7*	3.8*	8.3*	0.6*	3.9*	8.3*
AUTO	0.7*	4.3*	9.2*	0.6*	4.0*	8.5*	0.5*	3.6*	8.1*
VAR (AIC)	0.7*	4.4*	9.3*	0.8	4.2*	9.0*	0.8	4.4*	9.4*
VAR (SC)	0.7*	4.1*	8.7*	0.7*	4.0*	8.5*	0.6*	4.0*	8.4*
				All-pass ARMA (1,1), $n = 5{,}000$					
Q	1.1	5.1	10.0	0.9	4.8	9.6	1.0	4.9	9.8
Q^*	0.8	4.1*	8.4*	0.7*	3.9*	8.2*	0.8	4.2*	8.7*
GP	0.8	4.1*	8.4*	0.7*	3.9*	8.2*	0.7*	4.1*	8.6*
AUTO	1.0	4.7	9.4	0.8	4.4*	8.9*	0.8	4.3*	9.0*
VAR (AIC)	0.9	4.7	9.4*	0.7*	4.5*	9.2*	0.9	4.6	9.4*
VAR (SC)	0.8	4.3*	8.7*	0.7*	4.1*	8.5*	0.7*	4.1*	8.7*

Notes: The number of replications is 25,000. An asterisk denotes that the empirical rejection probability is significantly different at the 0.01 level from the nominal rejection probability, where the significance is evaluated using a 0.01 level two-sided asymptotic test.

In the first set of experiments, the time series are generated by $y_t = u_t + \theta u_{t-1}$, and, in the second, they are generated by $y_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}$. The processes generating the u_t's are those specified in Examples 1–4, and hence the u_t's are uncorrelated. The null hypothesis is imposed by setting $\theta = 0$ in the MA(1) case and $\theta_1 = \theta_2 = 0$ in the MA(2) case; with these restrictions, the y_t's are uncorrelated. The values of $\rho(1)$ range from 0.025 to 0.15. For the sake of brevity, results for the MA(2) case are presented for the case $\rho(1) = \rho(2) = \bar{\rho}$

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

TABLE 3. Empirical powers of 0.05 adjusted tests for an MA(1) process, $n = 1{,}000$

	H_1						H_5					
$\rho(1)$	.025	.050	.075	.100	.125	.150	.025	.050	.075	.100	.125	.150
GARCH (1,1) with normal errors												
Q^*	11.1	30.6	59.6	83.8	96.0	99.3	7.3	16.3	35.0	60.2	82.3	95.2
GP	11.1	30.6	59.6	83.8	96.0	99.3	7.2	16.2	35.1	60.8	83.2	95.4
AUTO	10.9	30.4	59.1	83.7	95.9	99.4	7.2	15.8	33.7	59.1	81.9	94.9
VAR (AIC)	10.9	30.4	59.0	83.3	95.7	99.2	7.1	15.5	32.8	57.2	79.9	93.7
VAR (SC)	10.9	30.6	59.4	83.7	96.0	99.4	7.4	16.3	35.0	60.4	82.8	94.9
GARCH (1,1) with chi-square (3) errors												
Q^*	7.6	23.4	50.4	78.2	94.0	99.0	5.8	11.0	24.3	46.7	72.0	90.2
GP	7.6	23.4	50.4	78.2	94.0	99.0	5.8	11.3	25.4	49.1	74.7	91.8
AUTO	7.8	23.9	51.4	78.4	93.7	98.9	5.8	11.6	25.8	48.9	73.8	90.2
VAR (AIC)	8.0	23.3	50.2	76.8	92.6	98.1	6.0	12.1	26.0	48.5	71.8	88.0
VAR (SC)	7.7	23.2	50.2	77.4	92.9	98.4	6.0	12.1	26.1	49.5	73.8	90.4
Bilinear												
Q^*	9.0	21.1	40.9	63.5	82.6	93.4	6.7	12.7	24.1	41.6	62.2	80.4
GP	9.0	21.1	40.9	63.5	82.6	93.4	6.6	12.8	24.9	43.2	64.5	82.4
AUTO	7.3	18.6	38.6	62.9	82.7	94.3	5.9	11.5	23.6	42.6	64.7	83.0
VAR (AIC)	7.1	18.0	37.6	61.8	82.1	93.8	6.0	10.9	21.5	38.8	60.4	79.9
VAR (SC)	7.1	18.0	37.5	60.8	81.1	93.2	6.3	12.3	24.5	43.9	66.3	84.2
All-pass ARMA (1,1)												
Q^*	11.9	34.4	65.6	88.2	97.8	99.8	7.7	17.7	38.7	66.0	87.3	97.3
GP	11.9	34.4	65.6	88.2	97.8	99.8	7.6	17.8	39.1	66.5	87.8	97.6
AUTO	11.6	33.8	64.7	87.7	97.7	99.8	7.7	17.6	38.1	65.2	86.7	97.1
VAR (AIC)	11.7	33.4	64.4	87.4	97.6	99.8	8.1	18.3	38.9	65.9	87.2	97.2
VAR (SC)	11.9	34.0	65.0	88.0	97.7	99.8	7.8	18.3	39.4	66.6	87.7	96.9

Notes: The number of replications is 25,000. The critical values are adjusted so that empirical rejection probabilities of the tests under the null are exactly 0.05. The empirical powers are in percentages.

where the values of $\bar{\rho}$ range from 0.010 to 0.085. The empirical powers are reported for the 0.05 tests of the null hypotheses H_1 and H_5 when $n = 1{,}000$. In the power comparisons, the critical values are adjusted so that empirical rejection probabilities of the tests under the null are exactly 0.05.

The empirical powers are reported in Table 3 for the MA(1) case and in Table 4 for the MA(2) case. The results show that the empirical powers are very similar for all the 0.05 tests, both for the MDS and non-MDS examples, when the tests are based on adjusted critical values. The empirical powers are also very similar for all the 0.01 tests and for all the 0.10 tests. However, this does not imply that the tests are equally satisfactory when they are based on chi-square critical values. The results in Section 4 show that the tests can suffer from distortions in the probability of making a type I error when they are based on chi-square critical values. This is illustrated by the substantial overrejections of the Q_K^*- and GP_K-tests in the case of the bilinear model.

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

TABLE 4. Empirical powers of 0.05 adjusted tests for an MA(2) process, $n = 1{,}000$

	H_1						H_5					
$\rho(1) = \rho(2)$	.010	.025	.040	.055	.070	.085	.010	.025	.040	.055	.070	.085
GARCH (1,1) with normal errors												
Q^*	8.6	30.3	65.4	90.6	98.7	99.9	6.2	15.8	38.9	70.6	92.1	98.9
GP	8.6	30.3	65.4	90.6	98.7	99.9	6.4	16.2	40.2	72.0	92.9	99.1
AUTO	8.6	31.0	66.3	90.9	99.0	99.9	6.4	16.1	39.5	80.0	93.3	99.0
VAR (AIC)	8.3	29.5	64.2	90.1	98.6	99.9	6.3	15.6	38.7	69.3	90.9	98.5
VAR (SC)	8.5	29.8	65.1	90.6	98.7	99.9	6.3	15.8	39.8	71.2	92.1	98.5
GARCH (1,1) with chi-square (3) errors												
Q^*	6.3	23.5	57.1	86.6	98.1	99.9	5.5	11.9	30.3	60.5	86.7	97.5
GP	6.3	23.5	57.1	86.6	98.1	99.9	5.4	12.3	31.8	62.8	88.4	98.2
AUTO	6.5	24.4	58.7	87.4	98.0	99.7	5.4	11.9	30.7	61.5	86.5	96.7
VAR (AIC)	6.2	23.3	56.0	85.0	97.0	99.3	5.7	12.6	31.3	60.1	84.2	95.1
VAR (SC)	6.2	23.2	56.3	85.5	97.3	99.4	5.5	12.5	31.5	61.0	86.4	96.7
Bilinear												
Q^*	7.5	21.2	45.6	73.2	91.0	97.9	6.1	12.6	27.5	51.0	75.6	91.8
GP	7.5	21.2	45.6	73.2	91.0	97.9	5.8	12.4	27.6	52.3	77.5	92.7
AUTO	6.1	19.0	44.5	73.5	91.8	98.5	5.5	11.7	27.4	53.4	78.5	93.7
VAR (AIC)	6.2	18.7	43.5	72.7	91.5	98.2	5.6	11.0	24.9	48.6	74.4	91.9
VAR (SC)	6.3	18.5	42.9	71.8	90.7	97.8	5.7	12.4	29.0	54.5	80.0	94.3
All-pass ARMA (1,1)												
Q^*	9.2	35.1	72.7	94.4	96.6	100	6.8	18.1	45.3	77.8	95.4	99.7
GP	9.2	35.1	72.7	94.4	96.6	100	6.7	18.1	45.6	78.4	95.8	99.7
AUTO	9.1	34.3	71.7	94.0	99.5	100	6.6	17.9	44.6	77.1	95.1	99.6
VAR (AIC)	9.1	34.1	71.4	94.0	99.5	100	7.0	18.6	45.5	77.8	95.3	99.6
VAR (SC)	9.2	34.6	72.1	94.2	99.6	100	6.8	18.6	45.9	78.3	95.2	98.9

Notes: The values of θ_1 and θ_2 are selected so that $\rho(1) = \rho(2) = \bar{\rho}$. The number of replications is 25,000. The critical values are adjusted so that empirical rejection probabilities of the tests under the null are exactly 0.05. The empirical powers are in percentages.

6. DISCUSSION

This paper presents an omnibus test of uncorrelatedness in the presence of statistical dependence. The proposed test statistic, $\tilde{Q}_K$, which can be viewed as an extension of the Box–Pierce Q_K statistic, is asymptotically chi-square distributed under the null. The finite sample performance of three automatic data-based versions of the $\tilde{Q}_K$-test are examined. The study first considers MDS examples, which are based on GARCH (1,1) models. In this case, the GP_K-test and the $\tilde{Q}_K$-test tended to provide satisfactory control over the probability of making a type I error when the sample size was $n = 1{,}000$, but control was unsatisfactory for all tests at $n = 200$. The study next considers examples based on two non-MDS models, namely, a bilinear model and an all-pass ARMA (1,1) model. As expected, the GP_K-test suffered from distortions in the type I error at $n = 1{,}000$ and 5,000. The distortions are substantial for the bilinear example and rel-

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

atively mild for the all-pass example. No version of the $\tilde{Q}_K$-test is satisfactory at $n = 1{,}000$ for both examples. The VAR (AIC) version works at $n = 5{,}000$ for the bilinear example but only marginally so for the all-pass ARMA (1,1).

The message here is that large samples are required for asymptotic theory to provide a reasonable approximation to the distribution of the test statistics when the null is true. The $\tilde{Q}_K$-tests are useful in many applications involving financial time series, but it does not appear that the same is true for most macroeconomic time series as a result of sample size limitations.

Finally, the results of a small power study using adjusted critical values show that the empirical powers are very similar for all the tests, both for the MDS and the non-MDS examples. In practice, the tests are based on chi-square critical values, not adjusted critical values. The tests differ in how well they control the probability of making a type I error when they are based on chi-square critical values. The argument in favor of the $\tilde{Q}_K$-tests is that they provide satisfactory control over the probability of making a type I error in both MDS and non-MDS settings, at least asymptotically.

There are several topics that merit further research. In many applications, it is of interest to test for zero autocorrelation in the errors of a regression model. Thus, one line of research is to investigate how the results in this paper are affected when the tests are based on estimated residuals. Another is to extend the results to multivariate time series.

REFERENCES

Akaike, H. (1973) Information theory and the extension of the maximum likelihood principle. In B. Petrov & F. Csaki (eds.), *Second International Symposium on Information Theory*, pp. 267–281. Budapest: Akailseoniai-Kiudo.

Andrews, D.W.K. & J.C. Monahan (1992) An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* 60, 953–966.

Bartlett, M.S. (1946) On the theoretical specification and sampling properties of autocorrelated series. *Journal of the Royal Statistical Society*, supplement 8, 27–41, 85–97.

Bera, A.K. & M.L. Higgins (1997) ARCH and bilinearity as competing models for nonlinear dependence. *Journal of Business and Economic Statistics* 15, 43–51.

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.

Breidt, F.J., R.A. Davis, & A. Trindade (1999) Least Absolute Deviation Estimation for All-Pass Time Series Models. Preprint, Colorado State University.

Davidson, J. (2000) When is a Time Series $I(0)$? Evaluating the Memory Properties of Nonlinear Dynamic Models. Preprint, Cardiff University, available at http://www.cf.ac.uk/carbs/econ/davidsonje/research.html.

Davidson, J. & R.M. De Jong (2000) Consistency of kernel estimators of heteroscedastic and autocorrelation covariance matrices. *Econometrica* 68, 407–424.

De Jong, R.M. & J. Davidson (2000) The functional central limit theorem and weak convergence to stochastic integrals, part 1: Weakly dependent processes. *Econometric Theory* 16, 621–642.

Den Haan, W.J. & A. Levin (1997) A practitioner's guide to robust covariance matrix estimation. In G. S. Maddala & C.R. Rao (eds.), *Handbook of Statistics: Robust Inference*, vol. 15, pp. 291–341. Amsterdam: North-Holland.

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press

Diebold, F.X. (1986) Testing for serial correlation in the presence of heteroskedasticity. In *Proceedings of the American Statistical Association, Business and Economics Statistics Section*, pp. 323–328.

Granger, C.W.J. & A.P. Andersen (1978) *An Introduction to Bilinear Time Series Models*. Gottingen: Vanenhoek and Ruprecht.

Granger, C.W.J. & T. Teräsvirta (1993) *Modelling Nonlinear Economic Relationships*. Oxford: University of Oxford Press.

Guo, B.B. & P.C.B. Phillips (1998) Testing for Autocorrelation and Unit Roots in the Presence of Conditional Heteroskedasticity of Unknown Form. Cowles Foundation for Research in Economics, Yale University.

Hamilton, K.G. & F. James (1997) Acceleration of RANLUX. *Computer Physics Communications* 101, 241–248.

Hannan, E.J. & C.C. Heyde (1972) On limit theorems for quadratic functions of discrete time series. *Annals of Mathematical Statistics* 43, 2058–2066.

He, C. & T. Teräsvirta (1999) Properties of moments of a family of GARCH processes. *Journal of Econometrics* 92, 173–192.

Lo, A.W. & A.C. 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., J.C. Nankervis, & N.E. Savin (2001) Testing for autocorrelation using a modified Box–Pierce *Q* test. *International Economic Review* 42, 187–205.

Newey, W.K. & K.D. West (1994) Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61, 631–653.

Phillips, P.C.B. & C.C. Lee (1994) An ARMA-Prefiltered Estimator of the Long Run Variance with an Application to Unit Root Test. Preprint, Yale University.

Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.

Romano, J.L. & L.A. Thombs (1996) Inference for autocorrelations under weak assumptions. *Journal of the American Statistical Association* 91, 590–600.

Schwarz, G. (1978) Estimating the dimension of a model. *Annals of Statistics* 6, 461–464.

Taylor, S. (1984) Estimating the variances of autocorrelations calculated from financial time series. *Applied Statistics* 33, 300–308.

Terdik, G. (1999*) Bilinear Stochastic Models and Related Problems of Nonlinear Time Series Analysis*. New York: Springer-Verlag.

https://doi.org/10.1017/S0266466602183083 Published online by Cambridge University Press