Testing non-linearity in financial returns using a modified Q test

Marián Vávra*
Advisor to the Governor and Research Department
National Bank of Slovakia

Zacharias Psaradakis
Department of Economics, Mathematics, and Statistics
University of London

June 3, 2013

JEL classification: C12, C15, C32, C46.
Key words: non-linearity testing, portmanteau Q test, auto-correlation, cross-correlation.

Abstract

A new version of the Q test, based on generalized residual correlations (i.e. auto-correlations and cross-correlations), is developed in this paper. The Q test fixes two main shortcomings of the Mcleod and Li Q (MLQ) test often used in the literature: (i) the test is capable to capture some interesting non-linear models, for which the original MLQ test completely fails (e.g. a non-linear moving average model). Additionally, the Q test significantly improves the power for some other non-linear models (e.g. a threshold moving average model), for which the original MLQ test does not work very well; (ii) the new Q test can be used for discrimination between simple and more complicated (non-linear/asymmetric) GARCH models as well.

1 Introduction

Testing for non-linearity in economic time series is of much practical importance in applied time series analysis. One of the most popular non-linearity tests is probably

*Correspondence to: Marián Vávra, National Bank of Slovakia, Imricha Karvasa 1, 813 25 Bratislava, Slovakia, E-mail: marian.vavra@nbs.sk.

the Mcleod-Li Q test (MLQ) developed by McLeod and Li (1983). The test is based on inspecting the correlation structure of squared residuals. Its popularity comes from the fact that the test is very intuitive and easy to calculate as compared to some other non-linearity tests, which require "deep" knowledge for a correct application.[1] Unfortunately, as shown in Vávra (2012), the MLQ test suffers from several shortcomings: (i) The MLQ test cannot detect some interesting non-linear models (e.g. a NLMA model); (ii) The MLQ test exhibits a relatively low power for some commonly applied non-linear time series models (e.g. a TAR model or a TMA model); (iii) Although the test is originally constructed as a conditional heterescedasticity test, it lacks the discrimination power against more advanced GARCH models (e.g. non-linear and/or asymmetric GARCH models) developed recently in the literature; (iv) Finally, the MLQ test suffers from the relatively high sensitivity to the configuration of model parameters and the distribution of innovations, see Vávra (2012) for Monte Carlo evidence.

The main task of this paper is to show that a new version of the Q test, based on generalized residual correlations (i.e. auto-correlations and cross-correlations), can easily fix the aforementioned shortcomings. The idea of using generalized correlations is not entirely new in the literature. For instance, Lawrance and Lewis (1985, 1987) analytically demonstrated the possible usefulness of cross-correlations for detecting non-linearity in time series analysis. However, they only use very specific models (e.g. a random coefficient model), for which derivation of a cross-correlation structure is analytically tractable. What is more, they focused only on inspecting individual cross-correlations, whereas this paper focuses on the portmanteau form of the test. A more efficient variance-stabilizing transformation for the Q test is implemented as well, which improves its properties.

This paper is organized as follows. Three Q tests are discussed in Section 2. A description of non-linear models and Monte Carlo setup are presented in Section 4. Finally, the results of an extensive Monte Carlo analysis are presented in Section 5. Section 6 is devoted to an empirical application of the proposed Q tests.

[1] A nice example of such a test is the neural network (NN) test developed by White (1989), where the user needs to specify the following test quantities: (a) the number of squashing functions; (b) the functional form of squashing functions; (c) a particular distribution for parameters of squashing functions; (d) the number of eliminated principal components. Even if the user sets all the parameters correctly, the power results of the NN test are affected by randomization of the parameters of squashing functions. Loosely speaking, the different application of the NN test, the different results even though the same time series is inspected.

2 Portmanteau tests

The idea of inspecting the auto-correlation structure as a tool for detecting non-linearity in time series analysis dates back to the influential work of Granger and Andersen (1978). They show that, provided that $\{X_t : t \in \mathbb{Z}\}$ is a sequence of a linear Gaussian stationary process, it holds that

$$\rho_k(X_t^2) = \rho_k^2(X_t), \quad \text{for} \quad k \in \mathbb{Z},$$

where ρ_k denotes the k-th theoretical auto-correlation coefficient. A simple proof of this relationship can be found in Maravall (1983, p. 69). A departure from the above result might indicate some form of non-linearity and/or non-normality.

Before we proceed to a testing procedure, we state an important assumption about a stochastic process under consideration. The assumption is of the crucial importance for setting the null hypothesis of linearity and for the derivation of a limiting distribution of the test statistic.

Assumption 1 *Let us assume $\{X_t : t \in \mathbb{Z}\}$ is a zero-mean real-valued finite-order ARMA(p,q) model given by*

$$X_t = \xi_1 X_{t-1} + \cdots + \xi_p X_{t-p} + \zeta_1 a_{t-1} + \cdots + \zeta_q a_{t-q} + a_t, \tag{1}$$

where $\{a_t : t \in \mathbb{Z}\}$ is a sequence of IID(0,σ^2) model innovations such that $\mathbb{E}(|a_t|^8) < \infty$. Let $\boldsymbol{\beta} = (\xi_1, \ldots, \xi_p, \zeta_1, \ldots, \zeta_q, \sigma)'$ be a $(p+q+1 \times 1)$ parameter vector, which is assumed to be in the interior of the parameter space

$$\begin{aligned}\mathbf{B} = \{\boldsymbol{\beta} \in \mathbb{R}^{p+q} \times \mathbb{R}_{++} : \xi(z) &= 1 - \sum_{i=1}^{p} \xi_i z^i \neq 0 \quad \textit{for all} \quad |z| \leq 1,\\ \zeta(z) &= 1 - \sum_{i=1}^{q} \zeta_i z^i \neq 0 \quad \textit{for all} \quad |z| \leq 1,\\ \xi(z) \quad &\textit{and} \quad \zeta(z) \quad \textit{have no root in common}\}\end{aligned}$$

□

Provided that all conditions of Assumption 1 are satisfied, then a given stochastic process $\{X_t\}$ is stationary, an appropriate model is identified and the true parameter vector $\boldsymbol{\beta}$ does not lie on the boundary of the parameter space $\mathbf{B}$. All these conditions are important for obtaining consistent estimates of unknown parameters and to ensure the validity of the asymptotic properties of unknown parameters. Note that some authors, for example, Box and Pierce (1970), Li (1992), and Li and Mak (1994), follow a conventional assumption about Gaussian innovations in model (1). The advantage of this approach is that all the moment requirements are implicitly satisfied. Another advantage is that uncorrelated Gaussian innovations immediately imply their independence, which is a very convenient property for testing the null hypothesis of linearity using a portmanteau Q test. On the other hand, this assumption might be too restrictive in practice. Therefore, we follow McLeod and Li (1983) and assume IID model innovations with a particular moment restriction.[2]

The theoretical generalized correlation function is defined as

$$\rho_{rs}(k) = \frac{\gamma_{rs}(k)}{\gamma_{rs}(0)} = \frac{\mathbb{E}\left[g_r(a_t)g_s(a_{t-k})\right]}{\sqrt{\mathbb{E}\left[g_r^2(a_t)\right]\mathbb{E}\left[g_s^2(a_t)\right]}}, \tag{2}$$

where $g_r(\cdot)$ and $g_s(\cdot)$ are assumed to be real-valued zero-mean continuous functions given by

$$g_r(a) = a^r - \mathbb{E}(a^r), \quad g_s(a) = a^s - \mathbb{E}(a^s), \quad \text{for} \quad r, s \in \{1, 2\}. \tag{3}$$

The functional form of $g_r(\cdot)$ and $g_s(\cdot)$ is used to simplify expressions about covariances and variances discussed later in this chapter.

The theoretical generalized covariance term $\gamma_{rs}(k)$ is estimated as follows

$$\hat{\gamma}_{rs}(k) = \frac{1}{T}\sum_{t=k+1}^{T} g_r(\hat{a}_t)g_s(\hat{a}_{t-k}), \quad \text{for} \quad k \in \{1, \ldots, m\}, \tag{4}$$

where $\hat{a}_t$ is the estimated residual after applying a linear ARMA filter. A consistent sample analogue of the $g_r(\cdot)$ and $g_s(\cdot)$ functions is given by

$$g_r(\hat{a}_t) = \hat{a}_t^r - \frac{1}{T}\sum_{t=1}^{T}\hat{a}_t^r, \quad g_s(\hat{a}_t) = \hat{a}_t^s - \frac{1}{T}\sum_{t=1}^{T}\hat{a}_t^s, \quad \text{for} \quad r, s \in \{1, 2\}. \tag{5}$$

[2]Alternatively, the null hypothesis of linearity can be specified for innovations being a martingale difference sequence. That means innovations are assumed to be uncorrelated, but not necessarily independent. There are, however, at least two difficulties with testing the null in this form. First, it rules out some important non-linear processes such as conditional volatility models often used in finance. Second, the limiting distribution of the Q test can differ substantially from a χ^2 distribution since the variance-covariance matrix of the sample auto-correlations depends on parameters of a given data generating process. As a result, the limiting variance-covariance matrix is no longer a diagonal matrix. In these cases, two possible solutions are available. Romano and Thombs (1996) and Horowitz et al. (2006) recommend to apply a bootstrap procedure for calculating critical values of the Q test, whereas Lobato (2001) and Lobato et al. (2002) develop the Q test with the corrected variance-covariance matrix.

Recall that $\hat{\gamma}_{rs}(k)$ is slightly downward biased in small samples.[3] The sample analog of $\gamma_{rs}(0)$ takes the following form

$$\hat{\gamma}_{rs}(0) = \sqrt{\left[\frac{1}{T}\sum_{t=1}^{T} g_r^2(\hat{a}_t)\right]\left[\frac{1}{T}\sum_{t=1}^{T} g_s^2(\hat{a}_t)\right]}, \tag{6}$$

where functions $g_r(\hat{a}_t)$ and $g_s(\hat{a}_t)$ are defined in (5).

The Q test is then given by

$$Q_{rs}(m) = \sum_{k=1}^{m} (T-k)\hat{z}_{rs}(k)^2, \tag{7}$$

where $\hat{z}_{rs}(k)$ is a transformed sample generalized correlation coefficient of the form

$$\hat{z}_{rs}(k) = \frac{1}{2}\log\left(\frac{1+\hat{\rho}_{rs}(k)}{1-\hat{\rho}_{rs}(k)}\right), \tag{8}$$

where $\hat{\rho}_{rs}(k) = \hat{\gamma}_{rs}(k)/\hat{\gamma}_{rs}(0)$ is the k-th sample correlation coefficient calculated by combining (4) and (6). For analytical reasons, slightly modified versions of the above defined quantities are also used in this chapter: $\dot{\gamma}_{rs}(k)$ denotes the sample generalized covariance based on observed sequence of innovations $\{a_t : t = 1, \ldots, T\}$. Other quantities, such as $\dot{\gamma}_{rs}(0)$, $\dot{\rho}_{rs}(k)$ or $\dot{z}_{rs}(k)$, are defined analogically. It can be shown that, under the null hypothesis, $\mathbb{E}(\dot{z}_{rs}(k)) = O(T^{-1})$ and $\mathrm{var}(\dot{z}_{rs}(k)) = (T-k)^{-1} + O(T^{-2})$, see Johnson et al. (1994, Vol. 2, p. 571). This fact justifies the scaling factor $(T-k)$ used in (7). As mentioned by Anderson (2003, p. 134), an interesting property of the logarithmic transformation is that the quantity $\dot{z}$ converges to the limiting normal distribution faster than $\dot{\rho}$ in general.[4] This fact implies that, under the null hypothesis and provided that we directly observe the sequence $\{a_t : t = 1, \ldots, T\}$, $\mathrm{var}(Q_{rs}(m)) = 2\sum_{k=1}^{m}(T-k)^2[\mathbb{E}(\dot{z}_{rs}^2(k))]^2 \approx 2m$, since $\mathbb{E}(\dot{z}_{rs}^2(k)) = (T-k)^{-1} + O(T^{-2})$ and $\mathrm{cov}(\dot{z}_{rs}^2(i), \dot{z}_{rs}^2(j)) \approx 0$ for integers $i, j \in \{1, \ldots, m\}$, such that $i \neq j$. Therefore, a more complicated variance-stabilizing transformation of the Q test, as in Ljung and

[3]See also Kendall and Ord (1973, p. 79) for a textbook example. The bias, however, disappears quite quickly, see Wei (1990, p. 19).

[4]Under the null hypothesis that $z = 0$, Konishi (1978)'s approximation of the distribution of the standard correlation coefficient $\dot{z}$ takes the following form

$$\mathbb{P}\left(\sqrt{N}(\dot{z} - z) \leq x\right) = \Phi(x) - 0.5\left(\frac{x^3}{6N}\right)\phi(x) + O(N^{-3/2}),$$

where $N = T - 3/2$ and T denotes the sample size. It is easy to see that the effect of the second term disappears relatively quickly with increasing N. Recall that the above approximation is valid even if innovations are not Gaussian.

Box (1978), does not have to be considered. The efficiency of the log-transformation was also confirmed in Kwan and Sim (1996a,b) by means of Monte Carlo experiments.

The test specification captures two very well known Q tests: (i) setting $r = s = 1$ leads to a test proposed by Box and Pierce (1970) and modified by Ljung and Box (1978); (ii) setting $r = s = 2$ leads to a test proposed by McLeod and Li (1983). The only difference between our specification of the Q tests and those proposed by other authors is that we use directly a more efficient variance-stabilizing transformation. The main focus of this chapter is on the following two specifications: (i) $r = 1$ and $s = 2$; (ii) $r = 2$ and $s = 1$. Note that the Q test can be theoretically defined for any integers $r, s \in \mathbb{N}$, but for high values, extremely high moment conditions must be satisfied. For instance, for $r = s = 2$, the Q_{22} test, an analogy to the MLQ test, requires the existence of the first eight moments to have a valid limiting distribution. Yet, this is in sharp contrast with empirical findings about economic time series, for which the maximum exponent, $\kappa = \sup_{k>0} \mathbb{E}(|X_t|^k) < \infty$, usually lies between 2 and 4, see Jansen and Vries (1991), Loretan and Phillips (1994), or Runde (1997). Anderson and Walker (1964) and Anderson (1991) show that, for linear time series models, the moment condition can be further relaxed provided that one imposes a stronger restriction on the parameters of a data generating process. Generally, similar moment restrictions might be difficult to obtain for non-linear time series models. For this reason, some authors recommend the use of the Q test with correlations based on absolute residuals rather than squared ones. It can be shown that the limiting distribution of the Q test is unchanged with the solely requirement of the existence of the first four moments, which is a relatively reasonable assumption, see Pérez and Ruiz (2003) for a discussion. Moreover, Ding et al. (1993) argue that for short-memory models, auto-correlation functions of absolute and squared asset returns are similar.

Finally, we state two theorems about the limiting properties of the above discussed quantities.

Theorem 1 *Under Assumption 1, the limiting distribution of a vector of sample correlations* $\hat{\mathbf{z}}_{rs} = (\hat{z}_{rs}(1), \ldots, \hat{z}_{rs}(m))'$ *is given by*

$$\sqrt{T}(\hat{\mathbf{z}}_{rs} - \mathbf{z}_{rs}) \xrightarrow{d} N(\mathbf{0}, \mathbf{I}),$$

for integers $(r, s) \in \{(1, 2), (2, 1), (2, 2)\}$ *and some integer* $m > 0$. □

Proof. *See Section A for a proof.* ■

Theorem 2 *Under Assumption 1, the limiting distribution of the Q tests is given by*

$$Q_{rs}(m) \xrightarrow{d} \chi^2(m),$$

for integers $(r, s) \in \{(1,2),(2,1),(2,2)\}$ *and some integer* $m > 0$. □

Proof. *See Section A for a proof.* ■

3 Monte Carlo setup

The statistical properties of the proposed Q test are examined using: (i) simple linear time series models: an autoregressive (AR) model and a moving average (MA) model; (ii) seven non-linear time series models: a threshold (TAR) model, a Markov switching autoregressive (MSAR) model, generalized autoregressive conditional heteroscedasticity (GARCH) model, a non-linear autoregressive conditional heteroscedasticity (NL-GARCH) model, a bilinear (BL) model, a non-linear moving average (NLMA) model, and finally, a threshold moving average (TMA) model. Although the list of non-linear time series models is not definitely exhaustive, we are strongly convinced that all the main classes of non-linear models are included. The models are summarized in Table 1. A complete set of model parameters can be found in Table 2.

Originally, T+100 observations are generated in each experiment, but the first 100 of them are discarded in order to eliminate the effect of the initial observations. The number of repetitions of all experiments is set to $R = 1000$. In all experiments, the generated series is filtered by an AR(p) model, where the lag order p is selected by the Bayesian information criterion (BIC) developed by Schwarz (1978). Following the arguments in Ng and Perron (2005), a modified version of the information criterion is used. Ng and Perron (2005) show, based on extensive Monte Carlo experiments, that the best method to give the correct lag order is that with a fixed efficient sample size. Therefore, the selection criterion is defined as follows

$$BIC_l = \log(\hat{\sigma}_l^2) + \frac{l \log(N)}{N},$$
$$\hat{\sigma}_l^2 = \frac{1}{N} \sum_{t=L+1}^{T} \hat{a}_{lt}^2,$$

where $l \in \{1, \ldots, L\}$, $N = T - L$, T is the sample size and the maximum lag order is constraint according to $L = [8(T/100)^{0.25}]$. The lag order for an AR(p) model is estimated by the following simple rule $\hat{p} = \min_{l \in \{1,\ldots,L\}}(BIC_l)$. Finally, the sample size

Table 1: List of data generating processes

M1: ARMA models:

$$Y_t = c + \phi Y_{t-1} + \sigma a_t,$$
$$Y_t = c + \theta a_{t-1} + \sigma a_t,$$

M2: A TAR model:

$$Y_t = (c_1 + \phi_1 Y_{t-1} + \sigma_1 a_t) I(Y_{t-1} \le 0) + (c_2 + \phi_2 Y_{t-1} + \sigma_2 a_t) I(Y_{t-1} > 0),$$

M3: A MSAR model:

$$Y_t = (c_1 + \phi_1 Y_{t-1} + \sigma_1 a_t) I(S_t = 1) + (c_2 + \phi_2 Y_{t-1} + \sigma_2 a_t) I(S_t = 2),$$

M4: A GARCH model:

$$Y_t = c + \phi Y_{t-1} + \epsilon_t, \quad \epsilon_t = a_t \sqrt{h_t},$$
$$h_t = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1},$$

M5: A NLGARCH model:

$$Y_t = c + \phi Y_{t-1} + \epsilon_t, \quad \epsilon_t = a_t \sqrt{h_t},$$
$$h_t = \omega + \alpha(|\epsilon_{t-1}| + \xi \epsilon_{t-1})^2 + \beta h_{t-1},$$

M6: A TMA model:

$$Y_t = c + \phi_1 a_{t-1} I(Y_{t-1} \le 0) + \phi_2 a_{t-1} I(Y_{t-1} > 0) + \sigma a_t,$$

M7: A BL model:

$$Y_t = c + \phi Y_{t-1} + \psi Y_{t-1} a_{t-1} + \sigma a_t,$$

M8: A NLMA model:

$$Y_t = c + \phi a_{t-1} + \psi a_t a_{t-1} + \sigma a_t,$$

Table 2: Model parameters

model	**parameters**
AR, MA	$c = 1$ $\sigma^2 = 1$ $\phi \in \{-0.8, -0.6, -0.4, -0.2, 0.2, 0.4, 0.6, 0.8\}$ $\theta \in \{-0.8, -0.6, -0.4, -0.2, 0.2, 0.4, 0.6, 0.8\}$
TAR, TMA MSAR	$c_1 = -0.25$, $c_2 = 0.25$ $\sigma_1^2 = 3$, $\sigma_2^2 = 1$ $\sigma^2 = 1$ (for TMA only) $p_{11} = 0.9$, $p_{22} = 0.7$ (for MSAR only) $p_1 = 0.5$ (for MAR only) $(\phi_1, \phi_2) \in \left\{\begin{matrix} (-0.8,-0.8) & (-0.8,-0.5) & (-0.8,-0.2) & (-0.8,0.2) & (-0.8,0.5) \\ (-0.8,0.8) & (-0.5,-0.8) & (-0.5,-0.5) & (-0.5,0.5) & (-0.5,0.8) \\ (-0.2,-0.8) & (-0.2,0.8) & (0.2,-0.8) & (0.2,0.8) & (0.5,-0.8) \\ (0.5,-0.5) & (0.5,0.5) & (0.5,0.8) & (0.8,-0.8) & (0.8,-0.5) \\ (0.8,-0.2) & (0.8,0.2) & (0.8,0.5) & (0.8,0.8) & \end{matrix}\right\}$
GARCH	$c = 1$ $\phi = 0.5$ $\sigma^2 = 1$ $\xi = -0.5$ (for NLGARCH only) $(\alpha, \beta) \in \left\{\begin{matrix} (0.05,0.3) & (0.05,0.4) & (0.05,0.5) & (0.05,0.6) & (0.05,0.7) \\ (0.05,0.8) & (0.10,0.3) & (0.10,0.4) & (0.10,0.5) & (0.10,0.6) \\ (0.10,0.7) & (0.10,0.8) & (0.15,0.3) & (0.15,0.4) & (0.15,0.5) \\ (0.15,0.6) & (0.15,0.7) & & & \end{matrix}\right\}$
BL	$c = 1$ $\sigma^2 = 1$ $(\phi, \psi) \in \left\{\begin{matrix} (-0.8,-0.2) & (-0.6,-0.2) & (-0.4,-0.2) & (-0.2,-0.2) & (-0.2,0.2) \\ (-0.4,0.2) & (-0.6,0.2) & (-0.6,0.4) & (-0.8,0.2) & (0.2,-0.2) \\ (0.2,0.2) & (0.4,-0.2) & (0.4,0.2) & (0.6,-0.2) & (0.6,-0.4) \\ (0.6,0.2) & (0.8,-0.2) & (0.8,0.2) & & \end{matrix}\right\}$
NLMA	$c = 1$ $\sigma^2 = 4$ $(\phi, \psi) \in \left\{\begin{matrix} (-0.2,0.2) & (-0.2,0.4) & (-0.4,0.2) & (-0.4,0.4) & (-0.6,0.2) \\ (-0.6,0.4) & (0.2,0.2) & (0.2,0.4) & (0.4,0.2) & (0.4,0.4) \\ (0.6,0.2) & (0.6,0.4) & & & \end{matrix}\right\}$

is set to $T \in \{200, 1000\}$.

We also report the Q tests with automatically selected lag order based on Escanciano and Lobato (2009).[5] The estimated lag order is selected by maximizing the following objective function

$$Q_{rs}(l)^* = Q_{rs}(l) - q_{rs}(l),$$

$$q_{rs}(l) = \begin{cases} p\log(N) & \text{if} \quad \max_{j\in\{1,\ldots,L\}} |\hat{\rho}_{rs}(j)| \leq \sqrt{c\log(N)/N}, \\ \\ 2p & \text{if} \quad \max_{j\in\{1,\ldots,L\}} |\hat{\rho}_{rs}(j)| > \sqrt{c\log(N)/N}, \end{cases}$$

where the constant $c = 2.4$ is recommended by Escanciano and Lobato (2009). Finally, the lag order of the Q tests is determined by the simple rule $\hat{m} = \max_{l\in\{1,\ldots,L\}}(Q_{rs}(l)^*)$.

4 Monte Carlo results

4.1 Size and statistical properties

The average rejection frequency is calculated for each Q test as follows

$$\mathcal{P}_i = \frac{1}{R}\sum_{j=1}^{R} I(\hat{\alpha}_j \leq \alpha),$$

where $i \in \{1, \ldots, K\}$ denotes the i-th particular parameter configuration of a given time series model; R is the number of repetitions set to $R = 1000$; $I(\cdot)$ is a standard indicator function taking 1 if $\hat{\alpha}_j \leq \alpha$ and 0 otherwise; α represents the statistical significance level set to 0.05, and $\hat{\alpha}$ is the estimated p-value of a given Q test. Subsequently, three quantities for each Q test are presented in the following tables: "*avg*" stands for the average rejection frequency of a given Q test over all parameter configurations of a given model, "*min*" and "*max*" indicate the minimum and maximum of the average rejection frequencies of the test over all parameter configurations of a given model. Formally, the statistics are defined as follows

$$avg = \frac{1}{K}\sum_{i=1}^{K}\mathcal{P}_i,$$

$$min = \min_{i\in\{1,\ldots,K\}}(\mathcal{P}_i),$$

$$max = \max_{i\in\{1,\ldots,K\}}(\mathcal{P}_i),$$

[5] Recall that a given procedure is proposed for the realizations of stochastic processes and not the filtered ones. The additional simulations show, however, that the procedure may be adopted for filtered processes as well.

where K is the number of parameter configurations for a given time series model inspected by Monte Carlo experiments: $K = 8$ for AR and MA models, $K = 24$ for a TAR, MAR, MSAR, EXPAR, TMA models, $K = 17$ for GARCH and NLGARCH models, $K = 18$ for a BL model, and $K = 12$ for a NLMA model.

Since the Q tests proposed above are new versions of a standard Q test, it is important to check how a good approximation the limiting χ^2 distribution is for the tests. Provided that a χ^2 distribution is a valid limiting distribution, then $\mathbb{E}(Q(m)) \approx m$, $\text{var}(Q(m)) \approx 2m$, and the variance-mean ratio $\text{var}(Q(m))/\mathbb{E}(Q(m)) \approx 2$ as the sample $T \to \infty$ and $m/T \to 0$. Monte Carlo results of the proposed Q tests for AR(1) and MA(1) processes and fixed lag order $m \in \{5, 10, 15\}$ can be found in Table 3. The table shows that the finite sample properties of the Q tests are in line with the limiting distribution, even for relatively small samples and different lag orders: the average value of the Q tests is very close to m and the variance to $2m$. Figure 2 depict the χ^2 density function accompanied by the lower and upper bound of the smoothed empirical densities of the Q tests for the sample size $T = 200$ and the lag orders $m \in \{5, 10, 15\}$.[6] The figure clearly confirms that the χ^2 distribution is a valid distribution for all the Q tests even in relatively small samples. Additionally, the figure also clearly shows that the χ^2 distribution is much better approximation for the Q tests based on cross-correlations (i.e. Q_{12} and Q_{21}) as compared to the Q test based on auto-correlations (i.e. Q_{22}).

Table 4 illustrates that the Q tests have good size properties for both AR and MA processes. Even for a relatively small sample $T = 200$, the empirical size of the Q tests is close to the nominal level 0.05. In addition, other descriptive statistics (*min* and *max*) indicate that the behaviour of the Q tests is good, regardless of the specification of the lag order m or the sample size T. For instance, the minimum value of the individual average rejection frequencies for both AR and MA models, denoted as *min*, is not smaller than 0.03 and the maximum value, denoted as *max*, does not exceed 0.08. The size properties of the Q tests for leptokurtic and asymmetric distributions of innovations lead to very similar results.[7]

4.2 Power results

Regime switching models: From Tables 5 – 6, it is clear that Q_{12} and Q_{21} tests significantly outperform the results of the normally used Q_{22} test for a TAR model. From detailed records, it can be concluded that all the Q tests have very good power provided that parameters of a TAR model lie in a specific range, $|\phi_2 - \phi_1| \geq 1$, with

[6] A simple reference bandwidth is used for smoothing the empirical density functions of the Q tests.
[7] Results are available upon request.

rather opposite signs and a probability of a (lower) regime $\pi \in (0.3, 0.7)$. The second result is not very surprising since if $\pi \to 0$ or $\pi \to 1$, one regime dominates the other and the process can be relatively well approximated by a simple AR(p) model, which negatively affects the power of the Q tests.

Surprisingly, rather different results are obtained for a MSAR model. For this class of models, a switching mechanism is independent of any DGP parameter, and therefore fully under control. We set a probability of a lower regime to $\pi = 0.5$ only for simplicity of Monte Carlo experiments. MAR models are detected very efficiently by the Q_{22} tests but only if $|\phi_2 - \phi_1| \geq 1$. The Q_{12} test is not informative for any parameter configuration under consideration, and the Q_{21} only for just a few parameter configurations.[8]

Conditional volatility models: It can be clearly concluded from Tables 5 – 6 that the Q_{22} test is very useful in detecting conditional volatility. In the case of a simple GARCH model, the Q_{12} and Q_{21} tests are not informative if model innovations are drawn from a Gaussian (symmetric) distribution. In the case of a NLGARCH model, both the Q_{21} and Q_{22} tests exhibit a good power. Again, the Q_{12} is not at all informative. However, note that the power results of the Q_{21} test depend on a combination of asymmetry of innovations and a non-linear component. Provided that model innovations are negatively skewed and a non-linear component exhibits negative asymmetry as well (as in the Monte Carlo setup in this chapter), then the power of the Q_{21} and Q_{22} tests is around 0.7 in the sample $T = 1000$.

Other models: In the case of a NLMA model, Tables 5 – 6 show that both the Q_{12} and Q_{22} tests have no power, whereas the Q_{21} exhibits a very good power, regardless of the parameter specification of a NLMA model. Completely opposite results come from a BL model, where the only non-informative test is the Q_{21} test. The Q_{12} and Q_{22} tests have a very good power regardless the model parameters. Very similar results are obtained for a TMA model, where the only informative test is the Q_{12} test. However, it is worth noting that the above results for NLMA and BL models are to some extent model dependent since these two classes are extremely flexible.

[8] It is also interesting to point out the sensitivity of the power results on the regime-variances. In the case of a TAR model, the Q_{21} and Q_{22} tests lose the power if $\sigma_1^2 = \sigma_2^2$. On the contrary, the power of the Q_{21} improves substantially for a MSAR model if $\sigma_1^2 = \sigma_2^2$.

4.3 Sensitivity analysis

Since the original MLQ test suffers from a high power variation, special attention is paid to the sensitivity of the power of the new Q tests to the parameter configurations of data generating processes. For better understanding, the individual Monte Carlo results are presented in the form of graphical images. Each point depicted in a given graphical image represents the estimated p-value of a given Q test for a given parameter configuration of a given time series model (x-axis) and a particular Monte Carlo replication (y-axis). For example, in the case of a TAR model, each graphical image summarizes the results over 24000 replications ($K = 24$ parameter configurations of a TAR model and $R = 1000$ replications). A color range (from black to white) is used to explicitly indicate the different magnitude of the statistical significance of the Q tests. The results, based on Gaussian innovations, can be found in Figure 1.

For example, from the results given by a NLMA model it can be seen that all the p-values of the Q_{21} test are less then the significance level 0.05 and the results are not sensitive to a parameter configuration of a NLMA model at all, whereas the results of the Q_{12} and Q_{22} tests are statistically insignificant, but not sensitive to a parameter configuration of a NLMA model as well. Another relevant example, illustrating the benefit of using the graphical images, is related to a GARCH model. The p-values of the Q_{12} and Q_{21} tests are statistically very insignificant, whereas the p-values of the standard Q_{22} test are statistically significant, but merely for the second half of parameter configuration (i.e. $\alpha \geq 0.1$ and $\beta \geq 0.3$). From the figure it can be also concluded that the power of the Q tests is quite sensitive for regime switching models such as TAR, MAR, and MSAR models, whereas for models such as BL and NLMA, the stability of the power properties of the Q test is excellent. All in all, the proposed Q tests provide the robust power results (i.e. rejecting and/or not rejecting linearity regardless of the parameter configuration) for almost 70 % cases.

5 Empirical application

In this section, the proposed Q tests are applied to a set of 22 asset returns: 5 exchange rate time series; 5 interest rate time series; 6 commodity time series; and finally 6 equity indices. We use average weakly returns spanning the period 1980 and 2010 (1620 observations). A complete description of indicators can be found Table 7.

In this exercise, we are particularly interested in the following two questions: (i) “Is it worth inspecting cross-correlations of asset returns from the non-linearity testing point of view? Or is it fully sufficient to rely only on the Q_{22} test (an analogy of the MLQ test)?"; (ii) "Is a simple GARCH model an adequate model for asset returns in general?".

The results are presented in Table 7. The results suggest the following: (i) The Q_{22} test rejects linearity for all 22 series of asset returns at the significance level 0.01, regardless of the lag order specification m of the Q tests. This result is not very surprising since the conditional volatility is a common stochastic feature of asset returns. It can be concluded that the use of the standard Q_{22}, or analogically the MLQ test, is fully sufficient from the linearity testing standpoint. In addition, the results also confirm previous findings in the literature that standard linear homoscedastic ARMA models are not adequate for modelling asset returns; (ii) A simple GARCH model might be considered as appropriate, provided that the null hypothesis of linearity is rejected by only the Q_{22} test. However, this is not the case in 14 out of 22 cases where the null hypothesis of linearity is rejected by at least one of the cross-correlation tests (i.e. Q_{12} and/or Q_{21}) at the nominal level 0.05. Put differently, almost 65 % of asset returns exhibit stochastic features incompatible with a simple GARCH process.

6 Conclusion

This paper has proposed a new version of the Q test based on generalized correlations (i.e. auto- and cross-correlations). It has been demonstrated that the proposed Q tests have very good size and power properties, and the limiting χ^2 distribution is an accurate approximation, regardless of the sample size T and the lag order m. In addition, the Monte Carlo results reveal that Q tests fix two main shortcomings of the Mcleod and Li Q (MLQ) test often used in the literature: (i) the test is capable to capture some interesting non-linear models, for which the original MLQ test completely fails (e.g. a NLMA). The Q tests significantly improves the power for some other non-linear models (e.g. a TAR and TMA), for which the original MLQ test does not work very well; (ii) the new Q test can be used for discrimination between simple and more complicated (non-linear/asymmetric) GARCH models as well. The empirical results, based on a set of 22 asset returns, confirm the usefulness of the new Q tests for empirical time series analysis.

References

Anderson, T. (1991). The asymptotic distributions of autoregressive coefficients. Technical Report 26, Stanford University.

Anderson, T. (2003). *An Introduction to Multivariate Statistical Analysis*. Wiley.

Anderson, T. and A. Walker (1964). On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process. *The Annals of Mathematical Statistics 35*(3), 1296–1303.

Box, G. and D. Pierce (1970). Distributions of residual autocorrelations in autoregressive integrated moving average time series models. *Journal of the American Statistical Assciation 65*(1), 1509–1525.

Brockwell, P. and R. Davis (1991). *Time Series: theory and methods*. Springer.

Ding, Z., C. Granger, and R. Engle (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance 1*(1), 83–106.

Escanciano, J. and I. Lobato (2009). An automatic Portmanteau test for serial correlation. *Journal of Econometrics 151*(2), 140–149.

Fuller, W. (1996). *Introduction to Statistical Time Series*. Wiley.

Granger, C. and A. Andersen (1978). *An Introduction to Bilinear Time Series Models*. Gottingen.

Horowitz, J., I. Lobato, J. Nankervis, and N. Savin (2006). Bootstrapping the Box-Pierce Q test: A robust test of uncorrelatedness. *Journal of Econometrics 133*(2), 841–862.

Jansen, D. and C. D. Vries (1991). On the frequency of large stock returns: Putting booms and busts into perspective. *The Review of Economics and Statistics 73*(1), 18–24.

Jiang, J. (2010). *Large Sample Techniques for Statistics*. Springer.

Johnson, N., S. Kotz, and N. Balakrishnan (1994). *Continuous Univariate Distributions*. Wiley.

Kendall, M. and J. Ord (1973). *Time Series*. Griffin London.

Konishi, S. (1978). An approximation to the distribution of the sample correlation coefficient. *Biometrika 65*(3), 654–656.

Kwan, A. and A. Sim (1996a). On the finite-sample distribution of modified portmanteau tests for randomness of a Gaussian time series. *Biometrika 83*(4), 938–943.

Kwan, A. and A. Sim (1996b). Portmanteau tests of randomness and Jenkins' variance-stabilizing transformation. *Economics Letters 50*(1), 41–49.

Lawrance, A. and P. Lewis (1985). Modelling and residual analysis of nonlinear autoregressive time series in exponential variables. *Journal of the Royal Statistical Society 47*(2), 165–202.

Lawrance, A. and P. Lewis (1987). Higher-order residual analysis for nonlinear time series with autoregressive correlation structures. *International Statistical Review 55*(1), 21–35.

Li, W. (1992). On the asymptotic standard errors of residual autocorrelations in nonlinear time series modelling. *Biometrika 79*(2), 435–437.

Li, W. and T. Mak (1994). On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity. *Journal of Time Series Analysis 15*(6), 627–636.

Ljung, G. and G. Box (1978). On a measure of fit in time series models. *Biometrika 65*, 297–303.

Lobato, I. (2001). Testing that a dependent process is uncorrelated. *Journal of the American Statistical Association 96*(455), 1066–1076.

Lobato, I., J. Nankervis, and N. Savin (2002). Testing for zero autocorrelation in the presence of statistical dependence. *Econometric Theory 18*(3), 730–743.

Loretan, M. and P. Phillips (1994). Testing the covariance stationarity of heavy-tailed time series. *Journal of Empirical Finance 1*(2), 211–248.

Maravall, A. (1983). An application of nonlinear time series forecasting. *Journal of Business & Economic Statistics 1*(1), 66–74.

McLeod, A. and W. Li (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis 4*(4), 269–273.

Ng, S. and P. Perron (2005). A note on the selection of time series models. *Oxford Bulletin of Economics and Statistics 67*(1), 115–134.

Pérez, A. and E. Ruiz (2003). Properties of the sample autocorrelations of nonlinear transformations in long-memory stochastic volatility models. *Journal of Financial Econometrics 1*(3), 420–444.

Rao, C. (1973). *Linear Statistical Inference*. Wiley.

Romano, J. and L. Thombs (1996). Inference for Autocorrelations under Weak Assumptions. *Journal of the American Statistical Association 91*(434), 590–600.

Runde, R. (1997). The asymptotic null distribution of the Box-Pierce Q-statistic for random variables with infinite variance an application to German stock returns. *Journal of Econometrics 78*(2), 205–216.

Schott, J. (2005). *Matrix Analysis for Statistics*. Wiley.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics 6*(2), 461–464.

Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley.

Vávra, M. (2012). Power properties of non-linearity tests. *Birkbeck Working Paper in Economics and Finance, No. 1205*.

Wei, W. (1990). *Time Series Analysis: univariate and multivariate methods*. Addison-Wesley.

White, H. (1989). An additional hidden unit test for neglected nonlinearity in multilayer feedforward networks. *Proceedings of the International Joint Conference on Neural Networks 2*, 451–5.

Yao, Q. and P. Brockwell (2006). Gaussian maximum likelihood estimation for arma models. *Journal of time series analysis 27*(6), 857–875.

A Proofs

A.1 Useful Theorems

Theorem 3 *Let Assumption 1 be satisfied, then the LS estimate* $\hat{\boldsymbol{\beta}}$ *has the following properties: (i)* $\hat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}$*; (ii)* $\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{d} N(\mathbf{0}, \mathbf{V})$. □

Proof. *See a proof to Theorem 8.4.1 in Fuller (1996, p. 432).* ■

Note that Yao and Brockwell (2006) obtained the same results for the ML estimator of ARMA parameters. □

Theorem 4 *Let* $\{Z_t : t \in \mathbb{Z}\}$ *be a sequence of IID(0,*σ^2*) innovations such that* $\mathbb{E}(|Z_t|^4) < \infty$*, then for some integer* $m > 0$ *we have that* $\sqrt{T}(\dot{\boldsymbol{\rho}} - \boldsymbol{\rho}) \xrightarrow{d} N(\mathbf{0}, \mathbf{I})$*, where* $\boldsymbol{\rho} = (\rho(1), \ldots, \rho(m))'$ *denotes a vector of auto-correlations and* $\dot{\boldsymbol{\rho}} = (\dot{\rho}(1), \ldots, \dot{\rho}(m))'$ *denotes a vector of sample auto-correlations.* □

Proof. *See a proof to Theorem 7.2.1 in Brockwell and Davis (1991, p. 221) with a restriction* $X_t = Z_t$*.* ■

Theorem 5 *Let* $\{X_t, Y_t : t \in \mathbb{Z}\}$ *be a sequence of pairs of variables. If* $|X_t - Y_t| \xrightarrow{p} 0$ *and* $Y_t \xrightarrow{d} Y$*, then* $X_t \xrightarrow{d} Y$ *as well. That is, the limiting distribution of* X_t *exists and is the same as that of* Y*.* □

Proof. *See a proof to 2c.4(ix) result in Rao (1973, p. 122).* ■

Proposition 1 *Let* $\{X_t : t \in \mathbb{Z}\}$ *be a sequence of random variables such that* $X_t = a + O_p(r_t)$ *where* $a \in \mathbb{R}$ *and* $0 < r_t \to 0$ *as* $t \to \infty$*. If* g *is a function with the first* s *derivatives, then*

$$g(X_t) = \sum_{j=0}^{s} \frac{g^{(j)}(a)}{j!}(X_t - a)^j + o_p(r_t),$$

where $g^{(j)}$ *denotes the* j*-th derivative of* g *and* $g^0 = g$*.* □

Proof. *See a proof to Proposition 6.1.5 in Brockwell and Davis (1991, p. 201).* ■

Proposition 2 *Let Assumption 1 be satisfied. Let us define $\dot{\gamma}_{rs}(k)$ as follows*

$$\dot{\gamma}_{rs}(k) = \frac{1}{T}\sum_{t=k+1}^{T} g_r(a_t)g_s(a_{t-k}),$$

for the lag order $k \in \{1, \ldots, m\}$ and some integers $m > 0$, $r > 0$, $s > 0$, and $g_r(\cdot)$ and $g_s(\cdot)$ functions are defined as follows

$$g_r(a_t) = a_t^r - \frac{1}{T}\sum_{t=1}^{T} a_t^r, \quad g_s(a_t) = a_t^s - \frac{1}{T}\sum_{t=1}^{T} a_t^s.$$

Then it holds that

$$\frac{\partial \dot{\gamma}_{rs}(k)}{\partial \beta_i} = O_p(T^{-1/2}),$$

for all $\beta_i \in \boldsymbol{\beta}$. □

Proof.

$$\begin{aligned}
\frac{\partial \dot{\gamma}_{rs}(k)}{\partial \beta_i} &= \frac{1}{T}\sum_{t=k+1}^{T}\left(\frac{\partial g_r(a_t)}{\partial \beta_i}\right) g_s(a_{t-k}) + \frac{1}{T}\sum_{t=k+1}^{T} g_r(a_t)\left(\frac{\partial g_s(a_{t-k})}{\partial \beta_i}\right),\\
&= O_p(T^{-1/2}) + O_p(T^{-1/2}),\\
&= O_p(T^{-1/2}),
\end{aligned}$$

since $g_r(\cdot)$ and $g_s(\cdot)$ are continuous functions in $\boldsymbol{\beta}$ and a sample average of (stationary) random variables is $O_p(T^{-1/2})$, see Jiang (2010, Ch. 3). ■

Theorem 6 *Let $X_t \xrightarrow{d} X$ and $Y_t \xrightarrow{p} c$, where c is a finite constant different from 0. Then it holds that $X_t/Y_t \xrightarrow{d} X/c$.* □

Proof. *See a proof to Slutsky Theorem in Serfling (1980, p. 19).* ■

Theorem 7 *Let $\{X_t : t \in \mathbb{Z}\}$ and X be random variables defined on a probability space and let g be a Borel-measurable function defined on $\mathbb{R}$. Suppose that g is continuous with probability 1. Then $X_t \xrightarrow{p} X$ implies that $g(X_t) \xrightarrow{p} g(X)$.* □

Proof. *See a proof to Continuous Mapping Theorem in Serfling (1980, p. 24).* ■

Proposition 3 *Let Assumption 1 with $\hat{\gamma}_{rs}(0)$ given by*

$$\hat{\gamma}_{rs}(0) = \sqrt{\left[\frac{1}{T}\sum_{t=1}^{T} g_r^2(\hat{a}_t)\right]\left[\frac{1}{T}\sum_{t=1}^{T} g_s^2(\hat{a}_t)\right]}, \tag{9}$$

where $\hat{a}_t$ is the estimated residual from model in (1), and functions $g_r(\hat{a}_t)$ and $g_s(\hat{a}_t)$ are defined in (5). Then it holds that $\hat{\gamma}_{rs}(0) \xrightarrow{p} \gamma_{rs}(0)$. $\square$

Proof. *Theorem 7 implies that $\hat{\gamma}_{rs}(0) \xrightarrow{p} \gamma_{rs}(0)$, provided that*

$$\frac{1}{T}\sum_{t=1}^{T} g_r^2(\hat{a}_t) \xrightarrow{p} \mathbb{E}\left(g_r^2(a_t)\right).$$

Note that the same applies to the $g_s^2(\cdot)$ function as well since both $g_r(\cdot)$ and $g_s(\cdot)$ functions are equivalent for $r = s$. Therefore, it is fully sufficient to base the proof on one of these two functions. In order to simplify the proof, the following notation is used

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} g_r^2(a_t) &= \frac{1}{T}\sum_{t=1}^{T} a_t^{2r} - \left(\frac{1}{T}\sum_{t=1}^{T} a_t^r\right)^2, \\ &= M_2 - M_1^2.\end{aligned}$$

Following arguments in a proof of Theorem 8.4.1 in Fuller (1996, p. 432), the proof consists of the following two steps:

(i) It directly follows from Strong Law of Large Numbers (SLLN), see Theorem B in Serfling (1980, p. 24), that

$$\begin{aligned}M_1 &\xrightarrow{as} \mathbb{E}(a_t^r), \\ M_2 &\xrightarrow{as} \mathbb{E}(a_t^{2r}),\end{aligned}$$

which implies that

$$\frac{1}{T}\sum_{t=1}^{T} g_r^2(a_t) \xrightarrow{as} \mathbb{E}\left(g_r^2(a_t)\right),$$

for any integer $r > 0$.

(ii) It follows from Theorem 3 that $\hat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}$, which implies that $a_t(\hat{\boldsymbol{\beta}}) \equiv \hat{a}_t \xrightarrow{p} a_t$. Then, since $g_r(\cdot)$ is a continuous function in $\boldsymbol{\beta}$, it holds that $g_r(\hat{a}_t) \xrightarrow{p} g_r(a_t)$ for any integer $r > 0$.

Combining the results from (i) and (ii), it follows that

$$\frac{1}{T}\sum_{t=1}^{T} g_r^2(\hat{a}_t) \xrightarrow{p} \mathbb{E}\left(g_r^2(a_t)\right),$$

for any integer $r > 0$. This completes the proof. $\blacksquare$

Theorem 8 *Let* $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I})$ *and* $\mathbf{C}$ *a symmetric matrix with* $\mathrm{rk}(\mathbf{C}) = r$*, then* $\mathbf{z}'\mathbf{C}\mathbf{z} \sim \chi^2(r)$. $\square$

Proof. *See a proof for Theorem 9.8 in Schott (2005, p. 378).* $\blacksquare$

A.2 Proof of Theorem 1

Let $\boldsymbol{\beta} = (\xi_1, \ldots, \xi_p, \zeta_1, \ldots, \zeta_q, \sigma)'$ denote a vector of true model parameters and let $\hat{\boldsymbol{\beta}}$ denote the LS and/or ML estimates. Expanding a sample generalized covariance $\hat{\gamma}_{rs}(k)$ by a first-order Taylor expansion gives

$$\hat{\gamma}_{rs}(k) = \gamma_{rs}(k) + \sum_i (\hat{\beta}_i - \beta)\frac{\partial \dot{\gamma}_{rs}(k)}{\partial \beta_i} + O_p(T^{-1}),$$

for the lag order $k \in \{1, \ldots, m\}$ and $(r, s) = \{(1, 2), (2, 1), (2, 2)\}$. It concludes from Theorem 3 and Proposition 2 that $(\hat{\beta}_i - \beta) = O_p(T^{-1/2})$ and $\partial\dot{\gamma}_{rs}(k)/\partial\beta_i = O_p(T^{-1/2})$, which immediately implies that the product of these two stochastic components is $O_p(T^{-1})$. Then it holds that

$$\hat{\gamma}_{rs}(k) = \gamma_{rs}(k) + O_p(T^{-1}).$$

Moreover, it concludes from Proposition 3 that $\hat{\gamma}_{rs}(0) \xrightarrow{p} \gamma_{rs}(0)$ for given integers r, s. Then Theorem 6 implies that $\gamma_{rs}(0)$ can be considered as a normalizing constant having no effect on the limiting distribution of $\hat{\rho}_{rs}(k)$. Then it holds that

$$\hat{\rho}_{rs}(k) = \frac{\hat{\gamma}_{rs}(k)}{\gamma_{rs}(0)} + O_p(T^{-1}).$$

See also McLeod and Li (1983, p. 271) or Li and Mak (1994, p. 629–631) for a discussion.

Under Assumption 1 and using a slightly modified Theorem 4, it can be shown that a vector of sample correlations is given by

$$\sqrt{T}(\dot{\boldsymbol{\rho}}_{rs} - \boldsymbol{\rho}_{rs}) \xrightarrow{d} N(\mathbf{0}, \mathbf{I}).$$

Note that the modification of Theorem 4 lies in the requirement of the existence of the first eighth moments of the random variable a to ensure the validity of the above limiting result. This condition is a part of Assumption 1. Combining results from Theorem 4 and Theorem 5, it easy to show that the limiting distribution of a vector of the estimated correlations is given by

$$\sqrt{T}(\hat{\boldsymbol{\rho}}_{rs} - \boldsymbol{\rho}_{rs}) \xrightarrow{d} N(\mathbf{0}, \mathbf{I}),$$

since $|\hat{\boldsymbol{\rho}}_{rs} - \dot{\boldsymbol{\rho}}_{rs}| \xrightarrow{p} \mathbf{0}$ due to the fact that estimated residuals/parameters are consistent, which directly follows from Theorem 3.

Under Assumption 1, Proposition 1 concludes that a first-order Taylor expansion of $\hat{z}_{rs}(k)$ around $\rho_{rs}(k)$ gives us the following expression

$$\hat{z}_{rs}(k) = \hat{\rho}_{rs}(k) + o_p(T^{-1/2}),$$

for the lag order $k \in \{1, \ldots, m\}$ and integer $m > 0$. It is now easy to see that vectors $\hat{\mathbf{z}}_{rs} = (z_{rs}(1), \ldots, z_{rs}(m))'$ and $\hat{\boldsymbol{\rho}}_{rs} = (\hat{\rho}_{rs}(1), \ldots, \hat{\rho}_{rs}(m))'$ have the same limiting distribution given by

$$\sqrt{T}(\hat{\mathbf{z}}_{rs} - \mathbf{z}_{rs}) \xrightarrow{d} N(\mathbf{0}, \mathbf{I}).$$

■

A.3 Proof of Theorem 2

Note that the proposed Q tests can be written into the form of a quadratic function given by

$$Q_{rs}(m) = \sum_{k=1}^{m} (T-k)\hat{z}_{rs}^2(k) = \hat{\mathbf{z}}_{rs}' \mathbf{C} \hat{\mathbf{z}}_{rs},$$

where $\hat{\mathbf{z}}_{rs} = (\hat{z}_{rs}(1), \ldots, \hat{z}_{rs}(m))'$ is an $(m \times 1)$ vector of the estimated correlations and $\mathbf{C}$ is an appropriate $(m \times m)$ symmetric matrix. The limiting χ^2 distribution of the Q_{rs} tests, for $(r, s) \in \{(1, 2), (2, 1), (2, 2)\}$, follows directly from Theorem 1 and Theorem 8, and the degrees of freedom follow from the fact that $\text{rk}(\mathbf{C}) = m$. ■

B Tables and figures

Table 3: Statistical properties of the Q tests: $N(0,1)$

			AR (#8)			MA (#8)		
sample	test	lag m	mean	var	var/mean	mean	var	var/mean
T=200	Q_{12}	5	5.0	9.8	1.96	5.0	10.1	2.02
		10	9.9	19.6	1.97	10.0	19.9	2.00
		15	14.9	29.4	1.97	14.9	29.8	2.00
	Q_{21}	5	5.0	9.7	1.96	5.1	10.3	2.03
		10	9.9	19.3	1.94	10.0	20.1	2.00
		15	14.9	28.2	1.90	15.0	30.4	2.02
	Q_{22}	5	5.0	9.9	1.99	5.0	10.5	2.12
		10	9.9	21.0	2.12	9.9	21.3	2.15
		15	14.9	32.8	2.20	14.9	32.9	2.21
T=1000	Q_{12}	5	5.0	10.1	2.02	5.0	10.0	2.00
		10	9.9	20.4	2.06	10.0	20.3	2.02
		15	15.0	30.9	2.07	15.1	30.4	2.01
	Q_{21}	5	5.0	10.2	2.03	5.0	9.9	1.97
		10	10.0	19.5	1.95	10.0	19.9	1.99
		15	15.0	29.6	1.98	15.0	30.1	2.01
	Q_{22}	5	5.0	10.2	2.06	5.0	10.0	1.98
		10	10.0	21.0	2.10	10.0	19.9	1.99
		15	14.9	31.6	2.11	15.0	30.4	2.03

[a] "AR (#8)" indicates that 8 different parameter configurations of an AR model are evaluated.

[b] "mean" stands for a sample mean of the Q test over all replications and parameter configurations, "var" stands for a sample variance value of the Q test over all replications and parameter configurations, "var/mean" denotes a variance-mean ratio. The significance level is set to $\alpha = 0.05$.

Table 4: Size of the Q tests: $N(0,1)$

			AR (#8)			MA (#8)		
sample	test	lag m	avg	min	max	avg	min	max
T=200	Q_{12}	5	0.050	0.037	0.063	0.049	0.043	0.054
		10	0.046	0.035	0.056	0.054	0.045	0.057
		15	0.045	0.038	0.052	0.049	0.043	0.054
		m	0.050	0.037	0.058	0.049	0.044	0.063
	Q_{21}	5	0.049	0.036	0.056	0.049	0.040	0.063
		10	0.051	0.044	0.057	0.044	0.034	0.057
		15	0.048	0.044	0.055	0.049	0.038	0.059
		m	0.049	0.035	0.062	0.047	0.039	0.055
	Q_{22}	5	0.049	0.037	0.057	0.054	0.040	0.063
		10	0.054	0.040	0.063	0.058	0.046	0.064
		15	0.056	0.039	0.062	0.058	0.053	0.071
		m	0.057	0.042	0.079	0.055	0.041	0.069
T=1000	Q_{12}	5	0.049	0.044	0.056	0.050	0.045	0.053
		10	0.048	0.037	0.056	0.050	0.033	0.062
		15	0.051	0.043	0.058	0.050	0.044	0.060
		m	0.053	0.039	0.073	0.054	0.044	0.073
	Q_{21}	5	0.049	0.036	0.060	0.050	0.039	0.063
		10	0.050	0.043	0.056	0.050	0.040	0.066
		15	0.048	0.038	0.055	0.051	0.042	0.061
		m	0.052	0.044	0.066	0.048	0.037	0.056
	Q_{22}	5	0.052	0.038	0.064	0.050	0.032	0.059
		10	0.051	0.040	0.068	0.051	0.042	0.066
		15	0.049	0.040	0.055	0.054	0.045	0.064
		m	0.054	0.041	0.067	0.057	0.050	0.068

[a] m denotes the automatically selected lag order based on Escanciano and Lobato (2009).

[b] "AR (#8)" indicates that 8 different parameter configurations of an AR model are evaluated.

[c] "avg" stands for the average rejection frequency of the Q tests over all parameter configurations of a given time series model, "min" and "max" denote a minimum and maximum of the average rejection frequencies over all parameter configurations. The significance level is set to $\alpha = 0.05$.

Table 5: Power properties of the Q tests: $N(0,1)$, $T = 200$

		TAR (#24)			**MSAR (#24)**			**GARCH (#17)**			**NLGARCH (#17)**			**TMA (#24)**			**BL (#18)**			**NLMA (#12)**		
test	lag	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max
Q_{12}	5	0.31	0.04	0.80	0.13	0.06	0.22	0.08	0.05	0.11	0.09	0.05	0.17	0.51	0.03	0.98	0.78	0.62	1.00	0.05	0.04	0.06
	10	0.25	0.05	0.69	0.11	0.04	0.18	0.07	0.04	0.12	0.08	0.04	0.18	0.44	0.04	0.94	0.66	0.49	0.99	0.05	0.03	0.07
	15	0.21	0.04	0.61	0.09	0.05	0.16	0.07	0.05	0.12	0.08	0.04	0.17	0.39	0.04	0.90	0.58	0.41	0.97	0.05	0.04	0.06
	m	0.36	0.04	0.93	0.15	0.06	0.27	0.08	0.06	0.12	0.10	0.05	0.18	0.59	0.04	1.00	0.85	0.54	1.00	0.05	0.04	0.07
Q_{21}	5	0.59	0.05	0.90	0.20	0.07	0.57	0.08	0.05	0.11	0.25	0.08	0.54	0.05	0.04	0.07	0.09	0.04	0.36	0.26	0.11	0.42
	10	0.50	0.05	0.80	0.17	0.06	0.52	0.07	0.05	0.11	0.21	0.06	0.50	0.05	0.04	0.06	0.08	0.04	0.32	0.19	0.08	0.32
	15	0.43	0.04	0.73	0.16	0.06	0.47	0.07	0.05	0.11	0.19	0.06	0.45	0.05	0.04	0.06	0.08	0.04	0.28	0.16	0.07	0.25
	m	0.63	0.06	0.98	0.22	0.08	0.59	0.09	0.06	0.14	0.26	0.10	0.48	0.06	0.04	0.09	0.09	0.04	0.38	0.43	0.20	0.68
Q_{22}	5	0.19	0.04	0.66	0.48	0.10	0.98	0.21	0.06	0.52	0.31	0.08	0.75	0.07	0.04	0.13	0.22	0.11	0.71	0.05	0.04	0.06
	10	0.15	0.04	0.53	0.43	0.09	0.97	0.19	0.07	0.47	0.28	0.07	0.70	0.06	0.04	0.12	0.19	0.09	0.65	0.05	0.04	0.06
	15	0.14	0.05	0.46	0.40	0.09	0.95	0.17	0.07	0.44	0.26	0.07	0.67	0.06	0.04	0.10	0.17	0.09	0.59	0.06	0.04	0.07
	m	0.22	0.05	0.82	0.53	0.11	0.98	0.24	0.07	0.52	0.33	0.09	0.72	0.09	0.05	0.19	0.27	0.16	0.81	0.07	0.06	0.09

[a] m denotes the automatically selected lag order based on Escanciano and Lobato (2009).

[b] "TAR (#24)" indicates that 24 different parameter configurations of a TAR model are evaluated.

[c] "avg" stands for the average rejection frequency of the Q tests over all parameter configurations of a given time series model, "min" and "max" denote a minimum and maximum of the average rejection frequencies over all parameter configurations. The significance level is set to $\alpha = 0.05$.

Table 6: Power properties of the Q tests: $N(0,1)$, $T = 1000$

		TAR (#24)			**MSAR (#24)**			**GARCH (#17)**			**NLGARCH (#17)**			**TMA (#24)**			**BL (#18)**			**NLMA (#12)**		
test	lag	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max
Q_{12}	5	0.65	0.04	1.00	0.26	0.10	0.49	0.08	0.05	0.16	0.11	0.05	0.29	0.78	0.04	1.00	1.00	1.00	1.00	0.05	0.04	0.06
	10	0.61	0.04	1.00	0.21	0.08	0.40	0.08	0.04	0.17	0.11	0.05	0.34	0.76	0.04	1.00	1.00	1.00	1.00	0.05	0.04	0.06
	15	0.59	0.03	1.00	0.19	0.07	0.36	0.08	0.04	0.16	0.11	0.05	0.32	0.74	0.04	1.00	1.00	1.00	1.00	0.05	0.04	0.06
	m	0.64	0.04	1.00	0.27	0.09	0.50	0.09	0.06	0.17	0.13	0.06	0.35	0.79	0.04	1.00	1.00	1.00	1.00	0.06	0.04	0.07
Q_{21}	5	0.89	0.07	1.00	0.41	0.09	0.89	0.09	0.05	0.17	0.68	0.20	0.99	0.08	0.04	0.16	0.19	0.05	0.90	0.77	0.49	1.00
	10	0.87	0.07	1.00	0.36	0.08	0.87	0.08	0.06	0.17	0.63	0.15	0.99	0.07	0.04	0.12	0.18	0.05	0.85	0.69	0.36	0.99
	15	0.86	0.07	1.00	0.33	0.08	0.84	0.08	0.05	0.17	0.58	0.14	0.99	0.07	0.04	0.11	0.17	0.04	0.83	0.65	0.30	0.97
	m	0.89	0.07	1.00	0.41	0.09	0.88	0.09	0.06	0.17	0.69	0.25	0.99	0.13	0.05	0.28	0.19	0.05	0.90	0.89	0.77	1.00
Q_{22}	5	0.44	0.05	1.00	0.84	0.44	1.00	0.63	0.18	1.00	0.75	0.24	1.00	0.19	0.04	0.59	0.69	0.52	1.00	0.05	0.04	0.06
	10	0.42	0.06	1.00	0.80	0.36	1.00	0.58	0.14	1.00	0.70	0.19	1.00	0.15	0.05	0.47	0.60	0.41	1.00	0.05	0.04	0.07
	15	0.40	0.05	1.00	0.78	0.29	1.00	0.54	0.13	1.00	0.67	0.17	1.00	0.13	0.04	0.40	0.55	0.34	1.00	0.05	0.04	0.07
	m	0.45	0.06	1.00	0.84	0.41	1.00	0.66	0.24	1.00	0.77	0.30	1.00	0.28	0.06	0.76	0.77	0.63	1.00	0.06	0.05	0.09

[a] m denotes the automatically selected lag order based on Escanciano and Lobato (2009).

[b] 'TAR (#24)" indicates that 24 different parameter configurations of a TAR model are evaluated.

[c] "avg" stands for the average rejection frequency of the Q tests over all parameter configurations of a given time series model, "min" and "max" denote a minimum and maximum of the average rejection frequencies over all parameter configurations. The significance level is set to $\alpha = 0.05$.

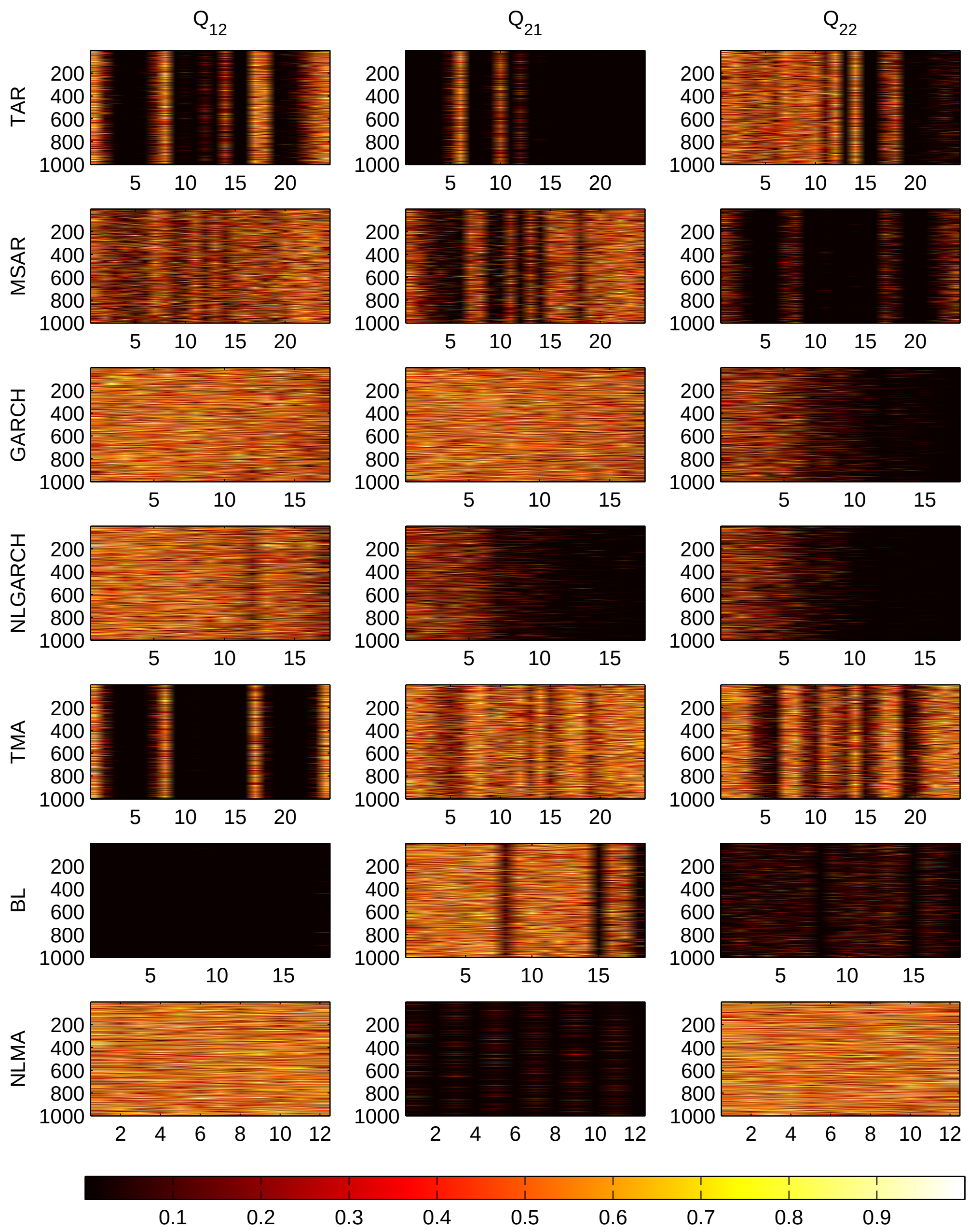
Table 7: Application of the Q tests

	$\mathbf{Q_{12}}$				$\mathbf{Q_{21}}$				$\mathbf{Q_{22}}$			
variable/lag	5	10	15	m	5	10	15	m	5	10	15	m
Exchange rates												
USDGBP	0.37	0.46	0.43	0.12	0.08	0.11	0.03	0.03	0.00	0.00	0.00	0.00
USDJPY	0.97	0.94	0.66	0.00	0.10	0.03	0.07	0.21	0.00	0.00	0.00	0.00
USDCAD	0.37	0.66	0.80	0.31	0.89	0.92	0.96	0.96	0.00	0.00	0.00	0.00
USDAUD	0.88	0.70	0.81	0.48	0.05	0.19	0.24	0.33	0.00	0.00	0.00	0.00
USDCHF	0.85	0.99	1.00	0.98	0.43	0.25	0.24	0.25	0.00	0.00	0.00	0.00
Interest rates												
USIR3M	0.62	0.93	0.83	0.89	0.99	1.00	0.89	0.97	0.00	0.00	0.00	0.00
UKIR3M	0.11	0.08	0.02	0.01	0.62	0.53	0.80	0.86	0.00	0.00	0.00	0.00
AUIR3M	0.74	0.47	0.72	0.89	0.90	0.84	0.97	0.95	0.00	0.00	0.00	0.00
CAIR3M	0.13	0.01	0.01	0.05	0.83	0.34	0.35	0.48	0.00	0.00	0.00	0.00
CHIR3M	0.88	0.93	0.97	0.98	0.31	0.28	0.50	0.82	0.00	0.00	0.00	0.00
Equity indices												
DJIA	0.53	0.50	0.12	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FTSE	0.10	0.24	0.13	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AUSE	0.49	0.16	0.13	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TSE	0.66	0.69	0.50	0.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TOPIX	0.17	0.02	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CHSE	0.59	0.70	0.59	0.63	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Commodities												
WHEAT	0.61	0.26	0.19	0.31	0.87	0.98	1.00	0.93	0.00	0.00	0.00	0.00
SOYBEAN	0.55	0.01	0.02	0.10	0.23	0.63	0.50	0.02	0.00	0.00	0.00	0.00
COFFEE	0.45	0.32	0.23	0.31	0.01	0.01	0.01	0.09	0.00	0.00	0.00	0.00
COTTON	0.26	0.36	0.62	0.67	0.41	0.50	0.26	0.25	0.00	0.00	0.00	0.00
FUEL	0.14	0.05	0.03	0.07	0.31	0.68	0.88	0.98	0.00	0.00	0.00	0.00
GOLD	0.60	0.56	0.38	0.62	0.41	0.70	0.53	0.81	0.00	0.00	0.00	0.00

[a] m denotes the automatically selected lag order based on Escanciano and Lobato (2009).

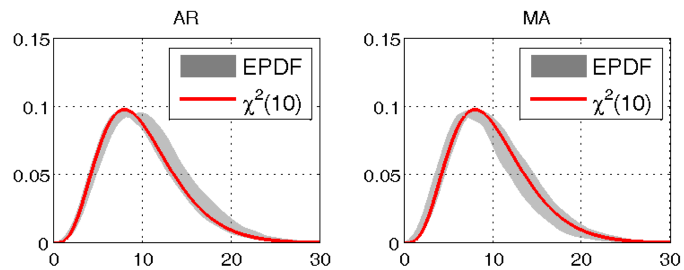
[b] Note that a particular transformation of each series is indicated in the brackets. USDGBP = the US dollar to British pound exchange rate ($\Delta \log$), USDJPY = the US dollar to Japanese yen exchange rate ($\Delta \log$), USDCAD = the US dollar to Canadian dollar exchange rate ($\Delta \log$), USDAUD = the US dollar to Australian dollar exchange rate ($\Delta \log$), USDCHF = the US dollar to Swiss frank exchange rate ($\Delta \log$), DJIA = the US Dow Jones Industrials Share Index ($\Delta \log$), UKFT = the UK FT All Shares Index ($\Delta \log$), TOPIX = Tokyo Stock Exchange Index ($\Delta \log$), TSE = the Toronto Stock Exchange Index ($\Delta \log$), AUSE = Australian Stock Exchange Index ($\Delta \log$), CHSE = the Swiss Stock Exchange Index ($\Delta \log$), WHEAT = Kansas wheat, hard, cents/bushel ($\Delta \log$), SOYBEAN = soybeans, yellow, cents/bushel ($\Delta \log$), COFFEE = Brazilian coffee beans cents/pound ($\Delta \log$), COTTON = cotton, cents/pound ($\Delta \log$), FUEL = fuel oil, cents/gallon ($\Delta \log$), GOLD = gold bullion, USD/troy ounce ($\Delta \log$), USIR = the US interest rates, 3M (Δ), UKIR = the UK interest rates, 3M (Δ), CAIR = the Canadian interest rates, 3M (Δ), AUIR = the Australian interest rates, 3M (Δ), CHIR = the Swiss interest rates, 3M (Δ).

Figure 1: Power images of the Q tests: $T = 1000$, $R = 1000$, $N(0,1)$

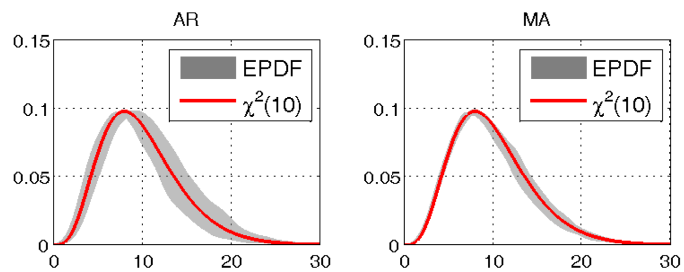


* Each point depicted in the graphical image represents the estimated p-value of a given Q test for a given parameter configuration of a given time series model (x-axis) and a given Monte Carlo replication (y-axis). The results of the Q tests are based on the automatically selected lag order m.

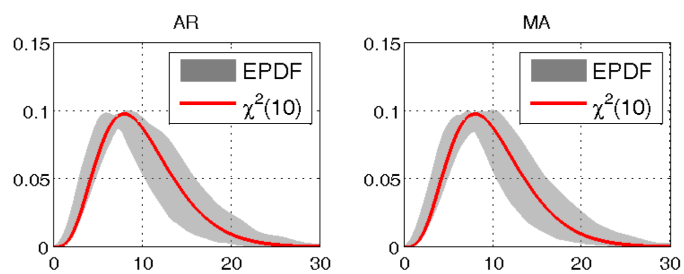
Figure 2: Accuracy of the smoothed empirical density functions of the $Q(10)$ tests to the limiting $\chi^2(10)$: $T = 200$, $R = 5000$



(a) $Q_{12}(10)$



(b) $Q_{21}(10)$



(c) $Q_{22}(10)$

* The empirical distributions are smoothed by a kernel smoothing procedure with a simple reference bandwidth for all parameters of AR and MA models $\phi, \theta \in \{-0.8, -0.6, -0.4, -0.2, 0.2, 0.4, 0.6, 0.8\}$. We present bands calculated from the highest and lowest smoothed empirical distribution functions in order to explicitly show parameter uncertainty of the finite sample distributions.