



BIROn - Birkbeck Institutional Research Online

Vavra, Marian (2012) Testing non-linearity using a modified Q test. Working Paper. Birkbeck College, University of London, London, UK.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/5955/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>
contact lib-eprints@bbk.ac.uk.

or alternatively

ISSN 1745-8587



School of Economics, Mathematics and Statistics

BWPEF 1204

Testing Non-linearity Using a Modified Q Test

Marian Vavra

Birkbeck, University of London

March 2012

Testing Non-linearity Using a Modified Q Test

Marian Vavra ^{*†}

Department of Economics, University of London, UK

February 28, 2012

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 also 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

There are many different non-linear time series models and non-linearity tests in the literature, see Tong (1990, Chapter 3 and 5) for a comprehensive summary. The most

^{*}Correspondence to: Marian Vavra, Department of Economics, University of London, Malet Street, WC1E 7HX, London, UK. E-mail: marian.vavra@gmail.com

[†]This paper is a part of my Ph.D. thesis. I would like to thank Professor Zacharias Psaradakis and Professor Ron Smith for very useful comments and suggestions leading to significant improvements of the paper. I would also like to thank Professor Olivier Linton and Professor Howell Tong from the London School of Economics since some ideas implemented in the paper follow from the Advanced Econometrics course, which I attended during M.Phil. studies. All remaining errors are solely mine.

popular test for non-linearity is probably the portmanteau Q test developed by McLeod and Li (1983), denoted as the MLQ test. 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 compared to some other non-linearity tests, which require “deep” knowledge for a correct application.¹ Moreover, no estimation of complicated non-linear models is required in the case of the Q test. Another advantage of the MLQ test is that it is routinely available in many statistical packages. However, Vavra (2012) finds two shortcomings of the MLQ test. First, the MLQ test cannot detect some interesting non-linear models (e.g. an exponential autoregressive (EXPAR) model or non-linear moving average (NLMA) model). Moreover, the test exhibits a relatively low power for some commonly applied non-linear time series models (e.g. a threshold autoregressive (TAR) model). Second, the MLQ test lacks significant improvements made in volatility modelling. For example, the test has no discrimination power against more advanced GARCH models (e.g. non-linear and/or asymmetric GARCH models, or GARCH models with asymmetric innovations) developed recently in the literature.

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 both of 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.

Our results reveal that the proposed Q tests, based on generalized residual correlations, significantly increase the power of the Q tests against several interesting non-linear time series models. It is also shown that the Q tests can be used for preliminary discrimination between some classes of non-linear time series models (e.g. linear and non-linear/asymmetric GARCH models).

¹A 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 parameters: (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 parameters of squashing functions. Loosely speaking, the different application of the NN test, the different results even though the same time series is used.

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.

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 ?. They show that, provided that $\{X_t : t \in \mathbb{Z}\}$ is a sequence of a linear Gaussian stationary process, it holds that

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

where ρ_k denotes the k -th theoretical auto-correlation. 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 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*

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

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

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

□

Provided that all conditions of Assumption 1 are satisfied, then a given stochastic process $\{X\}$ is stationary, an appropriate model is identified and the true parameter vector β 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 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 about 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

A theoretical generalized correlation function is defined as

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

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

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

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

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

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

²Alternatively, the null hypothesis about 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 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 a corrected variance-covariance matrix.

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

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

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

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

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

The Q test is then given by

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

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

$$\hat{z}_{rs}(k) = 0.5 \log \left(\frac{1 + \hat{\rho}_{rs}(k)}{1 - \hat{\rho}_{rs}(k)} \right), \quad (8)$$

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 paper: $\dot{\gamma}_{rs}(k)$ denotes sample generalized covariance based on observed sequence of innovations $\{a_t\}$. 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 $\text{var}(\dot{z}_{rs}(k)) = (T-k)^{-1} + O(T^{-2})$, 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 $\{a_t\}$, $\text{var}(Q_{rs}(m)) = 2 \sum_{k=1}^m (T-k)^2 [\mathbb{E}(\dot{z}_{rs}^2(k))] \approx 2m$, since

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

⁴Under 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 as well.

$\mathbb{E}(\hat{z}_{rs}^2(k)) = (T-k)^{-1} + O(T^{-2})$ and $\text{cov}(\hat{z}_{rs}^2(i), \hat{z}_{rs}^2(j)) \approx 0$ for integers $i, j \in \{1, \dots, m\}$, such that $i \neq j$. Therefore, a more complicated variance-stabilizing transformation of the Q test, as in Ljung and 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.

Test specification captures two very well known Q tests: (i) $r = s = 1$ leads to a test proposed by Box and Pierce (1970) and modified by Ljung and Box (1978); (ii) $r = s = 2$ 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 paper 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 $r, s \in \mathbb{N}$, 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 macroeconomic and/or financial 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 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 returns are quite 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 $\hat{\mathbf{z}}_{rs} = (\hat{z}_{rs}(1), \dots, \hat{z}_{rs}(m))'$ 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 Appendix 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 Appendix A for a proof. ■

3 Motivation for a new Q test

We demonstrate that not only can the proposed Q test be used for testing non-linearity itself, but it also captures recently developed non-linear and/or asymmetric GARCH models, or GARCH models with asymmetric innovations. Under some conditions, this property can be used for preliminary discrimination among GARCH models.

Let us assume $\{\epsilon_t : t \in \mathbb{Z}\}$ is an observed martingale difference sequence given by

$$\begin{aligned}\epsilon_t &= a_t \sqrt{h_t}, \\ h_t &= \omega + \alpha \epsilon_{t-1}^2,\end{aligned}$$

where $\omega > 0$, $0 < \alpha < 1$, and $\{a_t : t \in \mathbb{Z}\}$ is a sequence of IID(0,1) innovations. Then, a generalized covariance structure, for $k = 1$, is given by

$$\begin{aligned}\gamma_{12}(1) &= \mathbb{E} [\epsilon_t (\epsilon_{t-1}^2 - \sigma^2)] = \mathbb{E}(a_t) \mathbb{E}(a_{t-1}^2 \sqrt{h_t} h_{t-1}) = 0, \\ \gamma_{21}(1) &= \mathbb{E} [(\epsilon_t^2 - \sigma^2) \epsilon_{t-1}] = \alpha \mathbb{E}(a_{t-1}^3) \mathbb{E}(h_{t-1}^{3/2}) \neq 0, \quad \text{if } \mathbb{E}(a_{t-1}^3) \neq 0, \\ \gamma_{22}(1) &= \mathbb{E} [(\epsilon_t^2 - \sigma^2)(\epsilon_{t-1}^2 - \sigma^2)] = \alpha \mathbb{E}[(\epsilon_{t-1}^2 - \sigma^2)^2] \neq 0.\end{aligned}$$

Now, let us assume $\{\epsilon_t : t \in \mathbb{Z}\}$ is another martingale difference sequence given by

$$\begin{aligned}\epsilon_t &= a_t \sqrt{h_t}, \\ h_t &= \omega + \alpha \epsilon_{t-1}^2 + \xi \epsilon_{t-1},\end{aligned}$$

where $\omega > 0$, $0 < \alpha < 1$, $|\xi| < 1$, and $\{a_t : t \in \mathbb{Z}\}$ is a sequence of IID(0,1) innovations. This gives a generalized covariance structure, for $k = 1$, as follows

$$\begin{aligned}\gamma_{12}(1) &= \mathbb{E} [\epsilon_t (\epsilon_{t-1}^2 - \sigma^2)] = \mathbb{E}(a_t) \mathbb{E}(a_{t-1}^2 \sqrt{h_t} h_{t-1}) = 0, \\ \gamma_{21}(1) &= \mathbb{E} [(\epsilon_t^2 - \sigma^2) \epsilon_{t-1}] = \mathbb{E}(a_t^2) \mathbb{E}(h_t \epsilon_{t-1}) = \xi \mathbb{E}(\epsilon_{t-1}^2) + \alpha \mathbb{E}(\epsilon_{t-1}^3) \neq 0, \\ \gamma_{22}(1) &= \mathbb{E} [(\epsilon_t^2 - \sigma^2)(\epsilon_{t-1}^2 - \sigma^2)] = \mathbb{E}(a_t^2) \mathbb{E}(h_t \epsilon_{t-1}) = \alpha \mathbb{E}(\epsilon_{t-1}^2 - \sigma^2)^2 + \delta \mathbb{E}(\epsilon_{t-1}^3) \neq 0.\end{aligned}$$

The derived covariances of ARCH and quadratic ARCH models used above immediately show that inspecting auto-correlations and cross-correlations might be very useful not only for non-linearity testing, but also for some preliminary discrimination. For example, the results of an ARCH model prove that $\rho_{12}(1) = \gamma_{12}(1)/\gamma_{12}(0) = 0$ regardless of the distribution of innovations, while $\rho_{21}(1) \neq 0$, provided that model innovations are asymmetric (i.e. $\mathbb{E}(a_t^3) \neq 0$). Similar findings can be obtained for a quadratic ARCH model. In this case, however, the $\rho_{21}(1)$ term depends both on asymmetry of innovations (i.e. $\mathbb{E}(a_t^3)$) and on the non-linear component of the model (i.e. $\xi\mathbb{E}(\epsilon_t^2)$). Therefore, it can theoretically happen, at least theoretically, that both components off-set or trigger the effect. For this reason, results from the new Q tests should be interpreted with caution. The analytical results reveal that cross-correlations, especially the $\rho_{12}(k)$ terms, can be very useful for discriminating between: (i) simple (G)ARCH models and (G)ARCH models with asymmetric innovations; or (ii) under some conditions, simple (G)ARCH models and asymmetric and/or non-linear (G)ARCH models.

Since similar analytical results are rather difficult to obtain for other non-linear time series models, the usefulness of auto-correlations and cross-correlations is inspected by means of extensive Monte Carlo experiments.

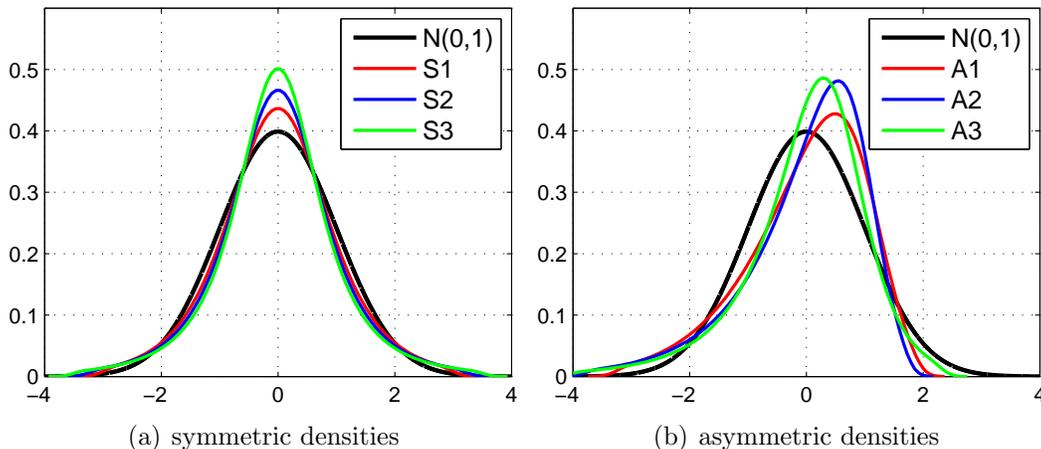
4 Time series models and Monte Carlo setup

4.1 Time series models

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) nine non-linear time series models: a threshold (TAR) model, an exponential autoregressive (EXPAR) model, a mixture autoregressive (MAR) 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 2. A complete set of model parameters can be found in Table 3.

The robustness of the proposed Q tests is examined against various distributions of innovations as well. In particular, apart form a Gaussian distribution, which serves as a benchmark for comparison, various model innovations coming from a generalized lambda distributions (GLD) are considered in this paper, see Randles et al. (1980).

Figure 1: Density functions

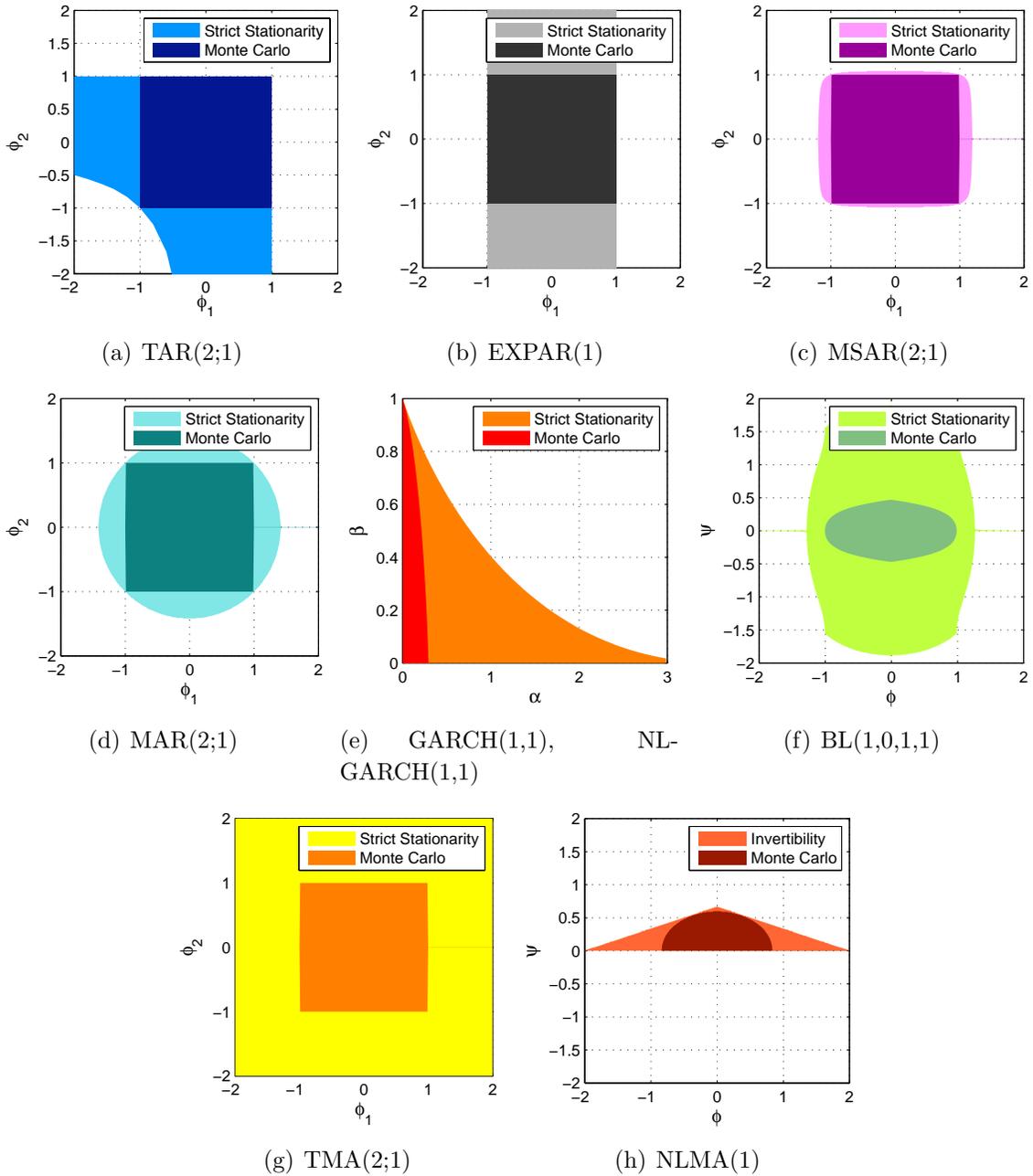


This family provides a wide range of distributions that are easily generated since they are defined in terms of the inverses of the cumulative distribution functions: $F^{-1}(u) = \lambda_1 + [u^{\lambda_3} - (1-u)^{\lambda_4}]/\lambda_2$, for $u \in [0, 1]$. This paper considers the following particular distributions: 3 distributions are symmetric, but leptokurtic, and 3 distributions are asymmetric, see Table 1 and Figure 1 for details. All generated innovations are normalized to have zero mean and unit variance. Note that parameters of all non-linearity models are designed in such a way to satisfy strict stationarity, 4th-moment stationarity and/or invertibility conditions, if necessary. The only exception is S3 specification of model innovations, see Table 1, for which the 4th-moment stationarity cannot be reached, at least for some non-linear time series models.

4.2 Monte Carlo setup

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.

Figure 2: Stationarity regions of non-linear models



* Strict stationarity regions are calculated based on the assumption that $a \sim N(0, 1)$, whereas Monte Carlo regions are calculated based on the intersection of 4-th moment stationarity and/or invertibility conditions for S1, S2, A1, A2, A3 specifications of distributions of model innovations. Only S3 specification leads to very restrictive parameter regions, and therefore, is not considered.

Table 1: Parameters of a generalized lambda distribution

	λ_1	λ_2	λ_3	λ_4	skewness	kurtosis	moment
S1	0.00000	-1.00000	-0.08000	-0.08000	0.0	6.0	12
S2	0.00000	-0.39791	-0.16000	-0.16000	0.0	11.6	6
S3	0.00000	-1.00000	-0.24000	-0.24000	0.0	126.0	4
A1	0.00000	-0.04306	0.02521	0.09403	-0.9	4.2	10
A2	0.00000	1.00000	-0.00750	-0.03000	-1.5	7.5	33
A3	0.00000	1.00000	-0.10090	-0.18020	-2.0	21.1	5

* Note that a standard normal distribution can be also approximated by a generalized lambda distribution with the following parameters: $\lambda_1 = 0$, $\lambda_2 = 0.1975$, $\lambda_3 = \lambda_4 = 0.1349$.

^a The maximum exponent of a given distribution.

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 $l \in \{1, \dots, L\}$, $N = T - L$, T is the sample size and the maximum lag order is constraint according to $L = \lceil 8(T/100)^{0.25} \rceil$. The maximum lag order is $L \in \{10, 14\}$ for samples $T \in \{200, 1000\}$.⁵ Finally, the lag order for an AR(p) model is estimated by the following simple rule $\hat{p} = \min_{l \in \{1, \dots, L\}} (BIC_l)$.

We also report the Q tests with automatically selected lag order based on Escanciano and Lobato (2009).⁶ The estimated lag order is selected by maximizing the following objective function

$$Q_l^* = Q_l - \pi_l,$$

$$\pi_l = \begin{cases} p \log(N) & \text{if } \max_{j \in \{1, \dots, L\}} |\hat{\rho}_{22}(j)| \leq \sqrt{c \log(N)/N}, \\ 2p & \text{if } \max_{j \in \{1, \dots, L\}} |\hat{\rho}_{22}(j)| > \sqrt{c \log(N)/N}, \end{cases}$$

⁵The sample size $T = 200$ is the common sample size in macro-econometrics, whereas $T = 1000$ is the standard sample in financial-econometrics. The results for $T = 500$ are available upon request from the author.

⁶Recall 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.

Table 2: List of non-linear models

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:

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

M3: An EXPAR model:

$$Y_t = c + (\phi_1 + (\phi_2 - \phi_1) \exp(-Y_{t-1}^2))Y_{t-1} + \sigma a_t,$$

M4: A MAR model:

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

M5: A MSAR model:

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

M6: A GARCH model:

$$Y_t = c + \phi Y_{t-1} + \epsilon_t, \quad \epsilon_t = a_t \sqrt{h_t},$$

$$h_t = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1},$$

M7: A NLGARCH model:

$$Y_t = c + \phi Y_{t-1} + \epsilon_t, \quad \epsilon_t = a_t \sqrt{h_t},$$

$$h_t = \omega + \alpha (|\epsilon_{t-1}| + \xi \epsilon_{t-1})^2 + \beta h_{t-1},$$

M8: A TMA model:

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

M9: A BL model:

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

M10: A NLMA model:

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

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

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 $\hat{m} = \max_{l \in \{1, \dots, L\}} (Q_l^*)$. The determined lag order \hat{m} is used for all three versions of the Q test. The reason for that lies in the fact that some parameters (e.g. constant $c = 2.4$) of the loss function Q^* are fine-tuned based on extensive Monte Carlo experiments. Provided we use the estimated cross-correlations (i.e. $\hat{\rho}_{12}(j)$ or $\hat{\rho}_{21}(j)$) in the above defined loss function, there is no guarantee that the fine-tuned parameters remain still “optimal”. Therefore, we leave this issue for further research.⁷

5 Monte Carlo results

5.1 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 $i \in \{1, \dots, K\}$ 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

$$\begin{aligned} avg &= \frac{1}{K} \sum_{i=1}^K \mathcal{P}_i, \\ min &= \min_{i \in \{1, \dots, K\}} (\mathcal{P}_i), \\ max &= \max_{i \in \{1, \dots, K\}} (\mathcal{P}_i), \end{aligned}$$

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

⁷However, it is interesting to point out that comparison of the $Q_{rs}(5)$, $Q_{rs}(10)$, $Q_{rs}(15)$, and $Q_{rs}(\hat{m})$ tests indicates that the effect of the lag order selection plays a minor role for the size and power properties, see Section 5 for MC results and a discussion.

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 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 $T \rightarrow \infty$ and $m/T \rightarrow 0$. 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 5. 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 4 depicts the χ^2 density function accompanied by the lower and upper bound of the smoothed empirical densities of the Q tests.⁸ 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}) compared to the Q test based on auto-correlations (i.e. Q_{22}).

Table 6 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: 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.⁹

5.2 Power results

Regime switching models: From Tables 7 and 8, 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 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 $\pi \rightarrow 0$ or $\pi \rightarrow 1$, one regime dominates the other and the process can be relatively well approximated by a simple AR(p) model, which

⁸A simple reference bandwidth is used for smoothing the empirical density functions of the Q tests.

⁹Results are available upon request.

negatively affects the power of the Q tests. Surprising results are obtained for an EX-PAR model, where none of the proposed Q tests exhibit any reasonable power even in large samples.

Surprisingly, rather different results are obtained for a MAR 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. Very similar results are obtained for a MSAR model as well.

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 MAR model if $\sigma_1^2 = \sigma_2^2$. For completeness, the sensitivity of the Q tests against the specification of a probability of switching of a MAR model is inspected as well. Two other probability specifications of a lower regime of a MAR model $\pi \in \{0.15, 0.85\}$ are considered. In both cases, the power of the Q tests decreases compared to the benchmark setup $\pi = 0.5$.

Conditional volatility models: It can be clearly concluded from Tables 7 and 8 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 very good power. Again, the Q_{12} is not at all informative, which is in line with theoretical results presented in Section 3. However, note that the power results of the Q_{21} test depend on a combination of asymmetry of innovations and a non-linear component, see Section 3 for theoretical justification. 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 paper), then the power of the Q_{21} and Q_{22} tests is around 0.9 in the sample $T = 1000$, see Table 10.

Other models: In the case of a NLMA model, Tables 7 and 8 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, see Tables 7 and 8. 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.

5.3 Sensitivity analysis

Parameters: Although Monte Carlo results confirm that the proposed Q tests can be useful in testing non-linearity in practice, it is important to point out very high variability of the power results: the minimum average rejection frequency over all parameter configurations, denoted as *min*, is close to the nominal level 0.05 for a vast majority of time series models, whereas the maximum average rejection frequency, denoted as *max*, is close to 1.0 for many time series models in large samples ($T = 1000$). For this reason, a special attention is paid to the sensitivity of the parameters of the data generating process. 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 3.

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 than 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, which is significantly more than

Innovations: The power properties of the proposed Q tests are examined against the following classes of distributions of innovations: a Gaussian distribution (N), symmetric

but leptokurtic (S), and asymmetric (A) innovations. The results are summarized in a coherent way in Table 4. The notation is as follows: “■” indicates that a given Q test exhibits a good power for a given non-linear process (i.e. $avg \geq 0.5$), whereas “□” indicates only a reasonable power of the Q test (i.e. $0.2 < avg < 0.5$), and “no-square” indicates almost no power of a given Q test (i.e. $avg \leq 0.2$). The table is reproduced from Monte Carlo results based on $T = 1000$ observations.

Table 4: Power properties of the Q tests: $T = 1000$, $R = 1000$

test	TAR			TMA			MAR			MSAR		
	N	S	A	N	S	A	N	S	A	N	S	A
Q_{12}	■	■	■	■	■	■				□	□	□
Q_{21}	■	■	■				□	□	□	□	□	□
Q_{22}	□	■	□	■	■	□	■	■	■	□	■	■
test	GARCH			NLGARCH			BL			NLMA		
	N	S	A	N	S	A	N	S	A	N	S	A
Q_{12}							■	■	■			
Q_{21}			□	■	■	■	□	□	□	■	■	■
Q_{22}	■	■	■	■	■	■	■	■	■			□

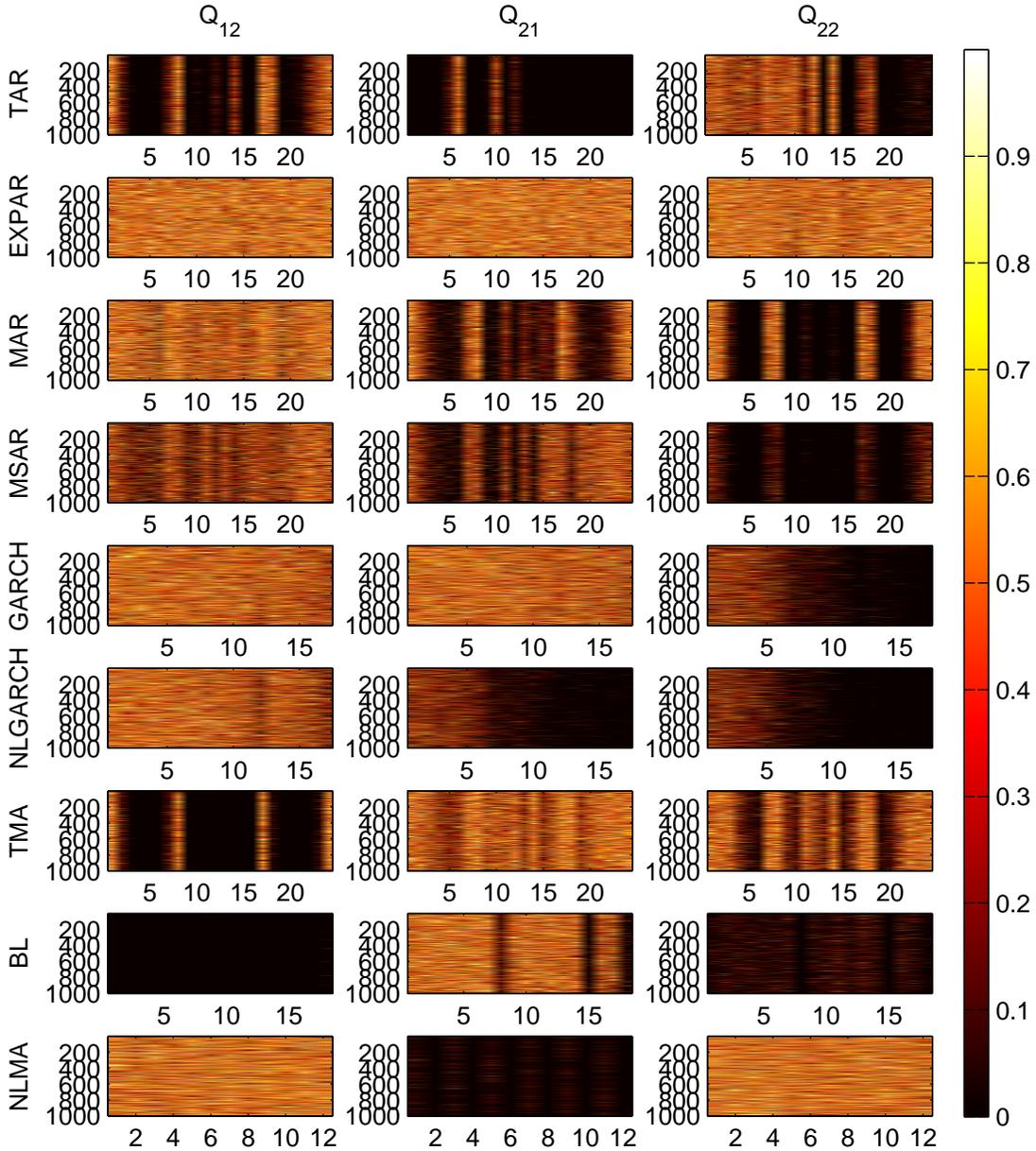
The table concludes the following. First, inspecting the cross-correlation structure can be useful tool supplementing the results from the auto-correlation structure. Second, the correlation structure of many non-linear models is rather different, which enables us to use the proposed Q tests for some preliminary discrimination among various classes of non-linear time series models. For example, it is easy to see a discrimination power for TAR and GARCH models, or BL and NLMA models.

6 Empirical application

In this section, the proposed Q tests are applied to a set of 22 financial time series: 5 exchange rate time series; 5 interest rate time series; 6 commodity time series; and finally 6 equity indices. We use average weekly returns from 1980 to 2010 consisting of 1620 observations.¹⁰ For a complete description of time series, see Table 11 in Appendix B.

¹⁰Weakly time series seem to serve as a good compromise between daily returns, which are too noisy and contaminated by jumps, and monthly returns, which are too aggregated.

Figure 3: Power images of the Q tests: $T = 1000$, $R = 1000$, $a \sim N(0, 1)$



* 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 .

In this exercise, we are particularly interested in the following three 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 the generalized correlation structure of various asset classes similar or rather different?”; (iii) “Is a simple GARCH model an adequate model for asset returns in general?”.

The results are presented in Table 12. It can be concluded from the table that the Q_{22} test rejects linearity for all 22 series of asset returns regardless of the lag order specification m of the Q tests. This finding clearly implies that an application of the standard Q_{22} , or analogically the MLQ test, seems to be fully sufficient from the linearity testing standpoint. The results clearly confirm previous studies that standard linear homoscedastic ARMA models are not adequate for modelling selected asset returns at all. Individual results suggest that the behaviour of the correlation structure of equity returns differ significantly from the correlation structure of other asset classes: the Q_{21} and Q_{22} tests are statistically significant for all equity returns regardless of the lag order m . However, this is not the case for other asset classes. On the other hand, when considering all three Q tests, the correlation structure of asset returns seems to be series dependent rather than asset class dependent. For example, values of the Q_{12} and Q_{21} tests are statistically significant in 11 and 8 cases out of 22. This means that, although cross-correlations seem not to give us additional information from the linearity testing point of view, they might be very useful for some preliminary discrimination. All in all, the results of the Q tests suggest the following: (a) due to the statistical significance of the Q_{22} test, but insignificance of the Q_{12} and also Q_{21} , a simple GARCH model seems to be appropriate for only 7 out of 22 asset returns - USDCHF, USDCAD, USIR3M, AUIR3M, CHIR3M, WHEAT, GOLD; (b) due to the statistical significance of the Q_{12} , Q_{21} and Q_{22} tests, a TARMA-GARCH model seems to be appropriate for 5 out of 22 asset returns - USDJPY, DJIA, FTSE, TOPIX, SOYBEANS; and (c) due to the statistical significance of only Q_{21} and Q_{22} tests, some non-linear and/or asymmetric GARCH model seems to be appropriate for 6 out of 22 asset returns - USDGBP, USDAUD, AUSE, TSE, CHSE, COFFEE. Note that the above classification of asset returns is just an example of the above described discrimination power of the Q tests.

7 Conclusion

It is demonstrated, both analytically and by means of Monte Carlo experiments, that inspecting residual auto-correlations and cross-correlations can be an useful, yet very simple, tool both for testing linearity and for some preliminary model discrimination.

Our results suggest that the proposed Q tests, using a more efficient variance-stabilizing transformation, have good size and power properties for all non-linear models but an EXPAR model. It is also shown that the power properties of the Q tests are sensitive to the parameter configuration for some non-linear models (e.g. regime-switching models). Therefore, the results should be interpreted with caution. In addition, the results reveal that the correlation structure of 9 considered non-linear models is rather different, This finding can be easily used for preliminary discrimination between some rival non-linear models such as TAR and GARCH models, or BL and NLMA models.

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, Walker A. 1964. On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process. *The Annals of Mathematical Statistics* **35**: 1296–1303.
- Box G, Pierce D. 1970. Distributions of residual autocorrelations in autoregressive integrated moving average time series models. *Journal of the American Statistical Association* **65**: 1509–1525.
- Ding Z, Granger C, Engle R. 1993. A long memory property of stock market returns and a new model. *Journal of Empirical Finance* **1**: 83–106.
- Escanciano J, Lobato I. 2009. An automatic Portmanteau test for serial correlation. *Journal of Econometrics* **151**: 140–149.
- Horowitz J, Lobato I, Nankervis J, Savin N. 2006. Bootstrapping the Box-Pierce Q test: A robust test of uncorrelatedness. *Journal of Econometrics* **133**: 841–862.
- Jansen D, Vries CD. 1991. On the frequency of large stock returns: Putting booms and busts into perspective. *The Review of Economics and Statistics* **73**: 18–24.
- Johnson N, Kotz S, Balakrishnan N. 1994. *Continuous Univariate Distributions*. Wiley.
- Kendall M, Ord J. 1973. *Time Series*. Griffin London.
- Konishi S. 1978. An approximation to the distribution of the sample correlation coefficient. *Biometrika* **65**: 654–656.
- Kwan A, Sim A. 1996a. On the finite-sample distribution of modified portmanteau tests for randomness of a Gaussian time series. *Biometrika* **83**: 938–943.
- Kwan A, Sim A. 1996b. Portmanteau tests of randomness and Jenkins' variance-stabilizing transformation. *Economics Letters* **50**: 41–49.
- Lawrance A, Lewis P. 1985. Modelling and residual analysis of nonlinear autoregressive time series in exponential variables. *Journal of the Royal Statistical Society* **47**: 165–202.

- Lawrance A, Lewis P. 1987. Higher-order residual analysis for nonlinear time series with autoregressive correlation structures. *International Statistical Review* **55**: 21–35.
- Li W. 1992. On the asymptotic standard errors of residual autocorrelations in nonlinear time series modelling. *Biometrika* **79**: 435–437.
- Li W, Mak T. 1994. On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity. *Journal of Time Series Analysis* **15**: 627–636.
- Ljung G, Box G. 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**: 1066–1076.
- Lobato I, Nankervis J, Savin N. 2002. Testing for zero autocorrelation in the presence of statistical dependence. *Econometric Theory* **18**: 730–743.
- Loretan M, Phillips P. 1994. Testing the covariance stationarity of heavy-tailed time series. *Journal of Empirical Finance* **1**: 211–248.
- Maravall A. 1983. An application of nonlinear time series forecasting. *Journal of Business & Economic Statistics* **1**: 66–74.
- McLeod A, Li W. 1983. Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis* **4**: 269–273.
- Ng S, Perron P. 2005. A note on the selection of time series models. *Oxford Bulletin of Economics and Statistics* **67**: 115–134.
- Pérez A, Ruiz E. 2003. Properties of the sample autocorrelations of nonlinear transformations in long-memory stochastic volatility models. *Journal of Financial Econometrics* **1**: 420–444.
- Randles R, Fligner M, Policello G, Wolfe D. 1980. An asymptotically distribution-free test for symmetry versus asymmetry. *Journal of the American Statistical Association* **75**: 168–172.
- Romano J, Thombs L. 1996. Inference for Autocorrelations under Weak Assumptions. *Journal of the American Statistical Association* **91**: 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**: 205–216.

- Schwarz G. 1978. Estimating the dimension of a model. *The Annals of Statistics* **6**: 461–464.
- Tong H. 1990. *Non-linear Time Series: a dynamical system approach*. Oxford University Press.
- Vavra M. 2012. *Testing and Modelling Non-linear Time Series*. Ph.D. Thesis, University of London.
- 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, Brockwell P. 2006. Gaussian maximum likelihood estimation for arma models. *Journal of time series analysis* **27**: 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) $\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{d} N(\mathbf{0}, \mathbf{V})$. \square

Proof. See a proof to Theorem 8.4.1 in ?, p. 432. \blacksquare

Remark 1 Note that Yao and Brockwell (2006) obtained the same results for the ML estimator of ARMA parameters. \square

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

Proof. See a proof to Theorem 7.2.1 in ?, p. 221 with a restriction $X_t = Z_t$. \blacksquare

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 . \square

Proof. See a proof to 2c.4(ix) result in ?, p. 122. \blacksquare

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 $0 < r_t \rightarrow 0$ as $t \rightarrow \infty$. 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 $g^0 = g$. \square

Proof. See a proof to Proposition 6.1.5 in ?, p. 201. \blacksquare

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

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

for the lag order $k \in \{1, \dots, m\}$ 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

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

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

Proof.

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

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 ?, 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 ?, 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 ?, p. 24. ■

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

$$\hat{\gamma}_{rs}(0) = \sqrt{\left[\frac{1}{T} \sum_{t=1}^T g_r^2(\hat{a}_t) \right] \left[\frac{1}{T} \sum_{t=1}^T g_s^2(\hat{a}_t) \right]}, \quad (11)$$

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

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

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 ?, 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 ?, 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

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

for any integer $r > 0$.

(ii) It follows from Theorem 3 that $\hat{\beta} \xrightarrow{p} \beta$, which implies that $a_t(\hat{\beta}) \equiv \hat{a}_t \xrightarrow{p} a_t$. Then, since $g_r(\cdot)$ is a continuous function in β , 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

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

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 $\text{rk}(\mathbf{C}) = r$, then $\mathbf{z}'\mathbf{C}\mathbf{z} \sim \chi^2(r)$. \square

Proof. See a proof for Theorem 9.8 in ?, p. 378. \blacksquare

A.2 Proof of Theorem 1

Let $\boldsymbol{\beta} = (\xi_1, \dots, \xi_p, \zeta_1, \dots, \zeta_q, \sigma)'$ 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

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

for the lag order $k \in \{1, \dots, m\}$ 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 $\partial \gamma_{rs}(k) / \partial \beta_i = O_p(T^{-1/2})$, 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

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

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 $|\hat{\boldsymbol{\rho}}_{rs} - \boldsymbol{\rho}_{rs}| \xrightarrow{p} \mathbf{0}$ 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 $k \in \{1, \dots, m\}$ and integer $m > 0$. It is now easy to see that vectors $\hat{\mathbf{z}}_{rs} = (z_{rs}(1), \dots, z_{rs}(m))'$ and $\hat{\boldsymbol{\rho}}_{rs} = (\hat{\rho}_{rs}(1), \dots, \hat{\rho}_{rs}(m))'$ 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 $\hat{\mathbf{z}}_{rs} = (\hat{z}_{rs}(1), \dots, \hat{z}_{rs}(m))'$ 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

Table 5: Statistical properties of the Q tests

sample	test	lag m	AR (#8)			MA (#8)		
			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 6: Size of the Q tests

sample	test	lag m	AR (#8)			MA (#8)		
			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 7: Power properties of the Q tests: $N(0,1)$, $T = 200$

test	lag	TAR (#24)			EXPAR (#24)			MAR (#24)			MSAR (#24)			GARCH (#17)			NLGARCH (#17)			TMA (#24)			BL (#18)			NLMA (#12)		
		avg	min	max	avg	min	max	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	0.05	0.03	0.06	0.09	0.04	0.23	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	0.05	0.03	0.07	0.08	0.03	0.21	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	0.05	0.03	0.06	0.07	0.03	0.20	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	0.05	0.04	0.10	0.10	0.03	0.22	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	0.05	0.04	0.07	0.18	0.04	0.45	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	0.05	0.03	0.06	0.15	0.03	0.39	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	0.05	0.04	0.07	0.13	0.04	0.36	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	0.05	0.03	0.07	0.21	0.05	0.49	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	0.05	0.04	0.07	0.45	0.04	1.00	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	0.05	0.04	0.07	0.42	0.05	1.00	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	0.06	0.04	0.07	0.40	0.05	1.00	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	0.05	0.04	0.08	0.51	0.04	1.00	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 8: Power properties of the Q tests: $T = 1000$

test	lag	TAR (#24)			EXPAR (#24)			MAR (#24)			MSAR (#24)			GARCH (#17)			NLGARCH (#17)			TMA (#24)			BL (#18)			NLMA (#12)		
		avg	min	max	avg	min	max	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	0.05	0.04	0.08	0.10	0.04	0.25	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	0.05	0.04	0.07	0.09	0.05	0.24	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	0.05	0.04	0.06	0.09	0.04	0.23	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	0.06	0.04	0.09	0.11	0.04	0.24	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	0.05	0.04	0.07	0.46	0.03	0.88	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	0.05	0.04	0.06	0.39	0.05	0.85	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	0.05	0.04	0.06	0.36	0.04	0.81	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	0.05	0.04	0.07	0.49	0.04	0.89	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	0.05	0.03	0.06	0.63	0.04	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	0.05	0.04	0.07	0.61	0.04	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	0.05	0.04	0.06	0.60	0.04	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	0.06	0.04	0.11	0.65	0.05	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 9: Power properties of the Q tests: $T = 200$

test	distr.	skew.	kurt.	TAR (#24)			EXPAR (#24)			MAR (#24)			MSAR (#24)			GARCH (#17)			NLGARCH (#17)			TMA (#24)			BL (#18)			NLMA (#12)		
				avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max
Q_{12}	N	0.0	3.0	0.36	0.04	0.93	0.05	0.04	0.10	0.10	0.03	0.22	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
	S1	0.0	6.0	0.42	0.04	0.95	0.06	0.04	0.08	0.13	0.03	0.29	0.16	0.07	0.28	0.11	0.06	0.20	0.13	0.07	0.27	0.63	0.04	1.00	0.94	0.69	0.98	0.06	0.04	0.07
	S2	0.0	11.6	0.44	0.05	0.95	0.06	0.04	0.10	0.15	0.05	0.28	0.17	0.06	0.33	0.13	0.08	0.20	0.15	0.07	0.27	0.64	0.04	1.00	0.95	0.70	0.99	0.06	0.05	0.08
	S3	0.0	126.0	0.46	0.06	0.94	0.06	0.03	0.10	0.18	0.04	0.34	0.19	0.05	0.32	0.15	0.09	0.24	0.17	0.10	0.28	0.65	0.06	1.00	0.95	0.76	0.99	0.06	0.04	0.07
	A1	-0.9	4.2	0.35	0.03	0.94	0.07	0.03	0.24	0.10	0.04	0.21	0.16	0.07	0.30	0.07	0.04	0.12	0.11	0.04	0.28	0.57	0.03	1.00	0.83	0.45	0.97	0.03	0.03	0.04
	A2	-1.5	7.5	0.35	0.03	0.96	0.09	0.03	0.37	0.12	0.04	0.28	0.18	0.07	0.33	0.07	0.04	0.11	0.13	0.04	0.32	0.56	0.03	0.99	0.83	0.44	0.96	0.03	0.02	0.03
	A3	-2.0	21.2	0.40	0.03	0.98	0.07	0.03	0.22	0.15	0.04	0.33	0.19	0.06	0.35	0.10	0.06	0.20	0.16	0.06	0.36	0.61	0.04	1.00	0.91	0.66	0.98	0.04	0.03	0.05
Q_{21}	N	0.0	3.0	0.63	0.06	0.98	0.05	0.03	0.07	0.21	0.05	0.49	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
	S1	0.0	6.0	0.48	0.07	0.93	0.05	0.04	0.07	0.21	0.04	0.48	0.21	0.07	0.56	0.10	0.06	0.16	0.23	0.11	0.40	0.08	0.05	0.13	0.15	0.08	0.45	0.30	0.13	0.48
	S2	0.0	11.6	0.42	0.07	0.87	0.05	0.04	0.07	0.22	0.04	0.49	0.22	0.07	0.59	0.13	0.07	0.21	0.23	0.10	0.38	0.09	0.05	0.15	0.21	0.11	0.43	0.26	0.11	0.41
	S3	0.0	126.0	0.36	0.07	0.80	0.05	0.04	0.08	0.23	0.05	0.51	0.22	0.07	0.58	0.14	0.08	0.21	0.22	0.11	0.37	0.11	0.05	0.17	0.27	0.19	0.44	0.21	0.10	0.32
	A1	-0.9	4.2	0.56	0.08	0.96	0.06	0.04	0.08	0.28	0.05	0.76	0.21	0.07	0.51	0.14	0.08	0.23	0.41	0.18	0.67	0.09	0.06	0.20	0.19	0.09	0.85	0.42	0.21	0.64
	A2	-1.5	7.5	0.47	0.08	0.89	0.07	0.06	0.10	0.32	0.05	0.89	0.23	0.08	0.50	0.18	0.09	0.31	0.44	0.22	0.67	0.12	0.06	0.24	0.34	0.15	0.95	0.38	0.18	0.58
	A3	-2.0	21.2	0.39	0.09	0.81	0.06	0.05	0.09	0.28	0.05	0.74	0.22	0.07	0.51	0.16	0.08	0.26	0.37	0.17	0.60	0.11	0.06	0.23	0.35	0.16	0.89	0.27	0.13	0.42
Q_{22}	N	0.0	3.0	0.22	0.05	0.82	0.05	0.04	0.08	0.51	0.04	1.00	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
	S1	0.0	6.0	0.25	0.05	0.82	0.06	0.04	0.09	0.55	0.04	1.00	0.54	0.08	0.98	0.27	0.11	0.54	0.34	0.11	0.69	0.23	0.05	0.64	0.66	0.54	0.97	0.07	0.04	0.08
	S2	0.0	11.6	0.27	0.05	0.83	0.06	0.04	0.09	0.56	0.04	1.00	0.55	0.09	0.98	0.28	0.11	0.54	0.34	0.12	0.65	0.32	0.07	0.80	0.79	0.68	0.99	0.07	0.05	0.08
	S3	0.0	126.0	0.30	0.07	0.85	0.06	0.04	0.10	0.58	0.05	1.00	0.54	0.07	0.98	0.28	0.13	0.52	0.33	0.14	0.58	0.40	0.08	0.91	0.89	0.82	1.00	0.07	0.05	0.09
	A1	-0.9	4.2	0.11	0.03	0.30	0.06	0.04	0.07	0.52	0.04	1.00	0.55	0.11	0.98	0.26	0.09	0.57	0.51	0.20	0.90	0.14	0.05	0.45	0.38	0.19	0.95	0.11	0.08	0.14
	A2	-1.5	7.5	0.12	0.02	0.46	0.07	0.04	0.10	0.55	0.04	1.00	0.56	0.10	0.98	0.28	0.12	0.54	0.54	0.24	0.88	0.19	0.06	0.64	0.56	0.37	0.98	0.11	0.08	0.14
	A3	-2.0	21.2	0.20	0.04	0.60	0.06	0.04	0.11	0.56	0.04	1.00	0.55	0.08	0.99	0.27	0.11	0.56	0.48	0.19	0.83	0.27	0.07	0.81	0.76	0.60	0.99	0.07	0.06	0.08

^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 10: Power properties of the Q tests: $T = 1000$

test	distr.	skew.	kurt.	TAR (#24)			EXPAR (#24)			MAR (#24)			MSAR (#24)			GARCH (#17)			NLGARCH (#17)			TMA (#24)			BL (#18)			NLMA (#12)		
				avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max	avg	min	max
Q_{12}	N	0.0	3.0	0.64	0.04	1.00	0.06	0.04	0.09	0.11	0.04	0.24	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
	S1	0.0	6.0	0.69	0.04	1.00	0.06	0.04	0.08	0.15	0.04	0.32	0.26	0.10	0.47	0.16	0.07	0.30	0.20	0.08	0.48	0.80	0.05	1.00	1.00	1.00	1.00	0.06	0.05	0.07
	S2	0.0	11.6	0.72	0.06	1.00	0.06	0.05	0.10	0.20	0.04	0.40	0.28	0.12	0.49	0.21	0.10	0.41	0.27	0.12	0.57	0.81	0.05	1.00	1.00	0.99	1.00	0.06	0.04	0.08
	S3	0.0	126.0	0.74	0.04	1.00	0.07	0.05	0.12	0.27	0.06	0.47	0.32	0.09	0.54	0.28	0.15	0.50	0.33	0.16	0.60	0.82	0.06	1.00	0.99	0.97	1.00	0.07	0.05	0.10
	A1	-0.9	4.2	0.64	0.04	1.00	0.19	0.03	0.77	0.11	0.04	0.28	0.31	0.15	0.60	0.09	0.05	0.20	0.21	0.05	0.61	0.78	0.03	1.00	1.00	0.99	1.00	0.03	0.03	0.04
	A2	-1.5	7.5	0.63	0.04	1.00	0.27	0.03	0.96	0.14	0.04	0.34	0.33	0.14	0.64	0.12	0.05	0.29	0.28	0.08	0.72	0.79	0.03	1.00	1.00	0.95	1.00	0.02	0.02	0.03
Q_{21}	A3	-2.0	21.2	0.69	0.05	1.00	0.19	0.04	0.70	0.21	0.05	0.42	0.35	0.11	0.63	0.20	0.09	0.43	0.36	0.12	0.73	0.81	0.06	1.00	0.99	0.89	1.00	0.04	0.03	0.05
	N	0.0	3.0	0.89	0.07	1.00	0.05	0.04	0.07	0.49	0.04	0.89	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
	S1	0.0	6.0	0.80	0.12	1.00	0.05	0.03	0.07	0.48	0.04	0.87	0.41	0.09	0.88	0.15	0.08	0.28	0.65	0.27	0.96	0.17	0.05	0.43	0.30	0.13	0.89	0.81	0.59	1.00
	S2	0.0	11.6	0.77	0.13	1.00	0.06	0.04	0.08	0.47	0.04	0.84	0.42	0.10	0.87	0.19	0.09	0.36	0.62	0.24	0.93	0.20	0.05	0.49	0.41	0.25	0.87	0.75	0.50	0.98
	S3	0.0	126.0	0.73	0.13	1.00	0.06	0.05	0.10	0.46	0.06	0.79	0.43	0.09	0.85	0.25	0.12	0.51	0.60	0.27	0.90	0.25	0.07	0.56	0.51	0.38	0.86	0.66	0.39	0.92
	A1	-0.9	4.2	0.84	0.13	1.00	0.07	0.05	0.10	0.51	0.05	1.00	0.44	0.09	0.76	0.34	0.13	0.68	0.87	0.57	1.00	0.17	0.07	0.59	0.50	0.18	1.00	0.87	0.71	1.00
Q_{22}	A2	-1.5	7.5	0.80	0.10	1.00	0.09	0.05	0.19	0.48	0.06	1.00	0.47	0.09	0.95	0.49	0.21	0.81	0.90	0.66	1.00	0.23	0.10	0.75	0.78	0.52	1.00	0.82	0.60	1.00
	A3	-2.0	21.2	0.77	0.13	1.00	0.09	0.06	0.20	0.49	0.06	1.00	0.46	0.09	0.89	0.42	0.19	0.70	0.86	0.57	1.00	0.27	0.10	0.77	0.78	0.47	1.00	0.75	0.48	0.97
	N	0.0	3.0	0.45	0.06	1.00	0.06	0.04	0.11	0.65	0.05	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
	S1	0.0	6.0	0.55	0.07	1.00	0.07	0.04	0.17	0.69	0.04	1.00	0.80	0.22	1.00	0.70	0.25	1.00	0.79	0.36	1.00	0.56	0.08	1.00	1.00	1.00	1.00	0.07	0.05	0.08
	S2	0.0	11.6	0.62	0.07	1.00	0.09	0.04	0.19	0.71	0.05	1.00	0.79	0.19	1.00	0.72	0.31	1.00	0.78	0.35	1.00	0.63	0.08	1.00	1.00	1.00	1.00	0.08	0.06	0.10
	S3	0.0	126.0	0.67	0.07	1.00	0.10	0.04	0.26	0.72	0.06	1.00	0.78	0.16	1.00	0.72	0.33	1.00	0.78	0.40	0.99	0.70	0.09	1.00	1.00	1.00	1.00	0.09	0.06	0.12
Q_{22}	A1	-0.9	4.2	0.37	0.05	0.92	0.07	0.04	0.15	0.66	0.05	1.00	0.87	0.46	1.00	0.70	0.26	1.00	0.92	0.64	1.00	0.34	0.07	0.99	0.87	0.68	1.00	0.33	0.15	0.52
	A2	-1.5	7.5	0.38	0.04	1.00	0.08	0.05	0.21	0.69	0.04	1.00	0.86	0.38	1.00	0.74	0.32	1.00	0.94	0.73	1.00	0.42	0.08	1.00	0.99	0.96	1.00	0.45	0.20	0.71
	A3	-2.0	21.2	0.50	0.04	1.00	0.09	0.04	0.28	0.71	0.06	1.00	0.82	0.24	1.00	0.74	0.33	0.99	0.93	0.68	1.00	0.56	0.08	1.00	1.00	1.00	1.00	0.22	0.10	0.35

^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 11: Description of weakly time series: sample 07/01/1980 – 31/12/2010

variable	description	transformation
Exchange rates		
USDGBP	US dollar to British pound exchange rate	$\Delta \log$
USDJPY	US dollar to Japanese yen exchange rate	$\Delta \log$
USDCAD	US dollar to Canadian dollar exchange rate	$\Delta \log$
USDAUD	US dollar to Australian dollar exchange rate	$\Delta \log$
USDCHF	US dollar to Swiss frank exchange rate	$\Delta \log$
Interest rates		
USIR3M	US interbank interest rates, 3M	Δ
UKIR3M	UK interbank interest rates, 3M	Δ
CAIR3M	Canadian interbank interest rates, 3M	Δ
AUIR3M	Australian interbank interest rates, 3M	Δ
CHFIR3M	Swiss interbank interest rates, 3M	Δ
Equity indices		
DJIA	US Dow Jones Industrials Share Index	$\Delta \log$
FTSE	UK FT All Shares Index	$\Delta \log$
TOPIX	Tokyo Stock Exchange Index	$\Delta \log$
TSE	Toronto Stock Exchange Index	$\Delta \log$
AUSE	Australian Stock Exchange Index	$\Delta \log$
CHSE	Swiss Stock Exchange Index	$\Delta \log$
Commodities		
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	$\Delta \log$

* Source: Thomson Reuters.

^a “ Δ ” denotes a first difference of a given series, “ $\Delta \log$ ” is an approximation to the growth rate of a given time series.

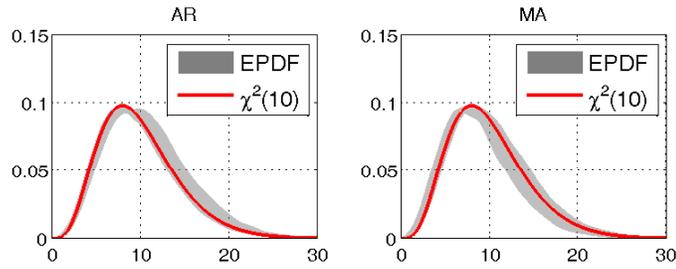
Table 12: Application of the Q tests

variable/lag	Q_{12}				Q_{21}				Q_{22}			
	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

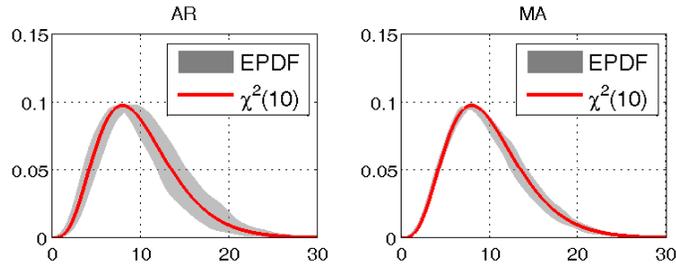
^a m denotes the automatically selected lag order based on Escanciano and Lobato (2009).

C Figures

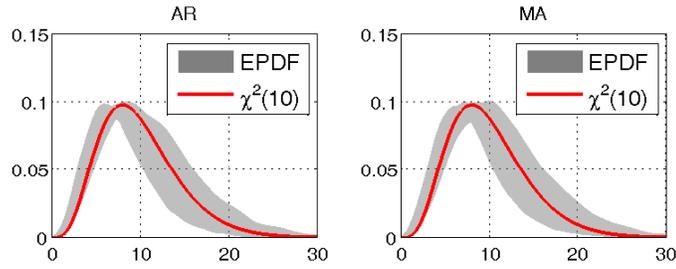
Figure 4: Accuracy of the smoothed empirical density functions of the $Q(10)$ tests to the limiting $\chi^2(10)$: $T = 200$, $R = 5000$



(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.