A BOOTSTRAP APPROACH TO TEST THE CONDITIONAL SYMMETRY IN TIME SERIES MODELS[*]

Alicia Pérez-Alonso

WP-AD 2006-18

Correspondence to: Alicia Pérez Alonso, Universidad de Alicante, Departamento de Fundamentos del Análisis Económico, Campus San Vicente del Raspeig, 99-E-03080, Alicante, España. E-mail: alicia@merlin.fae.ua.es.

IVIE working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

BOOTSTRAP APPROACH TO TEST THE CONDITIONAL SYMMETRY IN TIME SERIES MODELS

Alicia Pérez-Alonso

ABSTRACT

This paper discusses how to test for conditional symmetry in time series regression models. To that end, we utilize the Bai and Ng test. We also examine the performance of some popular (unconditional) symmetry tests for observations when applied to regression residuals. The tests considered include the coeficient of skewness, a joint test of the third and fifth moments, the Runs test, the Wilcoxon signed-rank test and the Triples test. An easy-to-implement symmetric bootstrap procedure is proposed to calculate critical values for these tests. Consistency of the bootstrap procedure will be shown. A simple Monte Carlo experiment is conducted to explore the finite-sample properties of all the tests.

1 Introduction

The problem of testing conditional symmetry in time series data is fundamental in both theoretical and empirical research. In the last few years considerable research has been devoted to model and forecast the conditional mean and the conditional variance of financial time series, that is, the return and risk of financial assets, respectively. The class of (Generalized) Autoregressive Conditional Heteroskedasticity ((G)ARCH) models, introduced by Engle [12] and Bollerslev [6], is the most widely used among economists and other applied practitioners to model time varying conditional variances. In essence, all empirical studies that assume conditional heteroskedasticity also use a quasi-maximum likelihood estimator (QMLE). If the likelihood is assumed to be Gaussian, the QMLE is known to be consistent if the conditional mean and the conditional variance are correctly specified. However, normality of innovations is frequently not a very realistic assumption for high-frequency financial time series because the resulting model fails to capture the kurtosis in the data. Alternative distributions for innovations are considered in the literature. For example, following Bollerslev [7], a popular choice is the standardized Student-t distribution. If the likelihood is assumed to be non-Gaussian, Newey and Steigerwald [27] show that consistency of a QMLE requires that both the assumed innovation density and the true innovation density are unimodal and symmetric around zero. Moreover, if conditional symmetry fails, an additional parameter is needed to ensure consistency of a non-Gaussian QMLE. The additional parameter accounts for the location of the innovation density. The reader may refer to the work of Franses and Van Dijk [14] for an extensive survey of the recent developments of modelling, estimation and hypothesis testing for time-varying conditional variance models.

Whether or not conditional symmetry holds is also an issue of interest for adaptive estimation. An adaptive estimator shares the asymptotic optimality properties of the maximum likelihood estimator, differing from it in that a nonparametric estimator of the score function of the log likelihood replaces the analytic expression that would be used if the actual functional form of the disturbance distribution was known. Bickel [5] shows that if the density function of the disturbance is symmetric about the origin, then the parameters of a linear regression model can be estimated adaptively. Newey [26] constructs adaptive estimators of linear regression parameters by a generalized method of moments (GMM) when the foregoing is true. The above results are extended to stationary autoregressive moving average (ARMA) process by Kreiss [23] and reduced-rank vector error correction models by Hodgson [21]. In the case of testing, the efficiency of the methods can be improved under the additional assumption of a symmetric error distribution, see for example Azzalini and Bowman [1] or Kulasekera and Wang [24]. Further, conditional symmetry is part of the stochastic restrictions on unobservable errors used in semiparametric modelling (see Powel [28] and references therein). The conditional symmetry restriction implies constant conditional mean and median, which is quite familiar in econometric theory and practice.

The conventional asymptotic theory of the bootstrap relies on Edgeworth expansions in order to prove the existence of asymptotic refinements. In many cases the efficiency of this method can be improved under the additional assumption of symmetry. Davidson and Flachaire [9] study

2

various versions of the wild bootstrap applied to a linear regression model with heteroskedastic errors. They show that when the error terms are symmetrically distributed about the origin, the wild bootstrap applied to heteroskedasticity consistent covariance matrix estimator based statistics benefits from better asymptotic refinements than when they are asymmetrically distributed. In particular, they found that the error in rejection probability (ERP) is at most of order $T^{-3/2}$ with symmetric errors and $T^{-1/2}$ with asymmetric errors, where T denotes the sample size. Comparable results are obtained by Hall [18] for the case of homoskedastic regression models. He shows that bootstrap tests on the slope parameters benefit from refinements in the case of unskewed error terms.

There is also a growing literature addressing the problem of conditional symmetry of macroeconomic time series related to asymmetries in business cycles. As discussed in Brunner [8], the assumption of Gaussian shocks places strong restrictions on the time series behaviour of economic fluctuations. Since the Gaussian distribution is symmetric about zero, the conditional density is symmetric about its conditional mean. Our notion of conditional symmetry is that, in an expansion (contraction), the probability of further expansion (contraction), relative to the conditional mean, is equal to the probability of a contraction (expansion). That is, positive shocks to the conditional mean are as likely as negative shocks. There is a substantial body of empirical evidence that suggests that business cycles expansions appear to be more persistent and less volatile than contractions. That is, economic time series behave asymmetrically over the business cycle; see e.g., DeLong and Summers [10], Hussey [22], Verbrugge [31] and Belaire-Franch and Contreras [3]. Thus, symmetry tests are an essential first step in practical model-building exercises since it is desirable to establish the validity or otherwise of the symmetry assumption before exploring more complicated business cycle structures.

Tests for symmetry have a long tradition in both Statistics and Econometrics. In this paper, we focus on the evaluation of several statistical testing procedures that can be used to test for conditional symmetry. In particular, we consider the nonparametric test for conditional symmetry of Bai and Ng [2]. The closely related problem of testing for (unconditional) symmetry was investigated by Wilcoxon (see [16]), Gupta [17], McWilliams [25] and Randles *et al.* [30] among others. It is not clear whether these tests can be extended to testing for conditional symmetry, since it has not yet been rigorously demonstrated that statistics computed by using regression residuals instead of the true errors have approximately the same distribution as those based on the errors. It is by no means obvious that this is so. However, for the case of tests of symmetry based on sample moments, we show that under standard regularity conditions that ensure asymptotic normality of moment estimators, the asymptotic null distributions of the tests do not change when replacing the unknown errors by well-behaved residuals. Another problem encountered when using real data is that, for finite samples, the distribution of the symmetry tests included in this study is still unknown. As a consequence, the true size of these tests often differs to a large extent from its nominal size based on asymptotic critical values. The main purpose of this paper is to investigate whether the bootstrap can be used to obtain improved finite-sample critical values.

The remainder of the paper is organized as follows. Section 2 details the class of nonlinear

dynamic processes under which we will work. In Section 3, we briefly review all the tests for conditional symmetry used in this paper. Section 4 describes the bootstrap method and establishes consistency property of the bootstrap for a linear regression model. Section 5 performs, for a wide variety of alternative symmetric and asymmetric distributions, Monte Carlo simulations to compare the finite-sample size and power of the tests when critical values are obtained using a bootstrap procedure with that we could achieve using the asymptotic theory. Concluding comments are presented in Section 6. Technical proofs of all results are deferred to Appendix A. Finally, tables are relegated to Appendix B.

2   The nonlinear dynamic model

Suppose that $\{(Y_t, X_t)\}$ is a strictly stationary discrete-time stochastic process with $Y_t \in \mathbb{R}$ and $X_t \in \mathbb{R}^d$, defined on some probability space $(\Omega, \mathcal{F}, P)$. Here, X_t is a vector containing both explanatory variables and lagged values of Y_t. That is, $X_t = (Z_t, Y_{t-1}, ..., Y_{t-p})'$, where $Z_t \in \mathbb{R}^{d-p}$ is a vector of some explanatory variables. Let Y_t and X_t be both defined based on a stationary process $\{V_t\}$ by

$$
\begin{aligned}
Y_t &= \Psi_Y(V_t, V_{t-1}, V_{t-2}, ...), \\
X_t &= (X_{t1}, ..., X_{td})' = \Psi_X(V_t, V_{t-1}, V_{t-2}, ...),
\end{aligned}
$$

where $\Psi_Y : \mathbb{R}^\infty \to \mathbb{R}$ and $\Psi_X : \mathbb{R}^\infty \to \mathbb{R}^d$ are two Borel measurable functions, respectively, and $\{V_t\}$ may be vector-valued. We see that $\{(Y_t, X_t)\}$ depends upon the infinite history of $\{V_t\}$. Let $r > 0$ be a positive real number. Following Gallant and White [15], we define $\{(Y_t, X_t)\}$ to be L_r-*near epoch dependent* (L_r-NED) with respect to a stationary process $\{V_t\}$, provided $E |Y_t|^r < \infty$ and $\upsilon_r(m) = E \left|Y_t - Y_t^{(m)}\right|^r + E \left\|X_t - X_t^{(m)}\right\|^r \to 0$ as $m \to \infty$, where $|\cdot|$ and $\|\cdot\|$ are the absolute value and the Euclidean norm of $\mathbb{R}^d$, respectively, $Y_t^{(m)} = \Psi_{Y,m}(V_t, V_{t-1}, ..., V_{t-m+1})$, $X_t^{(m)} = (X_{t1}^{(m)}, ..., X_{td}^{(m)})' = \Psi_{X,m}(V_t, V_{t-1}, ..., V_{t-m+1})$, and $\Psi_{Y,m}$ and $\Psi_{X,m}$ are $\mathbb{R}$- and $\mathbb{R}^d$-valued Borel measurable functions with m arguments involved, respectively. In particular, if $\upsilon_r(m) = O(m^{-a-\lambda})$ for some $\lambda > 0$ we say $\{(Y_t, X_t)\}$ is L_r-NED of size $-a$. The more negative $-a$ is, the more quickly the dependence of $\{(Y_t, X_t)\}$ on past values of V_t dies out. We will call $\upsilon_r(m)$ the stability coefficients of order r of the process $\{(Y_t, X_t)\}$. Since NED is only a measure of how $\{(Y_t, X_t)\}$ depends on $\{V_t\}$, we place no conditions here on the dependence properties of $\{V_t\}$.

We are interested in the conditional distribution of Y_t conditional on X_t. Conditional symmetry implies that the distribution of Y_t, given X_t, has a symmetric form about its conditional mean. That is to say, $f_t(y + \mu_t/X_t) = f_t(-y + \mu_t/X_t)$, where $f_t(\cdot/X_t)$ is the density of Y_t conditional on X_t, and $\mu_t = E[Y_t/X_t]$ is the conditional mean. We assume that the dynamic behaviour of Y_t is given by the general nonlinear time series regression model:

$$
Y_t = \mu(X_t, \theta) + \sigma(X_t, \theta)u_t, \qquad t = 1, 2, ..., T \tag{1}
$$

where $\mu(X_t, \theta)$ and $\sigma^2(X_t, \theta)$ are the conditional mean and the conditional variance of Y_t, respectively. The functional forms of $\mu : \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}$ and $\sigma : \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}$ are known except for $\theta \in \Theta \subset \mathbb{R}^k$, where Θ is the parameter space. $\{u_t\}_{t=1}^T$ are assumed to be independent and identically distributed (i.i.d.) zero-mean unit-variance unknown errors with u_t being independent of X_t for all t. $F_u(\cdot)$ is the cumulative distribution function (cdf) of u_t with density function $f_u(\cdot)$. Let $\widehat{\theta}$ be a root-T consistent estimator of the parameter vector θ. The estimated residuals are computed from the the estimated parameters. Then $\widehat{u}_t = (Y_t - \mu(X_t, \widehat{\theta}))/\sigma(X_t, \widehat{\theta})$. Unless otherwise stated, all summations considered here are taken from 1 to T, where T denotes the number of observations. Note that the general framework (1) encompasses linear regression models as a particular case.

Under model (1), conditional symmetry of Y_t is equivalent to the symmetry of u_t about zero, that is, $f_u(u) = f_u(-u)$ for all u. Therefore, the null hypothesis under test is that "H_0: u_t is symmetric about 0", versus the general alternative "H_1: u_t is not symmetric about 0". It is pointed out that conditional symmetry does not, in general, imply unconditional symmetry[1].

An example of a NED process less trivial than a finite moving average process is a simple AR(1) process (see Gallant and White [15], pp. 27-28). ARMA models of finite order with zeros lying outside the unit circle can be shown to be NED of arbitrarily large size, provided the parameters are chosen such that the stationarity as well as the invertibility condition is fulfilled and the innovations satisfy appropriate moment conditions. Infinite MA processes can also be shown to be NED under mild conditions on the moving average weights (see Wooldridge and White [32], example 3.3). As Hansen [20] has shown, strictly stationary GARCH processes are NED under mild regularity conditions. This framework also include the AR process with ARCH/GARCH errors, discussed in Engle [12], which is widely applied in financial econometrics. Consider the AR(1)-GARCH(1,1) process, in which observed data are generated as a realization of a stochastic compound process

$$Y_t = \gamma + \theta Y_{t-1} + e_t,$$
$$e_t = u_t h_t^{1/2}, \quad h_t = \lambda + \beta h_{t-1} + \alpha e_{t-1}^2,$$

with $\{u_t\}$ being i.i.d., so h_t is strictly stationary. If $|\theta| < 1$, it is well-known that this model can be expressed as

$$Y_t = \gamma/(1 - \theta) + \sum_{\tau=0}^{\infty} \theta^{\tau} e_{t-\tau}.$$

It can be shown that Y_t is NED of order r on the stationary process $\{e_t\}$, if for some $r \geq 2$, $E |e_t|^r \leq \Delta < \infty$, with stable coefficients

$$v_r(m) = E \left| Y_t - Y_t^{(m)} \right|^r = E \left| \sum_{\tau=m}^{\infty} \theta^{\tau} e_{t-\tau} \right|^r \leq |\theta|^m \sum_{\tau=0}^{\infty} |\theta|^{\tau} E |e_{t-m-\tau}|^r \leq |\theta|^m \Delta/(1 - |\theta|),$$

---

[1]To illustrate this, consider a $MA(1)$ process $Y_t = u_t - \theta u_{t-1}$ with u_t i.i.d. and $\theta = 1$. The unconditional distribution of Y_t is always symmetric with independence of whether or not $f_u(\cdot)$ is symmetric, since Y_t and $-Y_t$ have exactly the same distribution. However, the conditional distribution of Y_t on X_t (which inlcudes u_{t-1}) will be asymmetric in case u_t is asymmetric.

decaying at a geometric rate. The conditions to ensure that $E\left|e_t\right|^r < \infty$ are $E\left|u_t\right|^r < \infty$ for some $r \geq 2$, and $\beta + \alpha < 1$.

We next show that $e_t = u_t h_t^{1/2}$ is NED of order r on the stationary process $\{u_t\}$. By repeated substitution we have $h_t = \lambda + \lambda \sum_{k=1}^{\infty} \prod_{i=1}^{k} (\beta + \alpha e_{t-i}^2) = \lambda + \lambda \sum_{k=1}^{\infty} \prod_{i=1}^{k} z_{t-i}$, where $z_t = \beta + \alpha e_t^2$. Because under $\beta + \alpha < 1$, $\sup_{t \geq 1} E\left|z_t\right|^r \leq c < 1$ for some $r \geq 2$, it follows that

$$E\left|h_t\right|^r = \lambda + \lambda E\left|\sum_{k=1}^{\infty} \prod_{i=1}^{k} z_{t-i}\right|^r \leq \lambda + \lambda \sum_{k=1}^{\infty} E\left|\prod_{i=1}^{k} z_{t-i}\right|^r \leq \lambda(1 + c/(1-c)) < \infty,$$

by the Minkowski's inequality for infinite sums.

To see that h_t is NED on $\{u_t\}$, let $h_t^{(m)} = \lambda + \lambda \sum_{k=1}^{m-1} \prod_{i=1}^{k} z_{t-i}$. By Minkowski's inequality

$$v_r^h(m) = E\left|h_t - h_t^{(m)}\right|^r = \lambda E\left|\sum_{k=m}^{\infty} \prod_{i=1}^{k} z_{t-i}\right|^r \leq \lambda c^{m-1} \sum_{k=1}^{\infty} E\left|\prod_{i=1}^{k} z_{t-(m-1)-i}\right|^r \leq c^m \lambda/(1-c).$$

Thus h_t is NED of order r on $\{u_t\}$. By Theorem 4.2 of Gallant and White (1988), $e_t = u_t h_t^{1/2}$ is L_r-NED on $\{u_t\}$. This is also true for ARCH errors ($\beta = 0$).

3 Tests for conditional symmetry

We next describe the tests for symmetry considered in our Monte Carlo study. To test for conditional symmetry, tests are applied to regression residuals. Since these tests have been discussed extensively in the literature, their description here is relatively brief.

A classical test of symmetry is the test of skewness (see Gupta [17] for a review of this test). This test is developed for demeaned data, but the statistic has the same limiting distribution when applied to residuals from a simple linear regression model. It might be of interest to compare this test with a joint test of the third and fifth central moments. In principle, a joint test of more moments is possible, but the higher order-moments are difficult to estimate precisely. The potential advantage is to be more powerful than a test based on the third moment in isolation. A practical strategy would be to start with the skewness coefficient and consider joint tests of higher moments only if we do not reject H_0. Standard asymptotic results will lead to the derivation of both a joint test of the third and fifth central moments and the skewness coefficient. The proof is omitted here in order to save space, but is available upon request. An advantage of these tests is that they are intuitive and easy to compute. However, they present a number of limitations. First, the limiting distributions of the estimators are known and have a simple form for the case of ordinary least squares. Different estimation methods may yield different limiting distributions. Second, they are moment-based tests, which require the existence of the sixth and tenth moments, respectively. This is not satisfied by many useful distributions such as the student-t_5 or GARCH process. Finally, these tests are not consistent against alternatives which are asymmetric and yet have the third moment and/or the fifth moment equal to zero.

6

Bai and Ng [2] discuss how to test whether the regression residuals from a nonlinear time series regression model are symmetrically distributed. The test, which is based on martingale transformations, does not require the data to be stationary or i.i.d., and the dimension of the conditional variables can be infinite. The test is shown to be consistent and asymptotically distribution free, but its computation is rather intensive.

The literature on symmetry is large, an commonly used nonparametric tests are the Wilcoxon signed-rank test (for further details, see Gibbons and Chakraborti [16]), the Runs test of McWilliams [25], and the Triples test of Randles *et al* [30]. These test are asymptotically distribution-free for i.i.d. observations. In the present setting, we replace the unobservable errors by well-behaved residuals. Thus, the asymptotic distribution of these statistics is unknown.

4 Symmetric bootstrap

We consider the nonlinear regression model (1). Under the null hypothesis, we know that the population $\{u_1, ..., u_T\}$ is symmetric about zero. The tests under consideration were computed with estimated regression residuals when testing for symmetry of regression errors. Let $\widehat{T}_T = T_T(\widehat{u}_1, ..., \widehat{u}_T)$ denote the test statistic of interest, which is a function of the standardized residuals. By using standardized residuals, we are guaranteed that all model residuals have, at least, the same two first moments.

In this section, we consider a bootstrap procedure for approximating the distribution of the test statistic of interest, which is a function of the residuals, for testing on the symmetry about the mean of the underlying distribution of the errors. When bootstrapping any test statistic, our aim is to find a bootstrap distribution that mimics the null distribution of the data, even though the data may be generated by an alternative distribution. We propose a resampling scheme so that the null hypothesis is respected in the bootstrap data-generating process. That is, a resampling method that ensures the bootstrap distribution to be symmetric. To be precise, we define the bootstrap sample by $\Upsilon_T^* = \{(Y_t^*, X_t^*) : t = 1, 2, ..., T\}$, where $Y_t^* = \mu(X_t^*, \widehat{\theta}) + \sigma(X_t^*, \widehat{\theta})u_t^*$ and $X_t^* = (Z_t, Y_{t-1}^*, ..., Y_{t-p}^*)'$. Note that the exogenous explanatory variables are fixed in repeated samples, and $\widehat{\theta}$ is some estimate of the parameter vector θ. Bootstrap residuals $u^* = (u_1^*, ..., u_T^*)'$ were constructed by a two-stage procedure:

Stage 1: Construct recentred versions of the residuals $\widetilde{u}_t = \widehat{u}_t - T^{-1}\sum_t \widehat{u}_t$. Random signs are assigned to the centered residuals $\widetilde{u}_t$ according to independent realizations of a Rademacher random variable s_t, independent of $\widetilde{u}_t$, which takes values $+1$ and -1 with probability $1/2$ each. By doing that, we obtain a set of symmetrized residuals $\{s_1\widetilde{u}_1, ..., s_T\widetilde{u}_T\}$.

Stage 2: A random number device independently selects integers $i_1, ..., i_T$, each of which equals any value between 1 and T with probability $1/T$. We allow a single unit $s_t\widetilde{u}_t$ to appear more than once in the sample, that is to sample with replacement. Therefore, the bootstrap data set $\{u_1^*, ..., u_T^*\}$ consists of members of the original data set $\{s_1\widetilde{u}_1, ..., s_T\widetilde{u}_T\}$, some appearing zero times, some appearing once, some appearing twice, etc.

Each bootstrap sample Υ_T^* is then used to re-estimate the parameter vector θ. Let $\widehat{\theta}^*$ denote the bootstrap estimator of θ. The estimated residuals from the bootstrap sample are

$$\left\{ \widehat{u}_t^* = Y_t^* - \mu(X_t^*, \widehat{\theta}^*))/\sigma(X_t^*, \widehat{\theta}^*) : t = 1, 2, ..., T \right\}.$$

Using bootstrap residuals, we compute the bootstrap test statistic $\widehat{T}_T^* = T_T(\widehat{u}_1^*, ..., \widehat{u}_T^*)$.

Repeating this procedure B times gives a sample $\left\{ \widehat{T}_{Tb}^* : b = 1, ..., B \right\}$ of $\widehat{T}_T$ values. This sample mimics a random sample of draws of $\widehat{T}_T$ under the null hypothesis. In particular, we consider the problem of estimating the α-level critical value of the $\widehat{T}_T$ test from its empirical distribution. Let $\widehat{c}_{\alpha,B}^*$ denote the bootstrap estimate of the α-level critical value. Let $\widehat{T}_{T(1)}^* \leq \widehat{T}_{T(2)}^* \leq ... \leq \widehat{T}_{T(B)}^*$ denote the B realizations of $\widehat{T}_T$ arranged in order of increasing size, and suppose we choose B and ν such that $\nu/B = 1 - \alpha$. Since the B values of $\widehat{T}_{Tb}^*$ divide the real line into $B + 1$ parts, not B, then it makes sense to select $\widehat{c}_{\alpha,B}^* = \widehat{T}_{T(\nu+1)}^*$. For example, in the case of $\alpha = 0.05$ and $B = 1000$, this would involve taking $\widehat{c}_{\alpha,B}^* = \widehat{T}_{T(96)}^*$.

It is convenient to choose a single value of B at which to monitor the performance of all the tests. In this study, this is not possible since there are large differences between the run times of the tests. Different values of B were chosen, so as to make the run time of each test approximately the same. We set $B = 1999$ for the moment-based tests, the Wilcoxon signed-rank test and the Runs test, while $B = 999$ for the Bai and Ng test. For the Triples test, we carry out $B = 99$ bootstrap replications, which is the smallest value of B that is commonly suggested. The processing time becomes excessive when greater values are used, especially for $T \geq 100$. We will illustrate the finite sample performance of the bootstrap proposal of the paper by means of simulation in Section 4.

4.1 Asymptotic properties

To study the asymptotic properties of the proposed bootstrap, we need to state the underlying assumptions.

(A1) For some small $\delta > 0$ and some $r > 2$, the data generating process (DGP) (1) is $L_{2+\delta}$-NED on $\{Z_t, u_t\}$ of size $-2(r-1)/(r-2)$. The constant δ is specified in A2 below.

(A2) $E|Y_t|^{2+\delta} < \infty$ for some $\delta > 0$.

(A3) The errors u_t are i.i.d. random variables with zero mean, unit variance and $E|u_1|^4 < \infty$. The density of u_t is $f_u(\cdot)$ and the cdf $F_u(\cdot)$. Furthermore, u_t is independent of X_t.

(A4) $\mu(\cdot, \cdot)$ and $\sigma(\cdot, \cdot)$ are twice continuously differentiable with respect to the second argument with bounded derivatives. Additionally, there exists $\sigma_0 > 0$ such that $\sigma(\cdot, \cdot) > \sigma_0$.

(A5) The estimator $\widehat{\theta}$ satisfies $\sqrt{T}(\widehat{\theta} - \theta) = O_p(1)$.

(A6) $\mu(\cdot, \cdot)$ and $\sigma(\cdot, \cdot)$ are Lipschitz continuous with respect to the first argument, i.e., there exist a constant L_μ such that $|\mu(u, \theta) - \mu(v, \theta)| \leq L_\mu \|u - v\|$, and $\sigma(\cdot, \cdot)$ satisfies a similar inequality for a certain constant L_σ.

(A7) $\Psi_{X,m}$ is continuously differentiable with respect to the m arguments with bounded derivatives.

(A8) $m \to \infty$ with $m = o(T)$.

Assumptions A1 and A2 are related to the nonlinear process itself. Assumption A3 is concerned with the behaviour of the errors. The differentiability condition required in A4 is relatively standard in nonlinear estimations. A5 is a standard assumption, which ensures that the estimators are root-T consistent. Conditions A6-A8 are required for purely technical reasons.

We next provide a little theory for the convergence of the empirical distribution of standardized residuals under the symmetric bootstrap proposed above. The idea behind the bootstrap is to replace the true distribution function of the error term u_t by its empirical estimate. Let $\widehat{F}_T$ be the empirical distribution function of the recentred standardized residuals, putting mass $1/T$ on each $\widetilde{u}_t$, $t = 1, ..., T$. That is, the centered residuals are equally likely to appear in the bootstrap sample. Following Efron and Tibshirani [11], a bootstrap sample is defined to be a random sample of size T drawn from $\widehat{F}_T$, say $\widetilde{u}^* = (\widetilde{u}_1^*, ..., \widetilde{u}_T^*)'$. The start notation indicates that $\widetilde{u}^*$ is not the actual data set $\widetilde{u}$, but rather a randomized, or resampled, version of it.

We can construct the distribution $\widehat{G}_T$, which places mass $1/T$ at $s_t \widetilde{u}_t$, $t = 1, 2, ..., T$, where s_t is a Rademacher random variable, independent of $\widetilde{u}_t$. We use $\widehat{G}_T$ as the basis for our bootstrap resampling scheme. It is straightforward to prove that the distribution of the random variable $s_t u_t$ is symmetric about zero under both H_0 and H_1. Let G_u be its distribution function defined by

$$G_u(x) = \frac{1}{2}(1 - F_u(-x) + F_u(x))$$

It is pointed out that $G_u(x) = F_u(x)$ for every given x under the null hypothesis. Note that the symmetry of the bootstrap errors does not depend on whether the null hypothesis holds or not, although u_t does. That is, our bootstrap approximation to the null hypothesis is always valid even the data $\{(Y_t, X_t)\}_{t=1}^T$ were drawn from a population under which the null hypothesis does not hold. Therefore, the derived bootstrap tests automatically follow the first guideline set by Hall and Wilson [19]. Namely resampling should be done in a way that reflects the null hypothesis, even when the true hypothesis is distant from the null. As they pointed out, this ensures the reasonable power of the bootstrap test against the departure from the null hypothesis.

In order to investigate the asymptotic behaviour of the symmetric bootstrap, we use the *Mallows metric*[2] d_2 to show that the bootstrap errors u_t^* approximate the true errors u_t under H_0. There is one key result that make this metric a useful tool in proving asymptotic results for regression models. From Bickel and Freedman ([4]; Lemma 8.3), given distributions $F, F_1, F_{2,...}$, the condition $d_2(F_T, F) \to 0$ as $T \to \infty$ implies that the probability measures corresponding to F_T converge weakly to the measure corresponding to F.

Proposition 1 *Suppose that assumptions A1-A5 hold. Then, under H_0, $d_2(u_t, u_t^*) \to 0$ as $T \to \infty$.*

As next step, we show that Υ_T^* replicates the structure of (1), given the original data $\Upsilon_T =$

---

[2] The Mallows metric is defined by $d_2^2(X, Y) = d_2^2(G, H) = \inf \left\{ E[\|X - Y\|^2] : X \widetilde{} G, Y \widetilde{} H \right\}$, where the infimum is over all joint distributions of (X, Y) whose fixed marginal distributions are G and H respectively and where $\|.\|$ denotes the Euclidean norm on $\mathbb{R}$. See Bickel and Freedman ([4]; Section 8) for a detailed discussion of this metric.

$\{(Y_t, X_t), t = 1, ..., T\}$. For this purpose, we define $\widetilde{\Upsilon}_T = \left\{(\widetilde{Y}_t, \widetilde{X}_t), t = 1, ..., T\right\}$ as

$$\widetilde{Y}_t = \mu(\widetilde{X}_t, \theta) + \sigma(\widetilde{X}_t, \theta)\varepsilon_t, \qquad t = 1, 2, ..., T,$$

where $\widetilde{X}_t = \{Z_t, \widetilde{Y}_{t-1}, ..., \widetilde{Y}_{t-p}\}$ and $\{\varepsilon_t\}_{t=1}^T$ are conditionally i.i.d. random variables with the following properties. Given Υ_T, (i) ε_t has conditional distribution F_u; (ii) $d_2(\varepsilon_t, u_t^*) = d_2(u_t, u_t^*)$, (iii) $\widetilde{\Upsilon}_T$ is $L_{2+\delta}$-NED on $\{Z_t, \varepsilon_t\}$ for some $\delta > 0$. Here and in the following, a star appearing in E denotes expectation with respect to Υ_T^* conditional on the data Υ_T.

Proposition 2 *Suppose that assumptions A1-A8 hold. Then, under H_0,*

$$\sup_{1 \leq t \leq T} E^* \left| \widetilde{Y}_t - Y_t^* \right| = o_p(1) \qquad for\ T \to \infty.$$

The following Corollary, which show that $\left| \widetilde{Y}_t - Y_t^* \right| \to 0$ in mean for $T \to \infty$ given Υ_T, follows immediately from Proposition 2.

Corollary 1 *Suppose that assumptions A1-A8 hold. Then, under H_0,*

$$E^* \{ T^{-1} \sum_t \left| \widetilde{Y}_t - Y_t^* \right| \} \to 0 \qquad for\ T \to \infty.$$

Note that we do not prove that the conditional distribution of $\widehat{T}_T^*$ given Υ_T is asymptotically equal to the null-hypothesis distribution of $\widehat{T}_T$ since the asymptotic distribution of $\widehat{T}_T$ is unknown for some of the statistics under consideration.

5   A Monte Carlo study

In this section, we investigate the finite-sample properties of the symmetry tests of Section 2 by means of Monte Carlo simulation[3]. The aim of the experiments is two-fold. First, to investigate whether the bootstrap procedure proposed in Section 3 can be used to obtain improved finite-sample critical values with respect to the asymptotic theory, whenever this is available. Second, to identify the size and power properties of the test statistics under various scenarios, including linear, AR, MA and GARCH models. We first describe the data-generating processes (DGP) and the experimental design that is used in our simulations. A discussion of the results obtained in these simulation experiments follows.

5.1   Experimental design

The time series considered in our study are generated according to model (1), where functions $\mu(\cdot, \cdot)$ and $\sigma(\cdot, \cdot)$ are generated according to four basic types of DGPs:

DGP$_1$: $\mu(X_t, \theta) = \beta_0 + \sum_{i=1}^{k} Z_{it}\beta_i$, $(Z_{1t}, Z_{2t}, ..., Z_{kt})' \overset{i.i.d.}{\sim} N(0, I_k)$, and $\sigma(X_t, \theta) = \sigma = 1$;

---

[3] All the procedures for estimating the models described in this section were written in GAUSS programming language. Programs are available from the author upon request.

DGP$_2$: $\mu(X_t, \theta) = c + \rho Y_{-1}$ and $\sigma(X_t, \theta) = \sigma = 1$;

DGP$_3$: $\mu(X_t, \theta) = \mu + \phi u_{t-1}$ and $\sigma(X_t, \theta) = \sigma = 1$;

DGP$_4$: $\mu(X_t, \theta) = \gamma$ and $\sigma(X_t, \theta) = \{\alpha_0 + \alpha_1 \sigma(X_{t-1}, \theta)^2 + \alpha_2 \sigma(X_{t-1}, \theta)^2 u_{t-1}^2\}^{1/2}$.

DGP$_1$ is a linear regression model with an intercept component and k i.i.d. variables as regressors. Data are generated setting $\beta_0 = ... = \beta_k = 1$ and $k = 1, 4$. The reason for increasing the number of regressors is to observe the sensitiveness of the size and the power of the tests to the additional regressors. For an AR(1) specification, our simulation experiment is based on DGP$_2$. We set $c = 0$ and $\rho = 0.5, 0.8$. We denote by DGP$_3$ the MA(1) design. We set the constant regressor μ equal to zero and $\phi = 0.5, 0.8$. Finally, DGP$_4$ corresponds to a GARCH(1,1) model. In this framework, we set $\gamma = 1$ and $(\alpha_0, \alpha_1, \alpha_2) = (2, 0.5, 0.3)$. Also, we consider the model with $(\alpha_0, \alpha_1, \alpha_2) = (2, 0.9, 0.05)$, which is close to being an IGARCH(1,1). All parameter combinations considered were selected to make the results of our study comparable with those obtained by Bai and Ng [2], whenever this is possible.

For each DGP, we draw u_t from symmetric and asymmetric distributions to derive conditionally symmetric and asymmetric distributions for Y_t. To asses the size of the tests, we first generate u_t from the standard normal distribution and the student-t distribution with 5 degrees of freedom. To evaluate the power of the tests, we draw random variables from the exponential distribution and the chi-square with two degrees of freedom. We then consider another ten distributions, four symmetric and six asymmetric, from the generalized lambda family (GLF) discussed in Ramberg and Schmeiser [29]. The choice of all these distributions is motivated by the fact they are used in previous studies of testing symmetry and in consequence provide a benchmark for comparing size and power. In addition, they cover a wide range of values of third and fourth standardized moments. The GLF is easily generated since it is defined in terms of the inverse cumulative distribution function $F^{-1}(u) = \lambda_1 + \left[u^{\lambda_3} + (1 - u)^{\lambda_4}\right]/\lambda_2$, $0 < u < 1$, with mean and variance given by:

$$\mu = \lambda_1 + \left[(1 + \lambda_3)^{-1} - (1 + \lambda_4)^{-1}\right]/\lambda_2,$$
$$\sigma^2 = [(1 + 2\lambda_3)^{-1} - 2\beta(1 + \lambda_3, 1 + \lambda_4) + (1 + 2\lambda_4)^{-1} - ((1 + \lambda_3)^{-1} - (1 + \lambda_4)^{-1})^2]/\lambda_2^2,$$

where $\beta(\cdot, \cdot)$ denotes the beta function. The λ parameters defining the ten selected distributions are taken from Randles *et al.* [30] and are listed in Table 1, together with the associated skewness (η_3) and kurtosis (η_4) values. The distributions are arranged in ascending order of departure from symmetry[4]. To be under the assumptions of the regression model, all error distributions are standardized to have zero mean and unit variance. Among these distributions, the Student-t distribution with 5 degrees of freedom has finite variance, but does not have finite sixth and tenth moments. The generalized lambda distributions have finite qth moment if, and only if, $-1/q < min(\lambda_3, \lambda_4)$. All other distributions have finite sixth and tenth moments. This is aimed at checking how moment-based tests behave when data do not possess proper moments.

---

[4]The shapes of the GLF density functions are shown in McWilliams [25].

The experiments proceed by generating artificial time series of length T from (1) with $T \in \{50, 100, 200\}$. We have to estimate $k + 2$ parameters in DGP_1. The parameters of interest are estimated using ordinary least squares. Next, in DGP_i $(i = 2, 3)$ and DGP_4, we have three and four parameters to be estimated, respectively. In order to do that, we use maximum likelihood (ML) estimation. In the context of DGP_4, as Fiorentini et $al.$ [13] proposed, for estimation purposes we employ the analytic first and second derivatives of the log-likelihood instead of numerical approximations in order to benefit for computational reductions and avoid convergence problems. Finally, we compute the relevant test statistic $\widehat{T}_T = T_T(\widehat{u}_1, ..., \widehat{u}_T)$, which is based on the standardized residuals from estimation of (1).

Due to the computational demand required by some of the tests included in this study is very high, experiments were conducted using 500 replications for the Triples test, 1000 for the Bai and Ng test, and 2000 for the remaining tests. For each replication, we reject the null being tested at the nominal α-level, based on both bootstrap and asymptotic critical values, if the observed test $\widehat{T}_T$ is above $\widehat{c}^*_{\alpha,B}$ and exceeds the (1-α) quantile of the corresponding asymptotic distribution, respectively. We finally count the proportion of times that the null hypothesis is rejected for each test statistic using bootstrap- and asymptotic-based critical values. For non-symmetric alternatives, this proportion yields an estimate of the power of the test. In all cases power is not size-adjusted. On the other hand, the proportions from the symmetric distributions imply estimates of the Type I error. Since the results for tests performed at the 0.01, 0.05, and 0.10 significance levels are qualitatively similar and lead to the same conclusions about the relative merits of different tests, we focus on 5%-significance level tests[5].

5.2   Simulation results: a comparative study of symmetry tests

The reader has to consider that the nonparametric tests included in this study, that is, the Wilcoxon Signed-Rank test (WSR_T), the Runs test (R_T) and the Triples test (TR_T), are originally constructed for the problem of testing the unconditional symmetry of an i.i.d. sample of observations. We investigate the performance of these tests when testing for conditional symmetry. Under (1), conditional symmetry is equivalent to the symmetry of the error term about zero. Furthermore, at this point we do not provide an asymptotic distribution theory for these tests when unknown errors are replaced by well-behaved residuals. This is not the case of Bai and Ng test (CS_T) and moment-based tests ($\widehat{S}^3_T$ and $\widehat{S}^{3,5}_T$), whose corresponding asymptotic distributions are completely known. We implement a bootstrap version of all the tests. Tables 2 to 9 show the empirical size and empirical power of the various tests obtained using artificial time series generated according to DGP_1, DGP_2, DPG_3 and DGP_4. It should be pointed out that moment-based tests are only computed when the process is uncorrelated (DGP_1). We report empirical rejection rates (%) under the null and the alternative based on both asymptotic critical values as well as bootstrap critical values obtained from Monte Carlo trials. To establish heuristic comparisons, for the set of nonparametric tests we use the tabulated asymptotic critical values that will correspond to tests statistics

---

[5] Results at the 1% and 10% levels of significance are available upon request.

computed with "observable" errors[6]. This should be borne in mind when assessing the results. Based on the selected 5% nominal level (or size) of the tests, the empirical rejection frequencies should be around 5% under the null, while they should be around 100% under the alternative.

Table 2 presents the empirical size of the symmetry tests for DGP_1. A feature of the size properties of the $\widehat{S}_T^{3,5}$ statistic with asymptotic critical values is that it is consistently undersized with actual size about 2% for most of the cases investigated. The bootstrap does bring the empirical size of the test closer to its nominal level. The empirical size of this test is rather stable to an increase in the number of regressors. The performance of both moment-based tests under distributions S4 to S6 deserves further analysis, since sixth and tenth moments of these distributions do not exist. The empirical size of $\widehat{S}_T^{3,5}$ test in S3, for which sixth and tenth moments exist, is comparable to its size under distributions S4 to S6. Comparing the size of both moment-based statistics across distributions, shows that the $\widehat{S}_T^{3,5}$ test statistic rejects less often than the skewness coefficient under the null when using asymptotic critical values. The results are the reverse with bootstrap-based critical values except for distributions with kurtosis equal to 3.

We next consider the Bai and Ng test. Note that for all DGPs considered in Bai and Ng [2], the estimated regression model imposes $\sigma^2 = 1$. Under this circumstance, the variance is not treated as a parameter to be estimated, as it is assumed in (1), but a constant to be specify, and for this reason it is not possible to establish direct comparisons between their results and those of this paper[7]. Fixing $k = 1$, the size of the test based on asymptotic critical values is largely satisfactory for distributions S1 to S4 when $T = 50$. However, this result should be interpreted with caution, since when the sample increases to $T = 100$ empirical sizes fall drastically. This may suggest that this test presents inflated sizes for small samples under these distributions. Turning to fat-tailed distributions, we may see that S5 is slightly oversized for $T = 200$. The S6 case is more seriously oversized. This distortion increases with T. This reflects an efficiency loss in conducting the CS_T test based on asymptotic critical values in distributions with high kurtosis, since it tends to reject a true null too often. It should be pointed out, however, that oversizing is not so large as to render the test unattractive for applications. These results appear to be robust to increase the number of regressors to $k = 4$. On the other hand, the results from bootstrap critical values show that, in S5 and S6 cases, the bootstrap performs well, with sizes close to the nominal level for $T \geq 50$. Increasing the number of regressors moves bootstrap-based empirical sizes upward in small samples. Since size distortions may lead to misleading inference, to guard against possible over-rejection of the symmetry hypothesis, it is advisable to compute this test statistic with bootstrap-based critical values.

The actual sizes of the Runs test based on asymptotic critical values are close to the 5% nominal level even for $T = 50$. When $k = 1$, the Runs test with bootstrap-based critical values has poor

---

[6]All the tests conducted in this simulation study are one-tailed tests at the 0.05 level. The asymptotic critical values for the tests $\widehat{S}_T^{3,5}$ and CS_T are 5.99 and 2.20 respectively. For the remaining tests, the asymptotic critical value is 3.84.

[7]Models 1, 2, 4, 5 and 6 in Bai and Ng [2] correspond to distributions S1, S4, A5, A7 and A8 in this paper, respectively.

size. Interestingly, the performance of the test steadily deteriorates as T increases, its rejection frequencies being around 1% for $T = 200$. In unreported simulations, we found that the size distortions of the bootstrap test do disappear slowly as T increases further. The size properties of R_T for $k = 4$, on the other hand, seem largely satisfactory, being the performance of the bootstrap comparable to asymptotic values.

In the case of the Wilcoxon Signed-Rank test, the striking feature of the results is the quite severe size distortion of the test when asymptotic critical values are used for any of the distributions considered. Rejection frequencies are equal or close to zero, which illustrates the conservative nature of the test. These size distortions do not disappear as T increases. These results should be interpreted with caution. The poor size properties of WSR_T statistic might stem from incorrectly assuming the asymptotic distribution of the test is invariant to the replacement of errors by residuals. The use of bootstrap-based critical values instead of asymptotic ones corrects the differences between the empirical and nominal sizes.

Focusing on TR_T, for series of length $T = 50$ and $k = 1$, we may see that the empirical size of the test with asymptotic critical values tends to be much smaller than the nominal size, although the size distortions for S1 and S2 are considerably less than they are for distributions with kurtosis higher than normal (S3 to S6). When considering leptokurtic distributions (S3 to S6) and fixing $T = 50$, size increases substantially with k. A similar result holds for all the distributions when bootstrap critical values are used. Again, this is a small sample effect. For fixed T and k, the performance of the test improves with bootstrap critical values.

The size properties of the Bai and Ng test together with the nonparametric tests in DGP_2, DGP_3 and DGP_4 are reported, respectively, in Tables 3 to 5. Overall, the evidence from our simulations suggests that the relevant test statistics replicate the same patterns found for DGP_1. The results are quite robust even when the process is close to being nonstationary.

We now briefly review the performance of the tests under the eight alternatives of asymmetry. The results for each one of the four DGPs considered here are displayed in Tables 6 to 9, respectively. The Monte Carlo simulations reveal that the bootstrap performance is better or at worst equal to that of asymptotic critical values for all the tests under any DGP, with the exception of R_T. As intuition would suggest, this test is more successful with asymptotic critical values, given the size of the test based on bootstrap critical values is too small. Since CS_T does not hold its 5% level very well with asymptotic critical values, it is difficult to include it in any asymptotic power comparisons, since it high power might easily arise out of these inflated levels. Note also that WSR_T has disappointing power properties with asymptotic critical values. The reason for such poor performance, when compared to bootstrap critical values, is that the asymptotic-based empirical size of this test is extremely conservative (bearing in mind that the under-size in the WSR_T test computed with residuals is attributable to the use of critical values for the corresponding WSR_T test computed with errors). Therefore, unless stated otherwise, all power comparisons between tests reported hereafter are based on bootstrap critical values, which are more reliable.

Turning attention to comparisons between DGPs, results seem not to be affected by the DGP where the data come from. There are two exceptions to this rule in the case of DGP_4, for which

the bootstrap-based empirical power of WSR_T and R_T.tests undergoes a significant reduction for most of the distributions, no matter the sample size considered. It should be pointed out that this power distortion does not affect comparisons between tests described below. For a fixed DGP, the results of all the tests are quite robust across parameterizations for AR, MA and GARCH models.

Non-symmetry is detected with reasonable frequencies in nearly all cases. For a fixed distribution, power increases with T with both bootstrap and asymptotic critical values. For a fixed T, Bai and Ng test and the nonparametric tests exhibit monotonic power with the power of the tests increasing for increasing levels of asymmetry, except for A6, which is analyzed in depth below. For the moment-based tests, this monotonic behaviour is interrupted in A7 and A8, being their powers lower than in A5. This may reflect the sensitivity of these tests to the high kurtosis displayed by these distributions. The alternative of non-symmetry is detected with the lowest probability in the case A1, as we would expected given this distribution is rather close to symmetry. Distribution A6 is introduced to show the sensitivity analysis of the power to the kurtosis of the underlying distribution by comparing the behaviour of each particular test under A6 against alternative A5, which has the same asymmetry as A6 but a lower level of kurtosis[8]. From the empirical results, we can assert that R_T test is the most affected by the kurtosis of the underlying distribution. For the remaining tests, the differences in power between these two distributions decrease as T increases. It is noteworthy that A6 only outperforms, in terms of power, A1.

Overall, the WSR_T test clearly dominates the others on power for all the DGPs, being the differences in performance more remarkable when $T \leq 100$. It is followed by the CS_T test for DGP$_2$ to DGP$_4$. Under DGP$_1$, CS_T and $\widehat{S}_T^{3,5}$ statistics are the best performing competitors. Furthermore, these tests have complementary power, since $\widehat{S}_T^{3,5}$ performs better than CS_T for distributions A2 and A3, while CS_T appears to be more preferable for A7 and A8, which are thick-tailed distributions. For the remaining distributions, the power of both tests is comparable. The TR_T test on the other hand, which has good size properties for all the DGPs, has disappointingly low power. It's worth noting that the bootstrap power of R_T is not affected by the same erratic behaviour of the size. Its power increases substantially with T. Turning to the properties of moment-based tests in DGP$_1$, we may see that $\widehat{S}_T^{3,5}$ is more competitive in power than $\widehat{S}_T^3$ in all cases, especially when the sample size is moderate to large ($T = 100, 200$), which corresponds with our intuition.

To summarize, the following conclusions emerge when size and power performance are jointly considered. First, the simulations provide a strong case for the use of bootstrap critical values, especially for $T = 50$. The Runs test is the only one for which asymptotic-based critical values outperform bootstrap ones. For this test, the use of bootstrap critical values provides additional protection of the 5% level, but at a nontrivial cost in terms of power. Second, $\widehat{S}_T^{3,5}$ dominates $\widehat{S}_T^3$, since both moment-based tests have similar size properties, but a joint test of the third and fifth moments is more competitive in power. The obvious simplicity of this test and its robustness to the number of regressors in the model make it use attractive despite of its sensitivity to thick-

---

[8]To establish comparisons with A6, it would be also possible to use distribution A4, which has the same levels of asymmetry and kurtosis as A5.

tailed distributions. Third, considering the Monte Carlo results and the fact that WSR_T is easy to calculate, we recommend it over any of the competitors included in this study, being aware of the fact that its asymptotic distribution when replacing regression errors by residuals is unknown. The level of protection of the nominal level is higher with CS_T, but at a no minor reduction in power, especially for distributions with low skewness. Finally, note that the size and power advantages held by WSR_T using bootstrap critical values are limited to the specific class of null and alternative distributions considered in this study.

6   Conclusions

This paper investigates the finite sample properties of the Bai and Ng test commonly employed to detect conditional symmetry. We also explore the possibility of evaluating conditional symmetry by using some widely used tests for the unconditional symmetry of observations when the tests are applied to regression residuals. The tests investigated included the coefficient of skewness, a joint tests of the third and fifth moments, the Runs test, the Wilcoxon signed-rank tests and the Triples test. The limiting distribution of the conditional tests is only provided for moment-based tests. For this reason, the performance of a symmetric bootstrap to compute critical values for all the tests is discussed. The proposed symmetric bootstrap is easy to implement and is flexible enough to be adapted to a variety of nonlinear regression models. The potential of the methodology is illustrated using Monte Carlo simulation. The following general conclusions can be drawn from the results. First, the ability of the bootstrap to overcome the problem of oversizing observed for the Bai and Ng test when asymptotic critical values are used. Second, the size and power properties of the tests do not appear to be affected by the data-generating process for time series of relatively large length. Finally, the evidence from our simulations suggests that the Wilcoxon signed-rank test dominates the others in terms os size accuracy and power to detect non-symmetry.

References

[1] Azzalini, A. and Bowman, A., 1993. On the use of nonparametric regression for checking linear relationships. Journal of the Royal Statistical Society, Series B 55 549-559.

[2] Bai, J. and Ng, S., 2001. A consistent test for conditional symmetry in time series models. Journal of Econometrics, 103(1-2) 225-258.

[3] Belaire-Franch, J. and Contreras, D., 2002. A Pearson's test for symmetry with an application to the Spanish business cycle. Spanish Economic Review, 4(3) 221-238.

[4] Bickel, P.J. and Freedman, D.A., 1981. Some Asymptotic theory for the Bootstrap. The Annals of Statistics, 9(6) 1196-1217.

[5] Bickel, P.J., 1982. The 1980 Wald Memorial Lecures: On Adaptive Estimation. The Annals of Statistics, 10(3) 647-671.

[6] Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics, 31(3) 307-327.

[7] Bollerslev, T. 1987. A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. The Review of Economics and Statistics, 69(3) 542-547.

[8] Brunner, A.D., 1992. Conditional Asymmetries in Real GNP: A Seminonparametric Approach. Journal of Business & Economic Statistics, 10(1) 65-72.

[9] Davidson, R. and Flachaire, E., 2001. The Wild Bootstrap, Tamed at last. Queen's Institute for Economic Research Working Paper No. 1000.

[10] DeLong, J.B and Summers, L.H., 1986. Are Business Cycles Symmetrical? In Robert Gordon (ed.), American Business Cycles: Continuity and Change. National Bureau of Economic Research and University of Chicago Press, Chicago, 166-179.

[11] Efron, B. and Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman and Hall, London.

[12] Engle, R., 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50(4) 987-1007.

[13] Fiorentini, G., Calzolari, G. and Panattoni, L., 1996. Analytic Derivatives and the Computation of Garch Estimates. Journal of Applied Econometrics, 11(4) 399-417.

[14] Franses, P.H. and Van Dijk, D., 2000. Non-Linear Time Series Models in Empirical Finance. Cambridge University Press, Cambridge.

[15] Gallant, A.R. and White, H., 1988. A unified theory of estimation and inference for nonlinear dynamic models. Basil Blackwell. New York.

[16] Gibbons, J.D. and Chakraborti, S., 1992. Nonparametric Statistical inference. Mercel Dekker, Inc., 3rd edition.

[17] Gupta, M. K., 1967. An Asymptotically Nonparametric Test of Symmetry. Annals of Mathematical Statistics, 38(3) 849-66.

[18] Hall, P., 1992. The Bootstrap and Edgeworth Expansion. Springer Series in Statistics. New York: Springer Verlag.

[19] Hall, P. and Wilson, S.R., 1991. Two Guidelines for Bootstrap Hypothesis Testing. Biometrics, 47(2), 757-762.

[20] Hansen, B.E., 1991. GARCH(1,1) processes are near epoch dependent. Economics Letters, 36(2), 181-186.

[21] Hodgson, D.J., 1998. Adaptive Estimation of Error Correction Models. Econometric Theory, 14(1) 44-69.

[22] Hussey, R., 1992. Nonparametric evidence on asymmetry in business cycles using aggregate employment time series. Journal of Econometrics, 51(2) 217-231.

[23] Kreiss, J.P., 1987. On Adaptive Estimation in Stationary Arma Processes. The Annals of Statistics, 15(1) 112-133.

[24] Kulasekera, K.B. and Wang, J., 2001. A test of equality of regression functions using Gâteaux scores. Australian and New Zealand Journal of Statistics, 43(1) 89-100.

[25] McWilliams, T.P., 1990. A Distribution-free test for Symmetry based on a Runs Statistic. Journal of the American Statistical Association, 85(12) 1130-1133.

[26] Newey, W.K., 1988. Adaptive Estimation of Regression Models via Moment Restrictions. Journal of Econometrics, 38(3) 301-339.

[27] Newey, W.K. and Steigerwald, D.G., 1997. Asymptotic Bias for Quasi-Maximum-Likelihood Estimators in Conditional Heteroskedasticity Models. Econometrica, 65(3) 587-599.

[28] Powel, J.L., 1994. Estimation of semiparametric models. In R.F. Engle and D.L. MacFadden (eds), Handbook of Econometrics IV. Elseviere Science, Amsterdam, 2443-2521.

[29] Ramberg, J.S. and Schmeiser, B.W., 1974. An Approximate Method for Generating Asymmetric Random Variables. Communications of the ACM 17(2).

[30] Randles, R.H., Flinger, M.A., Policello, G.E. and Wolfe, D.A., 1980. An Asymptotically Distribution-Free Test for Symmetry versus Asymmetry. Journal of The American Statistical Association, 75(3) 168-172.

[31] Verbrugge, R.J., 1997. Investigating Cyclical Asymmetries. Studies in Nonlinear Dynamics and Econometrics, 2(1) 15-22.

[32] Wooldridge, J.M. and White, H., 1988. Some invariance principles and central limit theorems for dependent heterogeneous processes. Econometric Theory, 4(2), 210-230.

Appendix A

Proof. of Proposition 1: To proceed, corresponding to $\widehat{F}_T$ and $\widehat{G}_T$, let F_T and G_T denote the empirical distributions of $\{u_t\}_{t=1}^T$ and $\{s_t u_t\}_{t=1}^T$, respectively. By the triangular inequality,

$$d_2^2(u_t, u_t^*) = d_2^2(F_u, \widehat{G}_T) \leq d_2^2(F_u, G_T) + d_2^2(G_T, \widehat{G}_T).$$

Under H$_0$, the first term converges to 0 by Lemma 8.4 in Bickel and Freedman [4]. In order to obtain an upper bound for the second term, we consider particular random variables U_T and V_T, where $U_T = \{s_t u_t\}$ and $V_T = \{s_t \widetilde{u}_t\}$, $t = 1, ..., T$. Hence,

$$
\begin{aligned}
d_2^2(G_T, \widehat{G}_T) \;\leq\;& E^*(U_T - V_T)^2 = T^{-1}\sum_t (\widehat{u}_t - u_t - T^{-1}\sum_t \widehat{u}_t)^2 \\
\leq\;& 2T^{-1}\sum_t (\widehat{u}_t - u_t)^2 - 2(T^{-1}\sum_t (\widehat{u}_t - u_t))^2 + 2(T^{-1}\sum_t u_t)^2 \\
\leq\;& 2T^{-1}\sum_t (\widehat{u}_t - u_t)^2 + 2(T^{-1}\sum_t u_t)^2 \\
=\;& 2T^{-1}\sum_t \left(\frac{\mu(X_t,\theta) - \mu(X_t,\widehat{\theta})}{\sigma(X_t,\widehat{\theta})} + u_t\frac{\sigma(X_t,\theta) - \sigma(X_t,\widehat{\theta})}{\sigma(X_t,\widehat{\theta})}\right)^2 + 2(T^{-1}\sum_t u_t)^2 \\
\leq\;& 4T^{-1}\sum_t \frac{(\mu(X_t,\theta) - \mu(X_t,\widehat{\theta}))^2}{\sigma(X_t,\widehat{\theta})^2} + 4T^{-1}\sum_t u_t^2\frac{(\sigma(X_t,\theta) - \sigma(X_t,\widehat{\theta}))^2}{\sigma(X_t,\widehat{\theta})^2} + 2(T^{-1}\sum_t u_t)^2.
\end{aligned}
$$

For the first term on the right hand side (r.h.s.), we make use of the lower bound for $\sigma(\cdot,\cdot)$ and a Taylor expansion for the numerator. By adding and subtracting terms, it is bounded by

$$
4\sigma_0^{-2}T^{-1}\sum_t \left|(\partial\mu(X_t,\theta)/\partial\theta)'(\widehat{\theta} - \theta) + o_p(1)\right|^2
$$
$$
= O_p(T^{-1}) + o_p(1) = o_p(1) \text{ as } T \to \infty,
$$

which is obtained applying assumptions A4-A5.

Exactly along the same line, for the second term on the r.h.s,

$$
4T^{-1}\sum_t u_t^2\frac{(\sigma(X_t,\theta) - \sigma(X_t,\widehat{\theta}))^2}{\sigma(X_t,\widehat{\theta})^2} \leq 4(T^{-1/2}\sum_t u_t^2)(T^{-1/2}\sum_t \frac{(\sigma(X_t,\theta) - \sigma(X_t,\widehat{\theta}))^2}{\sigma(X_t,\widehat{\theta})^2}).
$$

Under assumption A3, by the Central Limit Theorem,

$$
T^{-1/2}\sum_t u_t^2 = O_p(1).
$$

The remainder is treated as follows,

$$
T^{-1/2}\sum_t \frac{(\sigma(X_t,\theta) - \sigma(X_t,\widehat{\theta}))^2}{\sigma(X_t,\widehat{\theta})^2} \leq \sigma_0^{-2}T^{-1}\sum_t \left|(\partial\sigma(X_t,\theta)/\partial\theta)'\sqrt{T}(\widehat{\theta} - \theta) + o_p(1)\right|^2,
$$

which is of order $o_p(1)$.

By the law of large numbers

$$
(T^{-1}\sum_t u_t)^2 = o_p(1),
$$

which completes the proof of Proposition 1. ∎

Proof. of Proposition 2: By the definition of $\widetilde{\Upsilon}_T$ and Υ_T^*

$$
\begin{aligned}
E^*\left|\widetilde{Y}_t - Y_t^*\right| \;=\;& E^*\left|\mu(\widetilde{X}_t,\theta) - \mu(\widetilde{X}_t^*,\widehat{\theta}) + (\sigma(\widetilde{X}_t,\theta) - \sigma(\widetilde{X}_t^*,\widehat{\theta}))u_t^* + \sigma(\widetilde{X}_t,\theta)(u_t - u_t^*)\right| \\
\leq\;& E^*\left|\mu(\widetilde{X}_t,\theta) - \mu(\widetilde{X}_t^*,\theta)\right| + E^*\left|\mu(\widetilde{X}_t^*,\theta) - \mu(\widetilde{X}_t^*,\widehat{\theta})\right| \\
& + E^*\left|\sigma(\widetilde{X}_t,\theta) - \sigma(\widetilde{X}_t^*,\theta)\right||u_t^*| + E^*\left|\sigma(\widetilde{X}_t^*,\theta) - \sigma(\widetilde{X}_t^*,\widehat{\theta})\right||u_t^*| \\
& + E^*\sigma(\widetilde{X}_t,\theta)\left|(\varepsilon_t - u_t^*)\right|.
\end{aligned}
$$

Following the same reasoning as in the proof of Proposition 1, the second and the fourth term converge to zero in probability. For the first term, we have from A6

$$E^* \left| \mu(\widetilde{X}_t, \theta) - \mu(\widetilde{X}_t^*, \theta) \right| \leq L_\mu E^* \left\| \widetilde{X}_t - \widetilde{X}_t^* \right\|.$$

Exactly along the same lines we obtain

$$E^* \left| \sigma(\widetilde{X}_t, \theta) - \sigma(\widetilde{X}_t^*, \theta) \right| |u_t^*| \leq E^* |u_1^*| L_\sigma E^* \left\| \widetilde{X}_t - \widetilde{X}_t^* \right\|.$$

Finally, by Proposition 1,

$$E^* \sigma(\widetilde{X}_t, \theta) \left| (\varepsilon_t - u_t^*) \right| = E^* \sigma(\widetilde{X}_t, \theta) E^* \left| (\varepsilon_1 - u_1^*) \right| = o_p(1).$$

Thus,

$$E^* \left| \widetilde{Y}_t - Y_t^* \right| \leq (L_\mu + E^* |u_1^*| L_\sigma) E^* \left\| \widetilde{X}_t - \widetilde{X}_t^* \right\| + o_p(1).$$

Under A1, we use the following approximation to the NED process $\widetilde{X}_t$ by the stationary process $\widetilde{X}_t^{(m)}$. That is,

$$\widetilde{X}_t = \widetilde{X}_t^{(m)} + (\widetilde{X}_t - \widetilde{X}_t^{(m)}) = \widetilde{X}_t^{(m)} + \eta_{X,t}^{(m)}, \tag{2}$$

where $E^* \{\eta_{X,t}^{(m)}\}^2 = E\{\eta_{X,t}^{(m)} | \Upsilon_T\}^2 = O(v_2(m))$ as $m \to \infty$. Note that $E^* \{\eta_{X,t}^{(m)}\}^2$ will never increase as $m \to \infty$. Thus, by (A8)

$$E^* \{\eta_{X,t}^{(T)}\}^2 \leq E^* \{\eta_{X,t}^{(m)}\}^2. \tag{3}$$

Letting $\varepsilon_t \equiv 0$ for $t \leq 0$, we have from (2) that for a given t

$$
\begin{aligned}
E^* \left\| \widetilde{X}_t - \widetilde{X}_t^* \right\| \quad &\leq \quad E^* \left\| \widetilde{X}_t - \widetilde{X}_t^{(T)} \right\| + E^* \left\| \widetilde{X}_t^{(T)} - \widetilde{X}_t^* \right\| \\
&\leq \quad (E^* \left\| \widetilde{X}_t - \widetilde{X}_t^{(T)} \right\|^2)^{1/2} + E^* \left\| \Psi_{X,T}(\varepsilon_t, \varepsilon_{t-1}, .., \varepsilon_1) - \Psi_{X,T}(u_t^*, u_{t-1}^*, ..., u_1^*) \right\| \\
&\leq \quad O(\sqrt{v_2(m)}) + E^* \left\| \partial \Psi_{X,T}(x)/\partial x \right\| \left\| \varepsilon_t - u_t^* \right\| = o_p(1),
\end{aligned}
$$

where the second inequality follows by the Liapounov's inequality, and the last inequality follows from (3), (A7) and Proposition 1. ■

Appendix B

Table 1. Distributions used in the Monte Carlo study

Distribution	λ_1	λ_2	λ_3	λ_4	η_3	η_4
Symmetric distributions						
S1		$N(0,1)$			0	3.0
S2	0	0.197454	0.134915	0.134915	0	3.0
S3	0	-1	-0.080000	-0.080000	0	6.0
S4		t_5			0	9.0
S5	0	-0.397912	-0.160000	-0.160000	0	11.6
S6	0	-1	-0.240000	-0.240000	0	126.0
Asymmetric distributions						
A1	-0.116734	-0.351663	-0.130000	-0.160000	0.8	11.4
A2	3.586508	0.043060	0.025213	0.094029	0.9	4.2
A3	0	-1	-0.007500	-0.030000	1.5	7.5
A4		exponential: $-ln(e),\ e \sim U(0,1)$			2.0	9.0
A5		χ_2^2			2.0	9.0
A6	0	-1	-0.100000	-0.180000	2.0	21.2
A7	0	-1	-0.001000	-0.130000	3.16	23.8
A8	0	-1	-0.000100	-0.170000	3.88	40.7

Table 2: Empirical Size of 5%-Level Tests for DGP$_1$

DGP$_1$: $Y_t = \beta_0 + \sum_{i=1}^{k} Z_{it}\beta_i + \sigma u_t$

$k = 1$

		$\widehat{S}_T^3$		$\widehat{S}_T^{3,5}$		CS_T		R_T		WSR_T		TR_T	
S1	50	4.30	(5.25)	2.75	(4.60)	4.30	(2.50)	4.35	(1.75)	0.00	(3.15)	2.40	(3.40)
	100	4.05	(4.65)	2.65	(4.10)	3.90	(3.30)	4.25	(1.60)	0.00	(3.95)	4.00	(5.40)
	200	4.65	(5.20)	2.40	(4.15)	3.70	(4.10)	4.50	(1.15)	0.00	(4.70)	4.50	(5.05)
S2	50	4.00	(4.80)	2.35	(4.70)	4.00	(3.20)	4.90	(2.45)	0.00	(3.40)	2.60	(3.80)
	100	4.50	(4.95)	3.10	(5.30)	2.40	(2.00)	4.85	(1.75)	0.00	(3.75)	2.40	(3.40)
	200	5.10	(5.45)	3.75	(4.65)	4.70	(4.90)	3.90	(0.70)	0.00	(4.60)	2.80	(3.90)
S3	50	3.50	(3.35)	1.85	(3.40)	4.80	(3.40)	4.85	(2.20)	0.00	(4.70)	1.40	(3.40)
	100	3.70	(3.70)	2.25	(4.00)	3.50	(2.80)	4.25	(1.70)	0.00	(5.00)	1.20	(3.80)
	200	3.55	(3.75)	2.60	(4.15)	5.60	(4.80)	4.50	(1.00)	0.00	(5.25)	2.20	(4.20)
S4	50	2.30	(2.55)	1.90	(3.25)	5.40	(3.60)	4.60	(2.20)	0.00	(4.10)	1.20	(3.20)
	100	2.70	(2.85)	2.05	(3.65)	5.20	(4.50)	4.60	(1.80)	0.00	(3.90)	3.20	(6.20)
	200	2.65	(3.15)	2.10	(3.90)	4.70	(3.70)	5.60	(1.20)	0.00	(4.50)	4.10	(5.50)
S5	50	3.60	(3.20)	1.80	(3.10)	6.00	(4.00)	5.85	(3.10)	0.15	(5.15)	1.40	(3.20)
	100	3.20	(2.90)	1.90	(3.60)	5.10	(3.30)	4.30	(1.40)	0.05	(4.75)	1.20	(3.80)
	200	2.65	(3.00)	2.10	(4.05)	7.40	(4.70)	4.55	(0.85)	0.20	(5.15)	2.10	(4.50)
S6	50	3.35	(2.65)	1.80	(2.75)	8.50	(4.80)	5.35	(2.35)	0.80	(5.80)	1.40	(3.20)
	100	3.00	(2.45)	1.70	(3.30)	9.80	(5.00)	5.05	(1.85)	1.00	(5.35)	1.20	(3.60)
	200	2.10	(2.50)	1.65	(3.80)	10.7	(5.20)	5.05	(1.20)	1.30	(5.35)	2.15	(4.05)

Note: The first number in each entry is empirical size based on asymptotic critical values, while the number in parentheses is empirical size based on bootstrap critical values.

Table 2: Empirical Size of 5%-Level Tests for DGP$_1$ (Continued)

DGP$_1$: $Y_t = \beta_0 + \sum_{i=1}^{k} Z_{it}\beta_i + \sigma u_t$

$k = 4$

		$\widehat{S}_T^3$		$\widehat{S}_T^{3.5}$		CS_T		R_T		WSR_T		TR_T	
S1	50	3.00	(4.10)	1.90	(4.00)	3.90	(3.10)	5.50	(4.70)	0.00	(2.90)	1.60	(4.80)
	100	3.35	(4.10)	1.90	(4.45)	3.00	(3.30)	4.10	(3.75)	0.00	(3.70)	4.20	(5.40)
	200	4.40	(4.80)	3.10	(5.00)	3.30	(4.10)	4.40	(4.00)	0.00	(4.60)	4.60	(4.95)
S2	50	2.90	(4.00)	1.75	(3.45)	4.10	(3.60)	4.65	(3.80)	0.00	(2.35)	3.00	(4.60)
	100	4.15	(4.65)	2.35	(5.05)	3.10	(3.40)	4.55	(4.20)	0.00	(3.85)	3.50	(4.90)
	200	5.00	(5.05)	3.70	(5.15)	3.50	(4.50)	4.80	(4.20)	0.00	(4.45)	4.60	(5.20)
S3	50	3.35	(4.10)	2.00	(3.85)	5.80	(5.10)	4.90	(4.20)	0.00	(4.50)	2.40	(4.40)
	100	3.75	(3.80)	2.20	(4.20)	4.10	(4.30)	4.70	(4.40)	0.00	(5.35)	2.20	(4.60)
	200	4.00	(4.30)	3.15	(4.55)	5.30	(4.70)	4.05	(3.70)	0.00	(4.80)	3.10	(4.35)
S4	50	2.80	(3.15)	1.55	(3.60)	4.90	(3.60)	4.50	(3.95)	0.00	(4.00)	2.40	(5.20)
	100	2.85	(2.95)	2.10	(3.85)	3.50	(3.90)	5.15	(4.55)	0.05	(5.00)	2.60	(5.80)
	200	3.05	(3.40)	2.75	(4.00)	5.00	(4.50)	4.70	(4.15)	0.05	(4.70)	3.60	(5.40)
S5	50	3.15	(3.40)	1.90	(3.75)	6.60	(5.60)	4.10	(3.45)	0.10	(5.35)	2.40	(4.60)
	100	3.55	(3.10)	1.95	(4.10)	4.80	(3.90)	4.65	(4.05)	0.10	(5.60)	2.20	(4.90)
	200	2.85	(3.50)	2.75	(4.25)	6.80	(5.50)	5.25	(4.70)	0.00	(4.70)	2.90	(5.10)
S6	50	3.05	(3.10)	1.90	(3.40)	8.00	(6.10)	4.65	(4.00)	0.45	(7.50)	2.40	(4.80)
	100	2.75	(2.90)	1.90	(3.60)	6.80	(4.80)	4.60	(3.85)	0.75	(6.10)	2.45	(4.55)
	200	2.35	(2.95)	2.25	(3.90)	9.40	(5.10)	4.95	(4.15)	0.80	(4.90)	3.00	(4.90)

Note: The first number in each entry is empirical size based on asymptotic critical values, while the number in parentheses is empirical size based on bootstrap critical values.

Table 3: Empirical Size of 5%-Level Tests for DGP$_2$

DGP$_2$: $Y_t = c + \rho Y_{t-1} + \sigma u_t$

		$\rho = 0.5$				$\rho = 0.8$			
		CS_T	R_T	WSR_T	TR_T	CS_T	R_T	WSR_T	TR_T
S1	50	3.90 (2.80)	4.75 (2.15)	0.00 (2.70)	2.20 (4.40)	3.70 (2.40)	4.85 (2.45)	0.00 (2.65)	1.20 (4.40)
	100	2.70 (2.40)	5.35 (1.95)	0.00 (3.40)	2.40 (4.60)	2.80 (2.90)	4.80 (1.70)	0.00 (4.05)	2.00 (4.25)
	200	3.00 (3.30)	4.50 (0.85)	0.00 (5.15)	2.60 (5.05)	3.40 (3.70)	3.75 (0.75)	0.00 (5.70)	2.25 (5.20)
S2	50	5.10 (3.50)	4.00 (2.05)	0.00 (3.10)	2.00 (4.40)	6.00 (4.10)	4.60 (2.20)	0.00 (3.00)	2.00 (4.20)
	100	4.30 (3.90)	4.20 (1.30)	0.00 (3.50)	1.90 (4.60)	3.90 (4.10)	4.50 (1.75)	0.00 (3.70)	2.00 (4.50)
	200	3.30 (3.90)	4.50 (1.25)	0.00 (4.40)	2.90 (5.20)	3.70 (4.10)	5.40 (1.10)	0.00 (4.75)	2.60 (5.25)
S3	50	5.40 (4.10)	4.60 (2.45)	0.05 (4.35)	1.60 (5.20)	6.00 (3.80)	5.55 (2.70)	0.05 (4.25)	1.80 (4.60)
	100	5.50 (4.50)	4.65 (1.60)	0.00 (4.25)	1.97 (4.80)	4.80 (4.00)	4.80 (1.75)	0.00 (4.15)	2.00 (4.60)
	200	4.70 (3.90)	4.55 (1.00)	0.00 (4.70)	2.14 (5.10)	4.50 (4.00)	4.85 (1.45)	0.00 (4.70)	2.10 (5.00)
S4	50	5.30 (3.60)	3.60 (1.75)	0.05 (3.75)	2.80 (4.40)	5.40 (3.70)	4.55 (2.20)	0.05 (3.80)	2.00 (3.60)
	100	4.10 (3.10)	5.15 (2.00)	0.00 (4.05)	2.60 (4.10)	4.60 (3.70)	4.55 (1.85)	0.00 (4.25)	2.40 (4.20)
	200	4.50 (4.00)	5.05 (1.00)	0.05 (4.40)	3.20 (4.80)	4.70 (4.20)	4.70 (0.80)	0.05 (4.55)	3.00 (4.50)
S5	50	7.40 (4.20)	4.80 (2.25)	0.25 (4.95)	1.60 (5.00)	6.40 (4.00)	4.95 (2.20)	0.30 (5.00)	2.00 (4.20)
	100	7.50 (5.20)	4.30 (2.00)	0.20 (4.60)	1.20 (4.25)	6.10 (4.40)	5.75 (1.85)	0.25 (4.50)	1.50 (4.80)
	200	6.40 (4.30)	4.90 (1.20)	0.20 (4.80)	1.99 (4.45)	5.90 (4.40)	4.60 (1.45)	0.20 (4.80)	2.00 (4.60)
S6	50	10.5 (5.30)	5.15 (2.25)	1.05 (5.70)	1.40 (4.80)	9.80 (4.70)	4.20 (1.60)	0.95 (6.10)	2.00 (4.60)
	100	11.8 (4.90)	5.25 (1.20)	1.55 (4.85)	1.35 (4.62)	10.2 (4.30)	4.65 (2.05)	1.50 (4.95)	1.80 (4.50)
	200	10.6 (4.90)	4.60 (1.00)	1.35 (4.60)	1.97 (4.56)	10.8 (4.40)	4.80 (1.20)	1.45 (4.65)	2.00 (4.70)

Note: The first number in each cell is empirical size based on asymptotic critical values, while the number in parentheses is empirical size based on bootstrap critical values.

Table 4: Empirical Size of 5%-Level Tests for DGP$_3$

DGP$_3$: $Y_t = \mu + \phi u_{t-1} + \sigma u_t$

| | | $\phi = 0.5$ | | | | | | | | $\phi = 0.8$ | | | | | | |
		CS_T		R_T		WSR_T		TR_T		CS_T		R_T		WSR_T		TR_T	
S1	50	4.20	(3.00)	4.52	(2.36)	0.00	(3.16)	2.61	(5.42)	5.21	(2.87)	4.94	(2.79)	0.21	(1.72)	2.58	(4.94)
	100	2.30	(2.10)	4.80	(2.20)	0.00	(3.30)	2.80	(5.20)	3.22	(3.32)	4.29	(2.27)	0.10	(2.32)	2.70	(5.00)
	200	3.10	(3.50)	4.85	(1.20)	0.00	(4.80)	3.10	(4.90)	3.00	(3.70)	4.65	(1.35)	0.00	(4.35)	3.20	(4.95)
S2	50	4.61	(3.20)	5.28	(3.02)	0.00	(2.51)	2.61	(4.21)	4.53	(3.24)	4.83	(2.34)	0.37	(1.85)	3.00	(4.28)
	100	4.40	(4.00)	4.05	(1.95)	0.00	(3.80)	2.70	(4.56)	3.65	(3.55)	3.93	(1.91)	0.10	(2.92)	2.90	(4.70)
	200	3.80	(3.90)	4.85	(1.20)	0.00	(4.10)	3.20	(4.60)	3.01	(3.61)	5.10	(1.30)	0.00	(4.25)	3.00	(4.40)
S3	50	4.71	(3.31)	3.36	(1.40)	0.05	(3.96)	2.20	(4.21)	4.57	(2.34)	5.43	(3.01)	0.26	(2.53)	2.13	(4.47)
	100	4.51	(4.00)	4.65	(2.05)	0.00	(4.15)	2.45	(4.00)	4.72	(4.02)	4.82	(2.51)	0.00	(3.47)	2.30	(4.10)
	200	4.90	(4.50)	4.10	(1.10)	0.00	(4.55)	2.95	(4.40)	6.01	(4.71)	5.50	(1.80)	0.00	(4.05)	2.64	(4.45)
S4	50	5.43	(3.82)	5.31	(2.81)	0.05	(3.71)	2.61	(5.81)	6.54	(4.01)	4.73	(2.44)	0.53	(2.76)	1.08	(3.46)
	100	4.50	(3.40)	4.56	(2.10)	0.10	(4.20)	3.10	(5.40)	5.74	(4.53)	4.13	(2.27)	0.20	(3.83)	2.10	(4.20)
	200	5.20	(4.40)	4.35	(0.90)	0.05	(5.00)	3.00	(5.40)	4.70	(3.70)	5.40	(1.25)	0.00	(4.50)	2.40	(5.00)
S5	50	6.43	(3.82)	4.17	(2.01)	0.15	(4.17)	1.80	(4.41)	4.36	(1.81)	4.56	(2.57)	0.52	(2.93)	1.70	(4.04)
	100	6.40	(4.10)	4.20	(1.40)	0.20	(4.45)	1.90	(4.60)	6.84	(4.33)	4.98	(2.11)	0.30	(3.78)	2.00	(4.35)
	200	6.70	(5.10)	4.40	(1.35)	0.20	(4.60)	2.20	(4.50)	6.80	(5.40)	4.30	(1.15)	0.25	(5.15)	2.00	(4.40)
S6	50	10.1	(4.51)	5.17	(2.46)	1.15	(4.92)	1.80	(4.40)	8.93	(3.93)	6.30	(3.10)	1.78	(4.88)	1.48	(3.38)
	100	9.02	(4.21)	3.91	(1.65)	1.70	(3.96)	2.00	(4.20)	9.87	(4.63)	4.00	(1.80)	1.60	(3.80)	1.70	(4.00)
	200	11.1	(5.20)	4.95	(1.05)	0.95	(4.90)	1.90	(4.70)	11.0	(4.81)	5.11	(1.30)	1.10	(4.76)	2.10	(4.90)

Note: The first number in each cell is empirical size based on asymptotic critical values, while the number in parentheses is empirical size based on bootstrap critical values.

Table 5: Empirical Size of 5%-Level Tests for DGP$_4$

DGP$_4$: $Y_t = \gamma + \sigma(X_t,\theta)u_t$; $\sigma(X_t,\theta) = \{\alpha_0 + \alpha_1\sigma(X_{t-1},\theta)^2 + \alpha_2\sigma(X_{t-1},\theta)^2 u_{t-1}^2\}^{1/2}$

		$(\alpha_0, \alpha_1, \alpha_2) = (2, 0.5, 0.3)$				$(\alpha_0, \alpha_1, \alpha_2) = (2, 0.9, 0.05)$			
		CS_T	R_T	WSR_T	TR_T	CS_T	R_T	WSR_T	TR_T
S1	50	4.90 (3.00)	4.30 (1.85)	0.00 (4.00)	2.80 (4.80)	5.10 (3.80)	5.15 (2.50)	0.00 (5.50)	2.60 (4.40)
	100	3.40 (2.90)	4.50 (1.55)	0.00 (6.70)	2.60 (5.10)	2.50 (2.30)	4.50 (2.70)	0.00 (10.0)	2.80 (5.20)
	200	3.10 (3.20)	5.25 (1.95)	0.00 (8.35)	3.00 (4.90)	2.80 (3.20)	5.35 (3.85)	0.00 (12.5)	3.00 (5.00)
S2	50	5.10 (3.60)	4.75 (1.80)	0.00 (4.90)	3.40 (5.00)	5.70 (4.00)	4.00 (2.15)	0.00 (5.65)	3.00 (5.00)
	100	3.20 (3.00)	3.90 (1.65)	0.00 (7.45)	3.10 (4.90)	4.10 (4.40)	4.10 (2.40)	0.00 (8.75)	3.20 (4.80)
	200	3.70 (4.00)	4.70 (2.40)	0.00 (8.35)	3.40 (5.20)	4.60 (4.80)	5.50 (3.70)	0.00 (11.6)	3.60 (5.10)
S3	50	6.90 (4.00)	3.85 (1.45)	0.15 (3.90)	2.20 (4.80)	7.20 (3.90)	4.65 (2.35)	0.30 (3.80)	2.00 (4.80)
	100	5.00 (3.50)	4.20 (1.50)	0.50 (5.70)	2.60 (5.15)	5.90 (3.70)	5.10 (2.95)	0.90 (5.60)	2.45 (5.40)
	200	5.40 (4.00)	4.75 (1.90)	0.10 (5.80)	2.50 (4.75)	6.00 (4.30)	4.45 (2.10)	0.40 (6.35)	2.30 (5.00)
S4	50	5.80 (2.90)	5.15 (2.40)	0.55 (3.25)	2.00 (5.20)	6.00 (3.20)	4.70 (2.20)	0.60 (4.45)	2.20 (5.40)
	100	5.00 (3.90)	4.75 (1.80)	0.20 (5.35)	2.20 (5.40)	5.90 (3.40)	5.40 (3.00)	0.55 (5.80)	2.00 (5.20)
	200	5.70 (4.10)	4.30 (1.60)	0.40 (6.25)	3.15 (4.90)	5.80 (3.30)	4.45 (2.65)	0.80 (6.60)	3.20 (4.95)
S5	50	8.80 (4.10)	5.00 (1.55)	0.65 (3.70)	1.80 (4.60)	8.90 (3.90)	4.75 (2.30)	1.10 (3.70)	1.80 (4.40)
	100	7.40 (4.20)	3.75 (1.05)	1.15 (5.20)	2.00 (4.80)	8.60 (3.90)	5.05 (1.55)	1.90 (4.75)	1.90 (4.80)
	200	6.50 (4.20)	4.25 (1.90)	1.15 (4.90)	2.45 (5.00)	8.90 (3.80)	4.70 (2.05)	2.15 (5.25)	2.30 (4.90)
S6	50	11.2 (4.20)	4.40 (1.15)	1.45 (4.10)	1.80 (4.60)	10.7 (3.50)	4.90 (1.45)	2.70 (3.90)	1.80 (4.20)
	100	11.5 (4.80)	4.65 (0.80)	3.25 (4.80)	1.90 (4.70)	11.7 (4.50)	5.05 (1.30)	4.40 (4.45)	2.00 (5.00)
	200	11.6 (4.00)	5.50 (2.00)	4.30 (4.90)	2.00 (4.50)	14.6 (4.20)	5.05 (1.85)	6.30 (4.30)	2.00 (4.60)

Note: The first number in each cell is empirical size based on asymptotic critical values, while the number in parentheses is empirical size based on bootstrap critical values.

Table 6: Empirical Power of 5%-Level Tests for DGP$_1$

DGP$_1$: $Y_t = \beta_0 + \sum_{i=1}^{k} Z_{it}\beta_i + \sigma u_t$

k = 1

		$\widehat{S}_T^3$		$\widehat{S}_T^{3,5}$		CS_T		R_T		WSR_T		TR_T	
A1	50	5.00	(4.75)	2.75	(4.10)	7.70	(4.80)	4.80	(2.05)	0.20	(8.40)	2.20	(3.20)
	100	6.60	(6.70)	4.55	(8.15)	11.9	(9.00)	4.45	(1.70)	0.25	(11.3)	2.60	(3.90)
	200	9.75	(11.5)	8.60	(11.8)	15.6	(12.0)	5.30	(1.60)	0.60	(18.4)	2.90	(4.10)
A2	50	33.3	(37.3)	25.2	(32.4)	19.5	(15.8)	7.40	()3.50	0.00	(36.7)	4.00	(6.20)
	100	68.3	(74.3)	67.5	(73.7)	46.4	(47.8)	12.2	()6.40	0.00	(78.9)	4.40	(7.40)
	200	95.6	(98.2)	98.2	(99.0)	81.4	(84.4)	20.8	()8.20	0.30	(99.4)	5.20	(9.00)
A3	50	37.9	(41.0)	32.3	()39.9	34.0	()31.1	11.5	()6.55	0.25	(55.6)	4.60	(7.80)
	100	64.5	(73.1)	79.4	()86.0	75.0	()74.2	23.0	()12.1	1.60	(91.7)	5.00	(10.4)
	200	86.0	(92.7)	99.0	()99.8	97.6	()98.2	44.1	()25.5	20.0	(100)	6.20	(12.5)
A4	50	55.5	(60.2)	67.5	(69.6)	72.5	(67.6)	46.3	(29.5)	2.50	(81.8)	9.60	(13.0)
	100	74.8	(82.1)	95.6	(97.7)	98.6	(98.2)	82.4	(65.9)	18.4	(99.2)	12.4	(20.0)
	200	89.8	(95.3)	99.9	(100)	100	(100)	99.2	(94.4)	76.5	(100)	19.8	(29.5)
A5	50	57.4	(61.1)	67.4	(69.7)	78.4	(70.7)	45.2	(28.9)	1.90	(83.8)	12.4	(15.8)
	100	76.5	(84.2)	95.2	(97.3)	98.8	(98.5)	82.1	(64.9)	17.8	(99.1)	24.5	(27.9)
	200	90.9	(95.9)	99.8	(100)	100	(100)	99.2	(94.7)	77.2	(100)	27.7	(34.5)
A6	50	13.9	(13.3)	8.75	(12.6)	17.9	(12.7)	6.65	(3.45)	1.05	(23.7)	3.00	(5.00)
	100	24.1	(26.6)	23.3	(31.1)	42.1	(37.4)	8.95	(4.70)	3.05	(49.2)	4.10	(5.60)
	200	43.8	(50.9)	54.3	(63.8)	74.0	(70.6)	14.5	(6.15)	14.6	(82.0)	4.90	(6.00)
A7	50	47.0	(48.4)	54.6	(56.9)	85.5	(80.1)	57.3	(36.3)	10.6	(83.7)	13.2	(16.0)
	100	58.0	(66.7)	83.0	(89.1)	99.8	(99.6)	89.4	(72.4)	47.7	(98.1)	23.0	(28.2)
	200	74.4	(82.4)	95.9	(98.3)	100	(100)	99.8	(97.8)	96.8	(100)	29.4	(36.0)
A8	50	44.0	(45.0)	51.5	(53.5)	88.7	(82.9)	61.1	(39.8)	15.9	(84.2)	13.8	(17.2)
	100	53.0	(61.5)	78.1	(84.3)	99.9	(99.8)	91.8	(77.6)	59.1	(97.4)	24.4	(29.8)
	200	67.4	(77.6)	93.9	(96.7)	100	(100)	99.9	(99.0)	98.3	(100)	30.0	(37.9)

Note: The first number in each entry is empirical size based on asymptotic critical values, while the number in parentheses is empirical size based on bootstrap critical values.

Table 6: Empirical Power of 5%-Level Tests for DGP$_1$ (Continued)

DGP$_1$: $Y_t = \beta_0 + \sum_{i=1}^{k} Z_{it}\beta_i + \sigma u_t$

$k = 4$

		$\widehat{S}_T^3$	$\widehat{S}_T^{3,5}$	CS_T	R_T	WSR_T	TR_T
A1	50	4.75 (4.90)	2.60 (4.90)	6.60 (5.90)	4.20 (3.05)	0.05 (9.65)	2.80 (4.60)
	100	6.05 (6.15)	3.95 (6.70)	7.90 (6.50)	4.90 (4.30)	0.05 (12.2)	2.60 (5.10)
	200	8.80 (9.85)	7.55 (11.6)	15.1 (12.8)	4.90 (4.15)	0.40 (18.6)	3.10 (4.90)
A2	50	28.8 (33.5)	20.7 (29.7)	15.3 (15.6)	7.20 (5.80)	0.05 (36.0)	4.20 (6.80)
	100	65.8 (73.2)	61.6 (70.5)	37.3 (41.3)	9.05 (8.50)	0.00 (77.8)	4.40 (6.40)
	200	94.9 (97.4)	97.0 (98.0)	80.6 (84.3)	18.8 (17.0)	0.20 (99.2)	5.00 (8.90)
A3	50	35.0 (39.2)	27.9 (38.6)	25.7 (25.5)	9.35 (8.05)	0.05 (53.8)	4.40 (8.60)
	100	64.8 (72.7)	76.0 (85.0)	66.3 (69.0)	18.9 (17.7)	1.15 (91.6)	5.10 (10.2)
	200	86.7 (92.5)	99.2 (99.7)	95.8 (96.7)	41.0 (38.5)	15.7 (100)	6.90 (19.0)
A4	50	52.9 (57.5)	56.9 (67.1)	62.7 (63.9)	25.7 (22.6)	0.70 (84.1)	8.20 (12.6)
	100	76.1 (83.2)	95.2 (98.3)	97.1 (97.7)	68.9 (65.7)	13.3 (99.7)	9.00 (14.5)
	200	91.7 (95.6)	99.9 (100)	100 (100)	96.8 (95.7)	73.3 (100)	14.9 (29.0)
A5	50	52.8 (57.8)	58.4 (68.7)	64.5 (64.6)	29.0 (25.6)	0.60 (83.9)	6.60 (9.20)
	100	76.4 (84.0)	95.1 (97.9)	97.6 (98.0)	65.8 (62.8)	11.7 (99.3)	10.3 (15.6)
	200	89.6 (94.8)	99.9 (100)	100 (100)	96.1 (95.3)	71.1 (100)	19.9 (20.0)
A6	50	13.3 (14.1)	7.85 (14.2)	14.7 (13.2)	6.80 (5.95)	0.90 (28.3)	3.60 (5.60)
	100	24.4 (27.0)	21.3 (30.4)	33.8 (32.3)	8.00 (7.30)	2.50 (50.2)	5.6 (7.60)
	200	44.8 (52.4)	54.4 (66.0)	70.4 (68.7)	12.5 (11.0)	11.6 (83.8)	6.70 (8.00)
A7	50	43.9 (47.6)	49.3 (60.1)	71.0 (69.9)	33.9 (29.0)	5.05 (86.4)	9.20 (12.8)
	100	58.3 (66.8)	82.4 (90.2)	99.3 (99.0)	74.6 (71.0)	40.4 (98.9)	15.6 (20.4)
	200	73.3 (82.0)	96.3 (98.7)	100 (100)	98.7 (98.2)	95.4 (100)	20.1 (33.4)
A8	50	41.1 (44.8)	46.2 (56.5)	75.3 (75.0)	37.1 (33.0)	7.70 (86.8)	9.00 (12.8)
	100	53.4 (61.8)	77.1 (86.0)	99.3 (99.2)	78.1 (74.5)	51.1 (98.7)	14.4 (21.2)
	200	67.5 (77.7)	93.1 (97.3)	100 (100)	98.9 (98.5)	97.9 (100)	21.2 (34.5)

Note: The first number in each entry is empirical size based on asymptotic critical values, while the number in parentheses is empirical size based on bootstrap critical values.

Table 7: Empirical Power of 5%-Level Tests for DGP₂

DGP₂: $Y_t = c + \rho Y_{t-1} + \sigma u_t$

		$\rho = 0.5$				$\rho = 0.8$			
		CS_T	R_T	WSR_T	TR_T	CS_T	R_T	WSR_T	TR_T
A1	50	7.10 (3.80)	5.00 (2.45)	0.30 (7.30)	1.20 (5.20)	6.70 (4.30)	4.75 (2.50)	0.35 (7.60)	3.00 (4.80)
	100	12.5 (8.60)	4.85 (1.70)	0.35 (11.6)	1.60 (5.40)	11.5 (7.90)	4.75 (1.80)	0.40 (11.5)	2.90 (4.10)
	200	15.4 (12.9)	5.75 (1.40)	0.40 (18.4)	1.80 (5.10)	14.6 (11.7)	3.95 (0.60)	0.35 (18.3)	3.10 (4.60)
A2	50	18.0 (15.2)	6.20 (2.90)	0.00 (38.3)	4.20 (7.60)	18.7 (16.2)	6.90 (4.85)	0.00 (36.5)	5.40 (8.00)
	100	46.3 (47.8)	11.4 (6.10)	0.00 (82.0)	4.40 (8.00)	46.0 (46.3)	11.2 (5.55)	0.00 (81.9)	5.30 (9.90)
	200	81.9 (83.7)	22.3 (8.95)	0.10 (99.5)	5.00 (10.1)	82.0 (84.3)	22.1 (9.85)	0.05 (99.5)	6.00 (11.1)
A3	50	32.7 (29.3)	11.9 (7.40)	0.45 (54.5)	6.80 (8.60)	31.8 (29.2)	11.5 (6.95)	0.35 (53.6)	6.80 (10.0)
	100	74.5 (74.9)	23.4 (13.2)	1.85 (93.8)	7.20 (10.2)	74.5 (74.8)	22.9 (13.8)	1.75 (93.6)	7.10 (11.7)
	200	97.1 (97.5)	46.1 (27.2)	18.2 (100)	8.30 (12.3)	97.6 (97.7)	43.3 (24.4)	18.2 (100)	6.90 (12.2)
A4	50	76.0 (70.5)	48.7 (31.1)	2.35 (82.5)	12.0 (15.4)	73.6 (68.8)	47.6 (31.0)	1.95 (83.0)	14.2 (17.0)
	100	98.8 (98.3)	83.0 (65.5)	17.3 (99.0)	15.0 (16.7)	98.8 (98.6)	84.3 (67.9)	16.5 (99.0)	17.8 (19.3)
	200	100 (100)	99.1 (94.4)	75.9 (100)	20.0 (20.4)	100 (100)	99.2 (94.8)	75.0 (100)	21.3 (24.5)
A5	50	78.5 (69.8)	48.6 (32.2)	2.45 (81.7)	13.4 (17.0)	75.9 (69.9)	45.0 (29.7)	2.70 (82.7)	15.8 (18.4)
	100	99.0 (98.8)	82.8 (65.0)	16.9 (99.1)	16.7 (20.5)	99.1 (98.7)	82.7 (65.5)	17.1 (99.3)	24.5 (21.6)
	200	100 (100)	99.1 (94.3)	76.5 (100)	25.4 (28.6)	100 (100)	99.5 (94.9)	76.6 (100)	28.7 (30.2)
A6	50	18.0 (12.9)	7.30 (3.90)	1.30 (24.9)	3.20 (6.60)	19.1 (13.5)	7.85 (3.40)	1.20 (25.4)	3.20 (5.80)
	100	43.7 (37.3)	8.05 (3.40)	3.25 (51.8)	3.40 (7.00)	43.7 (38.5)	8.65 (4.20)	3.50 (51.8)	4.00 (5.40)
	200	75.7 (71.2)	14.8 (5.50)	14.2 (83.7)	3.10 (7.10)	74.5 (70.5)	13.9 (4.85)	14.0 (83.2)	4.60 (6.00)
A7	50	86.2 (77.2)	58.6 (36.9)	11.2 (81.9)	13.0 (16.4)	84.6 (77.0)	54.9 (37.4)	10.7 (82.7)	17.0 (20.4)
	100	100 (99.9)	90.3 (74.2)	48.9 (97.8)	17.8 (22.9)	99.8 (99.6)	90.5 (74.9)	48.5 (97.9)	26.7 (25.9)
	200	100 (100)	99.8 (97.7)	96.5 (100)	29.0 (29.8)	100 (100)	99.8 (97.2)	96.5 (100)	31.4 (33.4)
A8	50	89.2 (79.3)	64.1 (40.4)	16.1 (81.5)	14.6 (17.4)	87.9 (79.3)	61.5 (40.4)	15.6 (82.4)	17.6 (20.6)
	100	100 (99.8)	93.7 (76.7)	60.4 (97.3)	20.2 (29.7)	100 (99.8)	93.4 (79.2)	59.5 (97.6)	29.8 (30.3)
	200	100 (100)	100 (98.8)	98.1 (99.9)	29.7 (31.4)	100 (100)	100 (97.9)	98.2 (99.9)	32.3 (394)

Note: The first number in each entry is empirical power based on asymptotic critical values, while the number in parentheses is empirical power based on bootstrap critical values.

29

Table 8: Empirical Power of 5%-Level Tests for DGP$_3$

DGP$_3$: $Y_t = \mu + \phi u_{t-1} + \sigma u_t$

| | | $\phi = 0.5$ | | | | | $\phi = 0.8$ | | | | |
		CS_T	R_T	WSR_T	TR_T		CS_T	R_T	WSR_T	TR_T
A1	50	5.31 (3.41)	5.16 (2.51)	0.20 (6.71)	1.80 (4.21)		8.20 (4.84)	5.97 (3.51)	0.53 (4.00)	1.52 (4.99)
	100	11.6 (7.60)	4.96 (2.20)	0.30 (11.0)	2.00 (5.40)		8.54 (7.04)	5.10 (2.30)	0.15 (10.6)	1.90 (5.10)
	200	14.7 (11.7)	4.98 (1.54)	0.15 (19.9)	2.10 (5.20)		15.6 (11.3)	4.90 (1.70)	0.20 (21.4)	2.15 (5.40)
A2	50	20.3 (17.8)	7.67 (4.66)	0.00 (31.5)	5.61 (9.02)		18.3 (15.3)	8.64 (5.36)	0.00 (30.4)	3.84 (6.82)
	100	46.1 (47.0)	10.8 (5.70)	0.00 (78.2)	6.00 (10.5)		42.3 (43.5)	11.2 (6.10)	0.00 (79.5)	4.50 (7.60)
	200	82.9 (83.5)	21.4 (9.35)	0.00 (99.3)	6.20 (11.3)		81.6 (83.6)	20.9 (9.69)	0.00 (98.5)	4.90 (10.4)
A3	50	37.0 (31.7)	11.1 (6.72)	0.25 (47.7)	4.83 (8.05)		33.1 (30.5)	13.9 (8.44)	0.30 (46.5)	3.20 (7.68)
	100	73.5 (73.5)	23.8 (14.5)	1.90 (91.4)	6.50 (10.9)		71.0 (70.5)	24.3 (13.7)	2.00 (92.3)	5.60 (12.5)
	200	97.3 (97.9)	43.9 (25.9)	18.3 (99.9)	7.10 (13.2)		97.3 (97.4)	44.5 (27.4)	19.4 (100)	6.75 (14.3)
A4	50	78.0 (69.6)	48.5 (32.3)	2.31 (76.3)	14.0 (16.6)		73.1 (65.9)	47.2 (33.7)	4.89 (49.8)	12.0 (16.4)
	100	99.2 (98.9)	83.5 (67.4)	18.6 (98.4)	16.7 (18.5)		98.4 (97.7)	84.2 (67.1)	17.9 (98.3)	15.7 (21.0)
	200	100 (100)	99.2 (95.2)	75.4 (100)	19.5 (23.1)		100 (100)	100 (95.5)	73.5 (100)	18.9 (25.3)
A5	50	75.1 (67.7)	46.6 (31.9)	2.30 (74.7)	12.1 (14.9)		76.7 (68.7)	46.2 (31.8)	3.77 (49.7)	11.3 (15.1)
	100	99.0 (98.9)	83.1 (65.5)	18.2 (98.7)	18.5 (20.6)		98.7 (98.7)	84.5 (65.0)	18.5 (99.4)	18.0 (25.7)
	200	100 (100)	99.1 (94.2)	76.5 (100)	23.4 (23.9)		100 (100)	99.7 (95.4)	77.9 (100)	22.1 (29.9)
A6	50	19.9 (14.2)	6.97 (3.86)	1.30 (22.8)	2.82 (5.43)		19.2 (13.9)	7.42 (4.51)	1.00 (23.4)	3.82 (5.52)
	100	43.1 (37.3)	8.47 (3.91)	3.96 (49.9)	3.00 (5.50)		42.8 (36.0)	9.45 (3.90)	2.98 (50.2)	4.10 (5.70)
	200	75.0 (69.7)	15.3 (5.25)	14.2 (83.9)	3.40 (5.90)		73.7 (69.4)	15.2 (5.10)	12.4 (85.6)	4.58 (6.10)
A7	50	85.7 (77.5)	55.0 (35.8)	11.5 (76.2)	13.6 (18.6)		83.6 (75.0)	57.5 (39.1)	10.6 (77.5)	15.1 (17.6)
	100	99.6 (99.4)	91.3 (73.9)	49.2 (97.9)	24.6 (29.7)		99.7 (99.1)	92.3 (74.5)	51.2 (98.9)	25.6 (28.9)
	200	100 (100)	99.9 (97.3)	97.5 (99.8)	31.6 (34.5)		100 (100)	100 (98.7)	96.4 (100)	30.1 (34.7)
A8	50	88.3 (79.1)	62.8 (39.4)	16.0 (76.7)	13.6 (17.0)		87.3 (78.6)	65.1 (49.9)	19.4 (77.8)	11.6 (16.0)
	100	100 (99.7)	93.3 (77.6)	59.1 (97.8)	28.7 (28.6)		99.8 (99.0)	97.6 (76.6)	58.9 (96.7)	27.8 (32.5)
	200	100 (99.9)	100 (99.8)	98.5 (100)	33.6 (35.9)		100 (100)	100 (100)	97.8 (100)	35.3 (42.3)

Note: The first number in each entry is empirical power based on asymptotic critical values, while the number in parentheses is empirical power based on bootstrap critical values.

Table 9: Empirical Power of 5%-Level Tests for DGP$_4$

DGP$_4$: $Y_t = \gamma + \sigma(X_t, \theta)u_t$; $\sigma(X_t, \theta) = \{\alpha_0 + \alpha_1\sigma(X_{t-1}, \theta)^2 + \alpha_2\sigma(X_{t-1}, \theta)^2 u_{t-1}^2\}^{1/2}$

		$(\alpha_0, \alpha_1, \alpha_2) = (2, 0.5, 0.3)$				$(\alpha_0, \alpha_1, \alpha_2) = (2, 0.9, 0.05)$			
		CS_T	R_T	WSR_T	TR_T	CS_T	R_T	WSR_T	TR_T
A1	50	8.40 (3.30)	4.15 (1.60)	0.65 (5.05)	2.00 (5.20)	9.30 (3.70)	4.45 (2.10)	1.40 (4.55)	2.00 (5.00)
	100	11.3 (6.30)	4.95 (1.55)	1.40 (8.90)	1.90 (5.40)	11.7 (6.80)	5.05 (2.15)	2.40 (8.05)	1.80 (5.60)
	200	16.0 (11.5)	5.65 (2.10)	1.50 (15.8)	2.20 (4.90)	16.5 (10.4)	5.30 (2.50)	3.00 (12.9)	1.90 (5.70)
A2	50	20.9 (17.5)	7.55 (3.80)	0.10 (21.1)	4.60 (9.00)	20.1 (17.5)	7.50 (4.35)	0.25 (18.8)	4.60 (8.20)
	100	46.1 (45.1)	11.9 (5.65)	0.40 (59.5)	5.50 (10.1)	45.8 (44.3)	10.8 (7.15)	0.70 (48.2)	7.10 (15.6)
	200	81.9 (83.0)	21.2 (12.3)	1.05 (86.8)	4.90 (12.3)	82.5 (82.3)	19.3 (15.1)	2.20 (75.3)	8.40 (18.4)
A3	50	36.6 (27.9)	11.4 (5.30)	0.70 (33.1)	5.60 (10.4)	34.4 (25.9)	10.5 (5.30)	1.30 (28.2)	6.00 (10.2)
	100	72.6 (68.8)	23.9 (11.8)	3.50 (70.9)	7.80 (13.4)	72.0 (65.9)	24.5 (15.7)	4.85 (58.3)	12.4 (25.3)
	200	96.5 (96.1)	41.4 (24.2)	19.0 (89.8)	9.00 (15.2)	96.4 (94.1)	41.3 (31.3)	21.9 (78.9)	16.7 (29.8)
A4	50	75.9 (64.1)	47.7 (24.3)	2.80 (57.3)	11.6 (14.4)	73.5 (61.3)	45.6 (25.6)	4.15 (49.7)	10.4 (13.2)
	100	97.8 (95.3)	82.8 (51.4)	15.7 (82.4)	19.6 (22.1)	98.2 (92.7)	80.3 (58.7)	18.1 (70.0)	12.4 (25.3)
	200	100 (99.7)	98.5 (81.8)	66.2 (93.5)	25.2 (29.5)	100 (99.2)	98.4 (91.4)	63.2 (84.8)	16.7 (29.8)
A5	50	77.0 (63.1)	46.1 (22.7)	3.20 (55.9)	12.4 (15.0)	74.3 (60.2)	44.1 (24.9)	4.55 (47.7)	11.2 (14.2)
	100	98.1 (96.0)	80.0 (50.2)	15.6 (82.4)	20.3 (27.3)	98.1 (93.7)	78.5 (56.8)	18.4 (70.0)	15.9 (29.6)
	200	100 (99.9)	98.9 (83.6)	66.0 (93.2)	26.9 (29.6)	100 (99.5)	98.5 (91.1)	62.6 (84.0)	25.6 (30.2)
A6	50	19.8 (10.3)	6.45 (2.65)	1.55 (14.6)	3.00 (8.60)	19.2 (9.40)	6.15 (2.15)	2.70 (12.2)	3.40 (7.20)
	100	41.2 (30.9)	9.85 (2.85)	4.60 (36.9)	4.10 (12.3)	41.3 (27.5)	9.30 (4.15)	6.35 (30.1)	4.10 (7.40)
	200	71.8 (59.9)	14.7 (5.40)	15.9 (63.1)	4.00 (13.7)	71.5 (57.1)	14.2 (7.15)	17.6 (52.0)	5.00 (8.10)
A7	50	82.9 (68.5)	56.4 (25.6)	10.5 (57.3)	15.2 (20.0)	80.8 (63.5)	51.6 (25.6)	12.4 (48.0)	14.0 (18.4)
	100	99.2 (96.1)	87.3 (52.2)	42.2 (80.0)	26.4 (30.6)	99.1 (92.1)	85.7 (58.4)	42.1 (67.5)	17.9 (30.6)
	200	100 (98.7)	99.4 (85.2)	88.5 (90.2)	30.6 (40.4)	100 (96.9)	99.4 (90.4)	85.6 (80.8)	27.8 (35.9)
A8	50	86.8 (71.2)	61.0 (27.0)	14.6 (58.0)	15.6 (21.6)	84.5 (65.5)	58.5 (28.3)	16.6 (49.5)	13.6 (19.0)
	100	99.3 (95.2)	90.7 (54.1)	52.4 (79.7)	28.9 (39.0)	99.3 (91.1)	89.7 (60.0)	51.2 (66.8)	25.9 (36.9)
	200	100 (97.9)	99.7 (84.5)	91.3 (89.0)	34.2 (45.2)	100 (96.1)	99.4 (90.5)	89.4 (79.2)	30.0 (42.5)

Note: The first number in each entry is empirical power based on asymptotic critical values, while the number in parentheses is empirical power based on bootstrap critical values.