



WORKING PAPER

2012:2

**Robust critical values for unit root tests for
series with conditional heteroscedasticity
errors: An application of the simple NoVaS
transformation**

Panagiotis Mantalos

ECONOMETRICS AND STATISTICS

ISSN 1403-0586

<http://www.oru.se/Institutioner/Handelshogskolan-vid-Orebro-universitet/Forskning/Publikationer/Working-papers/>

Örebro University
Swedish Business School
701 82 Örebro
SWEDEN

**Robust critical values for unit root tests for series with conditional
heteroscedasticity errors: An application of the simple NoVaS
transformation**

Panagiotis Mantalos

Departments of Statistics,

Swedish Business School at Orebro University, Sweden.

ABSTRACT

In this paper, we introduce a set of critical values for unit root tests that are robust in the presence of conditional heteroscedasticity errors using the normalizing and variance-stabilizing transformation (NoVaS) in Politis (2007) and examine their properties using Monte Carlo methods. In terms of the size of the test, our analysis reveals that unit root tests with NoVaS-modified critical values have actual sizes close to the nominal size. For the power of the test, we find that unit root tests with NoVaS-modified critical values either have the same power as, or slightly better than, tests using conventional Dickey–Fuller critical values across the sample range considered.

Keywords: Critical values, normalizing and variance-stabilizing transformation, unit root tests

JEL Classification Codes: C01, C12, C15

1. Introduction

Recently, there has been great interest by researchers and practitioners in the simultaneous presence of high-persistence and conditional heteroscedasticity. The main reason is their occurrence in many economic time series, including stock index and exchange rate series, both a common subject for unit root tests. Alongside burgeoning interest in the second moment (or variance), the first difference of economic variables—here, stock index returns and exchange rate changes—has also been the subject of a sizeable modelling literature in finance and related disciplines, with the most representative approach being the autoregressive conditional heteroscedasticity (ARCH) and generalized ARCH (GARCH) family of models.

Subsequent to the seminal work on ARCH/GARCH by Engle (1982) and Bollerslev (1986), the introduction of new models has continued apace, a common feature being that they estimate and allow for time-varying volatility. A number of suitable surveys of GARCH models are available, with Giraitis, Leipus and Surgailis (2006) providing a useful reference of existing work across the various models that have predictive ability for squared returns (or variance). Of this body of work, Pantula (1989) is one of the first studies investigating non-stationary univariate autoregressive models with a regular unit root and a first-order ARCH error, deriving the asymptotic distribution of the least squares estimator (LSE) of the unit root, the same as that given by Dickey and Fuller (1979).

Subsequently, Kim and Schmidt (1993) employed Monte Carlo simulation to show that the Dickey–Fuller (DF) tests tended to over-reject in the presence of GARCH errors. However, they also found that the problem was not very serious, the exception being near-integration of the GARCH process errors where the volatility parameter was not small. Ling

and Li (1998a) later investigated an autoregressive integrated moving average (ARIMA) model with GARCH (p, q) errors and derived the asymptotic distribution of the maximum likelihood estimation (MLE). They found that the asymptotic distribution of the MLE of various unit roots involves a series of bivariate Brownian motions and that the MLE of unit roots is more efficient than the corresponding LSE when GARCH innovations are present. Using these asymptotic distributions, Ling and Li (1998b) later constructed some new unit root tests and showed, using Monte Carlo simulation, that these tests could provide better performance than DF tests.

While there are differences across these and other studies, they all suggest that unit root tests require a set of robust critical values that apply in the presence of GARCH errors. For this purpose, this paper provides a simple and easy-to-apply algorithm using the normalizing and variance-stabilizing transformation (NoVaS) in Politis (2007) to produce robust critical values for ensuring DF tests have appropriate size and adequate power in the presence of conditional heteroscedasticity. However, because the subject of unit root tests is large, we limit the analysis here to investigating and demonstrating our procedure as it applies to unit root tests in the absence of serial correlation.

The remainder of the paper is organized as follows. Section 2 presents the simple NoVaS and the unit root test applying in the absence of serial correlation. In Section 3, we present our simulations and use these to establish the size and power properties of the tests. Section 4 provides a brief summary and the main conclusions.

2. NoVaS transformation and the unit root test

2.1. Simple NoVaS transformation

Politis (2007) first defined and analysed the properties of the NoVaS transformation. Let us consider the NoVaS more closely and thereby observe the motivation for our study.

Let ε_t be a sequence of random variables following the ARCH(q) model

$$\varepsilon_t = e_t \sqrt{h_t} \quad h_t = \omega + \sum_{i=1}^q a_i \varepsilon_{t-i}^2 \quad (1.1)$$

where $\omega > 0$, $a_i \geq 0$, $1 \leq i \leq q$ and the innovations e_t are i.i.d $N(0,1)$.

Note, we use the innovations e_t as a standard normal because we can interpret the quantity

$$\varepsilon_t^S = \frac{\varepsilon_t}{\sqrt{\omega + \sum_{i=1}^q a_i \varepsilon_{t-i}^2}} \quad (1.2)$$

as a “studentized” ε_t sequence. Moreover, this provides an intuitive explanation of the following NoVaS (2.3) quantity

$$u_{t,a} = \frac{\varepsilon_t}{\sqrt{W_{t,a}}} \text{ for } t = k+1, k+2, \dots, T \quad (1.3)$$

where $W_{t,a} = \gamma s_{t-1}^2 + a_0 \varepsilon_t^2 + \sum_{i=1}^k a_i \varepsilon_{t-i}^2$, $s_{t-1}^2 = \sum_{k=1}^{t-1} \varepsilon_k^2$, $\gamma \geq 0$, $a_i \geq 0$ for all $i \geq 0$ and

$$\gamma + \sum_{i=0}^k a_i = 1.$$

Equation (1.3) describes the NoVaS transformation proposed by Politis (2007) under which the data series ε_t is mapped to the new series $u_{t,a}$.

The practitioner needs to select the order $k(\geq 0)$ and the vector of non-negative parameters with the twin goals of normalization/variance-stabilization in mind. With a focus on that primary goal of normalization and variance stabilization, Politis (2007) introduced the so-called *simple* NoVaS transformation for choosing the order $k(\geq 0)$ of the parameters a_i in (2.3). Mantalos (2010) subsequently modified the algorithm for the simple NoVaS as follows.

Let $\gamma = 0$ and then $a_i = 1/(k+1)$ for all $0 \leq i \leq k$.

Select the order $k(> 0)$ such that the p-value of the Jarque–Bera statistic for the $u_{t,a}$ NoVaS-transformed series is maximized. In our analysis, we undertake this in a loop for $k = 1$ to 25.

Consider now that $u_{at} = \varepsilon_t / \sqrt{W_{t,a}}$ is the new NoVaS-transformed innovation. Then the transformed series has the following mean

$$E(u_{t,a}) = \frac{E(\varepsilon_t)}{E \sqrt{W_{t,a}}} \quad (1.4)$$

It is not difficult to show that, based on the assumption that the innovations e_t are i.i.d $N(0,1)$ (see equation (1.1)), that (1.4) becomes zero, that is

$$E(u_{t,a}) = \frac{E(\varepsilon_t)}{E \sqrt{W_{t,a}}} = 0 \quad (1.5)$$

As shown in the Appendix, the variance of the Simple NoVaS transformed series is

$$\text{Var } u_{t,a} = E(u_{at}^2) = \frac{E(\varepsilon_t^2)}{E W_{t,a}} = \frac{E(\varepsilon_t^2)}{E(\varepsilon_t^2)} = 1. \quad (1.6)$$

The skewness of the innovations e_t are i.i.d $N(0,1)$ is zero. Lastly, as also shown in the Appendix, the kurtosis is

$$kurt^z = \frac{E u_{at}^4}{E u_{at}^2} = \frac{E(\varepsilon_t^4)/E W_{t,a}^2}{E(\varepsilon_t^2)^2/E W_{t,a}^2} = 3 \frac{k}{k-2}. \quad (1.7)$$

As Equations (1.5) and (1.6) provide zero skewness and kurtosis close to three, the transformed series exhibits the same properties as an i.i.d $N(0,1)$ variable.

2.2. DF unit root tests in the absence of serial correlation

It is well known (Hamilton 1994) that when testing for unit roots it is important to specify the null and alternative hypotheses appropriately based on the type of the data at hand. For example, if the data do not display any increasing or decreasing trend, then we should not include a time trend in the test regression. Moreover, the inclusion of deterministic terms will also influence the asymptotic distribution of the unit root test statistics. Consequently, given the importance of the correct alternative hypothesis and the trend type, the data will determine the form of the test regression used. For this reason, and also because non-trending financial series, such as interest and exchange rates and spreads, exhibit non-zero means and GARCH errors, the first model we consider is the following.

2.2.1 Model 1: Constant term in the regression

The test regression is

$$y_t = c + \phi y_{t-1} + \varepsilon_t \quad (1.8)$$

and the hypotheses to be tested are

$$H_0 : \phi = 1 \Rightarrow y_t = I(1) \text{ without a drift} \quad (1.9)$$

$$H_A : |\phi| < 1 \Rightarrow y_t = I(0) \text{ with non-zero mean} \quad (1.10)$$

The test statistic we use is that proposed by Dickey–Fuller: $t_\phi = \frac{\hat{\phi} - 1}{SE(\hat{\phi})}$.

We remind ourselves here that ε_t follows an unknown ARCH/GARCH process that we could estimate, but which we do not require for the purpose of our test procedure.

The second model we consider is as follows.

2.2.2 Model 2: Time trend and constant term in the regression

The test regression is

$$y_t = c + \phi y_{t-1} + bt + \varepsilon_t \quad (1.11)$$

In this case, we assume that the data under the null hypothesis, with an example being a GDP series, follow a simple random walk with drift. That is

$$y_t = c + y_{t-1} + \varepsilon_t, \quad (1.12)$$

whereas under the alternative hypothesis it is an AR(1) model with $|\phi| < 1$ and trend.

However, even here we use the DF test statistic $t_\phi = \frac{\hat{\phi} - 1}{SE(\hat{\phi})}$. Note that this formulation is appropriate for time series such as asset prices, that is, trending time series with several possible kinds of GARCH errors.

2.3. NoVaS critical values for DF unit root tests in the presence of GARCH errors

The underlying idea and method we use in our procedure to produce robust critical values is simple and identical to the manner in which Dickey and Fuller (1979) produced their critical values for the unit root test, that is, by using Monte Carlo methods we can approximate the finite-sample distribution of the unit root test.

Model 1

Consider a sequence of T observations y_1, y_2, \dots, y_T for which we wish to test if there is a unit root in the series. In addition, we assume that our observed data do not display any increasing or decreasing trend while at the same time exhibiting a non-zero mean.

The first step is to estimate the test regression

$$y_t = c + \phi y_{t-1} + \varepsilon_t, \quad (1.13)$$

and calculate the DF test statistic $t_\phi = \frac{\hat{\phi} - 1}{SE(\hat{\phi})}$.

We perform the simple NoVaS transformation on the estimated residuals from (1.13)

$$\hat{\varepsilon}_t = y_t - \hat{c} - \hat{\phi} y_{t-1}, \text{ that it is, } W_{t,a} = a_0 \hat{\varepsilon}_t^2 + \sum_{i=1}^k a_i \hat{\varepsilon}_{t-i}^2.$$

Based on the simple NoVaS properties (2.5, 2.6, 2.7), $u_t = W_{t,a} * z_t$, is a good approximation of the error term in (2.13), where z_t , are T i.i.d random numbers from the standard normal distribution.

The next step is to calculate the new series under the null hypothesis $y_t^* = \sum_1^T u_t$.

We then estimate the regression $y_t^* = c^* + \phi^* y_{t-1}^* + e_t$ and afterwards the DF test statistic

$$t_\phi^* = \frac{\hat{\phi}^* - 1}{SE(\hat{\phi}^*)}.$$

Repeating steps 3–4 a large number of times, N_{MC} (in our Monte Carlo simulations, $N = 1000$), we are able to approximate the finite-sample distribution of the DF unit root test.

By taking the $(1-\alpha)$ quintile of the approximate finite-sample distribution of t_ϕ^* , we obtain the α -level “critical values” ($c_{1\alpha}^*$) and finally reject H_0 if $T_s \geq c_{1\alpha}^*$.

A Monte Carlo estimate of the p-value for testing is $P^* t_\phi^* \geq t_\phi$. We use this to study the size and power of the unit root test (with the assistance of the p-value plot).

Model 2

In the case of a trend

The first step is to estimate the test regression

$$y_t = c + \phi y_{t-1} + bt + \varepsilon_t, \quad (1.14)$$

and calculate the DF test statistic $t_\phi = \frac{\hat{\phi} - 1}{SE(\hat{\phi})}$.

We perform the simple NoVaS transformation on the estimated residuals from (1.14),

such that

$$\hat{\varepsilon}_t = y_t - \hat{c} - \hat{\phi} y_{t-1} - \hat{b}t. \quad (1.15)$$

That is, $W_{t,a} = a_0 \hat{\varepsilon}_t^2 + \sum_{i=1}^k a_i \hat{\varepsilon}_{t-i}^2$.

The next step is to calculate $u_t = W_{t,a} * z_t$, where z_t are as previously T i.i.d random numbers from the standard normal distribution.

The next step is to calculate the new series under the null hypothesis $y_t^* = \sum_1^T (\hat{c} + u_t)$.

We then estimate the regression $y_t^* = c^* + \phi^* y_{t-1}^* + b^*t + e_t$ and afterwards the DF test

$$\text{statistic } t_\phi^* = \frac{\hat{\phi}^* - 1}{SE(\hat{\phi}^*)}.$$

Steps 6–8 are the same as in Model 1.

3. Simulations

3.1. Monte Carlo

In this section, we perform a simple Monte Carlo experiment by generating data for the size of the test for Model 1 using

$$y_t = y_{t-1} + \varepsilon_t. \quad (1.16)$$

For the power test, we generate data using

$$y_t = 0,1 + 0,99y_{t-1} + \varepsilon_t. \quad (1.17)$$

The data-generating process for the size of the test for Model 2 is

$$y_t = 0,1 + y_{t-1} + \varepsilon_t, \quad (1.18)$$

For the power test, we generate data using

$$y_t = 0,1 + 0,99y_{t-1} + 0,08t + \varepsilon_t. \quad (1.19)$$

We study the performance of the NoVaS critical values with error terms *i.i.d* $N(0,1)$. We assume white noise and in all other cases the errors follow a GARCH (1, 1) model with the following data -generating process

$$\begin{aligned} \varepsilon_t &= e_t \sqrt{h_t} & h_t &= 0,001 + 0,199\varepsilon_{t-1}^2 + 0,800h_{t-1} \\ e_t &\sim i.i.dN(0,1) \end{aligned} \quad (1.20)$$

For each model, we perform 5,000 replications for the calculation of size and 1,000 replications for the power functions. Our results are based on 600, 1100, and 2100 replications, where the first 100 time series observations in each replication is discarded to avoid possible initial value effect. We calculate the estimated size of the test by counting how many times we reject the null hypothesis in repeated samples under conditions where the null is true.

3.2. The *p*-value plots

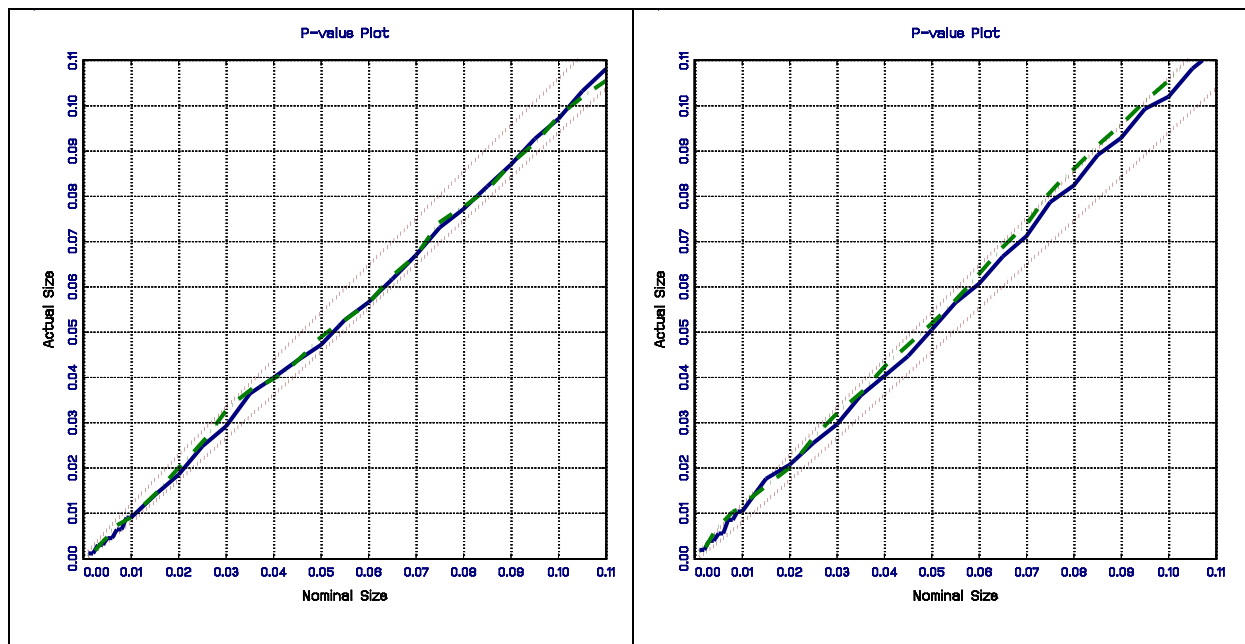
The conventional way of reporting the results of a Monte Carlo experiment is to tabulate the proportion of times we reject the null hypothesis in repeated samples under conditions where the null is true. Concerning the significance levels used when judging the properties of the tests, different studies have advocated both larger and smaller significance levels. For example, Maddala (1992) suggests significance levels of as much as 25% in diagnostic testing, while MacKinnon (1992) suggests going in the other direction.

To address this problem, in this study we use mainly graphical methods that potentially provide more information about the size and power of the tests. We use the simple graphical methods developed and illustrated in Davidson and MacKinnon (1998) as they are relatively easy to interpret. We employ the *p-value plot* to assess the size of the tests and *size–power curves* to evaluate the power of the tests. The graphs of the p-value plots and size–power curves are based on the empirical distribution function (EDF) of the p-values, denoted \hat{F}_{x_j} . For the p-value plots, if the distribution being used to compute the p_s terms is correct, each of the p_s terms should be distributed uniformly about (0,1). Therefore, the resulting graph should be close to the 45° line. The p-value plots also make it possible and easy to distinguish between tests that systematically over-reject or under-reject the null hypothesis and those tests that reject the null hypothesis near the right amount of time.

To judge the reasonableness of the results, we use a 95% confidence interval for the actual size (π_0) given $\pi_0 \pm 2\sqrt{\frac{\pi_0(1-\pi_0)}{N}}$, where N is the number of Monte Carlo replications. We consider any results that lie between these bounds as satisfactory. For example, if we consider 5,000 Monte Carlo replications and a nominal size of 5%, we define a result as reasonable if the estimated size lies between 4.39% and 5.61%.

Figure 1. Size of the unit root tests of Models 1 and 2 for 500 observations with white noise

Figure 1a. Model 1	Figure 1b. Model 2
--------------------	--------------------



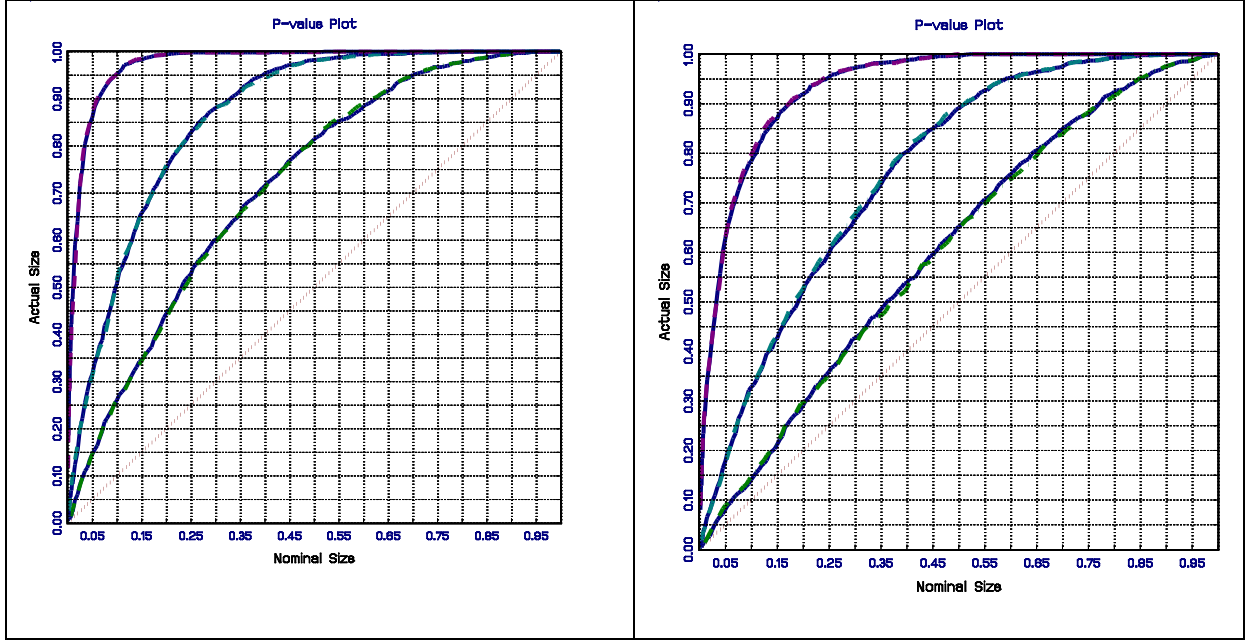
Solid line is unit root test with NoVaS critical values; dot–dash line is unit root test with DF critical values.

3.3. White noise case

We begin our study of the behaviour of our test procedure with the simple case of white noise errors, that is, series without GARCH errors. We show only the results for 500 observations given there is no difference between these and those for 1,000 or 2,000 observations. As shown in Figure 1a, in Model 1 the unit root test behaves well, with both the DF and NoVaS critical values indicating estimated sizes that are inside the confidence interval lines. As we observe the same behaviour for the unit root in Model 2, as shown in Figure 1b, we conclude that the NoVaS critical values work as well as we expect for the case of white noise.

Figure 2. Power of the unit root tests of Model 1 for 500 observations with white noise

Figure 2a. Model 1	Figure 2b. Model 2
--------------------	--------------------



Solid line is unit root test with NoVaS critical values; dot–dash line is unit root test with DF critical values.

Figure 2 depicts the power of the tests. The first thing we observe is that the unit root tests display the same power when using both the DF and NoVaS critical values. It is also obvious that there is a sample effect (the lowest curves are for 500 observations, followed by 1,000 observations, and the highest curves are for 2,000 observations). In sum, the unit root tests with NoVaS critical values assuming only white noise error work equally as well as the unit root tests with conventional DF critical values.

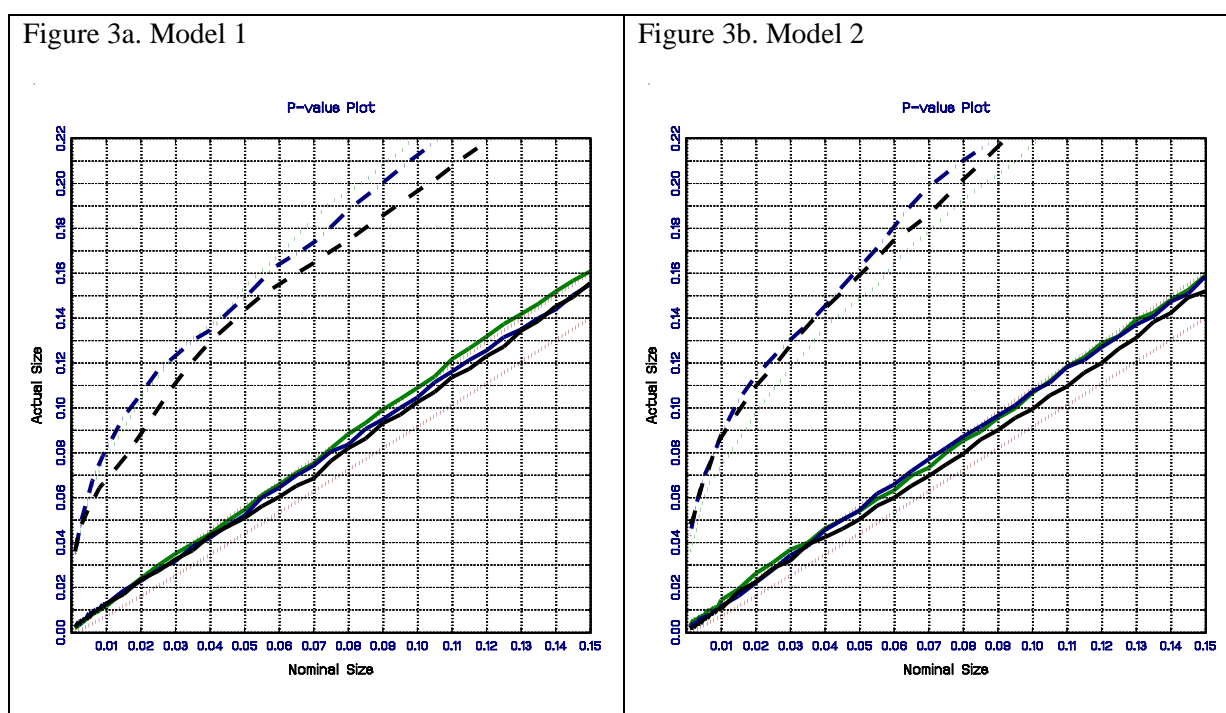
3.4. GARCH case

3.4.1. The size of the tests

In this subsection, we present the results of our Monte Carlo experiment concerning the size of the unit root test given GARCH errors. As shown in Figure 3a for Model 1 and Figure 3b for Model 2 for 500, 1,000 and 2,000 observations, the p-value plots make it possible to easily distinguish that unit root tests with DF critical values systematically over-reject the null

hypothesis, whereas unit root tests with NoVaS critical values reject the null hypothesis at about the right amount of time. Once again, we recall that we use the constant term in the regression for Model 1 and a time trend and constant term in the regression in Model 2. As shown, the unit root test with DF critical values over-rejects by almost three times the nominal size, that is, for a 5% nominal size and Model 1, we estimated sizes of 15.35% for 500 observations, 14.9% for 1,000 observations and for 14.35% for 2,000 observations. We observe a similar picture for the unit root test with DF critical values for Model 2, where for a 5% nominal size we estimated sizes of 14.85 % for 500 observations, 16.26% for 1,000 observations, and 15.9% for 2,000 observations. Overall, the unit roots tests with NoVaS critical values unlike those with DF critical values work well in the presence of GARCH errors.

Figure 3. Size of the unit root tests for 500, 1,000 and 2,000 observations with GARCH errors



Solid line is unit root test with NoVaS critical values; dotted, dot–dash, and dashed lines are unit root tests with DF critical values for 500, 1,000, and 2,000 observations, respectively.

3.4.2. The power of the tests

In this subsection, we analyse the power of the Wald and bootstrap tests using sample sizes of 500, 1,000, and 2,000 observations. We estimate the power function by calculating the rejection frequencies for 1,000 replications using Equation (1.17) for Model 1 and Equation (1.19) for Model 2. We employ the size–power curves to compare the estimated power functions of the alternative test statistics. We follow the same process to evaluate the EDFs denoted by $\hat{F}^{\oplus}(\mathbf{x}_j)$ using the same sequence of random numbers used to estimate the size of the tests.

Figure 4 displays the results using the size–power curves. As shown, the unit root tests with the DF critical values are now superior to the unit root tests with the NoVaS critical values. We also again observe a sample effect, such that the larger the sample, the greater the power of the tests. To ensure a fair comparison of the power of the unit tests with the DF and NoVaS critical values, we apply the size–power curves on a correct size-adjusted basis.

Figure 4 plots the size–power curves and the estimated power functions against the nominal size. By plotting the estimated power functions against the true size, that is $\hat{F}^{\oplus}(\mathbf{x}_j)$ against $\hat{F}(\mathbf{x}_j)$, we obtain size–power curves on a correct size-adjusted basis. As the unit root tests with DF and NoVaS critical values now share the same power, we can see they are even better for a small nominal size (less than 15%), as shown in Figure 5. The main finding for our power investigation is that unit root tests with NoVaS critical values generally perform adequately in both the white noise and GARCH error cases.

Figure 4. Power of the unit root tests for 500, 1,000, and 2,000 observations with GARCH errors

Figure 4a. Model 1

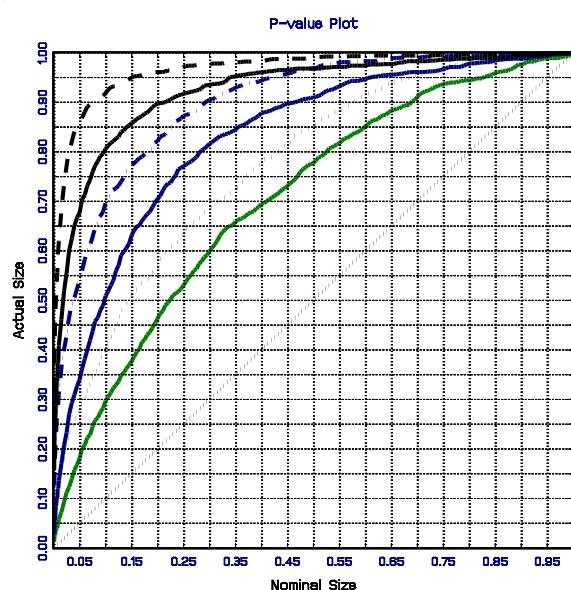
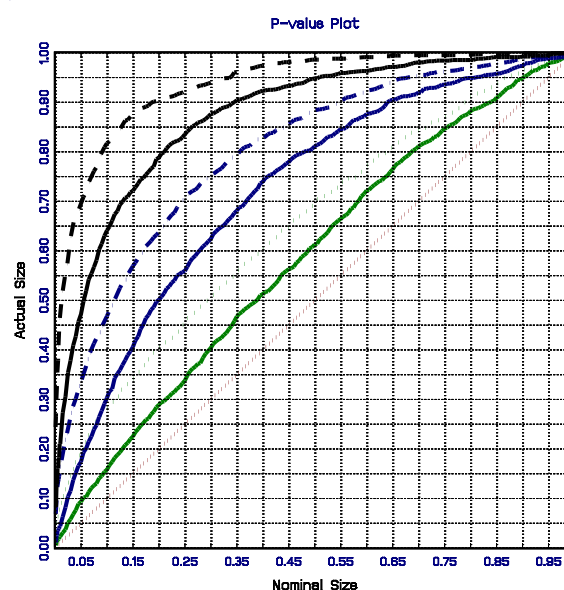


Figure 4b. Model 2



Solid line is unit root test with NoVaS critical values; dotted, dot-dash, and dashed lines are unit root tests with DF critical values for 500, 1,000, and 2,000 observations, respectively.

Figure 5. Size-adjusted power of the unit root tests for 500, 1,000, and 2,000 observations with GARCH errors

Figure 5a. Model 1

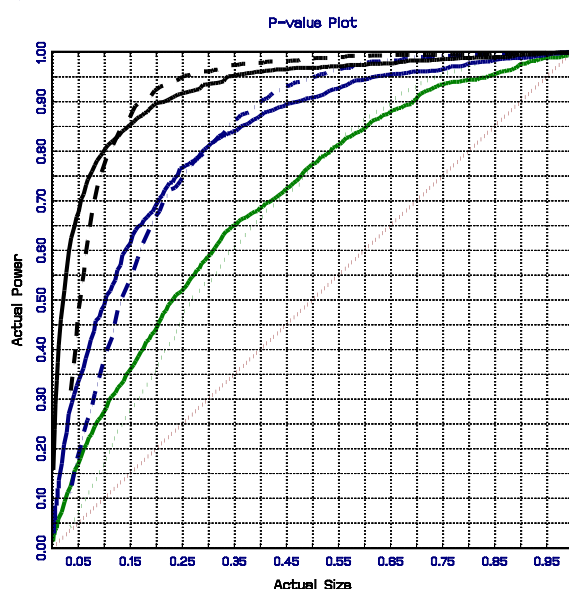
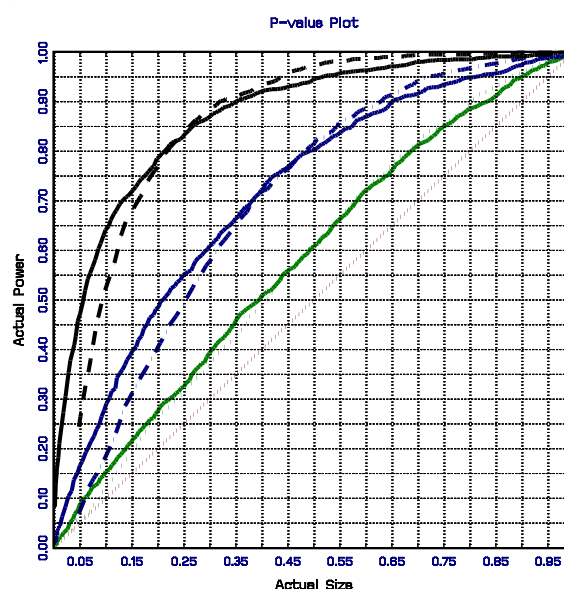


Figure 5b. Model 2



Solid line is unit root test with NoVaS critical values; dotted, dot–dash, and dashed lines are unit root tests with DF critical values for 500, 1,000, and 2,000 observations, respectively.

4. Brief summary and conclusion

In this section, we summarize the results of our investigation. The purpose of this study has been to study the unit root test in the presence of conditional heteroscedasticity errors with newly proposed critical values modified using the normalizing and variance-stabilizing transformation (NoVaS) method. The main conclusion is that NoVaS-modified critical values are robust, with respect to both white noise and conditional heteroscedasticity errors. Moreover, given they have an actual size that lies close to their nominal size and adequate power, it makes sense that we would select our NoVaS-modified critical values ahead of the conventional DF critical values, especially in the presence of conditional heteroscedasticity errors.

In summarizing our methodology, we studied the estimated size and power of the unit root test in the presence of conditional heteroscedasticity errors using two sets of critical values: namely, conventional DF values and our newly proposed NoVaS-modified values. In terms of the size of the tests, we used Monte Carlo methods to investigate their properties using 5,000 replications per model, with 500, 1,000, and 2,000 observations. We employed p-value plots to investigate the size of the tests and found that the unit root tests perform better with NoVaS-modified critical values across all of our samples. Importantly, when we consider the power results using the size–power curves, when considered on a correct size-adjusted basis, the unit root tests with NoVaS-modified critical values share the same power as tests using DF critical values or at least the difference is very small across all of our samples.

5. References

- T. Bollerslev, Generalized autoregressive conditional heteroscedasticity, *J. of Economet.* 31(1986), pp. 307–328.
- Davidson, R. and J.G. MacKinnon (1998). “Graphical methods for investigating the size and power of test statistics”. *The Manchester School*, 66, 1–26.
- Dickey, D. A, and Fuller W.A (1979). “Distribution of the Estimators for Autoregressive Time Series with a Unit Root.” *J. Amer. Statist. Assoc.* 74, 427–431.
- Engle, Robert F. (1982). “Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation”. *Econometrica*, 50, 987–1008.
- L. Giraitis, R. Leipus and D. Surgailis, D. *Recent advances in ARCH modelling*, in Long Memory in Economics, A. Kirman and G. Teyssiere, eds., Springer, Berlin, 2006.
- J. Hamilton, *Time Series Analysis*, Princeton University Press, 1994.
- Kim, K and Schmidt, P (1993). “Unit Root Tests with Conditional Heteroskedasticity.” *J. Econometrics* 59, 287–300.
- Ling, S and Li, W.K (1998a). “Limiting Distributions of Maximum Likelihood Estimators for Unstable ARMA Models with GARCH Errors.” *Ann. Statist.* 26, 84–125

Ling, S and Li, W.K (1998b) “Estimation and Testing for Unit Root Processes with GARCH(1,1) Errors” Technical report. Department of Statistics, University of Hong Kong.

MacKinnon, J. G. (1992). “Model Specification Tests and Artificial Regressions,” *Journal of Economic Literature*, 30, 102–146.

Maddala, G. S. (1992). *Introduction to Econometrics*, Second Edition, New York, Maxwell Macmillan.

Pantula, S.G. (1989). “Estimation of Autoregressive Models with ARCH Errors” *Sankhya B*, 50, 119–138

Politis, D.N (2007). “Model-free vs. model-based volatility prediction”, *J. Financial Econometrics*, 5(3), pp. 358–389.

6. Appendix

6.1 Mean, variance, kurtosis and skewness of NoVaS

Let us consider the model $y_t = E[y_t | \psi_{t-1}] + \varepsilon_t$ with $\varepsilon_t | \psi_{t-1} \sim f(0, h_t)$, where f is the unknown density function of the ε_t conditional on the set of past information ψ_{t-1} .

The first moment “the mean of NoVaS”

First we remind ourselves that for the simple NoVaS, $\gamma = 0$. Now for the simple NoVaS with one lag, that is, $W_{t,a} = a\varepsilon_t^2 + a\varepsilon_{t-1}^2$, such that we always take account that ε_t is variance stationary and $m_2 = E[\varepsilon_t^2]$, then the mean is

$$E[W_{t,a}] = [a\varepsilon_t^2 + a\varepsilon_{t-1}^2] = aE(\varepsilon_t^2) + aE(\varepsilon_{t-1}^2) = 2aE(\varepsilon_t^2) = 2am_2. \quad (1.21)$$

For two lags, the mean is $E[W_{t,a}] = 3am_2$, while for three lags it is $E[W_{t,a}] = 4am_2$, and finally for the general k lags, the mean is

$$E[W_{t,a}] = (k+1)am_2 \quad (1.22)$$

The second moment of NoVaS

For the simple NoVaS model with one lag, we have $W_{t,a} = a\varepsilon_t^2 + a\varepsilon_{t-1}^2$. By taking the squares of this and then the expectation, we have

$$\begin{aligned} E[W_{t,a}^2] &= E[a\varepsilon_t^2 + a\varepsilon_{t-1}^2]^2 \\ &= E[a^2\varepsilon_t^4 + a^2\varepsilon_{t-1}^4 + 2a^2\varepsilon_t^2\varepsilon_{t-1}^2] \Rightarrow E[a^2\varepsilon_t^4 + a^2\varepsilon_{t-1}^4 + 2a^2\varepsilon_t^2\varepsilon_{t-1}^2] \\ E[W_{t,a}^2] &= a^2E[\varepsilon_t^4] + a^2E[\varepsilon_{t-1}^4] + 2a^2E[\varepsilon_t^2\varepsilon_{t-1}^2]. \end{aligned} \quad (1.23)$$

If ε_t is fourth-order stationary with $m_4 = E \varepsilon_t^4$ and $m_2 = E \varepsilon_t^2$, then we have

$$E W_{t,a}^2 = a^2 E \varepsilon_t^4 + a^2 E \varepsilon_{t-1}^4 + 2a^2 E \varepsilon_t^2 E \varepsilon_{t-1}^2 = 2 a^2 m_4 + 2a^2 m_2^2 \quad (1.24)$$

In the same manner, we can show that the simple NoVaS model with two lags has

$$E W_{t,a}^2 = 3 a^2 m_4 + 3 2a^2 m_2^2, \quad (1.25)$$

for three lags

$$E W_{t,a}^2 = 4 a^2 m_4 + 6 2a^2 m_2^2 \quad (1.26)$$

and finally for the general k lags

$$E W_{t,a}^2 = (k+1) a^2 m_4 + k(k+1) a^2 m_2^2. \quad (1.27)$$

Now the variance is $Var W_{t,a} = E(W_{t,a}^2) - [E W_{t,a}]^2$.

Consequently, we have

$$Var(W_{t,a}) = (k+1) a^2 m_4 + k(k+1) a^2 m_2^2 - [(k+1) a^2 m_2^2] \quad (1.28)$$

$$Var(W_{t,a}) = a m_4 + k a m_2^2 - [m_2^2] = a m_4 - a m_2^2 = a (m_4 - m_2^2). \quad (1.29)$$

Kurtosis

Consider that $z_t = \varepsilon_t / \sqrt{W_{t,a}}$ is the new NoVaS transformed series. Then the kurtosis of this series is

$$kurt^z = \frac{E z_t^4}{E z_t^2^2} = \frac{E(\varepsilon_t^4) / E W_{t,a}^2}{E(\varepsilon_t^2)^2 / E W_{t,a}^2}. \quad (1.30)$$

We can see from (1.22) that the mean of the simple NoVaS is $E W_{t,a} = (k+1) a E(\varepsilon_t^2)$. It

is easy to see that for $a = 1 / (k+1)$ then (1.22) becomes

$$E W_{t,a} = E(\varepsilon_t^2). \quad (1.31)$$

Based on this, we have the kurtosis of the new transformed series as

$$kurt^z = \frac{E(\varepsilon_t^4)}{E W_{t,a}^2} \quad (1.32)$$

Recall that we derived the second moment of the NoVaS for the general k lags in

(1.27). Now given we expect that $\varepsilon_t / \sqrt{W_{t,a}}$ is i.i.d. normal, and $E \varepsilon_t^4 < \infty$, then

$$kurt W_{t,a} = \frac{E W_{t,a}^4}{E W_{t,a}^2} = 3 \Leftrightarrow E(\varepsilon_t^4) = 3E W_{t,a}^2 \quad (1.33)$$

Consequently, with for a simple analysis of (1.33) with the substitution of the $E W_{t,a}^2$,

and the right-hand side of (1.27), the kurtosis of the NoVaS transformed series

$$\text{becomes } kurt^z = \frac{3k(k+1)a^2}{1 - \left[3 \frac{k+1}{k+1} a^2 \right]} \quad (1.34)$$

Moreover, for $a = 1/(k+1)$, (1.34) becomes

$$kurt^z = \frac{3ka}{1 - \left[3a \right]} = \frac{3k/(k+1)}{k+1 - 3/(k+1)} = 3 \frac{k}{k-2} \quad (1.35)$$

The last part of this equation is the kurtosis of the simple NoVaS transformed series.