WORKING PAPER

2012: 4

TESTING FOR SKEWNESS IN AR CONDITIONAL VOLATILITY MODELS FOR FINANCIAL RETURN SERIES

Panagiotis Mantalos and Alex Karagrigoriou

Department of Statistics, University of Örebro, Sweden &

Department of Mathematics and Statistics, University of Cyprus, Cyprus

ECONOMETRICS AND STATISTICS

ISSN 1403-0586

TESTING FOR SKEWNESS IN AR CONDITIONAL VOLATILITY MODELS FOR FINANCIAL RETURN SERIES

Panagiotis Mantalos and Alex Karagrigoriou

Department of Statistics, University of Örebro, Sweden &

Department of Mathematics and Statistics, University of Cyprus, Cyprus

panagiotis.mantalos@oru.se, alex@ucy.ac.cy

ABSTRACT

In this paper a test procedure is proposed for the skewness in autoregressive conditional volatility models. The size and the power of the test are investigated through a series of Monte Carlo simulations with various models. Furthermore, applications with financial data are analyzed in order to explore the applicability and the capabilities of the proposed testing procedure.

Keywords: ARCH /GARCH model, kurtosis, NoVaS, skewness.

JEL Classification Codes: C01, C12, C15

1. INTRODUCTION

The moments and the variance of economic variables such as stock index returns and exchange rate changes have been the subject of a vast amount of works in financial literature with most representative modeling examples the generalized autoregressive conditional heteroscedasticity (GARCH) models. Since the introduction of the ARCH/GARCH model by Engle (1982) and Bollerslev (1986), numerous new models are being introduced and investigated. All these models are estimated and allow for time-varying volatility. A natural question that arises is the following: What about for time-varying skewness or kurtosis?

It is specifically known that the excess kurtosis makes extreme observations more likely than in the normal case, which means that the market gives higher probability to extreme observations than in the normal distribution. Moreover, the presence of a negative skewness has also interesting and practical implications: The effect of accentuating the left-hand side of the distribution, interpreted as that, the market gives higher probability to decreases than increases in asset pricing.

Some of the works that study that stock return distributions exhibit negative skewness and excess kurtosis, among others, are : Harvey and Siddique, 1999; Premaratne and Bera, 2000, Jondeau and Rockinger (2000) and León, Rubio and Serna, (2004).

Harvey and Siddique (1999) present a way to jointly estimate time-varying conditional variance and skewness. Premaratne and Bera (2000) have suggested capturing asymmetry and excess kurtosis with the Pearson type IV distribution. Similarly, Jondeau and Rockinger (2000) employ a conditional generalized Student-t distribution to capture conditional skewness and kurtosis. Finally, León, Rubio and Serna, (2004) jointly estimate time-varying volatility, skewness and kurtosis using a Gram-Charlier series expansion.

However, researchers before applying various models should ask the following question for capturing the kurtosis and skewness: Are the skewness and kurtosis explained by the second moment of those variables, that is, an ARCH/GARCH effect? Or is it because the underline distribution of the error term is not normal?

The purpose of the present work is to answer the skewness part of the question, by deriving a skewness test for series that have ARCH/GARCH effect. Such a test, in a easy and comprehensive way, can discriminate these two causes, namely the ARCH/GARCH or the non-normal underline distribution.

The test is developed with the implementation of the "Normalising and Variance-Stabilizing Transformation", known as NoVaS transformation, proposed and developed by Politis in a series of papers (2003, 2007).

The remainder of the paper is organized as follows. In Section 2 we evaluate the variance, the skewness and the kurtosis of the NoVaS transformed series and propose the new test procedure for the skewness of the error term. In Section 3, we present our simulations and use them to establish the size and power properties of the proposed test. Section 4 provides an application on financial data and a brief summary with the main conclusions.

2. SKEWNESS, KURTOSIS & THE SKEWNESS TEST

Let us consider the general model

$$y_t = E\left(y_t \mid \psi_{t-1}\right) + \varepsilon_t \qquad (2.1)$$

with $\varepsilon_t \mid \psi_{t-1} \sim f(0, h_t)$, where f is the unknown density function of ε_t conditional on the set of past information ψ_{t-1}. Let also the error $\{\varepsilon_t\}$ be a sequence of random variables following the ARCH(q) model given by:

$$\varepsilon_t = e_t \sqrt{h_t} \qquad h_t = \omega + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2 \tag{2.2}$$

where $\omega > 0$, $a_i \geq 0$ and the innovations e_t are i.i.d $N(0,1)$. Note that the standard normal assumption is used because the quantity

$$\varepsilon_t^S = \frac{\varepsilon_t}{\sqrt{\omega + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2}} \tag{2.3}$$

can be viewed as the *studentized* $\{\varepsilon_t\}$ sequence. Moreover it helps in understanding the NoVaS transformation.

It was the widely reported failure of the ARCH (and GARCH) models to predict squared returns one of the motivations for the introduction of the "Normalizing and Variance–Stabilizing Transformation", known as NoVaS, for financial returns series (see Politis, 2003; 2007). More specifically, Politis (2007) defined and analyzed the properties of the following NoVaS transformation:

$$z_{t,a} = \frac{\varepsilon_t}{\sqrt{\gamma s_{t-1}^2 + a_0 \varepsilon_t^2 + \sum_{i=1}^{k} a_i \varepsilon_{t-i}^2}} \text{ for } t = k+1, k+2, \cdots, T \tag{2.4}$$

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

Equation (2.4) describes the proposed NoVaS transformation under which the data series $\{\varepsilon_t\}$ is mapped to the new series $\{z_{t,a}\}$. The order $k\,(\geq 0)$ and the vector of nonnegative parameters have to be chosen by the practitioner with the twin goals of normalization/variance–stabilization in mind.

Additionally, Politis (2007) introduced also the so called *Simple* NoVaS transformation for choosing the order $k\,(\geq 0)$ and the parameters a_i in (2.4). The algorithm for the simple NoVaS is:

1. Let $\gamma = 0$ and take $a_i \equiv a = 1/(k+1)$ for all $0 \leq i \leq k$.
2. Pick the order $k\,(\geq 0)$ such that the absolute value of the excess kurtosis of $z_{t,a}$ is minimized.

Note that the simple NoVaS algorithm works in such a way that the coefficient a is always adjusted with the nearest integer k which gives the minimum excess kurtosis. Observe that due to step (1) of the algorithm, (2.4) reduces to:

$$z_{t,a} = \frac{\varepsilon_t}{\sqrt{a\sum_{i=0}^{k}\varepsilon_{t-i}^2 (\equiv W_{t,a})}} \text{ for } t = k+1, k+2, \cdots, T \ . \qquad (2.5)$$

Let $z_{t,\alpha} = \varepsilon_t \big/ \sqrt{W_{t,a}}$ be the new NoVaS transformed series. The following result provides the first 3 moments of the $W_{t,a}$ series (the proof is given in the appendix) and is used in Lemma 2 where we provide the moments of new NoVas transformed series $z_{t,\alpha}$.

Lemma 1. Assume that the sequence ε_t is variance stationary and $m_2 = E\left(\varepsilon_t^2\right)$. Then, the first moment of the $W_{t,a}$ series is:

$$E\left(W_{t,a}\right) = \left(k+1\right) a m_2 \ . \qquad (2.6)$$

If in addition the sequence ε_t is fourth-order stationary with $m_4 = E\left(\varepsilon_t^4\right)$ then the second and third moments of $W_{t,a}$ are:

$$E\left(W_{t,a}^2\right) = \left(k+1\right) a^2 m_4 + k(k+1) a^2 \left(m_2\right)^2 \qquad (2.7)$$

and

$$E(W_{t,a}^3) = \left(a m_4 + a k m_2^2\right) m_2 \, . \qquad (2.8)$$

Lemma 2: Under the same assumptions as in Lemma 1, the simple NoVaS Transformated series has zero mean and variance 1. Furthermore, the third and fourth moments for a general lag *k* are given by:

$$skew \equiv E\left(z_{t,a}^3\right) = \frac{E\left(\varepsilon_t^3\right)}{E\left(W_{t,a}^{3/2}\right)} = \sqrt{\frac{k-2}{k}} \frac{E\left(\varepsilon_t^3\right)}{E\left(\varepsilon_t^2\right)^{3/2}}$$

and

$$kurt\left(z_{t,a}\right) \equiv E\left(z_{t,a}^4\right) = \frac{m_4}{am_4 + ka\left[m_2\right]^2}.$$

Proof: Under the assumption that ε_t is variance stationary and $m_2 = E\left(\varepsilon_t^2\right)$ and with the use of (2.6) we have that the mean of the transformed series is

$$E\left(z_{t,a}\right) = \frac{E\left(\varepsilon_t\right)}{E\sqrt{W_{t,a}}} = \frac{E\left(\varepsilon_t\right)}{\sqrt{E(W_{t,a})}} \ .$$

Due to the fact that $E\left(W_{t,a}\right) > 0$ and based on (2.1) where $E\left(\varepsilon_t\right) = 0$ we have that $E\left(z_{t,a}\right) = 0$.

The second moment of simple NoVaS transformed series is given by:

$$E\left(z_{t,a}^2\right) = \frac{E\left(\varepsilon_t^2\right)}{E\left(W_{t,a}\right)} = \frac{E\left(\varepsilon_t^2\right)}{\left(k+1\right)aE(\varepsilon_t^2)} = 1 \qquad (2.9)$$

Where the last equality holds for $a = 1/\left(k+1\right)$. Thus, the NoVaS transformed variable $z_{t,a}$ has mean zero and variance one.

For the fourth moment of the simple NoVaS for general $\left(k\right)$ lags recall the assumption that ε_t is fourth-order stationary with $m_4 = E\left(\varepsilon_t^4\right)$ and $m_2 = E\left(\varepsilon_t^2\right)$. Then,

$$E\left(z_{t,a}^4\right) = \frac{E\left(\varepsilon_t^4\right)}{E\left(W_{t,a}^2\right)} \ . \qquad (2.10)$$

Using (2.7) and taking $a = 1/(k+1)$ we have

$$E\left(W_{t,a}^{2}\right) = am_4 + ka\left(m_2\right)^2 . \tag{2.11}$$

Combining (2.9) – (2.11) we have that the kurtosis of the simple NoVaS is:

$$kurt\left(z_{t,a}\right) = \frac{m_4}{am_4 + ka\left[m_2\right]^2} \tag{2.12}$$

Restricted (2.12) to 3, or near to 3 (by the NoVaS algorithm), it is easy to derive that:

$$m_4 = \frac{3k\left[m_2\right]^2}{k-2} . \tag{2.13}$$

Finally, for general (k) lags the third moment of the simple NoVaS is:

$$E\left(z_{t,a}^{3}\right) = \frac{E\left(\varepsilon_t^3\right)}{E\left(W_{t,a}^{3/2}\right)} . \tag{2.14}$$

Using (2.13) and (2.14) we get:

$$E(W_{t,a}^{3/2}) = \left\{\left[\left(\frac{a3k\left[m_2\right]^2}{k-2} + ak\left[m_2\right]^2\right)\right]\left(m_2\right)\right\}^{1/2} \tag{2.15}$$

$$= \left\{\left[\frac{ak\left[m_2\right]^2\left[k+1\right]}{k-2}\right]\left(m_2\right)\right\}^{1/2} = \sqrt{\frac{k}{k-2}}\left[m_2\right]^{3/2} . \tag{2.16}$$

Substituting (2.16) into (2.14) we finally obtain the third moment of the simple NoVaS transformed series:

$$E\left(z_{t,a}^{3}\right)=\frac{E\left(\varepsilon_{t}^{3}\right)}{E\left(W_{t,a}^{3/2}\right)}=\sqrt{\frac{k-2}{k}}\,\frac{E\left(\varepsilon_{t}^{3}\right)}{E\left(\varepsilon_{t}^{2}\right)^{3/2}}\,. \tag{2.17}$$

Expression (2.17) states the main result of our work because it reveals the relation between the skewness of the simple NoVaS and that of the original ε_t series. More specifically, it states that the skewness of the simple NoVaS transformed series $z_{t,\alpha}$ is $\frac{k-2}{k}$ times the skewness of the original untransformed series ε_t.

Let us focus on the third moment of the untransformed series in ARCH/GARCH form:

$$E\left(\varepsilon_{t}^{3}\right)=E(e_{t}^{3})E(h_{t}^{3/2})\equiv skew^{e}\left[h_{t}^{3/2}\right] \tag{2.18}$$

where e_t, h_t as in (2.2).

It is easily seen from (2.18) that if $skew^{e}=0$ then the skewness of the transformed series $skew^{z}$ is also zero. On the other hand, if $skew^{e}\neq 0$ then since (2.18) is not zero, so is the skewness of the NoVaS transformed series $skew^{z}$.

This last observation can be used to derive a skewness test for the error term in (2.1) as follows:

1. NoVaS transform the error $\{\varepsilon_t\}$ sequence.

2. Calculate the test statistic:

$$SKEW_z = T\frac{k}{k-2}(skew)^2 \tag{2.19}$$

where $skew$ is the sample skewness of the NoVaS transformed series given by:

$$skew = \frac{1/T\sum_{i=1}^{T}(z_t - \bar{z})^3}{\left[1/T\sum_{i=1}^{T}(z_t - \bar{z})^2\right]^{3/2}}. \tag{2.20}$$

Observe that the test statistic (2.20) under the null hypothesis follows a chi-square distribution with one degree of freedom χ_1^2. The null hypothesis is therefore, rejected for values of the test statistic that exceed the $100(1-\alpha)^{\text{th}}$ percentile $\chi_{1;a}^2$.

3. SIMULATIONS

In this section we provide the characteristics of the Monte Carlo experiment. We calculate the *estimated size* by observing the number of times the correct null hypothesis is rejected in repeated samples. By varying factors such as the number of observations (500, 1000, 2000 and 3000) we obtain a succession of estimated percentages of correct rate of selection under different conditions.

The Monte Carlo experiment has been performed by generating data according to the following GARCH(1,1) data generating processes:

$$\varepsilon_t = e_t \sqrt{h_t} \qquad h_t = \omega + \delta \varepsilon_{t-i}^2 + \beta h_{t-1},\ \omega = 1 - (\delta + \beta).$$

Models 1-5 are used to estimate the size of the test while, for the power we use Models 6 and 7. Note also that for Models 5-7 for generating the innovations e_t, we use the generalised lambda distribution suggested by Ramberg and Schmeiser (1974), which is an extension of Tukey's lambda distribution.

The inverse distribution function formula is

$$F^{-1}(u) = \lambda_1 + \left[u^{\lambda_3} - (1-u)^{\lambda_4} \right] / \lambda_2 \qquad (2.21)$$

where λ_1 and λ_2 represent the location and the scale parameters and λ_3 and λ_4 jointly determine the shape of the distribution. Under this formulation we are allowed to study the *skewness* test under different shapes. Table 1 summarizes the models with different values of the parameters considered.

The number of replications per model used is 5,000. 300 initial observations are discarded to remove initialization effects.The calculations were performed using GAUSS 8.

Table 1: Different models used for the performance of the *Skewness* test

	δ	β	Innovations e_t	λ_1	λ_2	λ_3	λ_4
Model 1	0,04	0,94	i.i.d $N(0,1)$				
Model 2	0,25	0,70	i.i.d $N(0,1)$				
Model 3	0,475	0,475	i.i.d $N(0,1)$				
Model 4	0,60	0,30	i.i.d $N(0,1)$				
Model 5	0,45	0,45	t-distr. 14 df	0.0000	0.05122	0.05122	0.078945
Model 6	0,04 0,475	0,94 0,475	skew=-0.30 kurt=3.0	0.3618	0.18590	0.09255	0.19910
Model 7	0,475	0,475	skew=-0.20 kurt=3.3	0.1687	0.10490	0.07651	0.14160

The steps of the simulation procedure are as follows:

1. Generate the GARCH(1,1) models as described before.
2. Use the *Simple* NoVaS to Normalize the generating series. A loop with lags from 1 to 35 are used, that is, 35 different *Simple* NoVaS transformations are used. Each time we calculate the kurtosis and the statistic for the kurtosis:

$$Kurt = T\left(kurtosis - 3\right)^2 / 24 \qquad (2.22)$$

Calculate the p-value of the statistic and the lag that gives the maximum p-value, plus 3, is the lag k that we use for the final (selected) *Simple* NoVaS transformation.

3. Calculate the skewness, its usual statistic and the p-value (using χ_1^2) for the original series:

$$Skew = T \cdot skewness^2 / 6 \tag{2.23}$$

while for the NoVaS Normalized series we use the k adjusted statistic:

$$SKEW_z = \frac{k}{k-2}\left(Skew_\varepsilon\right) \tag{2.24}$$

In this section we use simple graphical methods (Davidson and MacKinnon 1998), like the P-*value plot* and the *Size-Power curve* to study the size and the power of the proposed test. These graphs are based on the empirical distribution function (EDF) of the P-values which is denoted by $\hat{F}\ \ x_j$. For the P-value plots, if the distribution used to compute the p_s terms is correct, each of the p_s terms should be distributed uniformly on (0,1) and the resulting graph should be close to the 45^o line. For judging how reasonable the results are, we require that the estimated size should be bounded within the 95% confidence interval of the nominal size (two dot lines). For example, if one considers a nominal size of 5%and 5000 Monte Carlo replication, a result is defined as reasonable if the estimated size lies between 4.38% and 5.61%. All *Size Figures* show on the right hand side the truncated (up to 10% nominal level) P-value plots for the actual size of the *Skewness* tests.

3.1 The size and the power of the test

Figure 1 shows the results for Models 1 and 2 and n=500. The Dash line is the p-value for the skewness of the original series and Model 1 while the Dot line refers to Model 2. In both cases the test over-rejects the null hypothesis. Notice that the higher value of the β the smaller the over-rejection rate. On the other hand the proposed skewness test behaves well; the estimated size of the test is inside the 95% confidence interval (two dot lines); see the the right hand truncated figure.

Figure 1: Size of the test; Models 1 and 2; n=500 Observations

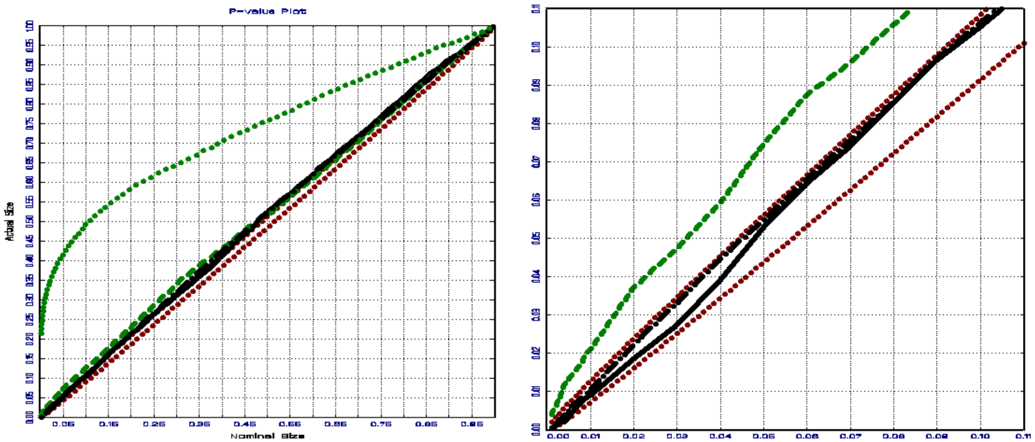


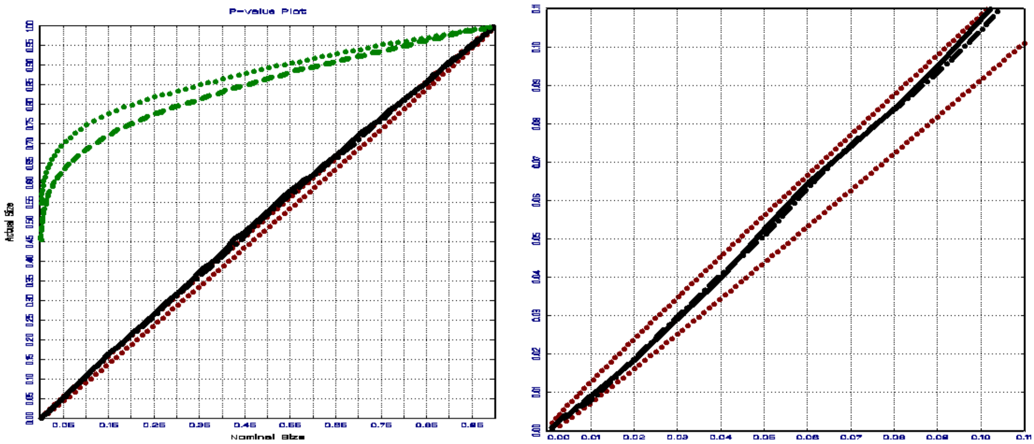
Figure 2 shows the results for Models 3 & 4 with n=500. The skewness test for the original series has estimated size that is far way from the nominal size (approx. 65% & 70%). Notice also that the higher the value of β the smaller the degree of over-rejection.

On the other hand, our skewness is within the 95% confidence bounds and rejects the null hypothesis about the right proportion of time.

For n=1000, 2000 & 3000 we notice that the GARCH dynamic increasing the over-rejection for the skewness test in the original series. In fact, the higher the number of the observations the higher the over-rejection.

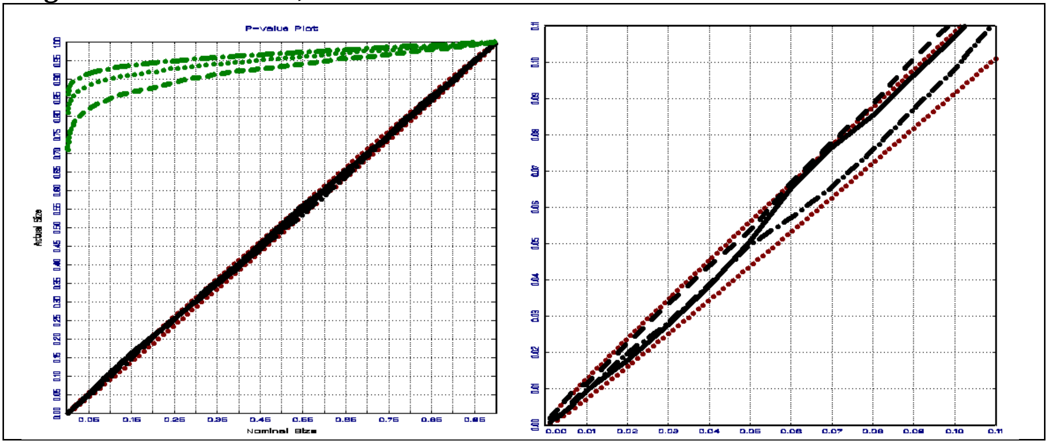
Again our proposed skewness test is very robust for the sample size.

Figure 2: Size of the test; Models 3 and 4; n=500 Observations



Finally, Figure 3 shows the results for Model 5 when the innovations are t-distributed with 14 df. Dash line represents the skewness test for 1000, while the Dot for 2000 and Dot-Dash for 3000 observations. The over-rejection is more than 80% larger than the estimated size, which is even higher than the one for the normal innovations. Once more the new test is robust and rejects about the right proportion of the time. .

Figure 3 Model 5 1000,2000 and 3000 Observations



Let us consider now the power of the test. In Figures 4 and 5 the *solid* curves represent the estimated power of the new test for n=500 (left part) and n=2000 (right part). The Dot-Dash curves are for n=1000 (left part) and n=3000 (right part). For the original series tests we have respectively the Dash and Dot curves.

Observe that the sample effect is significant. Indeed, the higher the sample size, the higher the power. The new test is so good that has higher power that the highest power at the 20% nominal level with n=3000 observations (Figure 10, right hand side).

For model 6 with balanced GARCH parameters, recall that tests based on the original series have size higher than 70%. Here, it turns out that the power is almost less than the size!

Figure 4 Power of the Test; Model 6

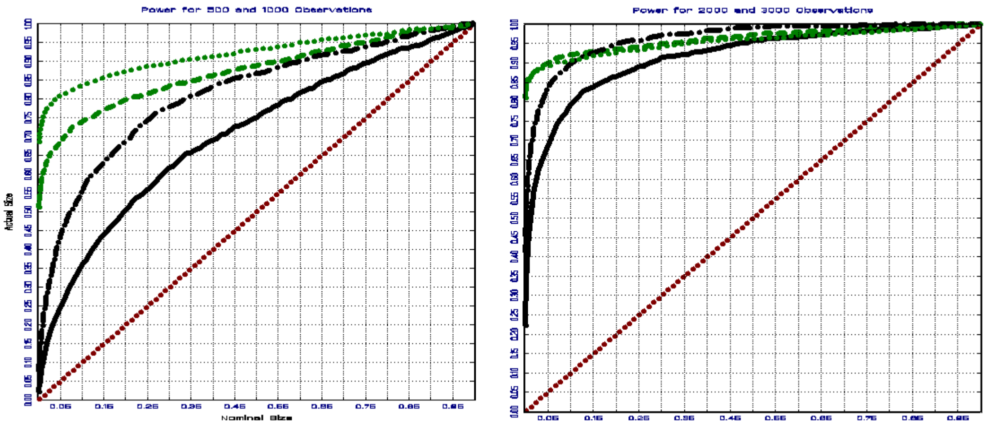
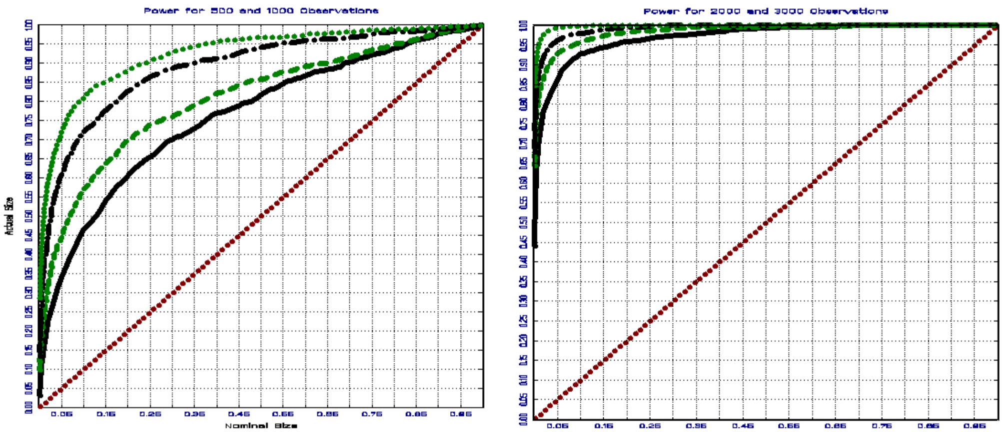


Figure 5: Power of the test; Model 7



In summary the size results clearly show that the Skewness test of the original series over-rejects the null hypothesis and is affected by both the sample size and the parameters of the GARCH process. On the other hand the proposed test is very robust in all cases examined with estimated size very near to the nominal one. The power results of the tests are also very impressive indicating the appropriateness of the proposed test.

4. APPLICATION TO FINANCIAL DATA

The Historical Prices in Fig. 6 & 7 are for the Swedish OMX and the German DAX index (source: http://finance.yahoo.com). In Fig. 8 & 9, we analyze the closing prices for BP and DELL. In each figure the 1st graph represents the original series and the 2nd the NoVaS transformed series. The solid curve is a standard normal variable that helps or comparative purposes.

Figure 6a: OMX Index Stockholm (6 Jul 2004-29 Nov-2010) (1633 observations)

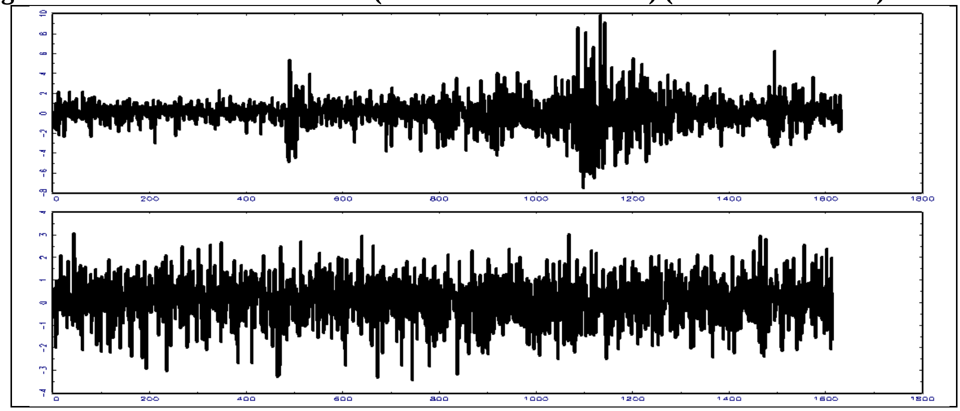


Figure 6b The cdf and pdf curve of the OMX series

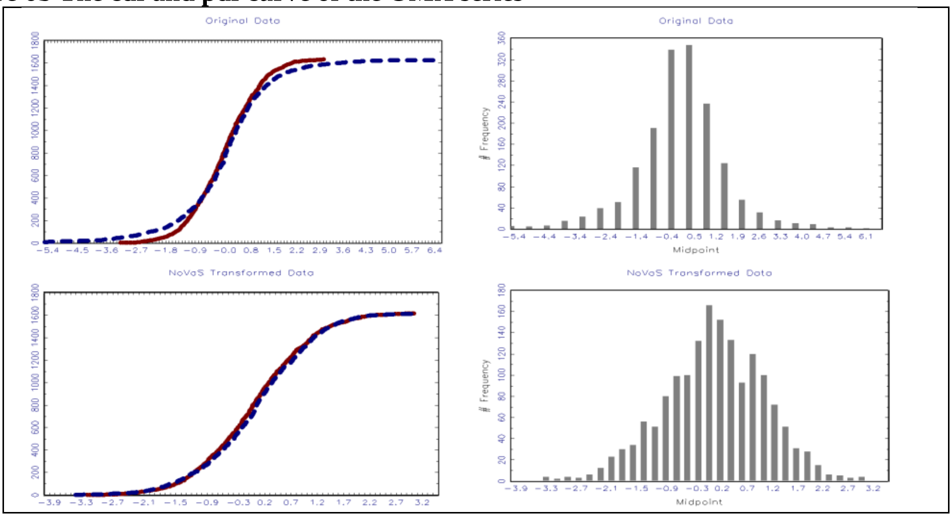


Figure 6b shows that the original series has positive skewness (0.16363007, p-value=0.00694) but after the NoVaS transformation (min=-3.4131, max=3.0741) the skewness is negative (-0.20848564, p-value=0.00118).

Figure 7a DAX (GDAXI) German Index (1 Feb 1995-20 Dec 2006)

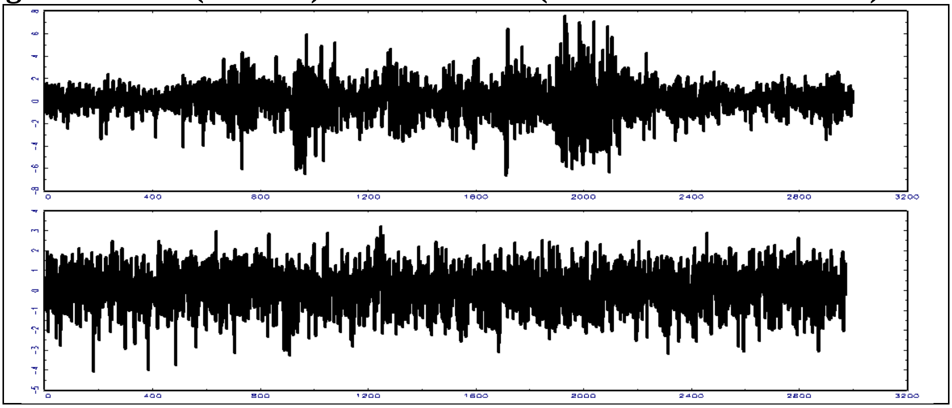


Figure 7b The cdf and pdf curve of the DAX series

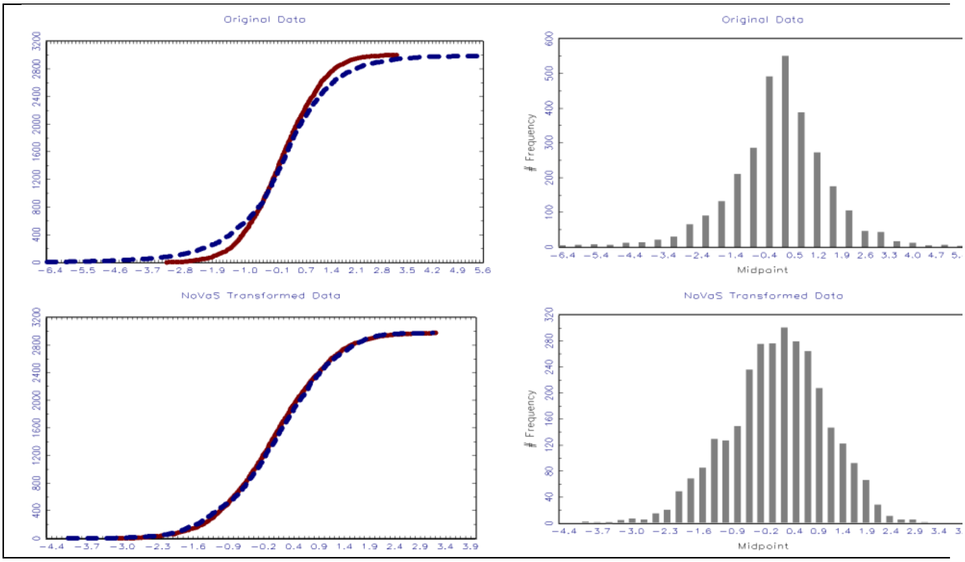


Figure 7b for the DAX index shows that both the original and the NoVaS transformed series (min=-4.071, max=3.2305) have negative skewness (before=-0.16284, p-value =0.00027, after=-0.261, p-value=1.95547e-008).

Figure 8a DELL (1 Jan 1999-29 Nov 2010)

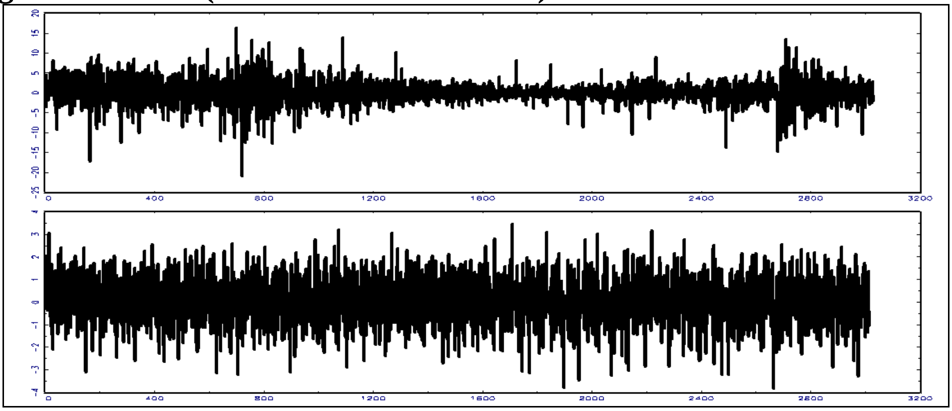


Figure 8b The cdf and pdf curve of the DELL series

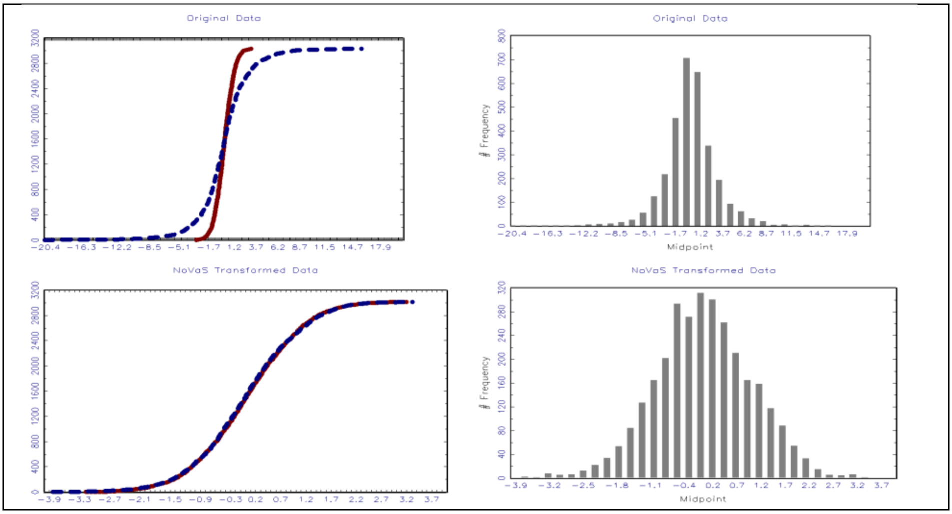


Figure 8b for the closing prices for DELL is a representative example of the NoVaS transformation. It shows that the negative skewness (-0.16466246, p-value =0.000216) of the original series after the NoVaS transformation (min=-3.8369, max=3.467) becomes significant equal to zero (-0.066567912, p-value 0.53).

Figure 9a BP (1 Jan 2003-29 Nov 2010)

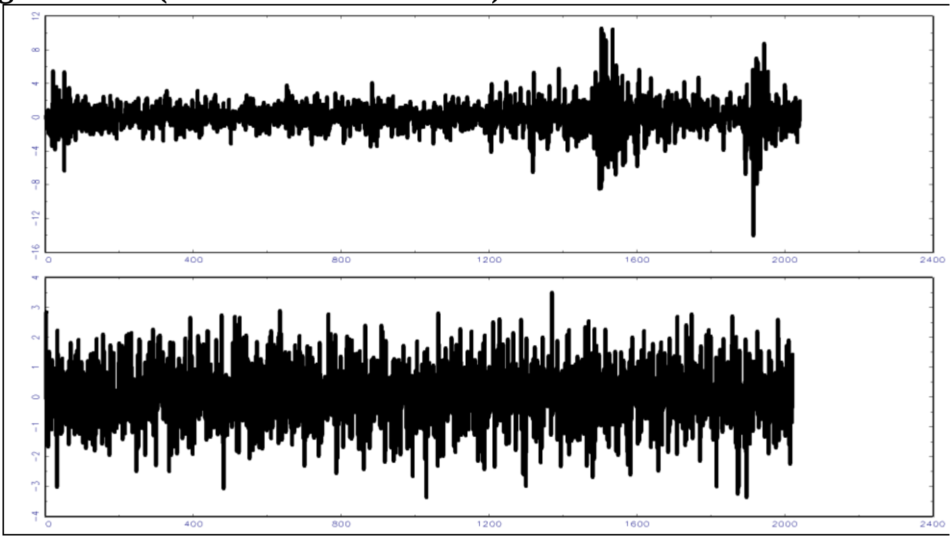


Figure 9b The cdf and pdf curve of the BP series

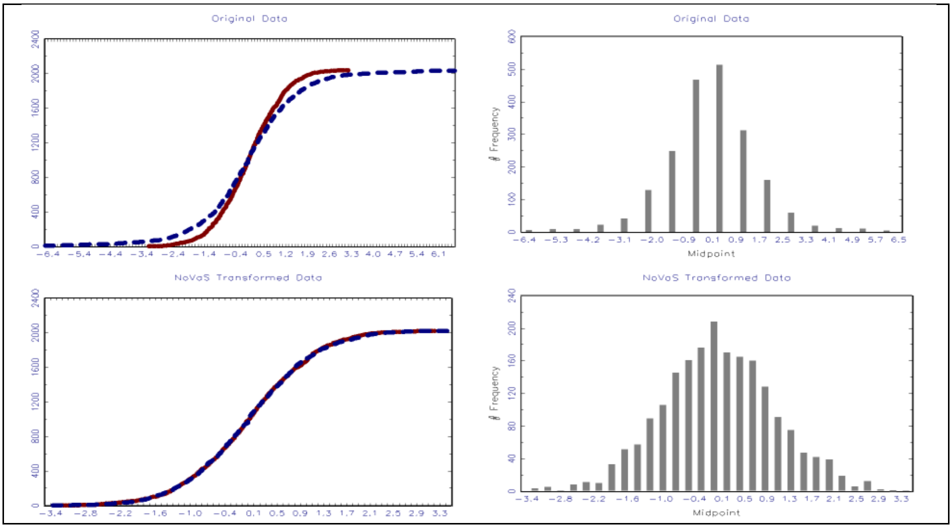


Figure 9b for the closing prices of BP shows that neither the original series nor the transformed NoVaS series (min=-3.37, max=3.5) shows any significant skewness (original=-0).

5. CONCLUSIONS

In this work we derived a skewness test for series that have ARCH/GARCH effect. A test, in a easy and comprehensive way, can distinguish between the causes for skewness, namely whether there is an ARCH/GARCH effect or the underline distribution is non-normal. The proposed test shows extremely good results in an extensive simulation study not only in terms of the size but also in terms of the power. The test has also been applied to a number of financial data with satisfactory results. This work provides the ground for investigating the complementary problem of the kurtosis which also can been used for similar purposes in financial return series.

REFERENCES

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity, J. of Econometics. 31, 307–328.

Engle, Robert F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica*, 50, 987–1008.

Harvey, C R. & Siddique, A, 1999. Autoregressive Conditional Skewness, Journal of Financial and Quantitative Analysis, Cambridge University Press, 34(04), 465-487.

Jondeau, E. & Rockinger, M., 2000. Conditional Volatility, Skewness, and Kurtosis: Existence and Persistence, Working papers 77, Banque de France.

Leon, A., G. Rubio and G. Serna (2004), Autoregressive Conditional Volatility,Skewness and Kurtosis, WP-AD 2004-13, Instituto Valenciano de Investigaciones Economicas.

Politis, D. N. (2003). A normalizing and variance-stabilizing transformation for financial time series, In: *Recent Advances and Trends in Nonparametric Statistics*, M. G. Akritas and D. N. Politis, (eds.), Elsevier (North Holland), 335–347.

Politis, D. N. (2007). Model-free versus Model-based Volatility Predictions, *Journal of Financial Econometrics*, 5, 358–389.

Premaratne, G. and Bera, A. (2000). Modeling Asymmetry and Excess Kurtosis in Stock Return Data, Working Paper 00-0123, University of Illinois.

Ramberg, J. and Schmeiser, B. (1974). An Approximate Method for Generating Asymmetric Random Variables, Communications of the Association for Computing Machinery, 17, 78-82.

APPENDIX

we derive below parts of the proof of Lemma 1:

I) The first moment of $W_{t,a}$

First recall that for the Simple NoVaS $\gamma = 0$ and consider the Simple NoVaS with one lag, namely,

$$W_{t,a} = a\varepsilon_t^2 + a\varepsilon_{t-1}^2 \quad . \tag{A.0}$$

Recall that ε_t is variance stationary and $m_2 = E\left(\varepsilon_t^2\right)$. Then the mean is:

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

For two lags the mean is $E\left(W_{t,a}\right) = 3am_2$, while for three lags is $E\left(W_{t,a}\right) = 4am_2$ and finally the general formula of the mean for (k) lags is:

$$E\left(W_{t,a}\right) = (k+1)am_2 \tag{A.1}$$

II) The second moment of $W_{t,a}$

For the Simple NoVaS model with one lag, by squaring both sides of (A.0) and then taking expectations we have:

$$E\left(W_{t,a}^2\right) = a^2 E\left(\varepsilon_t^4\right) + a^2 E\left(\varepsilon_{t-1}^4\right) + 2a^2 E\left(\varepsilon_t^2\right) E\left(\varepsilon_{t-1}^2\right).$$

If ε_t is 4-order stationary with $m_4 = E\left(\varepsilon_t^4\right)$ and $m_2 = E\left(\varepsilon_t^2\right)$ then we have:

$E\left(W_{t,a}^{2}\right)=\left(2\right)a^{2}m_{4}+2a^{2}\left(m_{2}\right)^{2}$. In a similar fashion we can show that for the Simple NoVaS model with two lags we have $E\left(W_{t,a}^{2}\right)=\left(3\right)a^{2}m_{4}+\left(3\right)2a^{2}\left(m_{2}\right)^{2}$. Also for three lags we have $E\left(W_{t,a}^{2}\right)=\left(4\right)a^{2}m_{4}+\left(6\right)2a^{2}\left(m_{2}\right)^{2}$ and finally for the $\left(k\right)$ lags:

$$E\left(W_{t,a}^{2}\right)=\left(k+1\right)a^{2}m_{4}+k(k+1)a^{2}\left(m_{2}\right)^{2} \ . \qquad \text{(A.2)}$$

III) *The third moment of* $W_{t,a}$

Following the same principle as in the case of the 2nd moment, it is easy to show that, the general formula for $\left(k\right)$ lags for the third moment is:

$$E(W_{t,a}^{3})=\left(a^{2}\left(k+1\right)E\left(\varepsilon_{t}^{4}\right)+a^{2}k\left(k+1\right)E\left(\varepsilon_{t}^{2}\right)E\left(\varepsilon_{t}^{2}\right)\right)\left(a\left(k+1\right)E\left(\varepsilon_{t}^{2}\right)\right). \qquad \text{(A.3)}$$

Equation (A.3) for $a=\dfrac{1}{k+1}$ and with $m_{4}=E\left(\varepsilon_{t}^{4}\right)$ and $m_{2}=E\left(\varepsilon_{t}^{2}\right)$ becomes:

$$E(W_{t,a}^{3})=\left(am_{4}+akm_{2}^{2}\right)m_{2} \ . \qquad \text{(A.4)}$$