Falsifying ARCH/GARCH Models using Bispectral Based Tests

Melvin Hinich

Applied Research Laboratories
The University of Texas at Austin
hinich@mail.la.utexas.edu *www.gov.utexas.edu/hinich*

Abstract

This paper shows that the Hinich (1982) bispectrum test for gaussianity and the Hinich and Rothman (1998) test for time reversibility can be used to falsify the null hypothesis that an autoregressive conditionally heteroskedastic model (ARCH) of its generalization (GARCH) generates nonlinear behavior in the variance of an observed time series. The term "falsify" means that the null hypothesis can be rejected with a given size using a nonparametric test based on the bispectrum where the data is trimmed to control the sizes. Rejecting the null hypothesis implies that the ARCH or GARCH model that is estimated from the data is not a complete statistical description of the dependence structure in the variance of the process.

Falsifying ARCH/GARCH mModels using Bispectral Based Tests

Melvin Hinich

Applied Research Laboratories
The University of Texas at Austin
hinich@mail.la.utexas.edu *www.gov.utexas.edu/hinich*

1 INTRODUCTION

This paper shows that the Hinich (1982) bispectrum test for gaussianity and the Hinich and Rothman (1998) test for time reversibility can be used to falsify the null hypothesis that an autoregressive conditionally heteroskedastic model (ARCH) of its generalization (GARCH) generates nonlinear behavior in the variance of an observed time series. The term. By “falsify” I means that the null hypothesis can be rejected with a given size using a nonparametric test based on the bispectrum where the data is trimmed to control the sizes. Rejecting the null hypothesis implies that the ARCH or GARCH model that is estimated from the data is not a complete statistical description of the dependence structure in the variance of the process.

Linear and nonlinear tTime series models have been widely employed in the literature to explain the dynamics of financial time series. Since its introduction 24 years ago, the applications of Autoregressive Conditional Heteroskedasticity (ARCH) model introduced by Engle (1982) or its extension Generalized Autoregressive Conditional Heteroskedasticity (GARCH) by Bollerslev (1986) in finance have become commonplace (for a survey see Bollerslev et al., 1992. This class of models relaxes the assumption of the classical linear regression model that the variance of the disturbance term is conditionally as well as unconditionally constant.

Since the GARCH generalization the number of empirical and theoretical developments in the field has exploded, with rapid development of applications and variants. This popularity is evidenced by the incorporation of GARCH estimation into major software packages (for a reviews of GARCH software see Brooks, 1997; McCullough and Renfro, 1999; and Brooks et al., 2001).

Let $\{x(t_n)\}$ denote an equally spaced sampled time series from a stationary random process $\{x(t)\}$ where $t_n = n\tau$. A zero mean ARCH/GARCH model for this time series is of the form $x(t_n) = \sqrt{h(t_n)}\ e(t_n)$ where $\{e(t_n)\}$ is a zero mean pure white noise process (i.i.d.) and $\{h(t_n)\}$ is a positive valued autoregressive moving average process whose inputs are lagged $e^2(t_n)$. For example an ARCH(q) model is of the form $h(t_n) = \alpha_0 + \sum_{k=1}^{q} \alpha_k e^2(t_{n-k})$ and a GARCH(p,q) model is $h(t_n) = \alpha_0 + \sum_{k=1}^{q} \alpha_k e^2(t_{n-k}) + \sum_{j=1}^{p} \beta_j h(t_{n-j})$. In most cases it is assumed that the $e(t_n)$ have a normal (gaussian) distribution but sometimes the assumed distribution is a Student's t.

The ARCH/GARCH type of non-linear time series model is claimed to be *the* model of a special type of non-linearity in the data generating process known as multiplicative non-linearity, or non-linear-in-variance, in which non-linearity affects the process through its variance (Hsieh, 1989). Although these models have been heralded as an accurate description of a number of important characteristics of financial data , as Hall, Miles and Taylor (1989) note, the ARCH parameterization of the conditional variance does not have any solid grounding in economic theory, but represents "a convenient and parsimonious representation of the data." Given the importance of these models in econometric time series it is important to be able to use a nonparametric statistical tool to falsify them. If it turns out that the ARCH/GARCH models lack a certain statistical property that has not been exploited then the time series community may create new nonlinear models that more accurately captures the complexity of the nonlinearity inherent in high frequency market data.

The bispectrum is defined in the next section. Since the $e^2(t_n)$ are independently distributed it will now be shown that all the bispectral values of $\{x(t_n)\}$ are zero as long as the distribution of the $e(t_n)$ is *symmetric*. The Hinich test for Ggaussianity is really a test of the null hypothesis that the bispectrum is zero for all bifrequencies and thus if the Hinich test rejects the null hypothesis then the ARCH/GARCH specification is falsified for any set of model parameters. This point was first advanced by Brock (1987) in an unpublished paper.

It will also be shown that the bispectrum of any ARCH or GARCH process is a real constant (its imaginary part is zero) for any distribution of $e(t_n)$ with finite moments. The Hinich-Rothman test of time reversibility is really a test for the null hypothesis that the imaginary part of the bispectrum is zero for all bifrequencies. Thus if the Hinich-Rothman test rejects this null hypothesis the ARCH/GARCH specification is falsified in a nonparametric manner.

The asymptotic properties of these two tests are valid for ARCH/GARCH models that have finite moments but the fat tails of especially the GARCH processes produce false rejections for moderate and even large sample sizes. The definition of the bispectrum is given in Section 2. Section 3 covers the estimation of the bispectrum and the large sample properties of the test statistics. Section 4 presents a discussion of the use of data trimming to control the sizes of the tests. Simulations are presented to support validity of trimming to obtain proper test sizes for sample sizes common for high frequency financial data.

2 THE BISPECTRUM

The bispectrum of a bandlimited random process $\{x(t_n)\}$ is

$$B_x(f_1,f_2)=\sum_{r=-\infty}^{\infty}\sum_{s=-\infty}^{\infty}b_x(r,s)\exp\left[-i2\pi\left(f_1t_r+f_2t_s\right)\right] \tag{2.1}$$

where $b_x(r,s)=Ex(t_n)x(t_{n+r})x(t_{n+s})$ is called the bicorrelation for lags r and s. The set of positive support of the bispectrum is the triangle $\Omega=\{0<f_1<f_o/2,\ f_2<f_1,\ f_1+f_2<f_o/2\}$ where f_o is the highest frequency component of the process (Hinich and Messer, 1995), which is called the band limit in the signal processing literature.

The conditional product $E\left[x(t_n)x(t_{n+r})x(t_{n+s})\,|\,e(t_m),m<n+s\right]$ is zero for all $0<r\le s$ and any ARCH or GARCH model since the $e(t_n)'s$ are independently distributed, implying that $b_x(r,s)=0$ for all $r\neq 0$ and $s\neq 0$. Thus $B_x(f_1,f_2)=E\left[h^{3/2}(t_n)e^3(t_n)\right]$. The bispectra of any ARCH or GARCH process is a *real constant* for all bifrequencies and thus if the Hinich-Rothman test rejects the null hypothesis that the imaginary part of the bispectrum is zero for all bifrequencies then the process can not be ARCH/GARCH.

In addition if $Ee^3(t_n)=0$ for a given ARCH/GARCH model then its bispectrum is zero for all bifrequencies. With the zero skewness assumption for the $e(t_n)'s$ then if the Hinich zero bispectrum test rejects the null hypothesis then it rejects the ARCH/GARCH model.

The bispectrum estimation method and the test statistics are presented in the next section.

3 BISPECTRAL ESTIMATION

The spectrum and bispectrum can be estimated using conventional nonparametric methods (Hinich and Clay, 1968). I prefer to use the frame averaging spectrum estimation method to illuminate the statistical issues of bispectrum normalization but the results will hold for any method that yields estimates that have similar asymptotic properties to the frame averaging method. Details of estimating the bispectrum from a sample of discrete-time observations of the process and the sampling properties of the estimate are presented in Brillinger (1965), Hinich (1982), Brockett et. al. (1988), Hinich and Patterson (1989) and (1992), and Hinich and Wolinsky (1988) and (2005).

Consider a sample $\{x(t_1),...,x(t_N)\}$ where $t_k = k\delta$. This sample is partitioned into $P = [N/L]$ non-overlapping frames of length $L\delta$ where the last frame is deleted if it has less than L observations. To simplify notation normalize the time unit by setting $\delta = 1$. The resolution bandwidth is then $f_1 = \frac{1}{L}$. If $P = N/L$ then the last undersized frame is not used to estimate the bispectrum.

The pth frame is $\{x_p(1),...,x_p(L)\} = \{x((p-1)L+1),...,x(pL)\}$ The discrete Fourier transform of the pth frame is $X_p(k) = \sum_{t=1}^{L} x_p(t)\exp\left(-i2\pi\frac{kt}{L}\right)$ and the periodogram of the pth frame is $\frac{1}{L}\left|X_p(k)\right|^2 = \frac{1}{L}X_p(k)X_p(-k)$. Since $N \approx LP$ the frame-averaged estimate of the spectrum at frequency $f_k = \frac{k}{L}$ is

$$\hat{S}(f_k) = \frac{1}{N}\sum_{p=1}^{P}\left|X_p(k)\right|^2 \tag{3.1}$$

Then $E\left[\hat{S}(f_k)\right] = S(f_k) + O\left(\frac{1}{L}\right)$ where the error term of order 1/L is due to the frame windowing of the spectrum. The variance of the estimate for large values of L and P is $\frac{1}{P}S^2(f_k)$.

The frame-averaged estimate of the bispectrum at the bifrequencies $\left(f_{k_1}, f_{k_2}\right)$ is

$$\hat{B}\left(f_{k_1}, f_{k_2}\right) = \frac{1}{N}\sum_{p=1}^{P} X_p\left(k_1\right) X_p\left(k_2\right) X_p\left(-k_1 - k_2\right) \quad (3.2)$$

Then $E\left[\hat{B}\left(f_{k_1}, f_{k_2}\right)\right] = B\left(f_{k_1}, f_{k_2}\right) + O\left(\frac{1}{L}\right)$ and the variance for large L and P is $\frac{L}{P} S\left(f_{k_1}\right) S\left(f_{k_2}\right) S\left(f_{k_1} + f_{k_2}\right)$.

The normalization of the estimated bispectrum is

$$\hat{\Gamma}\left(f_{k_1}, f_{k_2}\right) = \frac{\hat{B}\left(f_{k_1}, f_{k_2}\right)}{\sqrt{\hat{S}\left(f_{k_1}\right)\hat{S}\left(f_{k_2}\right)\hat{S}\left(f_{k_1} + f_{k_2}\right)}} \quad (3.3)$$

This normalization standardizes the variance of the bispectrum estimate using the estimated variance in place of the true variance. Let $\Gamma\left(f_{k_1}, f_{k_2}\right) = \left[S\left(f_{k_1}\right) S\left(f_{k_2}\right) S\left(f_{k_1} + f_{k_2}\right)\right]^{-\frac{1}{2}} B\left(f_{k_1}, f_{k_2}\right)$ .Let $L = N^{\varepsilon}$ where the bandwidth parameter e is in the interval $0 < \varepsilon < 0.5$. Then the real and imaginary parts of each $\sqrt{2N^{-1+2\varepsilon}}\left[\hat{\Gamma}\left(f_{k_1}, f_{k_2}\right) - \Gamma\left(f_{k_1}, f_{k_2}\right)\right]$ are asymptotically independent gaussianNormal variates with zero means and unit variances as $N \to \infty$ (Hinich, 1982). Moreover the $\hat{\Gamma}\left(f_{k_1}, f_{k_2}\right)$ are asymptotically independently distributed across the principal domain of the bifrequencies. The smaller the value of $ec\varepsilon$, the fewer the number of bifrequencies and thus the smaller the power the tests but the larger the number of frames and thus the faster the convergence to the asymptotic sampling properties.

Suppose that the null hypothesis is that the imaginary part of the bispectrum is zero and thus $\operatorname{Im}\Gamma\left(f_{k_1}, f_{k_2}\right) = 0$ for all bifrequencies, which is true for ARCH/GARCH models. The Hinich-Rothman *TR* test statistic is the sum S_{TR} of $2N^{-1+2\varepsilon}\left[\operatorname{Im}\hat{\Gamma}\left(f_{k_1}, f_{k_2}\right)\right]^2$ over the $L^2/16$ bifrequencies in the support set. This distribution of this sum is approximately central chi-squared with $M = L^2/16$ degrees of freedom for large *N*.

Now suppose that the null hypothesis is that the bispectrum is zero and thus $\Gamma\left(f_{k_1}, f_{k_2}\right) = 0$ for all bifrequencies, which is true for ARCH/GARCH models with $Ee^3\left(t_n\right) = 0$ as it is true for a gaussian process. The Hinich (1982) test statistic is the sum S_0 of $2N^{-1+2\varepsilon}\left|\hat{\Gamma}\left(f_{k_1}, f_{k_2}\right)\right|^2$ over the $L^2/16$ bifrequencies and its large sample distribution is central chi-squared with $2M$ degrees of freedom. For both tests if the null hypothesis is false then the statistics have noncentral chi-square distributions and the tests are one sided.

The program BISPEC is available on my website (www.gov.utexas.edu/hinich). This program transforms the tests statistic using the cumulative distribution function of the test statistics under the appropriate null hypothesis. For example the Hinich zero bispectrum test statistic is $Y_0 = F_{2M}\left(S_0\right)$ where $F_{2M}\left(s\right)$ is the cdf of a central chi–square density with $2M$ degree of freedom. Thus Y_0 has a uniform $(0,1)$ distribution under the null hypothesis. The null hypothesis is rejected if the *p*-value $p = 1 - F_{unif}\left(Y_0\right)$, where F_{unif} is the cdf of the uniform $(0,1)$ distribution, is deemed to small by the analyst.

The Hinich-Rothman TR test statistic is $Y_{TR} = F_M\left(S_{TR}\right)$. The null hypothesis is rejected if $p = 1 - F_{unif}\left(Y_{TR}\right)$ is deemed too small.

The sampling properties of the Hinich bispectral based tests of normality and linearity as well as the Hinich and Rothman TR test are large sample results based on the asymptotic normal distribution's mean and variance. The validity of any asymptotic result for a finite sample is always an issue in statistics. The rate of convergence to normality depends on the size of the cumulants of the observed process.

All data is finite since all measurements have an upper bound to their magnitudes. If the data is leptokurtic as is the case for stock returns and exchange rates then the cumulants are large. Trimming the tails of the empirical distribution of the data is an effective statistical method approach to limit the size of the cumulants in order to get a more rapid convergence to the asymptotic distribution. Trimming of time series in order to improve the validity of the use of the asymptotic properties of the test is discussed next.

4 TRIMMING TIME SERIES DATA

Trimming data to make sample means less sensitive to outliers has been used in applied statistics for years. Trimming is a simple data transformation that make statistics based on the trimmed sample more normally distributed. Transforming data is a technique with a long pedigree, it can be dated back at least to Galton (1879) and McAliser (1879) at the dawn of modern statistics. Subsequently, Edgeworth (1898), and Johnson (1949), among others, have contributed to our understanding of this technique for examining data.

Suppose one wants to trim the upper and lower $\kappa/2$ % values of the sample $\{x(t_1),...,x(t_N)\}$. Order the data and find the $\kappa/200$ quantile $x_{\kappa/200}$ and the $1-\kappa/200$ quantile $x_{1-\kappa/200}$ of the order statistics. Then set all sample values less than the $\kappa/200$ quantile to $x_{\kappa/200}$ and set all sample values greater than $1-\kappa/200$ quantile to $1-\kappa/200$ quantile. The remaining $(100-\kappa)\%$ data values are not transformed.

Assume that the $e(t_n)'s$ have a symmetric density. Then tThe bicorrelations of the trimmed sample are *zero* for all lags using the same conditioning argument as was used to show that the bicorrelations of a ARCH/GARCH are zero. Thus the imaginary part of the bispectrum of the trimmed sample is zero for all bifrequencies. Also the real part is zero for all bifrequencies since if $Ee^3(t_n)=0$. Simulations will be presented to determining the trimming level for threewo GARCH(1,1) models and two ARCH(1) models previously used in from the econometric literature.

5 SIMULATIONS USING GARCH(1,1) AND ARCH(1) MODELS

Two GARCH(1,1) models and two ARCH(1) models were used in the simulations. Brooks et.al. (2001) use the following GARCH(1,1) model parameters as a benchmark for their comparison of econometric software packages: $\alpha_0=0.0107613$, $\alpha_1=0.153134$, $\beta_1=0.805974$. This model is designated GARCH 1 in the tables below. The simulations are designed to learn how much trimming is needed to get sizes that match the size derived from the asymptotic sampling theory.

The second first is the GARCH(1,1) model in Example 2 of Horowitz, et. al. (2006) where $\alpha_0=0.001$, $\alpha_1=0.05$, $\beta_1=0.9$ using the notation above. This model is called GARCH 12.

The program BISPEC is available on my website (www.gov.utexas.edu/hinich). This program transforms the tests statistic using the cumulative distribution function of the test statistics under the appropriate null hypothesis. For example the Hinich zero bispectrum test statistic is $Y_0=F_{2M}(S_0)$ where $F_{2M}(s)$ is the cdf of a central chi–square density with $2M$ degree of freedom. Thus Y_0 has a uniform $(0,1)$ distribution under the null hypothesis. The Hinich-Rothman TR test statistic is $Y_{TR}=F_M(S_{TR})$. For both tests if the null hypothesis is false then the statistics have a noncentral chi-square distribution and the tests are one sided.

Using a number of different trimming levels simulations were run using these is models with 50,000 replications and a sample size of N = 50,000. The distribution of the pseudorandom pure white $\left\{e\left(t_n\right)\right\}$ used to generating the simulations are: normal, double tailed exponential and Student t with 6 degrees of freedom. Twelve quantiles (0.50, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 0.99, 0.995, and 0.999) were estimated from the 50,000 replications to determine how close they were to the large sample quantiles of the two statistics computed from trimmed series using different trimming levels. The top three quantiles are the important ones since each test rejects the null hypothesis if the test statistic is close to one.The 1% and 5% sizes of the two test statistics were estimated from the 50,000 replications computed from trimmed series using different trimming levels in order to determine how close they were to the large sample sizes for the two test statistics.

The Ttwo values of the bandwidth parameter were used to estimate the bispectra are;, ε ec = 0.33 and $ec\varepsilon$ = 0.4. The differences between the quantiles computed for the two bandwidth parameters were small and so only the ε ec = 0.4 will be shown.

Table 1 presents Tthe results for the two zero bispectrum tests using the five trimming levels $\kappa = 2\%,\ 5\%,\ 10\%,\ 15\%,\ \&\ 20\%$ for the GARCH 1 and GARCH 2 models with normal pure white $\left\{e\left(t_n\right)\right\}$ forusing the a trimming levels of $\kappa = 2\%,\ 5\%,\ 10\%,\ 15\%,\ \&\ 20\%$ for the two tests using the GARCH 1 model are presentedshown in Table Figures 1 and 2. Table 2 presents the results for the TR test using the same setup. The sizes are controlled for a 10% data trimming level and the tests become conservative for a 15% trimmingtop three sample quantiles are very close to their true values for this trimming level.

Table 3 and 4 presents the results for the two tests using the trimming levels $\kappa = 10\%,\ 15\%,\ \&\ 20\%$ for the GARCH 1 and GARCH 2 models with double tailed exponential pure white $\{e(t_n)\}$. The kurtosis of the exponential is 6 and thus a double tailed exponential density had fatter tails than is generally believed. For these time series the sizes are also conservative for a 15% data trimming level.

Table 5 presents the results for the two tests using the trimming levels $\kappa = 10\%\ \&\ 15\%$ for the GARCH 1 model where the $\{e(t_n)\}$ have a Student t with 6 degrees of freedom. Once again the 15% trimming level yields conservative sizes.

The sample quantiles for the two tests at the same trimming level are nearly the same for all the simulations that I ran. Thus only the TR test results will be shown for the following GARCH and ARCH models.

Brooks et.al. (2001) use the following GARCH(1,1) model parameters as a benchmark for their comparison of econometric software packages: $\alpha_0 = 0.0107613,\ \alpha_1 = 0.153134,\ \beta_1 = 0.805974$. The result for a trimming level of $\kappa = 15\%$ of the simulations of this model called GARCH 2 for the *TR* is shown in Figure 3. The fat tails of this model are larger than for GARCH 1 and thus it takes more trimming to get sizes that match the size derived from the asymptotic sampling theory.

To test trimming levels for ARCH models two ARCH(1)'s used by Becker and Hurn (2004) were simulated. The ARCH 1 parameters are $\alpha_0 = 1$ and $\alpha_1 = 0.8$. The ARCH 2 parameters are $\alpha_0 = 1$ and $\alpha_1 = 0.5$. The results for the trimming levels $\kappa = 5\%\ \&\ 10\%$ and the normal and Student t (6) $\{e(t_n)\}$ a trimming levelare of $\kappa = 10\%$ for ARCH 1 is presented shown in Tables 6.Figure 4. For both the ARCH 2 models aa $\kappa = 5\%$ trimming is sufficient to yield proper sizes for the tests for the normal $\{e(t_n)\}$(Figure 5). For the Student t (6) case the 10% trimming yields conservative sizes but the 5% trimming is not sufficient to keep the sizes at or below their target values.

Trimming at the 15% level makes the two tests conservative for all the models used but will reduce their power of the test to falsify the ARCH/GARCH hypothesis. It is important to keep in mind that not rejecting the null hypothesis is *not* the same as accepting it. If the tests reject the null after trimming then one has confidence that the null is rejected. One approach to justifying this confidence is to fit an ARCH or GARCH model assuming a symmetric density for the $\{e(t_n)\}$ and then using my simulation program to determine the proper trimming level.

The power of the bispectrum test for trimmed output of a new type of nonlinear model is presented in the next section.

6 DETECTING A NONLINEAR AR(2) PROCESS

Consider the nonlinear AR(2) model

$$x(t_n) + a_1\left(x(t_{n-2})\right)x(t_{n-1}) + a_2\left(x(t_{n-2})\right)x(t_{n-2}) = \sigma u(t_n) \tag{4.1}$$

where

$$a_2\left(x(t_{n-2})\right) = \exp\left(-c\left(1 + \delta_c x^2(t_{n-2}) + \delta_c^2 x^2(t_{n-3}) + \delta_c^3 x^2(t_{n-4}) + \delta_c^4 x^2(t_{n-5})\right)\right), \tag{4.2}$$

$$a_1\left(x(t_{n-2})\right) = -2a_2\left(x(t_{n-2})\right)\cos 2\pi f\left(1 + \delta_f x^2(t_{n-2}) + \delta_f^2 x^2(t_{n-3}) + \delta_f^3 x^2(t_{n-4}) + \delta_f^4 x^2(t_{n-5})\right)$$

A simulation with 50,000 replications and N = 50,000 was run using this model for the parameter values c = 0.2, f = 0.4, δ_c = 0.01, and δ_f = 1.0. A section of the output from the model for these parameters is shown in Figure 6. Using a trimming level of $\kappa = 15\%$ the simulated power of the time reversibility bispectrum test is $\beta = 91.43$ for a significance level of $\alpha = 1\%$. The power for a significance level of $\alpha = 0.1\%$ is $\beta = 0.837$. This high power is somewhat surprising since the sample skewness of the trimmed output is $\gamma = 0.0028$. The sample quantiles are shown in Figure 7.

7 INTRADAY STOCK RETURNS EXAMPLE

The two bispectrum based tests were computed for a times series of intraday rates of return of Coca Cola (NYSE symbol KO) for the period from January 2, 1980 to August 30, 1985. This data file was one of thirty returns analyzed for nonlinear structure by Hinich and Patterson (1989). These rates of return were constructed from the actual trade prices by a method that produced unaliased ten minute averages for each trading day. There are 36 such aggregated rates for each trading day yielding N = 51,622 data51,585 points. The frame length used was L = 72 implying an exponent value *ec* = 0.39. Thus there the number of full frame is yielding 716 full frames and thus the actual sample size used was N = 51,585.

The p-values of the test were computed to six decimal places. For a trimming level of $\kappa = 15\%$ the p-value of the zero bispectrum test was $p = 0.000000$ and the p-value of the TR test was $p = 0.000003$. The tests are still statistically significant for a trimming level of $\kappa = 25\%$ since the p-values were 0.000001 for the zero bispectrum test and 0.001383 for the TR test..

8 CONCLUSION

Data trimming has been shown to control the distortion of fat tailed time series on the sizes of the zero bispectrum tests that can be used to falsify ARCH/GARCH models in a nonparametric fashion. The results presented in this paper show that the trimming level depends on the nature of the model used. How is a user to determine the proper trimming level to use the bispectrum based tests to see if a fitted ARCH or GARCH is falsified? One way is to use my simulation program

SIMARCH that I used for the simulations used for this paper. The program source code and its executable are in the folder CUMSPEC on my webpage. The folder also contains the bispectrum program.

REFERENCES

Becker, R. and A.S. Hurn (2004). "Testing for Nonlinearity in Mean in the Presence of Heteroskedasticity," School of Economics and Finance, Queensland University of Technology

Brockett, P.L , M.J. Hinich and D. M. Patterson (1988). "Bispectral Based Tests for Detection of Gaussianity and Linearity in Time Series," *Journal of the American Statistical Association*, **83** (403), "Applications and Case Studies", 657-664

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics* **31**, 307-327

Bollerslev, T., R.Y. Chou, and K.F. Kroner (1992). "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics* **52** (5), 5-59

Brooks, C. (1997). "GARCH Modeling in Finance: A Review of the Software Options," *Economic Journal* **107**, 1271-1276

D.R Brillinger. (1965). "An Introduction to Polyspectra," *Annals of. Mathematical Statistics.* **36**, 1351-1374

Brock, W.A. (1987). "Nonlinearity in Finance and Economics," Department of Economics, The University of Wisconsin.

Brooks, C., S.P. Burke, and G. Persand (2001). "Benchmarks and the Accuracy of GARCH Model Estimation," *International Journal of Forecasting* **17**, 45-56

Edgeworth, F.Y. (1898). "On the Representation of Statistics by Mathematical Formulae," *Journal of the Royal Statistical Society*, **61**, 670-700

Engle, R.F. (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, **50**, 987-1007

Galton, F., (1879). "The Geometric Mean in Vital and Social Statistics," *Proceedings of the Royal Society of London*, **29**, 365-367

Hall, S.G., D.K. Miles, and M.P. Taylor (1989). "Modeling Asset Prices with Time-Varying Betas," *The Manchester School* 57 (4), 340-356

Hinich, M.J. (1982). "Testing for Gaussianity and Linearity of a Stationary Time Series," *Journal of Time Series Analysis* 3, 169-176

Hinich. M.J. and C.S. Clay [1968]. "The Application of the Discrete Fourier Transform in the Estimation of Power Spectra, Coherence and Bispectra of Geophysical Data," *Reviews of Geophysics*, **6**(3), 347-363

Hinich, M.J. and H. Messer (1995). "On the Principal Domain of the Discrete Bispectrum of a Stationary Signal," *IEEE Trans. on Signal Processing,* **43** (9)*,* 2130-2134

Hinich, M.J and D. M. Patterson (1989). "Evidence of Nonlinearity in the Trade-by-Trade Stock Market Return Generating Process," (with D. M. Patterson), in *Economic Complexity, Chaos, Sunspots, Bubbles, and Nonlinearity*, W. Barnett, J. Geweke, and K. Shell (eds.), Cambridge University Press, Chapter 16, 383-409

Hinich, M.J and D. M. Patterson (1992). "A New Diagnostic Test of Model Inadequacy which uses the Martingale Difference Criterion," *Journal of Time Series Analysis*, **13** (3), 233-252

Hinich, M.J and P. Rothman (1998). "Frequency-Domain Test of Time Reversibly," *Macroeconomic Dynamics,* **2** (1), 72-88

Hinich, M.J. and M. Wolinsky (1988). "A Test for Aliasing using Bispectral Analysis," *Journal of the American Statistical Association*, 83 (402), "Theory and Methods", 499-502

Hinich, M.J. and M. Wolinsky (2005). "Normalizing Bispectra," *Journal of Statistical Planning and Inference* **130**, 405-411

Hinich, M.J and J. Dalle Molle (1995). "Trispectral Analysis of Stationary Random Time Series," *Journal of the Acoustical Society of America*, **97** (5), Part 1, 2963-2978

Horowitz, J. L., I.N. Lobato, J.C. Nankervis and N.E. Savin. (2006*).* "Bootstrapping the Box–Pierce *Q* test: A robust test of Uncorrelatedness," *Journal of Econometrics,* In Press Available online 22 August 2005**133** (*2*), 841-862

Hsieh, D.A. (1989). "Testing for Nonlinear Dependence in Daily Foreign Exchange Rates," Journal of Business **62**, 339-368

Johnson, N.L., (1949). "Systems of Frequency Curves Generated by Means of Translation," *Biometrika*, **36**, 149-176

McAlister, D., (1879). "The Law of the Geometric Mean,*" Proceedings of the Royal Society of London*, **29**, 367-376

McCullough, B.D., and C.G. Renfro (1999). "Benchmarks and Software Standards: A Case Study of GARCH Procedures," *Journal of Economic and Social Measurement* **25**, 59-71

Trimming Level	GARCH 1 1% Size	GARCH 1 5 % Size	GARCH 2 1% Size	GARCH 2 5% Size
κ = 2 %	27.78 %	51.38 %	22.19 %	45.03 %
κ = 5 %	7.04 %	20.70 %	5.56 %	17.46 %
κ = 10 %	1.12 %	5.45 %	0.87 %	4.58 %
κ = 15 %	0.30 %	1.90 %	0.25 %	1.61 %
κ = 20 %	0.10 %	0.83 %	0.08 %	0.74 %

Zero Bispectrum Test   Normal Model   c = 0.4   N = 50,000

Table 1

Trimming Level	GARCH 1 1% Size	GARCH 1 5 % Size	GARCH 2 1% Size	GARCH 2 5% Size
κ = 2 %	14.83 %	33.40 %	11.85 %	29.20 %
κ = 5 %	4.43 %	14.91 %	3.63 %	12.83 %
κ = 10 %	1.18 %	5.42 %	1.00 %	4.69 %
κ = 15 %	0.47 %	2.53 %	0.44 %	2.33 %
κ = 20 %	0.21 %	0.1.53 %	0.18 %	1.43 %

Time Reversibility Test   Normal Model   c = 0.4   N = 50,000

Table 2

Trimming Level	GARCH 1 1% Size	GARCH 1 5 % Size	GARCH 2 1% Size	GARCH 2 5% Size
κ = 10 %	3.80 %	13.33 %	3.16 %	11.77 %
κ = 15 %	0.66 %	3.49 %	0.56 %	3.06 %
κ = 20 %	0.19 %	1.22 %	0.16 %	1.10 %

Zero Bispectrum Test Double Tailed Exponential c = 0.4 N = 50,000

Table 3

Trimming Level	GARCH 1 1% Size	GARCH 1 5 % Size	GARCH 2 1% Size	GARCH 2 5% Size
κ = 10 %	2.67 %	10.24 %	2.33 %	9.19 %
κ = 15 %	0.78 %	3.83 %	0.67 %	3.51 %
κ = 20 %	0.30 %	1.86 %	0.29 %	1.76 %

Time Reversibility Test Double Tailed Exponential c = 0.4 N = 50,000

Table 4

Tests	GARCH 1	GARCH 1
	1% Size	5 % Size
Bispectrum = 0		
T(6) - κ = 10%	2.43 %	9.39 %
T(6) - κ = 15%	0.46 %	2.69 %
Time Reverse		
T(6) - κ = 10%	1.92 %	7.90 %
T(6) - κ = 15%	0.58 %	3.28 %
Both Tests t(6) Model	***c = 0.4 N = 50,000***	***c = 0.4 N = 50,000***

Table 5

Tests	ARCH 1 1% Size	ARCH 1 5 % Size	ARCH 2 1% Size	ARCH 2 5% Size
Bispectrum = 0				
Normal - κ = 5%	0.84 %	4.34 %	0.54 %	3.15 %
Normal - κ = 10%	0.21 %	1.39 %	0.16 %	1.15 %
Time Reverse				
Normal - κ = 5%	0.98 %	4.78 %	0.73 %	3.80 %
Normal - κ = 10%	0.39 %	2.14	0.34 %	1.86 %
Bispectrum = 0				
T(6) - κ = 5%	2.23 %	8.75 %	1.41 %	6.31 %
T(6) - κ = 10%	0.30 %	2.08 %	0.24 %	1.63 %
Time Reverse				
T(6) - κ = 5%	1.70 %	7.59 %	1.25 %	5.91 %
T(6) - κ = 10%	0.47 %	2.61 %	0.36 %	2.22 %

5% & 10% Trimming Levels for both Tests c = 0.4 N = 50,000

Figure 6

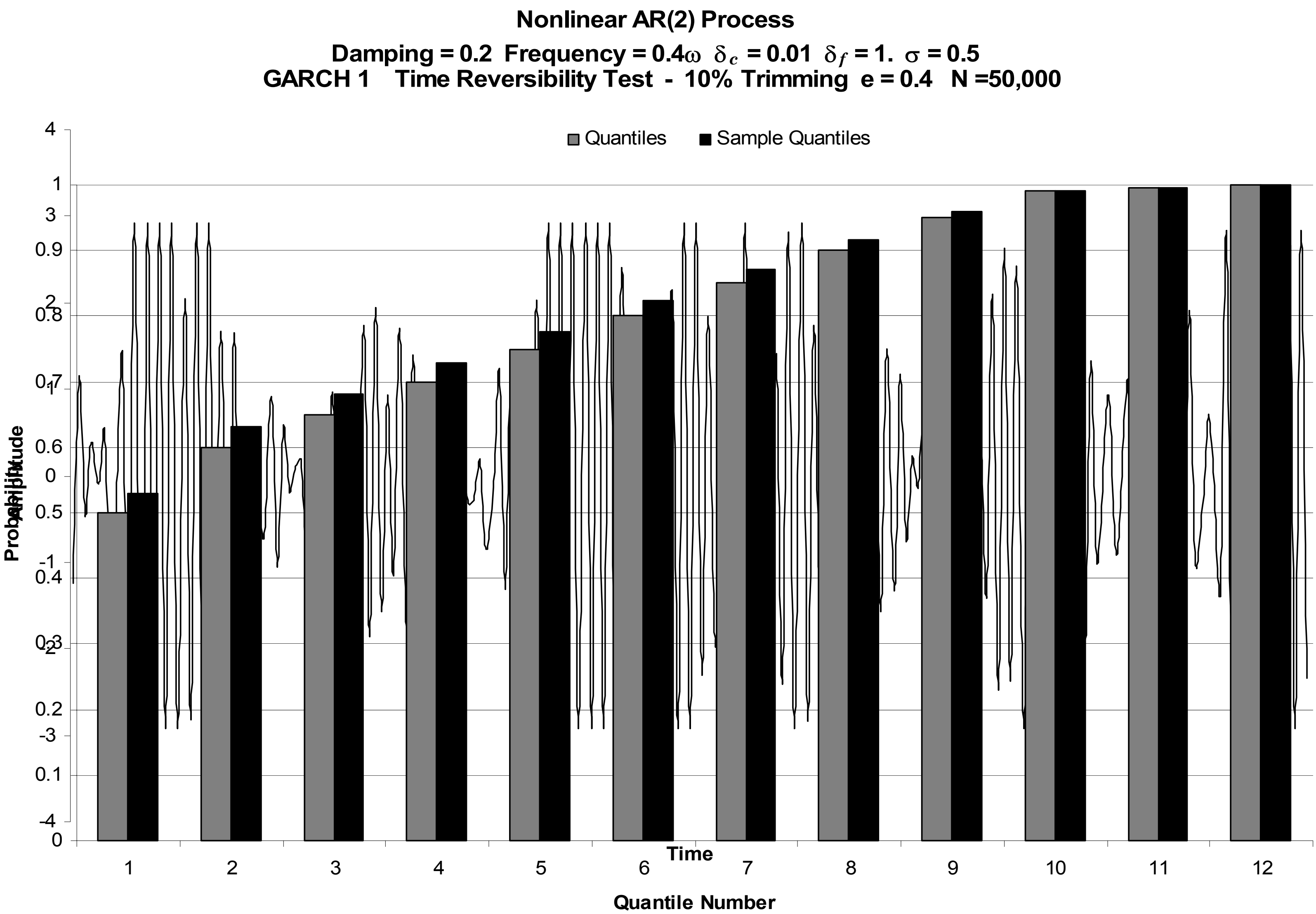

Nonlinear AR(2) Process
Damping = 0.2 Frequency = 0.4ω δc = 0.01 δf = 1. σ = 0.5
GARCH 1 Time Reversibility Test - 10% Trimming e = 0.4 N =50,000
Quantiles
Sample Quantiles
Probability
Amplitude
Time
Quantile Number
1
2
3
4
5
6
7
8
9
10
11
12

Figure 1

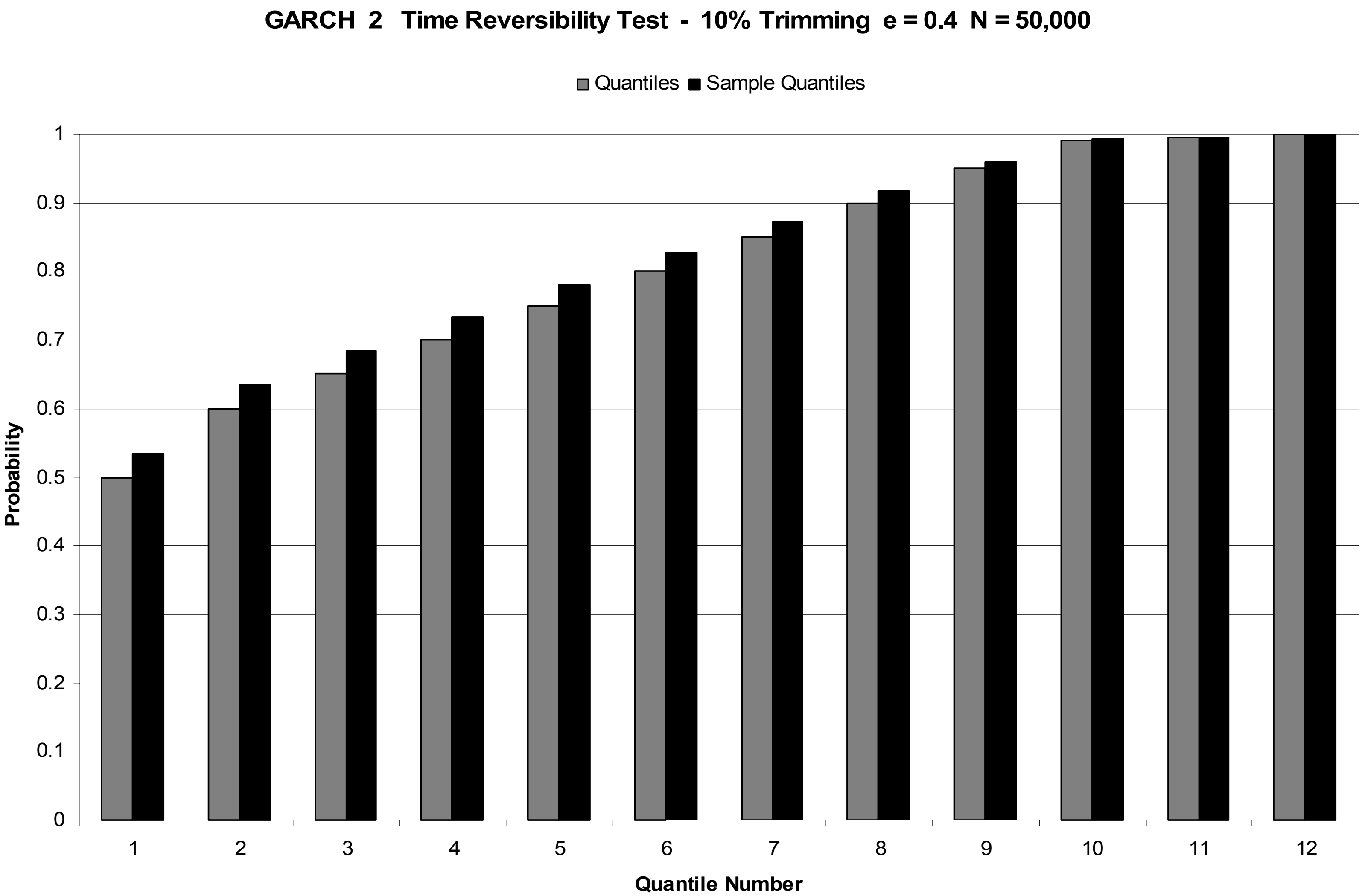

GARCH 2 Time Reversibility Test - 10% Trimming e = 0.4 N = 50,000
Quantiles
Sample Quantiles
Probability
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1
2
3
4
5
6
7
8
9
10
11
12
Quantile Number

Figure 2

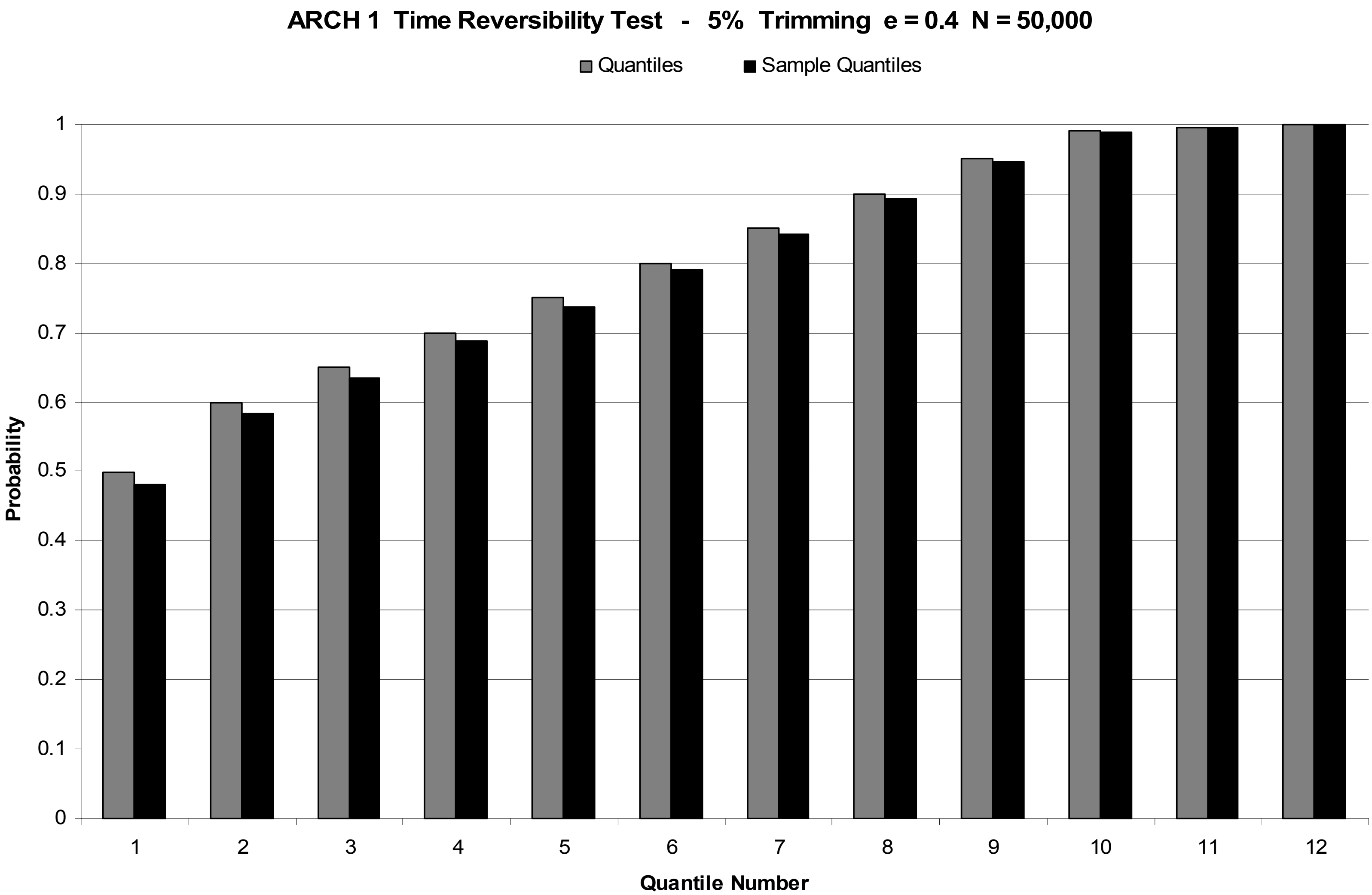

ARCH 1 Time Reversibility Test - 5% Trimming e = 0.4 N = 50,000
Quantiles
Sample Quantiles
Probability
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1
2
3
4
5
6
7
8
9
10
11
12
Quantile Number

Figure 3

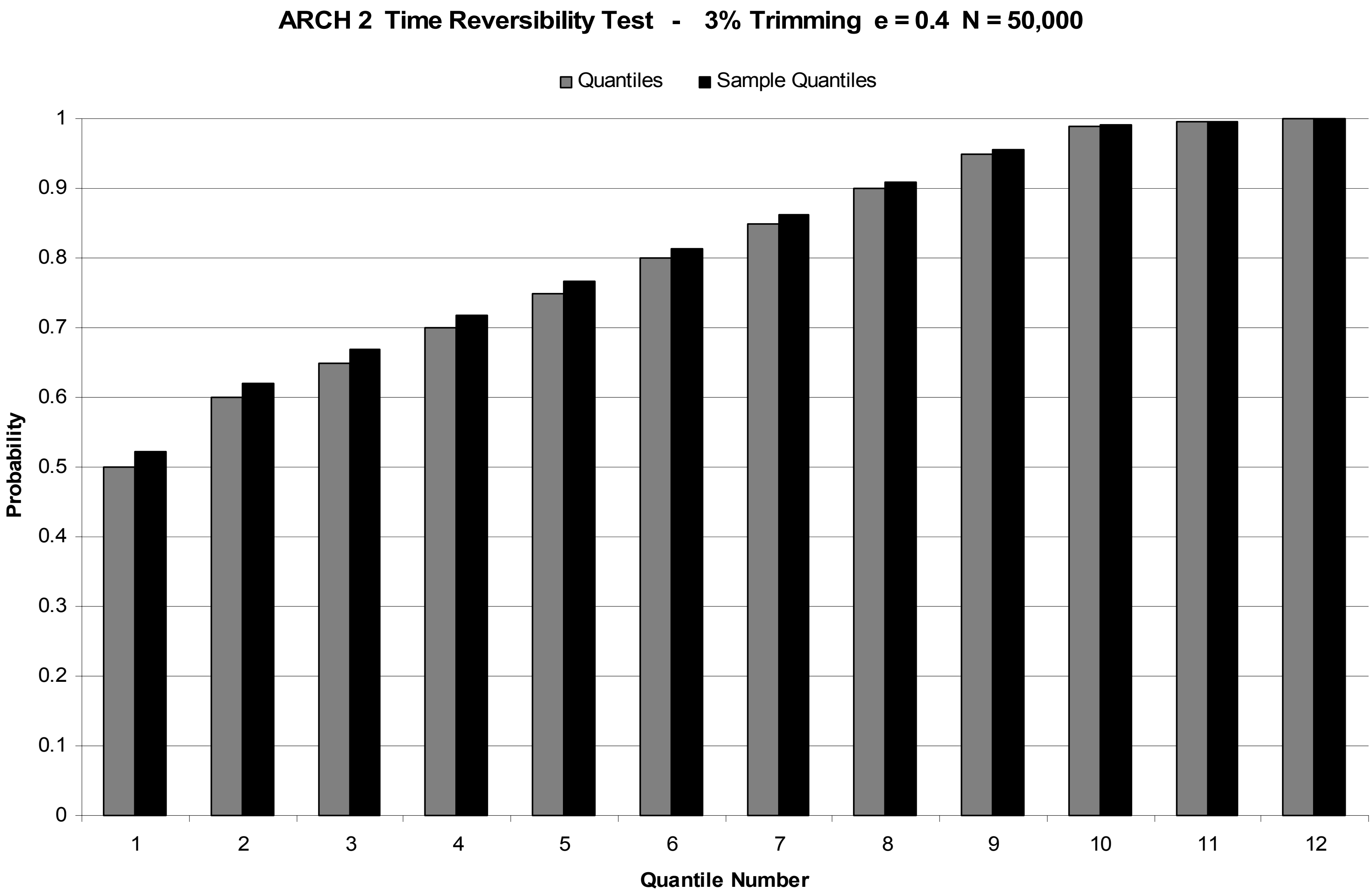
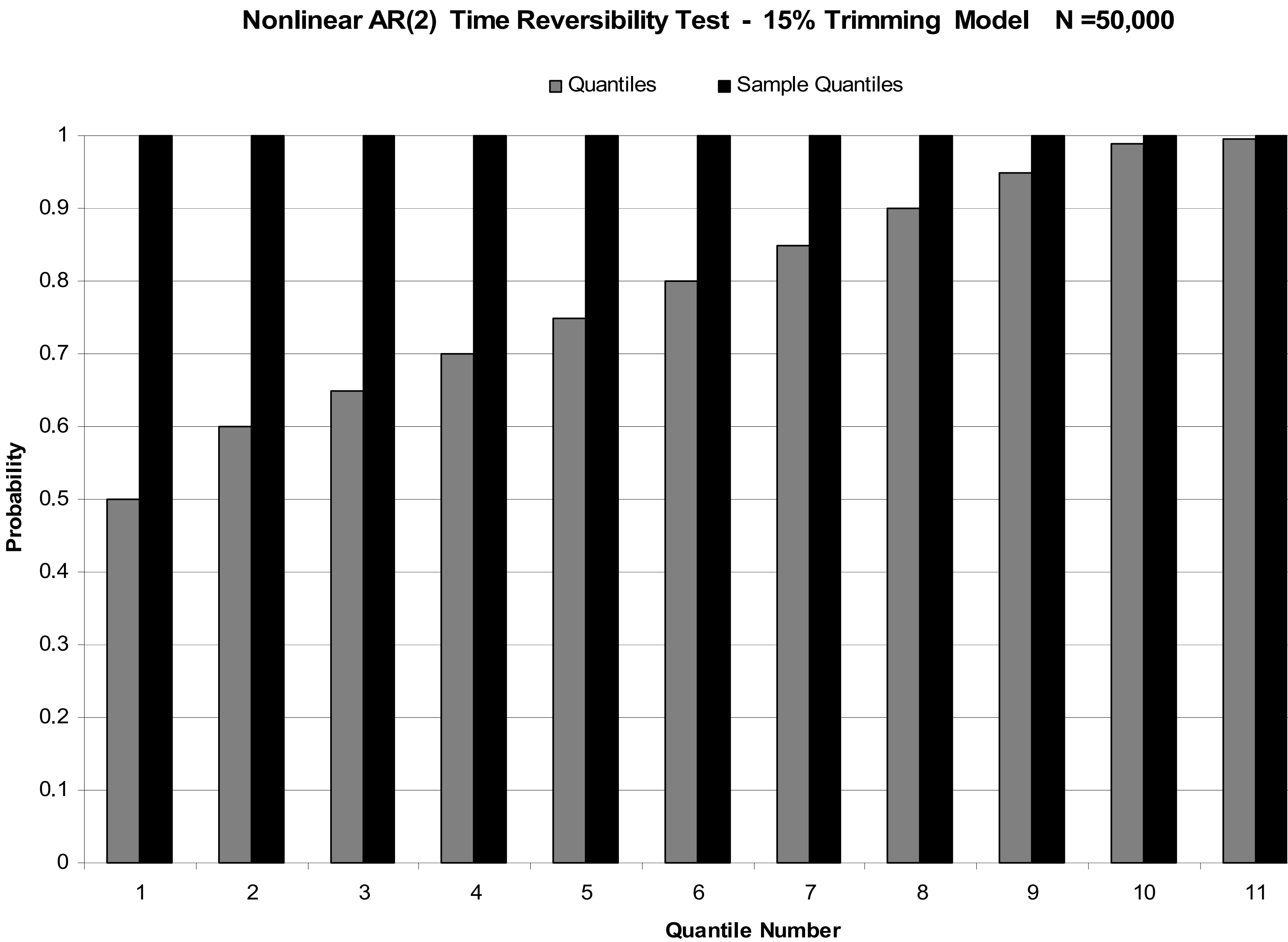


Figure 41

Figure 5



Nonlinear AR(2) Time Reversibility Test - 15% Trimming Model N =50,000
Quantiles
Sample Quantiles
Probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
8
9
10
11
Quantile Number

Figure 6