A Single-Blind Controlled Competition Among Tests for Nonlinearity and Chaos; Further Results

Maria Lemos
Department of Economics
Rua Itapoan 20 - Sao Cristovao
Rio De Janeiro
R. J. - CEP: 20911-450
BRAZIL
Mlemos2@UniverCidade

Houston H. Stokes*
University of Illinois at Chicago
Department of Economics (MC 144)
College of Business Administration
601 South Morgan Street, Room 2103
Chicago, Illinois 60607-7121
312-996-0971
hhstokes@uic.edu

Revised: July 6, 2000

JEL Classifications: C12, C15, C22, and C45. **Keywords**: competition, chaos, nonlinearity

Abstract: The size and power of the Hinich (1982) test is investigated using Group I and Group II models suggested by Lee et al (1993). Simulations were conducted over a range of admissible blocksizes. A test statistic discussed in Stokes (1991, 1997) that is an average of all test statistics between the upper and lower admissible blocksizes is evaluated. This test statistic has the distinct advantage that is does not require a blocksize assumption. Using a variety of blocksizes, the Hinich (1982) nonlinearity test is applied to the five test series from Barnett-Gallant-Hinich-Jungeilges-Kaplan-Jensen (1997). The Hinich (1982) test correctly detected whether there was nonlinearity in the series in 4 out of 5 cases.

*Note: Correspondence should be addressed to Houston H. Stokes, Department of Economics (M/C 144), University of Illinois at Chicago, 601 South Morgan Street, Chicago, IL 60607. E-mail address hhstokes@uic.edu. The authors appreciate comments of Mel Hinich, Douglas Patterson and others who have seen this paper in draft form. We are responsible for any errors that remain.

1. Introduction

The Hinich (1982) nonlinearity test is invariant to linear filtering of the data and thus does not require that the series be prewhitened prior to being tested.[1] For this reason alone, it makes an attractive test for the residuals of models that may or may not be white noise. Of considerable interest is the performance of this test over a range of alternative models. To this end, the size and power of the of the Hinich (1982) test is investigated using Group I and Group II models suggested by Lee et al (1993) who investigated other nonlinearity tests. The Hinich (1982) test was not studied due to software problems at that time. The simulations were conducted over a range of admissible blocksizes. A test statistic discussed in Stokes (1991, 1997), that is an average of all test statistics between the upper and lower admissible blocksizes, is evaluated. This test statistic has the distinct advantage that is does not require the user to make a blocksize assumption that the usual Hinich (1982) procedure requires.

Barnett-Gallant-Hinich-Jungeilges-Kaplan-Jensen (1997), here in after BGHJKJ (1997), designed a single-blind, controlled competition to test for nonlinearity in five generated series of 2,000 observations each. The objective of their paper was to determine what five highly regarded tests for nonlinearity would find in a situation where the testers did not know the structure of the series. Series I was a chaotic Feigenbaum recursion. Series II was a GARCH process. Series III was a nonlinear moving average and series IV was an ARCH process. Series V was an ARMA process. The exact model for each series is given in BGHJKJ (1997) and will not be given here.

The present authors obtained the data for the project from the Barnett web page. The Hinich (1982) Fortran program, as implemented in B34S, was used to perform the tests. While

[1] Ashley-Patterson-Hinich (1986, p. 174) and Patterson-Ashley (2000 page 44) prove the invariance theorem. Other completing nonlinearity tests require that the series be white noise.

the BGHJKJ (1997) paper clearly gives the impression that the Fortran program referred to in Hinich (1982) was used, the actual Fortran code used in the testing procedure was <u>not</u> the original Patterson (1983) program but rather a newer variant that used a block averaging procedure, which Hinich feels was equivalent to the older approach. Our paper applies the older Fortran program[2] to test the series and, in addition, performs the test for a range of blocksizes to remove the possibility that the results of the test are influenced by the blocksize selected. In order to avoid the possibility that the results reported here are tainted by the fact that we had seen the prior paper, only the default Hinich (1982) test settings in B34S®[3] have been used.

After briefly discussing the Hinich (1982) test in section 2, the performance of the Hinich test against a number of different types of nonlinearity discussed by Lee et all (1993) is presented in section 3. Analysis of the BGHJKJ (1997) data and brief conclusions are given in sections 4 and 5 respectively.

2. Discussion of Hinich Test

The Hinich (1982) test is based on the Brillinger (1965) proof that the skewness function, $\Gamma(f_1,f_2)$ at frequency pairs (f_1,f_2), is constant over all frequencies $(f_1, f_2) \in \Omega$ if $\{x(t)\}$ is linear and zero over all frequencies if $\{x(t)\}$ is Gaussian. The skewness function, $\Gamma(f_1,f_2)$ at frequency pairs (f_1,f_2), is defined in terms of the bispectrum $B_{xxx}(f_1,f_2)$ and power spectrum $S_x(f_1)$ as

$$\Gamma^2(f_1,f_2) = |B_{xxx}(f_1,f_2)|^2 / S_x(f_1)S_x(f_2)S_x(f_1+f_2). \tag{2.1}$$

[2] Patterson-Ashley (2000) use this older program.
[3] The B34S® Data Analysis Program is documented in Stokes (1991 and 1997). A free student version of this program containing the Hinich (1982) test and other time series capability can be downloaded from www.uic.edu/~hhstokes/hhstokes/b34s.htm. B34S contains a modified version of the Patterson (1983) BISPEC program which was used in Hinich (1982). The Barnett data is in member BARNETT in the file b34sdata.mac which is distributed with the B34S.

The Brillinger (1965) proof suggests that once a consistent estimator of the bispectrum is calculated, nonlinearity and Gaussianity tests can be performed, using a sample estimator of the skewness function, $\Gamma(f_1,f_2)$. For the sample $\{x(0),x(1),...,x(N-1)\}$, define $F_{xxx}(j,k)$ as an estimate of the bispectrum of $\{x(t)\}$ at the frequency pair (f_j,f_k), where $f_k=k/N$ for each integer k, as follows:

$$F_{xxx}(j,k) = X(f_j)X(f_k)X^*(f_{j+k})/N\ , \tag{2.2}$$

where $X(f_j) = \sum_{t=0}^{N-1} x(t)\exp(-i\omega_j t)$, where $\omega_j = 2\pi n/N$ for n=0,1,…,N-1 and the * denotes complex conjugate. $F_{xxx}(j,k)$ must be smoothed to form a consistent estimator. Hinich (1982) uses $\langle B_{xxx}(m,n)\rangle$ to denote an estimator of $B_{xxx}(m,n)$, which is obtained by averaging over adjacent frequency pairs of $F_{xxx}(j,k)$:

$$\langle B_{xxx}(m,n)\rangle = M^{-2} \sum_{j=(m-1)M}^{mM-1} \sum_{k=(n-1)M}^{nM-1} F_{xxx}(j,k)\ . \tag{2.3}$$

Equation 2.3 illustrates that the Hinich test involves using an estimator of the bispectrum, $\langle B_{xxx}(f_m,f_n)\rangle$, that is the average value of the $F_{xxx}(j,k)$ over a square of M^2 points and thus is critically dependent on the assumption of the appropriate M value to use. $\langle B_{xxx}(f_m,f_n)\rangle$ is a consistent and asymptotically complex normal estimator of the bispectrum $B_{xxx}(f_1,f_2)$, if the sequence (f_m,f_n) converges to (f_1,f_2) (see Hinich, 1982). The above discussion suggests that the key assumption needed to run the Hinich (1982) test involves the selection of M. The larger (smaller) M, the smaller (larger) the finite sample variance and the larger (smaller) the sample bias. Ashley-Patterson-Hinich (1986, 171) found that empirically larger (smaller) M values converged slower (faster) to the asymptotic distribution for the linearity test but faster (slower) for the Gaussinity test. Because of this trade-off, there is no one unique M that is appropriate to use in performing nonlinearity and Gaussianity tests. Hinich (1982) has suggested that a good value for M is the square root of the number of observations ($T^{0.5}$). As a lower bound, select M = $(T/3)^{0.5}$. The B34S® implementation of the Hinich test, which is documented in Stokes (1991, 1997), automatically does a grid search over these admissible M values to test the sensitivity of the results to the M value chosen. An average of the linearity test for all admissible M values is also given. In summary, when M is large, the bandwidth is large, the variance is reduced and the resolution of the tests is small since there are too few terms for the linearity test. If M is small, there is a large number of terms to sort for the linearity test, the variance may be too large and the chi-square approximation used for the linearity test may not be good.

Ashley-Patterson-Hinich (1986) presented an equivalence theorem that proved that the Hinich bispectral linearity test statistic is invariant to linear filtering of the data. Patterson-Ashley (2000) expands on these results. The importance of this equivalence theorem is that it proves that the linearity test can be either applied to the raw series or to the residuals of a linear model. No prewhitening model need be applied to the series before the test is applied. An additional important implication of the theorem is that if x(t) is found to be nonlinear, then the residuals of a linear model of the form y(t) = f(x(t)) will be nonlinear since the nonlinearity in x(t) will pass through any linear filter.

3. Performance of Hinich Test

To examine the power properties of the Hinich test against different types of nonlinearities, we considered essentially the same models used by Lee et al (1993)[4]. The present work complements their Monte Carlo study as they have not included the bispectrum test in their analysis on the grounds that it "proved to be too expensive to use in the Monte Carlo simulation."

---

[4] Instead of considering an MA(2) as one of the models in their Block II, we simulated a GARCH(1,1) model. All the AR models were generated using B34S® which, in turn, uses the IMSL routines ftgen, ggnml and ggubs. All nonlinear models considered were generated using GAUSS and the data then moved to B34S for analysis.

The models they studied were divided into two groups and for each simulated series the innovations follow the unit-normal distribution.

Group I

Model 1: Autoregressive (AR)

$$y_t = 0.6y_{t-1} + \varepsilon_t$$

Model 2: Bilinear (BL)

$$y_t = 0.7y_{t-1}\varepsilon_{t-2} + \varepsilon_t$$

Model 3: Threshold Autoregressive (TAR)

$$y_t = \begin{array}{ll} 0.9y_{t-1} + \varepsilon_t & \text{for } |y_{t-1}| \le 1, \\ -0.3y_{t-1} + \varepsilon_t & \text{for } |y_{t-1}| > 1 \end{array}$$

Model 4: Sign Autoregressive (SGN)

$$y_t = \text{sgn}(y_{t-1}) + \varepsilon_t \qquad \text{where} \quad \text{sgn}(x) = 1 \text{ if } x>0, \ =0 \text{ if } x=0, \ = -1 \text{ if } x<0.$$

Model 5: Nonlinear Autoregressive (NAR)

$$y_t = (0.7|y_{t-1}|) / (|y_{t-1}| + 2) + \varepsilon_t$$

Group II

Model 1: GARCH(1,1)

$$v_t = \varepsilon_t h_t^{0.5} \qquad \text{and} \qquad h_t = 0.1 + 0.8h_{t-1} + 0.1v_{t-1}^2$$

Model 2: Heteroskedastic MA(2) (HMA)

$$y_t = \varepsilon_t - 0.4\varepsilon_{t-1} + 0.3\varepsilon_{t-2} + 0.5\varepsilon_t\varepsilon_{t-2}$$

Model 3: Nonlinear MA (NLMA)

$$y_t = \varepsilon_t - 0.3\varepsilon_{t-1} + 0.2\varepsilon_{t-2} + 0.4\varepsilon_{t-1}\varepsilon_{t-2} - 0.25\varepsilon^2_{t-2}$$

Model 4: AR(2)

$$y_t = 0.4y_{t-1} - 0.3y_{t-2} + \varepsilon_t$$

Model 5: Bilinear AR (BAR)

$$y_t = 0.4y_{t-1} - 0.3y_{t-2} + 0.5\, y_{t-1}\varepsilon_{t-1} + \varepsilon_t$$

Model 6: Bilinear ARMA (BARMA)

$$y_t = 0.4y_{t-1} - 0.3y_{t-2} + 0.5\, y_{t-1}\varepsilon_{t-1} + 0.8\varepsilon_{t-1} + \varepsilon_t$$

As the Hinich test is sensitive to the choice of M, in our analysis, the simulations were conducted over a range of admissible blocksizes. For reasons of space, only blocksizes 25, 31 and 45 and the mean of the statistics for 25 – 45 were reported. As suggested by Hinich (1982), the upper bound value of M should be equal to the square root of the number of observations, $M=T^{0.5}$ and the lower bound should be equal to $M=(T/3)^{0.5}$. In addition, we used a blocksize suggested by Ashley-Patterson-Hinich (1986) equal to $M=T^{0.45}$. Finally, we considered a test statistic given by the B34S® and discussed in Stokes (1991, 1997) that is an average of all test statistics between the upper and lower blocksizes, which will be called $M=\mu$. For a sample size of 2,000, the blocksizes are 45, 25 and 31, respectively.

The size of the test was calculated for both the asymptotic theory and the bootstrap considering the different blocksizes mentioned above.[5] As a first step, the empirical 5% critical values were calculated from an AR(1) model with ϕ taking the values of -0.9, -0.6, -0.3, 0.0, 0.3, 0.6 and 0.9. The bootstrap consisted of 3,000 replications of sample size 2,000 for each linear model. Actually, 2,200 observations were calculated for each model, with 200 being dropped to reduce the effect of the starting values. These critical values are reported in Table 1 and are, in general, well below the asymptotic 5% critical value of 1.64 for M=25,31 and 45.

Table 1. Empirical 5% Critical Values[a]

Blocksize(M)	**-0.9**	**-0.6**	**-0.3**	**0.0**	**0.3**	**0.6**	**0.9**
μ (mean)	0.121	0.672	0.789	0.874	0.811	0.636	-0.096
25	-0.022	1.042	1.386	1.446	1.364	1.209	-0.035
31	0.550	1.257	1.516	1.563	1.437	1.224	0.437
45	1.228	1.585	1.675	1.707	1.632	1.521	1.132

[a] 5% critical values obtained from 3,000 simulations of an AR(1), with ϕ equal to -0.9, -0.6, -0.3, 0.0, 0.3, 0.6 and 0.9.

To obtain the size of the test, 1,000 new replications of each AR(1) model with sample size 2,000 were constructed and the 5% critical values obtained in the simulation stage were used. The results are reported in Table 2. The first number corresponds to the empirical size and the number in parentheses reflects the asymptotic one. Note that for the μ test statistic, we only report the size obtained with the empirical critical values as its asymptotic distribution should be different than N(0,1).

[5] This approach combines aspects of the traditional approach of Hinich (1982) with the newer approaches as suggested by Patterson-Ashley (2000) that involve the bootstrap.

Table 2. Size of Hinich test[a,b]

Blocksize (M)	-0.9	-0.6	-0.3	0.0	0.3	0.6	0.9
μ (mean)	**1.0**	4.3	**6.9**	3.9	4.5	4.5	5.5
25	4.0 **(0.0)**	4.6 **(2.0)**	5.2 **(3.1)**	3.6 **(2.2)**	4.8 **(3.1)**	3.7 **(0.8)**	4.0 **(0.2)**
31	**3.0** **(0.0)**	3.9 **(3.0)**	5.2 (4.1)	4.3 (3.8)	4.7 **(3.4)**	4.2 **(2.3)**	4.7 **(0.4)**
45	**3.2** **(1.0)**	3.9 **(8.0)**	**7.1** **(7.3)**	4.2 (4.7)	6.1 (5.9)	4.5 (3.6)	4.3 **(1.8)**

[a]Power (%) was computed for each $y_t = \phi y_{t-1} + \varepsilon_t$ using the 5% critical values obtained from simulations of an AR(1) model with ϕ equal to -0.9, -0.6, -0.3, 0.0, 0.3, 0.6 and 0.9.
[b]The theoretical 5% critical value is 1.64 for M=25, 31 and 45.

The 95% confidence interval of the sizes around the null hypothesis of 5% is 3.6 to 6.4% for 1,000 replications. Results significantly different from 5% are highlighted in Table 2. The results obtained here are in line with those of Ashley-Patterson (1998) with respect to the bispectral test for Gaussian innovations.[6] For the sample size of 2,000 observations, the asymptotic size estimates usually lie outside the 95% confidence interval around 0.05. That is the case for all the AR(1) models used in the exercise. Conversely, the simulations-based size estimates lie in the mentioned interval except for the linear models, where ϕ is equal to -0.9 and -0.3. The critical values obtained through bootstrapping seem to be more appropriate for the sample size in hand than relying on the asymptotic values.

Next, we report the power of the Hinich test for the models in Groups I and II as suggested by Lee et al (1993). For each model, 1,000 replications of sample size 2,000 were simulated. For Group I the power was computed using the 5% critical values from the AR(1) model simulation, with ϕ=0.6. For Group II, we computed the empirical critical values from

[6] Patterson-Ashley (2000) extends the 1998 results.

3,000 replications of sample size 2,000 of Model 4, an AR(2) model. To calculate the size of the test, another 1,000 replications of this model were simulated using the same sample size. In Tables 3 and 4 the numbers in parentheses reflect the power of the test considering the 5% theoretical critical value. Table 3 reports the power of the Hinich test against the models in Group I. The Hinich test proved very powerful against the type of nonlinearity present in the bilinear model. For the other three models in Group I, its performance was not good, even though the power was much better when the empirical critical values were used.

Table 3. Group I Models: Power of Hinich Test[a,b]

	M=μ	M=25	M=31	M=45
MODEL 1 (AR(1))	4.5%	3.2% (0.8%)	4.2% (2.3%)	4.5% (3.6%)
MODEL 2 (BL)	100%	100% (100%)	100% (100%)	99.9% (99.9%)
MODEL 3 (TAR)	8.0%	7.5% (3.3%)	7.0% (3.8%)	6.0% (5.3%)
MODEL 4 (SGN)	5.1%	4.9% (2.3%)	5.8% (1.8%)	6.2% (5.0%)
MODEL 5 (NAR)	21.3%	9.1% (4.6%)	12.0% (6.3%)	13.7% (11.1%)

[a]Power was computed using 5% critical value obtained from simulations of an AR(1) model as in $y_t=0.6y_{t-1}+\varepsilon_t$ (Model 1). The power of the test using the 5% theoretical critical values is reported in parentheses. The sample size considered in the simulations was equal to 2,000, with 1,000 replications of each model.

[b]The 5% critical values obtained from simulations for the Hinich test were equal to 0.636,1.209, 1.224 and 1.521 for M=μ, 25, 31 and 45, respectively. The theoretical 5% critical value is 1.64 for the Hinich test.

In Table 4 the power of the Hinich test against the models in Group II is reported. For these models the Hinich test performed well in every case, especially for M=μ..

Table 4. Group II Models: Power of Hinich Test[a,b]

	M=μ	M=25	M=31	M=45
MODEL 1 (GARCH)	96.8%	80.5% (73.5%)	82.0% (73.9%)	78.5% (76.4%)
MODEL 2 (HMA)	100%	100% (100%)	100% (100%)	97.3% (97.9%)
MODEL 3 (NLMA)	100%	99.8% (99.6%)	99.9% (99.8%)	98.3% (98.5%)
MODEL 4 (AR(2))	4.3%	4.0% (2.6%)	4.5% (3.4%)	3.7% (4.7%)
MODEL 5 (BAR)	100%	100% (100%)	100% (100%)	100% (100%)
MODEL 6 (BARMA)	100%	100% (100%)	100% (100%)	99.2% (99.6%)

[a]Power was computed using 5% critical value obtained from simulations of an AR(2) model as in $y_t=0.4y_{t-1}-0.3y_{t-2}+\varepsilon_t$ (Model 4). The power of the test using the 5% theoretical critical values is reported in parentheses. The sample sizes considered in the simulations were equal to 2,000, with 1,000 replications of each model.
[b]The 5% critical values obtained from simulations for the Hinich test were equal to 0.852, 1.360, 1.522 and 1.812 for M=μ, 25, 31 and 45, respectively. The theoretical 5% critical value is 1.64 for the Hinich test.

The results reported in Tables 3 and 4 indicate that, for the models in this study, the Hinich test had greater power to detect departures from linearity if one considered M=μ instead of M=25, 31 or 45. This is more pronounced for the GARCH model (Model 1-Group II) and for the nonlinear AR (Model 5-Group I). The only exception was provided by the Sign model (Model 4-Group I). In terms of the empirical sizes based on the simulations, the Hinich test for the average of the admissible values for M was not statistically different than 5% in 5 out of 7 of the linear models studied. This was also the case when M=45. It was not possible to detect from these simulations any pattern such that the size of the Hinich test for M=μ should be greater than when any specific value for M is used. The above simulations suggest it would be appropriate to apply the Hinich (1982) test to the BGHJKJ (1997) data. As mentioned earlier, to avoid any bias in our results from having known the distribution of the test series, only the default Hinich (1982) settings have been used.

4. Results Using the Data From BGHJKJ (1997)

The results from running the B34S® implementation of the Hinich (1982) test using the default setup are given in Table 5. The BGHJKJ (1997) results are shown as Λ. Because of the invariance theorem, the Hinich linearity test does not require that any autocorrelation be removed from the series, which is appropriate for series I and V, which have a great deal of autocorrelation. The bispectrum was smoothed with a cosine bell in the frequency domain using the Hinich recommended smoothing cosin width of 3*M, since we could not rule out autocorrelation in the series analyzed. Comparing our runs with the Hinich test results in BGHJKJ (1997), for series I (Feig) linearity is strongly accepted. This may be due to such series having widely scattered spikes on the bispectrum. For series II (GARCH) now linearity is rejected for all M values. The test statistics are way above the critical values given in Table 1. For series III (NLMA) linearity is again rejected for all M values resorted and also M=μ (6.94). For series IV (ARCH) again linearity is rejected for all M values reported, including the mean (5.61). These findings show that the older Hinich (1982) program is able to detect the nonlinearity that was built into series II, III and IV. Only in chaotic series I was the nonlinearity not detected. For series V (ARMA) linearity is accepted. In summary, our results show that the Hinich test correctly rejected linearity for series II, III and IV in contrast with the results reported by BGHJKJ (1997) which failed to detect nonlinearity. Notice that if we had considered the empirical critical values from the bootstrap reported in Table 1, we would be able to reject the null of linearity even more strongly. In other words, the 5% asymptotic critical value of 1.64 is conservative.

Table 5. Hinich (1982) Linearity Test Statistics

	M=μ	M=25	M=31	M=45	Λ
Series 1 (FEIG)	-13.14	-18.12	-14.39	-9.76	-12.15
Series 2 (GARCH)	4.55	3.08	4.02	3.82	-0.61
Series 3 (NLMA)	6.94	7.23	8.52	4.22	1.84
Series 4 (ARCH)	5.61	5.07	4.98	4.00	0.41
Series 5 (ARMA)	-13.12	-18.10	-14.38	-9.75	-12.03

M = block size, μ = the mean of the statistics. A positive value ≥ 1.64 implies a significant rejection of linearity at the 5% significance level. Λ= values reported by BGHJKJ (1997).

5. Conclusions

The performance of the Hinich test and a proposed modification involving the mean of the statistics calculated over the admissible blocksizes was studied using Monte Carlo analysis of the models suggested by Lee et al (1993). The original Hinich (1982) nonlinearity test program was able to detect the absence or presence of nonlinearity correctly in 4 out of 5 cases studied by BGHJKJ (1997). The only time the test failed to detect nonlinearity was for the chaotic Feigenbaum recursion, which appears to have a relatively flat bispectrum, with irregular and widely spaced spikes. If the series were longer, the nonlinearity might have been detected. All Hinich tests were run over the admissible grid of blocksizes. The fact that the Hinich test was able to detect the nonlinearity in the GARCH and ARCH cases is encouraging since it is not supposed to be particularly powerful in detecting nonlinearity in these types of models.

6. References

Ashley, R., Douglas Patterson (1998). “Nonlinear Model Specification/Diagnostics: Insights from a Battery of Nonlinearity Tests”. Virginia Tech Working Paper, E98-06.

Ashley, R., Douglas Patterson, and Melvin J. Hinich (1986). “A Diagnostic Test for Nonlinear Serial Dependence in Time Series Fitting Errors.” *Journal of Time Series Analysis* 7, no. 3: 165-177.

Barnett, William. A. Ronald Gallant, Melvin Hinich, Jochen Jungeilges, Daniel Kaplan, and Mark Jensen (1997). “A Single-Blind Controlled Competition Among Tests for Nonlinearity and Chaos”, *Journal of Econometrics*, 82, 157-192 .

Brillinger, D. R. (1965) “An Introduction to Polyspectrum.” *Annals of Mathematical Statistics*, 36: 1351-1374.

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

Lee, Tae-Hwy, Halbert White, and Clive Granger (1993). “Testing for Neglected Nonlinearity in Time Series Models.” *Journal of Econometrics*, 56, 269-290.

Patterson, Douglas (1983). “BISPEC: A Program to Estimate the Bispectrum of a Stationary Time Series.” *American Statistician*, 37 (4), 323-324.

Patterson, Douglas and Richard A. Ashley (2000). *A Nonlinear Time Series Workshop: A Toolkit for Detecting and Identifying Nonlinear Serial Dependence*, Kluwer Academic Publishers, Boston.

Stokes, Houston H. (1997). *Specifying and Diagnostically Testing Econometric Models*, 2nd ed, New York, Quorum Books, First edition 1991.