Testing for Nonlinearities in Economic and Financial Time Series

Maurice Peat
Maxwell Stevenson

ISSN: 1036-7373
TESTING FOR NONLINEARITIES IN ECONOMIC
AND FINANCIAL TIME SERIES

by

Maurice Peat
Max Stevenson

School of Finance and Economics
University of Technology, Sydney,
P.O. Box 123
BROADWAY N.S.W. 2007
ABSTRACT

Cyclical asymmetry has been recognised as a nonlinear phenomenon in numerous recent studies examining various economic and financial time series. If the nonlinear phenomena can be modelled by a nonlinear stochastic structure like the bilinear (BL), exponential autoregressive (EAR), smooth transition autoregressive (STAR), or self-exciting threshold autoregressive (SETAR) types, then we need tests to enable us to identify these various nonlinear models. In this paper we suggest modifications to the Tsay (1991) general test for identifying nonlinearities of the BL, EAR, STAR, and SETAR types as they occur in time series. Our testing procedure is simulated to determine its empirical properties.
1. INTRODUCTION

This paper is concerned with modelling economic and financial time series. In particular we are interested in testing for different types of stochastic nonlinearity which may provide a superior alternative to the linear model.

The need to be able to test for various types of stochastic nonlinearity stems from the numerous claims in the economics and finance literature of the superiority of nonlinear models for modelling univariate time series data.

Cyclical asymmetry has been recognised as a nonlinear phenomenon in numerous recent studies examining various economic and financial time series. For example, in business cycle theory, economists since Mitchell (1927) and Keynes (1936) have questioned the symmetric nature of the business cycle. Asymmetry of the cycle is reflected in sharp retractions during downturns in the economy as opposed to gradual upswings during recoveries. In practice, such time irreversibility is unlikely to be captured by a Gaussian linear time series model like the autoregressive moving average models of Box and Jenkins (1976). Using the postwar U.S. quarterly unemployment rate series, Rothman (1991) was the first to correctly identify asymmetric evidence of the business cycle within a Markov chain context. Brock and Sayers (1988) used the Grassberger - Procaccia (1983) dimension calculation algorithm to support the assertion that the unemployment rate series is generated by a chaotic attractor. Strong evidence of nonlinear behaviour was found by Luukkonen and Teräsvirta (1991) when they tested unemployment rate series for linearity against several nonlinear alternatives which included the smooth threshold autoregressive (STAR) class. Using U.S. monthly aggregate and sectoral unemployment rates, Ham and Sayers (1990) found strong evidence of nonlinearity of the self exciting threshold autoregressive (SETAR) type, while for monthly U.K. aggregate and regional data, Jones and Stevenson (1992), and Jones, Manning, and Stevenson (1993), have reported superior fits of SETAR models over linear alternatives for the aggregate series as well as those for most regions. Rothman (1994) uses an exponential autoregressive (EAR) model, as introduced by Ozaki (1980), to reveal a better fitting model to U.S. quarterly unemployment data than models from the ARMA \((p,q)\) class.
There is abundant empirical evidence that nonlinear dependence in daily or weekly changes in log exchange rates is suitably described by either an autoregressive conditional heteroscedasticity (ARCH) model as suggested by Engle (1982), or by the more generalised version introduced by Bollerslev (1986) called the GARCH model. In both models there is linear dependence in the conditional mean of exchange rate changes, with the ARCH term capturing the most apparent type of nonlinear dependence, namely, that large (small) exchange rate changes are followed by large (small) changes in either sign [Baillie and Bollerslev (1989), Diebold and Nerlove (1989), Engle and Bollerslev (1986), and Hsieh (1989)]. Recent evidence points to a nonlinear dependence in the conditional mean of exchange rate changes. Diebold and Pauly (1988) studied a portfolio balance model and found that changes in conditional variance affected the conditional mean of exchange rate changes. Both Hsieh (1989), who used daily exchange rate changes for five currencies between 1974-1983, and Kugler and Lenz (1990), with weekly data for ten currencies in the period 1979-1989, found nonlinear dependence in some series which was not of the pure ARCH or GARCH type by applying the correlation dimension-based independence test [the BDS-test] proposed by Brock, Dechert, and Scheinkman (1987). Engle and Hamilton (1990) have found evidence of regime shifts leading to long swings in exchange rates. Kräger and Kugler (1993) found statistically significant threshold affects for five currencies using weekly data over the last ten years, while Chiarella, Peat, and Stevenson (1994) detected nonlinearity of the bilinear (BL), EAR, and threshold varieties, in a study of eight European and Pacific Basin currencies using monthly and weekly data from 1978 to 1993, as well as daily observations between 1983 and 1992.

Many linear time-series models have been reported in the empirical finance literature to describe volatilities of stock returns [e.g. Poterba and Summers (1986), French, Schwert, and Stambaugh (1987), Schwert (1990), and Schwert and Seguin (1990)]. Features of volatility series point to linear time series models not being the best descriptors. These include empirical evidence showing stock returns exhibiting clusters of outliers, thus implying evolution of the series over time in a nonlinear fashion. The ARCH model of Engle (1982), the GARCH model of Bollerslev (1986), and the EGARCH model of Nelson (1991), are the most commonly used nonlinear time series models which are used for modelling volatility series [see Bollerslev, Chou, and Kroner (1992) for a review of the
applications of these models]. However, models from the ARCH-family are only one of a number of time series models which can capture the nonlinear characteristics of stock volatilities. One of the goals of the Cao and Tsay (1992) study was to consider directly a volatility process associated with monthly stock returns, and to test for possible nonlinearity of the process. They used the general nonlinearity test of Tsay (1991). The purpose of this test was to detect the presence of nonlinear structure in the series by controlling for various types of stochastic nonlinearity. Their results indicated that the transformed volatility series studied were nonlinear, and that a threshold autoregression (TAR) model was likely to capture some nonlinear characteristics of the series.

A number of studies have reported asymmetry, or nonlinear behaviour, of industrial production time series. These studies include that of Ashley and Patterson (1989) who use a bispectral nonlinearity test to discover nonlinearity in monthly growth rates of the index of industrial production for the U.S. manufacturing sector, and Brock and Sayers (1988) who reject the null hypothesis of linearity with the BDS-test. Teräsvirta and Anderson (1991) fitted logistic (LSTAR) and exponential smooth transition autoregressive (ESTAR) models to industrial production indices from thirteen OECD countries. Evidence of possible asymmetry in the West German business cycle was found by Luukkonen and Teräsvirta (1991) when they tested quarterly, seasonally unadjusted logarithmic OECD data. Krager (1993) examined the net production index of the producing sector for the Federal Republic of Germany and found evidence in favour of threshold nonlinearity.

In the light of recent efforts by researchers to test for nonlinearity in economic and financial timeseries data, there is an increased need for tests capable of identifying different types of nonlinearity. In this paper we modify aspects of the general test of Tsay (1991) to form tests for BL, EAR, and STAR type nonlinearity. By including some of the existing tests, we propose a battery of tests suitable for detecting most of the nonlinear models considered in the literature.

The remaining part of the paper is organised as follows. In the next section we describe the battery of tests we propose for identifying various types of nonlinear structure for modelling time series. Following that, we examine the power of these tests to identify the
appropriate type of nonlinearity via a simulation study. Finally, we conclude by summarising our results as well as outlining future directions in our research.

2. TESTS FOR IDENTIFYING VARIOUS TYPES OF NONLINEAR STRUCTURE IN TIME SERIES DATA

Both the Tsay (1989) and (1991) tests are based on the idea of taking a threshold autoregressive (TAR) model and transforming it into a regular change-point regression problem. By regressing residuals from a recursive estimation procedure on groups of added variables ("nuisance parameters") which are characteristic of various nonlinear models, nonlinearity is identified via an F-test for significance of parameters from all the different variable groups.

To transform the TAR model into a change-point problem an arranged autoregression is used. Assuming a two regime TAR model, with a given delay value, d, the flow of data is arranged from the smallest to the largest on the basis of each delayed observation, $Y_{t-d}$. The result is an arranged data set such that $Y_t$ is shuffled to ensure that every sequence $Y_n, Y_{t+1}, \ldots, Y_{t+d}$ from the original sequence is arranged so that $Y_{t+d}$ is in ascending order. A recursive estimation procedure (known as "local estimation") follows which results in the sequential estimation of a linear model for a fixed-length, but moving window, across the arranged data set. An indication of threshold nonlinearity results if stability of each of the coefficients over part of the data set is then followed by a smooth transition to another stable set. The smooth transition results from the overlapping window used in the estimation procedure. A scatterplot of the arranged data would indicate the threshold value by a step in the data. In reality, this is not necessarily a clean step with all data obviously in one regime before the step, and in the other after. As such, using a scatterplot of the arranged data is only an approximate method of detecting the threshold value.

Assume we have a time series $Y_1, Y_2, \ldots, Y_T$. Tsay's (1991) test involves, firstly, calculating a series of standardised predictive residuals, $\{\tilde{e}_t\}$, from the arranged
autoregression (order p and delay d) of the series Y_t. As p observations are used in prior calculations in obtaining the predictive residuals, the remaining time series is the last T-p observations which we rename as Y_1, Y_2, ..., Y_n where n = T-p. To detect various kinds of nonlinearity, the predictive residuals, \{\hat{e}_i\}, are then regressed on the regressors (1, Y_{t-1}, ..., Y_{t-p}) plus other added variables (explained below) which indicate various forms of nonlinearity. If the associated residuals, \{\hat{e}_i\}, from this second regression are asymptotically white noise, then the (arranged) Y_t is a linear AR(p) process. If the F-statistic points to significance of the coefficients of the regressors (1, Y_{t-1}, ..., Y_{t-p}), then a threshold model is suggested. The procedure to this point is known as the Tsay (1989) test and is denoted as TAR-F. To control for smooth threshold autoregressive (STAR) type of nonlinearity, Tsay adds the variables \{G(Y_{t-d}), Y_{t-1}G(Y_{t-d})\} where Z_{t-d} = (Y_{t-d} - \bar{Y}_d)/S_d ,

with \bar{Y}_d the sample mean, and S_d the sample variance of Y_{t-d} , while G(\cdot) is the cumulative distribution function of the standard normal variable. To detect linearity of the bilinear (BL) type he adds the variables \{Y_{t-1}\hat{e}_{t-1}, ..., Y_{t-p}\hat{e}_{t-p}\} and \{\hat{e}_{t-1}\hat{e}_{t-2}, ..., \hat{e}_{t-p}\hat{e}_{t-p-1}\} to the right hand side of the second regression. For exponential autoregressive (EAR) type nonlinearity the added variable is \{Y_{t-1}\exp(-Y_{t-1}^2/\gamma)\} with \gamma a normalisation constant and usually equal to \max[Y_{t-1}]. For this second regression the test statistic is the usual F-statistic under the null hypothesis of an AR(p) model for Y_t. It follows the F-distribution with degrees of freedom [3(p+1), n-b-3(p+1)], where b equals the number of observations lost in the recursive estimation process.

This test is a general test for nonlinearity, it cannot provide information on the specific functional form of the nonlinearity involved. We are interested in generating a set of tests which have as their alternate a specific class of nonlinear model, rather than the general alternate of nonlinear.

To approach this problem we will outline a sequence of tests against four specific forms of
nonlinearity. The tests for BL and EAR nonlinearity are new variants of the Tsay (1991) general nonlinearity test but where the data is unarranged. They are LM tests using added variables characteristic of the class of nonlinearity being tested for. There are two tests for STAR type nonlinearity. One is an existing test while the other is a new test, again a variant of the Tsay (1991) general test but with the data unarranged. It relies on an F-statistic associated with an added variable (as used by Tsay) to capture this class of nonlinear behaviour. Existing tests for nonlinearity of the threshold and autoregressive conditional heteroscedasticity types are also included.

(1) Bilinear (BL):

The general bilinear model can be written as

\[ Y_t = \sum_{i=1}^{p} a_i Y_{t-i} + \sum_{j=1}^{q} b_j e_{t-j} + \sum_{k=1}^{r} \sum_{l=1}^{s} b_{kl} Y_{t-k} e_{t-l} + e_t. \]  

(1)

Suppose that we have an observed series \( Y_t, \ t = 1, \ldots, T \). In this case we wish to test the null hypothesis that in (1)

\[ H_0: \ b_{kl} = 0 \quad \text{against} \quad H_1: \ \text{at least one} \ b_{kl} \neq 0. \]

This test, denoted by \( S_{BL} \), uses standard regression methods and would be carried out as follows:

(1) Regress \( Y_t \) on \( \{1, Y_{t-1}, \ldots, Y_{t-p}\} \), recover the residuals \( e_t, \ (t = p+1, \ldots, T) \) and compute

\[ SSE_0 = \sum_{t=p+1}^{T} e_t^2. \]

(2) Regress \( e_t \) on \( \{1, Y_{t+i}, Y_{t+i} e_{t+i} \ i = 1, \ldots, p\} \) recover the residuals \( r_t \) and compute
\[ \text{SSE}_1 = \sum_{t=p+1}^{T} r_t^2. \]

(3) Compute the test statistic
\[ S_{BE} = T(\text{SSE}_0 - \text{SSE}_1)/\text{SSE}_0. \]
Under the null hypothesis this statistic is asymptotically distributed as \( \chi^2 \) with \( p \) degrees of freedom (Harvey 1990, p 172).

(2) Exponential Autoregression (EAR):

The general exponential autoregressive (EAR) model can be written as
\[ Y_t = \sum_{i=1}^{p} a_i Y_{t-i} + \sum_{j=1}^{p} b_j e^{-\gamma r_{t-1}} \cdot Y_{t-j} + e_t. \]
(2)

In this case to test for EAR nonlinearity we wish to test
\[ H_0: \quad b_j = 0 \quad \forall j \quad \text{against} \]
\[ H_a: \quad \text{at least one } b_j \neq 0. \]

This test which, we denote as \( S_{EA} \), is also conducted by standard LM methods.

(1) Regress \( Y_t \) on \( \{1, Y_{t-1}, \ldots, Y_{t-p}\} \), recover the residuals \( e_t \), \( (t=p+1, \ldots, T) \) and compute
\[ \text{SSE}_0 = \sum_{t=p+1}^{T} e_t^2. \]

(2) Regress \( e_t \) on \( \{1, Y_{t-i}, Y_{t-i} \exp(Y_{t-i}/\gamma), i = 1, \ldots, p\} \) where \( \gamma = \max|Y_{t-i}|. \) Recover the residuals \( r_t \) and compute
\[ \text{SSE}_1 = \sum_{t=p+1}^{T} r_t^2. \]
(3) Compute the test statistic

\[ S_{EA} = T(SSE_0 - SSE_1)/SSE_0. \]

Under the null hypothesis of linearity this statistic is distributed as \( \chi^2 \) with \( p \) degrees of freedom.

(3) Smooth Transition Autoregressive Models (STAR):

The general STAR model can be written as

\[ y_t = \sum_{i=1}^{p} a_i Y_{t-i} + \sum_{j=1}^{p} b_j Y_{t-j} F(z) + cF(z) + e_t \]

To test for this class of nonlinear model we propose two tests. The first is the Luukkonen, Saiikkonen and Terasvirta (1988) S3 statistic, which is a LM test based on a Taylor expansion of the general STAR model. The added variables in this case are cross products of the lagged variable and cubed terms, i.e. \( Y_{t-i} Y_{t-j} \) and \( Y_{t-i}^3 \).

The test procedure is:

(1) Regress \( Y_t \) on \( \{1, Y_{t+i}, i = 1, \ldots, p\} \), recover the residuals \( e_t \), \( (t=p+1, \ldots, T) \) and compute

\[ SSE_0 = \sum_{t=p+1}^{T} e_t^2. \]

(2) Regress the \( e_t \) on \( \{1, Y_{t+i}, Y_{t+i}, Y_{t+j}, Y_{t+j}^3, i = 1, \ldots, p, j = 1, \ldots, p\} \) recover the residuals \( r_t \) and calculate

\[ SSE_1 = \sum_{t=p+1}^{T} r_t^2. \]

(3) Compute the test statistic
\[ S3 = T(SSE_0 - SSE_1)/SSE_0. \]

This statistic is distributed \( \chi^2 \) with \( 1/2 \ p(p + 1) + p \) degrees of freedom.

A second test for STAR non-linearity uses added variables of the cumulative normal type. This test is conducted by standard LM methods.

This test procedure is:

1. Regress \( Y_t \) on \( \{1, Y_{t-1}, \ldots, Y_{t-p}\} \) recover the residuals \( e_t \) and compute
\[
SSE_0 = \sum_{t=p+1}^{T} e_t^2
\]

2. Regress the \( e_t \) on \( \{1, Y_{t-1}, G(z_{t-1}), Y_{t-1}.G(z_{t-1}), i=1, \ldots, p\} \)

where

\( G(.) \) is the cumulative distribution function of the standard normal random variable.

\[ z_{t-1} = (Y_{t-1} - \bar{Y}_t)/S_t \]

\( \bar{Y}_t \) is the sample mean of \( Y_{t-1} \)

\( S_t \) is the standard deviation of \( Y_{t-1} \)

Recover the residuals \( r_t \) and calculate
\[
SSE_1 = \sum_{t=p+1}^{T} r_t^2
\]

3. Compute the test statistic
\[ S_{CN} = T(SSE_0 - SSE_1) / SSE_0 \]
This statistic is distributed $\chi^2$ with (p+1) degrees of freedom.

(4) **Threshold Nonlinearity (SETAR):**

The general SETAR model can be written as

$$Y_t = \begin{cases} 
c_1 + \sum_{i} a_i P_{t-i} & Y_{t-d} \leq th \\
c_2 + \sum_{i} b_i P_{t-i} & Y_{t-d} > th 
\end{cases}$$

To test for this form of nonlinearity we propose the use of the Tsay (1989) threshold test, which we denote as the TAR-F test.

The test is carried out as follows.

(1) Fit an arranged autoregression of order p to the data, recover the standardised residuals $e_i$, ($t=b+p+1, ..., T$) and compute

$$SSE_0 = \sum_{t=b+p+1}^{T} e_i^2.$$  

(2) Regress the $e_i$ on $\{1, Y_{t_i}, i = 1, ..., p\}$ recover the errors $r_i$ and compute

$$SSE_1 = \sum_{t=b+p+1}^{T} r_i^2.$$  

In this step the first $b = T/10$ errors are dropped from the regression.

(3) Compute the statistic

$$F = \frac{(SSE_0 - SSE_1)/(p+1)}{SSE_1/(T-d-b-p-h)}$$

where $h = \max(1, p + 1 - d)$.  

11
This statistic is distributed as an F distribution with $p + 1, T - d - b - p - h$ degrees of freedom.

(5) **Autoregressive Conditional Heteroscedasticity (ARCH):**

The work of Weiss (1986) argued that bilinearity may be misspecified as ARCH. We include a standard ARCH test in our simulation work to evaluate the extent of this misspecification.

The test is carried out as follows:

1. Regress $Y_t$ on $\{1, Y_{t-1}, ..., Y_{t-p}\}$, recover the residuals $e_t, (t=p+1, ..., T)$ and compute

   $$SSE_0 = \sum_{t=p+1}^{T} e_t^2$$

2. Regress $e_t^2$ on $\{1, e_{t-1}^2, \hat{Y}_t^2, Y_{t-1}^2, i=1 ... p\}$

   where

   $$\hat{Y}_t$$ is the least squares estimate of $Y_t$ from step 1.

   Recover the residuals $r_t, (t=p+1, ..., T)$ and calculate

   $$SSE_1 = \sum_{t=p+1}^{T} r_t^2$$

3. Compute the test statistic

   $$S_{\text{ARCH}} = T(SSE_0 - SSE_1) / SSE_0$$

Under the null hypothesis of linearity this statistic is distributed as $\chi^2$ with $2p+1$
degrees of freedom.

Some notes on the procedures:

(i) The order of the models can be chosen using an information criterion like the normalised Akaike Information Criterion (AIC).

(ii) It is possible that more than one form of nonlinearity is found to be significant using this procedure. This is to be expected as most of the specific forms of nonlinearity that we are testing for can be written as special cases of the other forms. Where more than one form is identified, we suggest fitting all significant specifications and, using the AIC, select a preferred model.

(iii) Use the Kalman filter, or recursive least squares, to estimate the recursive residuals, \( \hat{\epsilon}_t \), in the threshold test.

3. COMPARISON OF TESTS FOR VARIOUS FORMS OF STOCHASTIC NONLINEARITY

We now apply the proposed tests to simulated data sets\(^1\); the aim being to see if the proposed tests can determine the appropriate form of nonlinearity. We are also interested in how the tests perform in the classification of alternative forms of nonlinearity.

All the results reported are based upon 1000 replications, with 100 observations. The autoregressive order of the models is selected by AIC, the maximum order allowed is four. For each realisation, 3100 data points are generated with zero starting points, and only the last 100 are used as observations. The error terms in all the models are normal (0, 1) random variables generated from the RNDNS function in GAUSS.

---

\(^1\) The nonlinear models used to simulate the data are those used in the Tsay (1991) paper. These models were chosen to enable comparisons to be made between our simulation results and those of Tsay (1991) when the nonlinearity test was the same.
Tables 1 to 4 give the empirical frequencies of rejecting a linear model when the generating mechanism is:

(1) bilinear,
(2) exponential autoregressive,
(3) smooth transition autoregressive, and
(4) threshold nonlinear.

The tests used in the simulation are the six tests outlined in the previous section.

From Tables 1 to 4 we can make the following observations concerning the ability of each of the tests used to detect different types of nonlinearity, as well as conclusions as to what model building strategy underpins the use of individual tests.

3.1 The Bilinear (BL) Test

This is a new test designed to test for BL behaviour in the data and, as such, offers no direct comparison with the results from the Tsay (1991) study.

From Table 1 (Row 2), Table 3 (Row 4), and Table 4 (Row 3), we note that the BL test has low power in detecting a linear model. From Table 1 we observe the test to have power in correctly identifying the BL model, while from Table 2 we further observe it to be a good test for not incorrectly picking an EAR model. The test did incorrectly pick one specification of the LSTAR model around 60% of the time (Row 3, Table 3), but clearly rejected this model in the other two cases.

Before analysing the results concerning the BL test for the SETAR models in Table 4, it is important to understand the linkage of BL models to SETAR models, and whether this linkage holds for each of the models given in Row 1, Row 2, and Row 4. Firstly, the BL model can be regarded as a "natural" nonlinear extension of the ARMA model. Brockett (1976) showed that the BL model can approximate arbitrarily a "well-behaved" Volterra series relationship over a finite period by suitable choice of parameters. With Volterra
series providing an important representation for nonlinear models [see Priestley (1988), Chapter 2], BL models can be seen as a nonlinear parallel to ARMA models which arbitrarily approximate any general linear relationship between output, \( \{Y_t\} \), and noise, \( \{e_t\} \). Secondly, the SETAR model is a piecewise linear approximation to a true nonlinear process [see Tong (1990)]. Priestley (1988) shows that there is a tenuous link between BL and SETAR models. He shows that the SETAR model can possibly be re-expressed as a BL model using input which is determined by a feedback mechanism related to lagged output plus noise. We now examine whether such a linkage exists for the SETAR models in Rows 1, 2, and 4 of Table 4.

Consider the SETAR model given in Row 2 of Table 4:

\[
Y_t = \begin{cases} 
2 + 0.5Y_{t-1} + e_t, & Y_{t-1} \leq 1.0 \\
0.5 - 0.4Y_{t-1} + e_t, & Y_{t-1} > 1.0 
\end{cases}
\]

For input, \( U_t \), output, \( Y_t \), and noise, \( e_t \), the above model can be re-expressed as the BL model:

\[
Y_t = 2 + 0.5Y_{t-1} - (1.5 + 0.9Y_{t-1})U_{t-1} + e_t,
\]

where \( U_{t-1} \) is determined by the feedback mechanism,

\[
U_{t-1} = \begin{cases} 
0, & Y_{t-1} \leq 1.0 \\
1, & Y_{t-1} > 1.0 
\end{cases}
\]

Similarly, the model in Row 4 of Table 4:

\[
Y_t = \begin{cases} 
0.5Y_{t-1} + e_t, & Y_{t-1} \leq 0 \\
-0.5Y_{t-1} + e_t, & Y_{t-1} > 0 
\end{cases}
\]

can be re-expressed by the BL model:

\[
Y_t = 0.5Y_{t-1}U_{t-1} + e_t,
\]

where \( U_{t-1} \) is determined by the feedback mechanism,
\[ U_{t-1} = \begin{cases} 
1, & Y_{t-1} \leq 0 \\
-1, & Y_{t-1} > 0 
\end{cases} \]

For the model in Row 4 of Table 4:

\[ Y_t = \begin{cases} 
1 - 0.5Y_{t-1} + e_t, & Y_{t-1} \leq 0 \\
-1 - 0.5Y_{t-1} + e_t, & Y_{t-1} > 0 
\end{cases} \]

there is no BL representation involving an input feedback mechanism like that of the other two. In other words, there is no representation for \( Y_t \) involving a term in \( U_t \) \( Y_{t-1} \) on the right hand side that will always realise \(-0.5Y_{t-1}\) for \( Y_{t-1} \leq 0 \) and \( Y_{t-1} > 0 \), yet result in a constant of the same value but different sign.

It follows from Table 4 that the BL test has low power in detecting a SETAR model that cannot be represented as a bilinear model (Row 1), but correctly picks SETAR models that can be represented as BL models (Row 2 and Row 4).

In conclusion, to account for the possibility that a BL model can be represented as a SETAR model, if the BL test suggests a bilinear model then, according to the Ramsey and Chesher (1976) modelling strategy, both BL and SETAR models should be fitted.

3.2 The Exponential Autoregressive (EAR) Test

This is a new test which does not allow comparison with Tsay's (1991) study.

The test has low power in detecting a linear model (see Tables 1, 3, and 4) which indicates reasonable Type 1 error. As well, it exhibits low power in detecting a BL model.

It seems an excellent test for indicating EAR nonlinearity. There is some minor loss of power as the \( \beta \) coefficient is increased significantly.
This test incorrectly classified LSTAR nonlinearity. However, this may not be strange when one considers that the weighting function in such models is based on the exponential function which is also generic to the EAR generating process.

Except for the case where the only difference in the SETAR model over the regimes was in the constants (Row 1, Table 4), the EAR test displayed low power in detecting SETAR nonlinearity.

While EAR models behave somewhat similarly to SETAR models, the smooth change of the coefficients of the EAR model relative to the discrete changes of the SETAR implies that the EAR model is more closely related to the STAR class. Consequently following the model selection criterion of Ramsey and Cheshir (1976), if the EAR test suggests such a model it is advisable to also fit an LSTAR model and perhaps other models from the STAR class.

3.3 The S3 Test

This test was proposed by Luukkonen et al. (1988) to identify nonlinear models from the STAR class. It was incorporated into the Tsay (1991) study and denoted as AUG-F.

The S3 test has reasonable Type 1 error in that it has low power in identifying a linear model. This is evident from Tables 1, 3, and 4. This is consistent with what Tsay (1991) found in his study. This test is good at identifying BL nonlinearity, while poor at detecting EAR behaviour. Again, both these results are consistent with Tsay's (1991) findings.

Not surprisingly, the S3 test has power at detecting simulated data from an LSTAR model; a result corroborated by Tsay's (1991) study. Also, it is relatively powerful at detecting models from the SETAR class. We found the test had power for identifying threshold nonlinearity for simulated data from all three SETAR models (Rows 1, 2, and 4) in Table 4, compared to Tsay (1991) who found a similar result for data simulated from the models in Rows 1 and 2, but not 4.
With power for identifying BL, STAR, and SETAR type nonlinearity, we regard the S3 test more of a general test of stochastic nonlinearity. While the test can indicate a model from the STAR class, however it is advisable to fit BL and SETAR models as well. The relationship that exists between the STAR and SETAR classes of models, whereby the latter imposes a discrete jump between regimes for data points within a time series rather than a smooth transition between regimes via an appropriate ramping function in the case of the former, would account for the power of this test to detect nonlinearity of these types.

3.4 The $S_{CN}$ Test

This test is new and can be seen as a variant of the Tsay (1991) test, however it offers no means of comparison with results from that study.

Like the previous tests it has reasonable Type 1 error with low power of identifying a linear model when the data is generated from a nonlinear process.

It has power at detecting BL, LSTAR, and SETAR nonlinearity, while for relatively low values of $\beta$ has some power of identifying EAR type nonlinearity. We conclude that with the exception of perhaps EAR models, it is a good test for detecting other classes of nonlinearity. As such we regard it as a general test of stochastic nonlinearity.

3.5 The TAR-F Test

This test is a threshold test and was proposed by Tsay (1989).

As for all the previous tests it has low power of detecting a linear model and, therefore, has Type 1 errors within the significance bounds.

Consistent with the Tsay (1991) study, this test is powerful at detecting BL and SETAR types of nonlinearity. For the latter type of nonlinearity, both Tsay's (1991) study and that being reported here found evidence of this threshold test being able to correctly classify
that type in two of the three SETAR models in Table 4.

The TAR-F test has low power against EAR nonlinearity as with two of the three LSTAR models in Table 3. In all respects the results of this study coincide with those of Tsay (1991).

In conclusion, while recognising the TAR-F test is essentially a threshold test, we recommend that as well as fitting a SETAR model it is advisable to consider models from the BL and STAR classes.

3.6 The $S_{CH}$ Test

This test was designed to detect autoregressive conditional heteroscedasticity or ARCH effects. This type of test was not included in Tsay's (1991) study and, therefore, we have no basis of comparison with Tsay's (1991) study and the results reported here in Tables 1 to 4.

The test exhibits low power in detecting a linear model, with Type 1 errors within the significance bounds.

From Table 1 we observe that the $S_{CH}$ test is a very good test at identifying BL type nonlinearity. While we have discovered ARCH effects in the simulated data we know that the stochastic process generating such data is BL and not from an underlying ARCH model. This result has importance given the numerous studies of stock market efficiency which have pointed to the intertemporal dependence in the variance of stock market returns and the need to model them using ARCH and GARCH models [see Bollerslev and Hodrick (1992)]. If the underlying stochastic process is not ARCH but a simple BL, then this may explain the inability of the ARCH - type models to provide superior univariate forecasts of returns than those from simple ARMA alternatives. Identification of BL nonlinearity with this test is not entirely surprising, as it seems to confirm the empirical findings of Rothman (1991) and DeGooijer (1989) who suggest that the underlying stochastic process for stock returns is best modelled by a BL model.
The $S_{CH}$ test also appears to demonstrate power at detecting EAR nonlinearity (see Table 2). This result lends credibility to the Chiarella, Peat, and Stevenson (1993) findings that the underlying process for exchange rate changes was nonlinear dependence of the BL, EAR, and threshold varieties, rather than the much touted ARCH or GARCH types.

This test displays low power of detecting LSTAR and SETAR nonlinearity.

Our conclusion regarding results from using this test is that if a model from the autoregressive conditional heteroscedastic class is indicated, then models from the BL and EAR classes should also be considered.

5. CONCLUSIONS

In this paper we have discussed modifications to the Tsay (1991) general test for identifying nonlinearities of the BL, EAR, and STAR varieties. The major motivation behind our modifications is to refine the ability to test for BL and EAR nonlinearity in particular. By simulating data from BL and EAR models, and then applying the tests knowing what type of nonlinearity is present in the data, we find our modified tests perform creditably.

We also included existing tests for nonlinearity of the threshold variety along with a autoregressive conditional heteroscedacity test, to form a battery of tests. We generated data used in our simulation study from the same data generating models used in the Tsay (1991) study. For the test used in both that study and the one that is the focus of this paper, we find extremely consistent results across both studies. Concerning the common tests, our conclusions obviously coincide with those of Tsay (1991). Of further interest was the confirmation that ARCH effects may not be signalling that the underlying model is from that class, but rather from either the BL or EAR classes.
REFERENCES


Harvey, A. (1990), The Econometric Analysis of Time Series, LSE Press.


Table 1: Bilinear Models

Empirical frequencies of rejecting a linear model based on 5% and 10% critical values. The generating models are:

(a) \[ Y_t = 0.5Y_{t-1} + \beta Y_{t-1}e_{t-1} + e_t. \]

(b) \[ Y_t = e_t + 0.5e_{t-1} + \beta e_{t-1}^2. \]

<table>
<thead>
<tr>
<th>Model</th>
<th>$\beta$</th>
<th>Cri Val</th>
<th>$S_{BL}$</th>
<th>$S_{EA}$</th>
<th>S3</th>
<th>$S_{CN}$</th>
<th>TAR-F</th>
<th>$S_{CH}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>-0.6</td>
<td>5%</td>
<td>670</td>
<td>384</td>
<td>936</td>
<td>947</td>
<td>921</td>
<td>872</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10%</td>
<td>740</td>
<td>483</td>
<td>958</td>
<td>974</td>
<td>951</td>
<td>908</td>
</tr>
<tr>
<td>a</td>
<td>0.0</td>
<td>5%</td>
<td>39</td>
<td>48</td>
<td>29</td>
<td>27</td>
<td>33</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10%</td>
<td>82</td>
<td>96</td>
<td>56</td>
<td>70</td>
<td>62</td>
<td>73</td>
</tr>
<tr>
<td>a</td>
<td>0.6</td>
<td>5%</td>
<td>660</td>
<td>412</td>
<td>934</td>
<td>945</td>
<td>729</td>
<td>868</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10%</td>
<td>747</td>
<td>480</td>
<td>949</td>
<td>973</td>
<td>810</td>
<td>908</td>
</tr>
<tr>
<td>b</td>
<td>-0.6</td>
<td>5%</td>
<td>566</td>
<td>131</td>
<td>919</td>
<td>928</td>
<td>800</td>
<td>558</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10%</td>
<td>664</td>
<td>211</td>
<td>955</td>
<td>963</td>
<td>868</td>
<td>659</td>
</tr>
</tbody>
</table>

Table 2: Exponential Autoregressive Models

Empirical frequencies of rejecting a linear model based on 5% and 10% critical values. The generating models are:

\[ Y_t = [\phi + \beta \exp(-Y_{t-1})]Y_{t-1} + e_t. \]

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>$\beta$</th>
<th>Cri Val</th>
<th>$S_{BL}$</th>
<th>$S_{EA}$</th>
<th>S3</th>
<th>$S_{CN}$</th>
<th>TAR-F</th>
<th>$S_{CH}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>10</td>
<td>5%</td>
<td>64</td>
<td>1000</td>
<td>241</td>
<td>553</td>
<td>226</td>
<td>897</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10%</td>
<td>135</td>
<td>1000</td>
<td>371</td>
<td>712</td>
<td>344</td>
<td>955</td>
</tr>
<tr>
<td>0.3</td>
<td>20</td>
<td>5%</td>
<td>78</td>
<td>998</td>
<td>294</td>
<td>207</td>
<td>180</td>
<td>868</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10%</td>
<td>122</td>
<td>998</td>
<td>429</td>
<td>267</td>
<td>242</td>
<td>949</td>
</tr>
<tr>
<td>0.3</td>
<td>100</td>
<td>5%</td>
<td>97</td>
<td>723</td>
<td>209</td>
<td>164</td>
<td>224</td>
<td>423</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10%</td>
<td>134</td>
<td>870</td>
<td>292</td>
<td>222</td>
<td>280</td>
<td>565</td>
</tr>
</tbody>
</table>
Table 3: Logistic Star Models

Empirical frequencies of rejecting a linear model based on 5% and 10% critical values. The generating models are:

\[ Y_t = 1.0 - 0.5Y_{t-1} + (\beta_0 + \beta_1 Y_{t-1})G(\alpha Y_{t-1}) + \epsilon_t, \]

with \[ G(z) = \frac{\exp(z)}{1 + \exp(z)}. \]

<table>
<thead>
<tr>
<th>( \beta_0 )</th>
<th>( \beta_1 )</th>
<th>( \alpha )</th>
<th>Cri Val</th>
<th>( S_{BL} )</th>
<th>( S_{EA} )</th>
<th>S3</th>
<th>S_C</th>
<th>T_AR-F</th>
<th>S_CH</th>
</tr>
</thead>
<tbody>
<tr>
<td>-4.0</td>
<td>-0.4</td>
<td>2.0</td>
<td>5%</td>
<td>325</td>
<td>870</td>
<td>908</td>
<td>976</td>
<td>345</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10%</td>
<td>469</td>
<td>938</td>
<td>954</td>
<td>989</td>
<td>459</td>
<td>212</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0</td>
<td>2.0</td>
<td>5%</td>
<td>50</td>
<td>826</td>
<td>536</td>
<td>656</td>
<td>169</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10%</td>
<td>119</td>
<td>901</td>
<td>685</td>
<td>783</td>
<td>290</td>
<td>117</td>
</tr>
<tr>
<td>2.0</td>
<td>-0.4</td>
<td>2.0</td>
<td>5%</td>
<td>555</td>
<td>811</td>
<td>665</td>
<td>744</td>
<td>744</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10%</td>
<td>667</td>
<td>879</td>
<td>780</td>
<td>831</td>
<td>829</td>
<td>197</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>2.0</td>
<td>5%</td>
<td>46</td>
<td>46</td>
<td>51</td>
<td>55</td>
<td>47</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10%</td>
<td>97</td>
<td>98</td>
<td>96</td>
<td>101</td>
<td>84</td>
<td>78</td>
</tr>
</tbody>
</table>
Table 4: SETAR Models

Empirical frequencies of rejecting a linear model based on 5% and 10% critical values. The generalising models are:

\[ Y_t = \begin{cases} 
\phi_0 + \phi_1 Y_{t-1} + e_t & \text{if } Y_{t-1} \leq W \\
\beta_0 + \beta_1 Y_{t-1} + e_t & \text{if } Y_{t-1} > W
\end{cases} \]

<table>
<thead>
<tr>
<th>( \phi_0 )</th>
<th>( \phi_1 )</th>
<th>( \beta_0 )</th>
<th>( \beta_1 )</th>
<th>W</th>
<th>Cri Val</th>
<th>( S_{BL} )</th>
<th>( S_{EA} )</th>
<th>S3</th>
<th>( S_{CN} )</th>
<th>TAR-F</th>
<th>( S_{CN} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>-0.5</td>
<td>-1.0</td>
<td>-0.5</td>
<td>0.0</td>
<td>5%</td>
<td>10%</td>
<td>58</td>
<td>953</td>
<td>590</td>
<td>724</td>
<td>100</td>
</tr>
<tr>
<td>2.0</td>
<td>0.5</td>
<td>0.5</td>
<td>-0.4</td>
<td>1.0</td>
<td>5%</td>
<td>10%</td>
<td>767</td>
<td>469</td>
<td>971</td>
<td>997</td>
<td>968</td>
</tr>
<tr>
<td>0.0</td>
<td>0.5</td>
<td>0.0</td>
<td>0.5</td>
<td>0.0</td>
<td>5%</td>
<td>10%</td>
<td>40</td>
<td>49</td>
<td>42</td>
<td>40</td>
<td>34</td>
</tr>
<tr>
<td>0.0</td>
<td>0.5</td>
<td>0.0</td>
<td>-0.5</td>
<td>0.0</td>
<td>5%</td>
<td>10%</td>
<td>553</td>
<td>199</td>
<td>584</td>
<td>673</td>
<td>558</td>
</tr>
</tbody>
</table>

27