European Regional Science Association
41st Congress, Zagreb, 29 August – 1 September 2001

# Testing for Non-Linear Dependence in Univariate Time Series

## An Empirical Investigation of the Austrian Unemployment Rate

Wolfgang Koller
Research Institute for European Affairs
Vienna University of Economics and Business Administration
Email: wolfgang.koller@wu-wien.ac.at

Manfred M. Fischer
Department of Economic Geography and Geoinformatics
Vienna University of Economics and Business Administration
Email: manfred.fischer@wu-wien.ac.at

**Abstract**

In recent years interest has been growing in testing for stochastic non-linearity in macroeconomic time series. There are several inference procedures available. But not much is known about their behaviour on real world small-sized settings. This paper surveys some of these tests. Their performance is compared using monthly Austrian unemployment data that cover the period January 1960 to December 1997. It is found that the test procedures surveyed are complementary rather than competing. Several useful guidelines are provided for applying the increasingly complex test procedures in practice.

# 1 Introduction

Univariate time series modelling is a major interest of (regional) economists. These models have the advantage that behavioural patterns can be predicted simply by analysing the past history of a variable, reflecting these patterns. The most important aspect of building such a model is learning about the intrinsic time patterns of a variable or its underlying generating process.

Model identification is the key to model building and the most difficult stage in the iterative identification–estimation–diagnosis model building strategy. Historically, only a few time series tests have been available to assist in this respect and these were mostly tests for linear models. In recent years there has been a growing interest of economists in non-linear models, including autoregressive conditional heteroscedastic [ARCH] models, threshold autoregressive [TAR] models and bilinear models.

Discussion of non-linearity is made more complicated by the usual problems with macroeconomic time series. The data are usually discrete in time, contain often high levels of measurement with unknown properties, are affected by temporal aggregation and may have been filtered to remove a seasonal component. Many macroeconomic time series also have long memory properties, including deterministic and stochastic trends, which also have to be considered in the model identification stage.

Non-linearity is the necessary condition for non-linear modelling. Today several non-linear identification tests are available (for example, see Cromwell, Labys, and Terraza [1994]). Although the existence of these tests enables us to model univariate time series more adequately, a larger task meets us in attempting to decide which of several test procedures for detecting non-linearities, in fact, should be employed. The purpose of this contribution is to provide some practical guidelines in applying these increasingly complex test procedures in order to identify univariate time series models. Our goal is to survey some of these tests proposed in the time series literature and to apply and compare them on a real macroeconomic time series. The comparison is based on monthly unemployment rates in Austria[1]. The monthly sample is from January 1960 to December 1997. There are 456 observations in the data set.

The section that follows gives some background information by introducing some definitions and the inference methods selected to test for stochastic non-linearity. The tests are described in greater detail then in section 3. Section 4 follows with the presentation of empirical results. Some conclusions are drawn in the final section.

# 2 Background

## 2.1 Some basic definitions

The definition of a time series begins with the notion of a stochastic process that is defined as an ordered set of random variables indexed by time: $x_1, x_2, \ldots, x_T$. An observed time series is a realisation of some underlying stochastic process. In this sense, the relationship between realisation and process in time series analysis is analogous to the relationship between sample and population in cross-section analysis. A realisation [or time series] is used to build a model of the process that generated the series.

A most important property of a time series is that it be stationary. A time series is said to be stationary in the wide sense, or second-order stationary, when the mean, variance and covariance of the process are constant, but not higher order moments. Because most real world data are non-stationary, performing transformations of the data and testing for stationarity should conincide. The most common transformation is that of differencing, that is, subtracting a past value of a variable from its current value. If $x_t$ is a zero mean third-order stationary time series, then the mean $\mu = E(x_t) = 0$, the second order covariance $c_{xx}(q) = E(x_{t+q}x_t)$ and the third order covariances $c_{xxx}(q,r) = E(x_{t+q}x_{t+r}, x_t)$ are independent of $t$. If $c_{xx}(q) = 0$ for all non-zero $q$, the series is white noise. A white noise series in which $x_1, x_2, \ldots, x_T$ are independent random variables. Gaussian white noise series are necessarily pure white noise series.

In addition to stationarity, whiteness and pure whiteness, another often assumed characteristic of a time series is linearity, which may be defined in different ways. We conform here to the more conventional definition that a linear stochastic process is a linear filter of independent and identically distributed (iid) inputs. For example, an autoregressive moving average [ARMA] process is a finite order linear process.

Non-linearity may appear in different forms. Additive non-linear dependence arises through persistence in conditional mean of the process. Examples of such processes are the threshold autoregressive [TAR] models (see Tong and Lim [1980]), the exponential autoregressive models (see Ozaki [1980]) and the bilinear models (see Granger and Andersen [1978]). Multiplicative non-linear dependence arises when the source of non-linearity is in the variance of the process. Examples of such processes are the ARCH models (see Engle [1982]) and the generalizations of ARCH models (for example Bollerslev [1986], Engle, Lilien, and Robins [1987], and Sentana [1995]).

Non-linear stochastic models have the potential of improving forecasts and thus of providing stronger candidates against which to compare models found by a less purely statistical research strategy. But it is important to note that only in the case of additive non-linear dependence non-linear models can be utilized to generate improved point predictions, while in the case of multiplicative non-linear dependence they can be exploited to construct superior prediction intervals.

## 2.2 The tests selected

We use four inference methods to test for stochastic non-linearity with the Austrian unemployment rate data: a test originally proposed by Brock, Dechert, and Scheinkman [1987] and described to more detail in Brock et al. [1996], henceforth BDS test, a test introduced by McLeod and Li [1983], the socalled McLeod-Li test, a test developed by Hsieh [1989], the socalled Hsieh test, and a test suggested by Teräsvirta, Lin, and Granger [1993], henceforth the Teräsvirta-Lin-Granger test.

The BDS test does not provide a direct, but an indirect test for non-linearity. It is a test for independence and can be used to test for residual non-linear structure, after linear structure has been removed from the data through prewhitening. The BDS test produces a viable test of linearity against the omnibus alternative of non-linearity when the data are prefiltered by ARMA fit (Barnett et al. [1997]). We use the test for this purpose.

Also the McLeod-Li test is an indirect test and based on the fact that by fitting a linear model to the data, the inherent non-linearity has been swept into the residuals. While the BDS test makes use of the concept of correlation integral, the McLeod-Li test applies a standard Ljung-Box-Pierce Portmanteau test for serial correlation to the squared residuals from ARMA representation. The test is sensitive against multiplicative, less so against additive non-linearity.

Once it is established that some type of non-linearity exists, the Hsieh-test discriminates between additive and multiplicative non-linearity. The test requires to set up multiplicative non-linearity as the null hypothesis which implies that the third order correlation coefficients equal zero. The correlation is based on the residuals from a linear specification. Once again, the test is based on the assumption that the inherent non-linearity has been swept into the residuals.

The Teräsvirta-Lin-Granger test has the useful property that if the null hypothesis of linearity is rejected it will provide a non-linear model specification that is potentially relevant for forecasting. This non-linear model produced by the test should, however, not be accepted as the true model but only as a useful approximation. The question of how to form better approximations is still very much an open question. In the next section the tests will be described in some more detail.

# 3 Test descriptions

## 3.1 The Brock-Dechert-Scheinkman [BDS] test

The test developed by Brock, Dechert, and Scheinkman [1987] has been applied in macroeconomic time series modelling (Brock and Sayers [1988] and Peel and Speight [1998]) and elsewhere (see, for example, Craig, Kohlhase, and Papell [1991]). It is a test that examines the underlying probability structure of a time series searching for any kind of dependence. It was inspired by the Grassberger-Pocaccia correlation integral (see Grassberger and Procaccia [1983]), but it is a test for any kind of structure in a series, linear stochastic, non-linear stochastic, or deterministic chaos.

It is set up as follows. Let $u_t$ be a sequence of residuals of length $T$. Define the embedded subvector as

$$u_t^m = (u_t, \dots, u_{t-m+1}), t = 1, 2, \dots, T - m + 1 \tag{1}$$

The choice of the embedding dimension $m$ for the dimensionality of the vector is subjective. But note that $m-1$ data points are lost because it is required that all vectors have equal length.

The dependence of $u_t$ is analysed by means of the concept of the correlation integral, a measure that examines the distances between points, say $u_t^m$ and $u_s^m$, in the above $m$-dimensional space. For each embedding dimension $m$ and choice of the metric bound $\varepsilon$ the correlation integral $C(\varepsilon,m,T)$ is defined as the fraction of close points $(u_t^m, u_s^m)$:

$$C(\varepsilon,m,T) = \frac{2}{T_m(T_m-1)} \sum_{t<s} I_\varepsilon(u_t^m, u_s^m) \tag{2}$$

where $T_m = T - m + 1$, $t$ and $s$ range from 1 to $T-m+1$ in the summation and are restricted such that $t < s$. $I_\varepsilon(u_t^m, u_s^m)$ is an indicator function which equals 1 if $\|u_t^m - u_s^m\| < \varepsilon$, where $\|.\|$ is the sup norm over the subvector. The sup norm is given by $\|u\| = max_{1<i\leq m}|u_i|$.

Brock, Dechert, and Scheinkman [1987] show convergence in distribution for statistics of the form

$$C(\varepsilon,m,T) - C(\varepsilon,1,T)^m \quad \text{lim dist} \quad N(0,\sigma^2(\varepsilon,m)) \tag{3}$$

with

$$\sigma^2(\varepsilon,m) = 4[4K^m + 2\sum_{j=1}^{m-1} K^{m-j}C^{2j} + (m-1)^2C^{2m} - m^2KC^{2m-2}]. \tag{4}$$

$C$ and $K$ can be consistently estimated by $C(\varepsilon,1,T)$ and

$$K(\varepsilon,T) = \frac{1}{T_m(T_m-1)(T_m-2)} \sum_{t\neq s, t\neq r, r\neq s} I_\varepsilon(u_t^m, u_s^m) I_\varepsilon(u_s^m, u_r^m) \tag{5}$$

Thus, they suggest as the test statistic

$$BDS(\varepsilon,m,T) = \frac{T_m^{\frac{1}{2}}(C(\varepsilon,m,T) - C(\varepsilon,1,T)^m)}{\sigma(\varepsilon,m)} \tag{6}$$

for some selected $m$ and $\varepsilon$. The statistic is divided by the asymptotic standard deviation so it is distributed asymptotic normal with mean 0 and variance 1 under the null of independent, identically distributed $u_t$'s.

The BDS statistic is a function of two arguments: the embedding dimension $m$ and the metric bound $\varepsilon$ (i.e., the maximum difference between pairs of observations counted in computing the correlation integral). The values of the two arguments are finite and arbitrary in the definition of the test statistic. But an important relation exists between the two and the small sample properties of the statistic. For a given $m$, $\varepsilon$ can not be too small because $C(\varepsilon,m,T)$ will capture too few points, nor should $\varepsilon$ be too large in order to prevent $C(\varepsilon,m,T)$ from involving too many data points (Cromwell, Labys, and Terraza [1994]). In practice, $m$ is typically chosen over the range of 2 to 15, and $\varepsilon$ to lie between 0.5 and 2 standard deviations of the time series to be tested. Care must be undertaken in interpretation because tests performed for different values of $m$ and $\varepsilon$ may give contradictory information.

Under the null hypothesis, Brock, Dechert, and Scheinkman [1987] show that for $T$ large, $BDS(\varepsilon,m,T)$ will be normally distributed with mean 0 and a variance that is a complicated function of $m$ and $\varepsilon$. The null hypothesis of independence is rejected if $BDS(\varepsilon,m,T)$ is large. The definition of large should depend on the sample size. The test procedure can be performed in four steps:

(i) Choose a grid of values for $m$ and $\varepsilon$,

(ii) Compute $BDS(\varepsilon, m, T)$ as given by (6),

(iii) For a selected significance level $\alpha$, the critical values to test the null hypothesis of independence are based on the number of observations divided by the selected embedded dimension $m$:

- if $(T - m + 1)/m > 200$ use the standard normal distribution for the critical value $\tau$,
- if $(T - m + 1)/m \leq 200$ use the critical value $\tau$ from tables in Brock, Hsieh, and LeBaron [1991].

(iv) Reject the null hypothesis of independence if $|BDS(\varepsilon, m, T)| > \tau$

The BDS test does not provide a direct test for non-linearity, because the distribution of the test statistic is not known, either in finite samples or asymptotically, under a null hypothesis of non-linearity. The asymptotic distribution is known under the null of independence. Thus, the hypothesis of non-linearity is nested within the alternative hypothesis that includes both linear and non-linear processes. If all linear possibilities have been removed by fitting the best possible linear model, the BDS test can be utilized to test the residuals for remaining non-linear dependence.

If the null hypothesis is rejected then the alternative hypothesis implies the existence of non-linear dependence, and one can proceed to employ other tests to resolve the question if this non-linear dependence is of additive or multiplicative kind or a mixture of both kinds.

## 3.2 The McLeod-Li test

It was noted in Granger and Andersen [1978] that for a linear stationary process

$$\mathrm{corr}(x_t^2, x_{t-k}^2) = (\mathrm{corr}(x_t, x_{t-k}))^2 \text{ for all } k \tag{7}$$

and, thus, departures from this would indicate non-linearity. The McLeod-Li test is a Box-Ljung-Pierce Portmanteau test for non-linear dependence (McLeod and Li [1983]) that is conducted by examining the Ljung-Box-Pierce statistic of the squared residuals from an ARMA representation[2]. The test procedure can be conducted in four steps:

(i) Select lag length $k$ based on the sample frequency and estimate the autocorrelation function of $u_t^2$:

$$r_{uu}(k) = \frac{\sum_t (u_t^2 - \sigma^2)(u_{t+k}^2 - \sigma^2)}{\sum_t (u_t^2 - \sigma^2)^2} \tag{8}$$

where $\sigma^2$ is the variance of $u_t^2$.

(ii) Compute the Ljung-Box-Pierce (Box and Jenkins [1970]) statistic for the first $k$ autocorrelations of $u_t^2$ to test the null hypothesis of independence:

$$Q_{uu}(k) = T(T+2)\sum_{i=1}^{k}\frac{1}{(T-i)}r_{uu}^2(i) \tag{9}$$

(iii) For a selected significance level $\alpha$, find the critical value $\tau$ for testing the null hypothesis using the chi-square distribution, with $k$ degrees of freedom.

(iv) Reject the null hypothesis of linear dependence if $Q_{uu}(k) > \tau$.

Rejection of the null hypothesis indicates the existence of non-linear dependence. The main disadvantage of the test procedure is that the selection of $k$ is entirely based on the researcher's knowledge of the memory of the process, that is, the correlation between the current and previous periods. Besides this problem of lag selection the test examines the presence of serial correlation of $u_t^2$ only under the alternative. This depends on the assumption that the data are distributed normally and are stationary. If one or both of these assumptions are incorrect, then the power of the test decreases.

The Mc-Leod-Li test is not a direct test for either multiplicative or additive non-linearity, since the distribution of the test statistic is not known – either in finite samples or asymptotically – under a null hypothesis of multiplicative or additive non-linearity. The asymptotic distribution is known under the hypothesis of linear dependence. The hypotheses of multiplicative and additive non-linearity are nested within the alternative hypothesis.

In conventional statistical methodology, one tests a hypothesis by equating it with the null hypothesis or with the total alternative hypothesis, not by using the power of the test to try to discriminate between subsets of the alternative hypothesis. Monte Carlo evidence (see Lee, White, and Granger [1993]) shows that the McLeod-Li test has a strong power against multiplicative non-linearity, but less power against other forms of non-linearity. Thus under the above non-standard approach the McLeod-Li test may be expected to provide evidence for the presence of non-linearity in conditional variance[3].

## 3.3 The Hsieh test

Hsieh [1989] proposed a test to discriminate between additive and multiplicative non-linearity. Let $u_t$ be a vector of residuals, $u_t$, of a linearly filtered series $x_t$. Multiplicative non-linearity implies that the conditional expectation of the residuals given past lags of the variable, $x_t$, and the residuals, $u_t$, is zero:

$$\mathrm{E}(u_t|x_{t-1},\ldots,x_{t-k},u_{t-1},\ldots,u_{t-k}) = 0 \tag{10}$$

Additive non-linearity implies that the same conditional expectation is non-zero.

The test requires to set up multiplicative non-linearity as the null hypothesis. This implies that the third-order correlation coefficient, $\rho_{uuu}(r,s)$, equals 0 for all $r,s > 0$. The test is implemented as follows:

(i) Define the third-order moments of the residuals

$$\rho_{uuu}(r,s) = \mathrm{E}(\frac{u_t u_{t-r} u_{t-s}}{\sigma_u^3}) \tag{11}$$

and estimate $\rho_{uuu}(r,s)$ by the statistic

$$r_{uuu}(r,s) = \frac{(1/T)\sum_t u_t u_{t-r} u_{t-s}}{[(1/T)\sum_t u_t^2]^{1.5}} \tag{12}$$

(ii) In order to test the null hypothesis that $u_t$ posesses a multiplicative non-linearity compute the test statistic $H(r,s)$ given by

$$H(r,s) = \frac{\sqrt{T} r_{uuu}(r,s)}{\sqrt{V(r,s)}} \tag{13}$$

with

$$V(r,s) = \frac{(1/T)\sum_t u_t^2 u_{t-r}^2 u_{t-s}^2}{[(1/T)\sum_t u_t^2]^3} \tag{14}$$

Hsieh [1989] has shown that – with the null and other auxiliary assumptions derived from the central limit theorems for martingale differences – $\sqrt{T} r_{uuu}(r,s)$ is asymptotically normally distributed with zero mean and variance consistently estimated by (14). Thus, $H(r,s)$ follows a standard normal distribution with zero mean and variance one.

(iii) For a selected significance level $\alpha$, find the critical value $\tau$ for testing the null hypothesis of multiplicative non-linearity, using the standard normal distribution.

(iv) Reject the null hypothesis, if $|H(r,s)| > \tau$.

If the null hypothesis of multiplicative non-linearity is rejected then the alternative hypothesis implies the existence of additive non-linearity. The test has the disadvantage that several lags have to be tested and that the choice of lags $r$ and $s$ for the test statistic is ambiguous. The test is evaluated for a grid of values of $r$ and $s$, and one looks to the majority of the test results to sopport an outcome (Cromwell, Labys, and Terraza 1994).

## 3.4 The Teräsvirta-Lin-Granger test

The test suggested by Teräsvirta, Lin, and Granger [1993] is an indirect test for non-linearity.[4] The authors consider the specific non-linear model

$$x_t = \pi' w_t + \phi(\gamma' w_t) + u_t \tag{15}$$

where $w_t = (1, x_{t-1}, \ldots, x_{t-p})'$ denotes a vector of dependent variables including a constant, $\pi' = (\pi_0, \pi_1, \ldots, \pi_p)$ and $\gamma' = (\gamma_0, \gamma_1, \ldots, \gamma_p)$ are parameter vectors, and $u_t \sim iid(0, \sigma^2)$ is an error term. Set $\phi(\gamma' w_t) = \theta \psi(\gamma' w_t)$ where

$$\psi(\gamma' w_t) = (1 + \exp(-\gamma' w_t))^{-1} - 1/2. \tag{16}$$

Then, (15) can be interpreted as a non-linear autoregressive model of order $p$ in which the intercept $\pi_0 + \theta\psi(\gamma' w_t)$ is time varying and changes smoothly. Note that $\gamma = 0$ leads to the linear model $x_t = \pi' w_t$. Thus, the null hypothesis is

$$H_0 : \gamma = 0. \tag{17}$$

Note that model (15) is not identified under the null (17), but under the alternative. This motivates Teräsvirta, Lin, and Granger [1993] to replace $\phi$ in (15) by a Taylor expansion around $\gamma = 0$ in order to derive an applicable test for (17).

Thus, model (15) becomes

$$x_t = \tilde{\pi}' w_t + \sum_{i=1}^{p} \sum_{j=i}^{p} \delta_{ij} x_{t-i} x_{t-j} + \sum_{i=1}^{p} \sum_{j=i}^{p} \sum_{k=j}^{p} \delta_{ijk} x_{t-i} x_{t-j} x_{t-k} + \tilde{u}_t \tag{18}$$

where $\delta_{ij} = d_{ij}\theta\gamma_i\gamma_j\gamma_0$ with $d_{ij} = 1/36$ if $i = j$ and $d_{ij} = 1/18$ otherwise, and $\delta_{ijk} = d_{ijk}\theta\gamma_i\gamma_j\gamma_k$ with $d_{ijk} = 1/36$ if $i = j = k$, $d_{ijk} = 1/18$ if $i = j$, $j = k$ or $i = k$, and $d_{ijk} = 1/6$ otherwise. The null hypothesis corresponding to (17) is

$$H_0' : \delta_{ij} = 0, \delta_{ijk} = 0 \qquad i = 1, \ldots, p; j = i, \ldots, p; k = j, \ldots, p. \tag{19}$$

The test procedure can be performed in the following steps:

(i) Select the order $p$ of the autoregressive process by a conventional selection criterion, regress $x_t$ on $1, x_{t-1}, \ldots, x_{t-p}$ and compute the residuals $\hat{u}_t$ and the sum of the squared residuals $SSR_0 = \sum_t \hat{u}_t^2$.

(ii) Regress $\hat{u}_t$ on $1, x_{t-1}, \ldots, x_{t-p}$ and $m$ auxiliary regressors corresponding to the non-linear terms in (18). Compute the residuals $\hat{v}_t$ and the sum of squared residuals $SSR = \sum_t \hat{v}_t^2$.

(iii) In order to test the null hypothesis of linearity $H_0'$ according to (19), compute the test statistic

$$TLG = \frac{(SSR_0 - SSR)/m}{SSR/(T - p - 1 - m)} \tag{20}$$

(iv) For a selected significance level $\alpha$ find the critical value $\tau$ for testing the null hypothesis using the $F$-distribution with $m$ and $T - p - 1 - m$ degrees of freedom.

(v) Reject the null hypothesis of linearity if $TLG > \tau$.

The test possesses the useful property that if the null hypothesis of linearity is rejected it will provide a non-linear model that is potentially useful for forecasting. Of course, this non-linear model produced by the test procedure should not be accepted as being the true model but only as a useful approximation. The question of how to construct better approximations is still an open question (see Granger [1991]).

Table 1: Summary statistics for Austrian unemployment rates, $x_t$ and $\Delta_{12}x_t$

| Time Series | $N$ | Mean | Variance | Minimum | Maximum |
|---|---|---|---|---|---|
| Raw data series, $x_t$ | 456 | 3.7253 | 4.5531 | 0.8 | 9.2 |
| Seasonal differenced series, $\Delta_{12}x_t$ | 444 | 0.0975 | 0.2411 | -1.9 | 1.7 |

# 4 Data and empirical results

## 4.1 Raw and differenced data series

In recent years, much attention has been paid to identifying the appropriate time series characteristics of various macroeconomic aggregates. An interesting example in this respect is the empirical investigation of unemployment. Empirical studies in this area typically rely on linear specifications, the standard approach involving the estimation of ARMA processes. This approach reflects the view that macroeconomic data may be adequately described by stable linear processes driven into recurrent oscillations by successive random shocks (Peel and Speight [1998]). But a number of theoretical models have been prepared that suggest that unemployment may be a non-linear process (see, for example, Burgess [1992]).

This motivates to analyse the above test procedures on unemployment data, using the Austrian definition of the unemployment rate. The monthly sample is from January 1960 to December 1997 (456 data points). Figure 1 provides a visual representation of the time series with the y-axis defined as the unemployment rate and the x-axis as the time index. From an eyeball inspection of the plotted series, it seems obvious that this series is non-stationary and seasonal. In fact, the series level appears to increase in annual steps in the 1980s and 1990s. The Augmented Dickey-Fuller, the Phillips-Perron, the Kwiatkowski-Phillips-Schmidt-Shin and the Dickey-Hasza-Fuller test confirm that the series is seasonally non-stationary. Seasonal differencing, i.e. applying the seasonal filter $\Delta_s x_t = x_t - x_{t-s}$, with $s = 12$, makes the series stationary. The seasonally differenced series is plotted in Figure 2. Table 1 presents summary statistics for both the original and the seasonally differenced series.

Our primary concern is to detect non-linear departure from a linear process. The BDS test, the McLeod-Li test and the Hsieh test – in contrast to the Teräsvirta-Lin-Granger test – require the extraction of linear structure by the use of an estimated filter. Typically, an AR($p$) model is fitted to the series and the test then applied to the estimated residuals. For the purpose of model identification selection criteria such as the Akaike Information Criterion [AIC] are usually employed (see Brockwell and Davis [1991]). Minimisation of the AIC suggests model order $p = 26$ for the seasonally differenced series. Furthermore, the model contains no intercept, reflecting the fact, that unemployment rates do not tend to rise or fall in the long run.

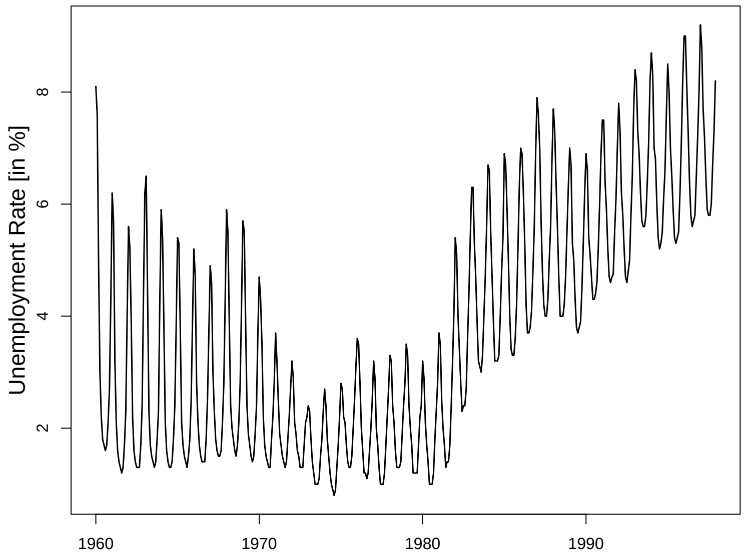


Figure 1: Monthly observations of the Austrian unemployment rate, from January 1960 to December 1997

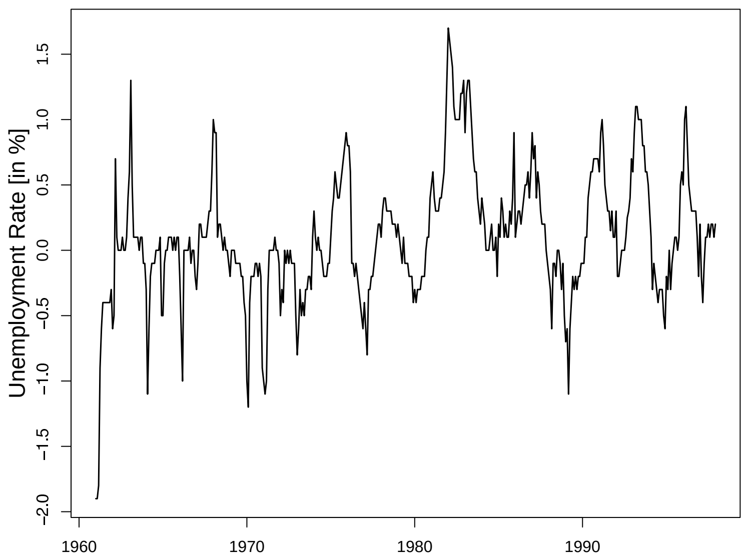


Figure 2: Seasonally differenced series, $\Delta_{12}x_t$, Austrian unemployment rates

## 4.2 Empirical results

The following is a summary of the results obtained by applying the BDS test, the McLeod-Li test, the Hsieh test, and the Teräsvirta-Lin-Granger test, with each test judged relative to the null hypothesis that it is designed to test.

### Results with the BDS test

The first test used is the BDS test that examines the underlying probability structure of the time series searching for any kind of dependence in the series. The BDS statistic will reject any deviation from independence. Evidence from the AR(26) model shows that the seasonally differenced series is not independent. In this situation the BDS statistic can be used as a test for residual non-linear structure, after linear structure has been removed by fitting the AR(26) model.

The BDS-statistic has been computed over a grid of embedding dimension ($m = 2,\ldots,15$) and $\varepsilon$'s ($\varepsilon = 0.5\sigma_u, 1.0\sigma_u$) where $\sigma_u$ is the standard deviation of the residual time series). The results are summarized in Table 2. Rejection of the null hypothesis of independence indicates structure beyond the fitted linear model. When $(T-m+1)/m$ exceeded 200 we used the standard normal distribution for the critical value, otherwise we looked up the critical value from tables in Brock, Hsieh, and LeBaron [1991] to assess the significance.

The results obtained are unambiguous. The rejection of the null is extremely strong for $\varepsilon = 0.5\sigma_u$ and $\varepsilon = 1.0\sigma_u$. Much of the Monte Carlo research that has been published on the BDS test (see, for example, Brock, Hsieh, and LeBaron [1991]) has emphasized the potential dependence of the properties of the test on the a priori linear filter. Thus, we considered a sparsely specified ARMA model with two AR coefficients at lags 1 and 12 and three MA coefficients at lags 11, 12 and 13 as an alternative linear filter. But the results did not change when varying the linear filter.

If the null hypothesis of the BDS test is rejected, other tests should be used to allow the class of relevant alternatives to be narrowed down. If the null hypothesis is accepted then there would be little point to continue further, since an acceptance of linearity by the BDS test is a strong result.

### Results with the McLeod-Li test

The McLeod-Li test is a Portmanteau test for non-linear dependence that examines the Ljung-Box-Pierce statistic of the squared residuals from an AR(26) representation. The results with the test, displayed in Table 3, for $k = 5,\ldots,26$, provide clear evidence against the null hypothesis of linearity. The strength of this conclusion is evident from the fact that the critical value of the test at the 0.05 level is reached in all displayed cases. This result corroborates the inference that the data contain non-linearities, and in particular provides a strong indication for non-linearity in conditional variance.

Table 2: BDS($\varepsilon, m, T$) for $m = 2, \ldots, 15$ and four different values of $\varepsilon$: Data are prewhitened by AR(26) fit

| $m$ | $\varepsilon = 0.0905$ | $\varepsilon = 0.1810$ |
|---|---|---|
| 2 | 5.6946 (0.0000)*** | 6.3186 (0.0000)*** |
| 3 | 6.2463 (0.0000)*** | 6.5120 (0.0000)*** |
| 4 | 6.5737 (0.0000)*** | 6.7165 (0.0000)*** |
| 5 | 6.3121 (0.0000)*** | 6.3626 (0.0000)*** |
| 6 | 6.1554 (0.0000)*** | 5.6167 (0.0000)*** |
| 7 | 5.5415 (0.0000)*** | 4.5242 (0.0000)*** |
| 8 | 4.7303 (0.0000)*** | 3.8954 (0.0001)*** |
| 9 | 3.7160 (0.0002)*** | 3.2697 (0.0011)** |
| 10 | 6.5993 (0.0000)*** | 3.0386 (0.0024)** |
| 11 | 16.2679 (0.0000)*** | 3.3001 (0.0010)*** |
| 12 | 35.6241 (0.0000)*** | 3.6136 (0.0003)*** |
| 13 | 73.1271 (0.0000)*** | 4.6322 (0.0000)*** |
| 14 | 140.1451 (0.0000)*** | 6.2985 (0.0000)*** |
| 15 | 239.2324 (0.0000)*** | 8.2771 (0.0000)*** |

Note: The BDS statistic is asymptotically standard normal under the null hypothesis of independence. If $(T - m + 1)/m \leq 200$ we use the critical values from tables in Brock, Hsieh, and LeBaron [1991]. Prob-values are included in parentheses. ***, ** and * denote significant values at the 0.1%, 1% and 5% confidence levels, respectively.

Table 3: McLeod-Li test: Ljung-Box-Pierce statistics, residuals of an AR(26) model

| $k$ | Q($k$) | $k$ | Q($k$) |
|---|---|---|---|
| 5 | 15.2538 (0.0093)** | 16 | 55.1627 (0.0000)*** |
| 6 | 22.1388 (0.0011)** | 17 | 59.5608 (0.0000)*** |
| 7 | 26.9507 (0.0003)*** | 18 | 66.2256 (0.0000)*** |
| 8 | 29.2521 (0.0003)*** | 19 | 72.2936 (0.0000)*** |
| 9 | 29.4720 (0.0005)*** | 20 | 72.8171 (0.0000)*** |
| 10 | 29.7238 (0.0009)*** | 21 | 72.8940 (0.0000)*** |
| 11 | 35.6877 (0.0002)*** | 22 | 72.9240 (0.0000)*** |
| 12 | 45.5349 (0.0000)*** | 23 | 81.3158 (0.0000)*** |
| 13 | 54.9075 (0.0000)*** | 24 | 95.1295 (0.0000)*** |
| 14 | 55.0969 (0.0000)*** | 25 | 96.3822 (0.0000)*** |
| 15 | 55.1211 (0.0000)*** | 26 | 97.0221 (0.0000)*** |

Note: The test statistic Q($k$) is chi-square-distributed with $k$ degrees of freedom. Prob-values are included in parentheses. *** and ** denote significant values at the 0.1% and 1% confidence levels, respectively.

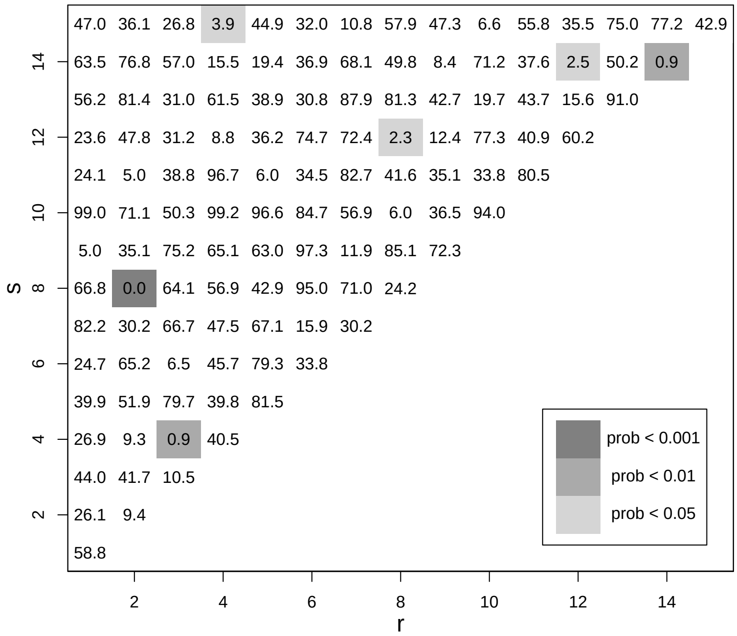


Figure 3: Hsieh test, residuals of an AR(26) model (prob-values $\times 100$)

**Results with the Hsieh test**

The Hsieh test discriminates between additive and multiplicative non-linearity once it is established that some type of non-linearity exists in the data. Again, the test is based on the assumption that by fitting an AR(26) model to the seasonally differenced time series, the inherent non-linearity has been swept into the residuals. The test is used to evaluate the null hypothesis of multiplicative non-linearity, using the standard normal distribution. The test is a third moment test which has the disadvantage that several lags have to be tested and that the selection of lags $r$ and $s$ for the test statistic is ambiguous. Thus, the test results are displayed in Figure 3 for a grid of $r, s = 1, \ldots, 15$.

The figure shows asymptotic prob-values for the test, a low prob-value suggesting rejection of the null hypothesis of multiplicative non-linear dependence in favour of the alternative hypothesis of additive non-linearity. The test produced only 6 rejections out of 120 cases. The results of the Hsieh test can be considered only as extremely weak evidence against the null hypothesis of multiplicative non-linearity.

**Results with the Teräsvirta-Lin-Granger test**

The final test used is the Teräsvirta-Lin-Granger test which does not need to extract linear structure by the use of an estimated filter. The disadvantage of this test is that the researcher has to fix $p$, and that the size of the regressions can grow quite large very rapidly, that is, the number $m$ of regressors used in the auxiliary regressions of the test procedure increases more than proportionally with $p$.

The test procedure starts with selecting the order $p$ of the autoregressive process. In doing so, we followed the rule of minimizing AIC (see Brockwell and Davis [1991]). The resulting AR model is of order 26, leading to $m = 2,266$ regressors in the non-linear auxiliary regression. Thus, computation of the TLG-statistic is not feasible anymore in this situation. To make the testing applicable to form, we modify the test procedure as follows:

(i) Select the order $p$ of the autoregressive process by a conventional selection criterion, regress $x_t$ on $1, x_{t-1}, \dots, x_{t-p}$ and compute the residuals $\hat{u}_t$ and the sum of the squared residuals $SSR_0 = \sum_t \hat{u}_t^2$.

(ii) Select a smaller subset, $(x_{t-q_1}, x_{t-q_2}, \dots, x_{t-q_n} \mid q_1, q_2, \dots, q_n \leq p)$, of $(x_{t-1}, \dots, x_{t-p})$ for appropriate $q_1, q_2, \dots, q_n$.

(iii) Regress $\hat{u}_t$ on $1, x_{t-1}, \dots, x_{t-p}$ and $\tilde{m}$ auxiliary regressors corresponding to the the second-order and third-order expansions of $x_{t-q_1}, x_{t-q_2}, \dots, x_{t-q_n}$. Compute the residuals $\hat{v}_t$ and the sum of squared residuals $SSR = \sum_t \hat{v}_t^2$ and the test statistic

$$\widetilde{TLG} = \frac{(SSR_0 - SSR)/\tilde{m}}{SSR/(T - p - 1 - \tilde{m})} \tag{21}$$

(iv) For a selected significance level $\alpha$ find the critical value $\tau$ for testing the null hypothesis using the $F$-distribution with $\tilde{m}$ and $T - p - 1 - \tilde{m}$ degrees of freedom.

(v) Reject the null hypothesis of linearity if $\widetilde{TLG} > \tau$.

Note that the modification of the test procedure has the advantage, that the number $\tilde{m}$ of regressors used in the non-linear part of the auxiliary regression is now in a reasonable range. It is straightforward to show that this modification does not affect the statistical properties of the test procedure.

The results with the modified Teräsvirta-Lin-Granger test are summarized in Table 4, for various subsets $(x_{t-1})$, $(x_{t-2})$, $(x_{t-10})$, $(x_{t-1}, x_{t-10})$, $(x_{t-1}, x_{t-2}, x_{t-10})$ and $(x_{t-1}, x_{t-2}, x_{t-10}, x_{t-12})$. The null hypothesis of linearity is rejected, suggesting non-linearity in the mean. The rejection is very strong except in the second case.

Figure 4 shows asymptotic prob-values for the modifed test for a grid of $q_1, q_2 = 1, \dots, 15$ with $q_1 > q_2$. A low prob-value suggests rejection of the null hypothesis of linearity in favour of the alternative hypothesis of non-linearity in the mean. This is unambiguously the case for $(q_1, q_2)$ with $q_1 = 1, 10$ and 12. This result underlines that it may be sufficient to consider specific lag combinations $(q_1, q_2)$ to reject the null hypothesis of linearity.

# 5 Conclusions

We find some consistency in our inferences across the methods of inference, although there are some clear differences among the power functions of the tests. It may be possible that

Table 4: The modified Teräsvirta-Lin-Granger test: The test statistic $\widetilde{TLG}$ for several sets $(x_{t-q_1}, \ldots, x_{t-q_n})$

| $q_1$ | $q_2$ | $q_3$ | $q_4$ | $\tilde{m}$ | $T-p-1-\tilde{m}$ | $\widetilde{TLG}$ |
|---|---|---|---|---|---|---|
| 1 | | | | 2 | 415 | 13.9272 (0.0000)*** |
| 2 | | | | 2 | 415 | 4.2322 (0.0152)* |
| 10 | | | | 2 | 415 | 11.4322 (0.0000)*** |
| 1 | 10 | | | 7 | 410 | 8.4546 (0.0000)*** |
| 1 | 2 | 10 | | 16 | 401 | 5.0293 (0.0000)*** |
| 1 | 2 | 10 | 12 | 30 | 387 | 3.6521 (0.0000)*** |

Notes: The test statistic $\widetilde{TLG}$ of the modified Teräsvirta-Lin-Granger test is asymptotically $F$-distributed with $\tilde{m}$ and $T-p-1-\tilde{m}$ degrees of freedom. Prob-values are included in parentheses. ***, ** and * denote significant values at the 0.1%, 1% and 5% confidence levels, respectively.

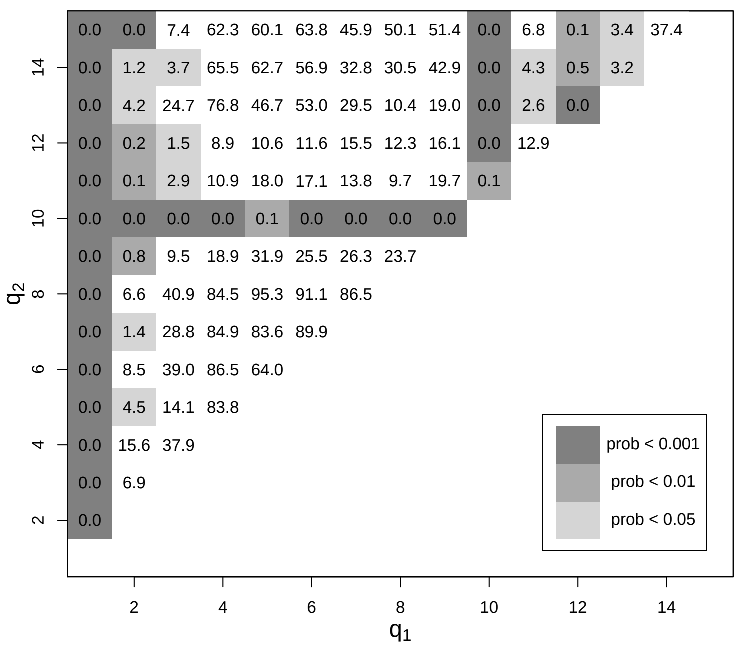


Figure 4: The modified Teräsvirta-Lin-Granger test (Prob-values $\times 100$): The case of $(x_{t-q_1}, x_{t-q_2}), q_1, q_2 \leq 15, q_1 < q_2$

greater robustness across inference procedures might be obtained at much greater sample size that are, however, rare in practice. Some of the test procedures can be considered as complementary rather than competing. None of the tests uniformly dominates the others. Using all of them jointly may produce deeper insights into the nature of the non-linearity that may exist in the data.

The BDS test is an omnibus test that tests linearity against all possible alternatives. Monte Carlo simulations indicate that the test is very sensitive to departures from linearity, but also emphasize that it may depend on the linear filter used for prewhitening. The highly significant test results and, thus, rejection of the null hypothesis may partly be caused by multiplicative non-linearity present in the data.

Only if non-linearity is rejected with the BDS test or if the test leads to ambiguous results it becomes reasonable to make use of other more focused tests for non-linearity. The McLeod-Li test is sensitive against multiplicative non-linearity, less so against additive non-linearity. The Hsieh test and the Teräsvirta-Lin-Granger test are specifically designed to detect additive non-linearity. The Teräsvirta-Lin-Granger test has high power to distinguish among non-linear processes that are non-linear in the mean and those that are not [such as ARCH processes]. But the Hsieh test should be run before proceeding to the modified Teräsvirta-Lin-Granger test because of the a priori knowledge required in choosing $q_1, q_2, \ldots, q_n$. Note that simply rejecting linearity is not likely to exhaust the useful information available in the data about non-linearity.

It is important to emphasize that we cannot be sure that there are not other features of the unemployment series that lead to the observed results. Especially, the possible presence of ARCH effects cannot be ruled out. ARCH effects may have two effects in general. First, they may cause the size of the test statistic to be incorrect while still resulting in a diagnostic bounded in probability under the null hypothesis, as is the case of the Teräsvirta-Lin-Granger test. Second, they may directly lead to rejection despite linearity in the mean. Lee, White, and Granger [1993] suggest two strategies that can be undertaken in this situation. The first strategy may be followed to remove effects of the first type by using a heteroskedasticity consistent matrix operator in calculating the test statistic. The second strategy involves specifying the form of the ARCH effect and jointly modelling non-linearity in the mean and heteroskedasticity when performing the tests. Joint modelling is necessary because using an ARCH with a linear filter may bias the test against the alternative. Note that the use of ARCH with a non-linear model may make one incorrectly not find actual non-linearity in the case of the BDS test. So it might be necessary to revise this test procedure, not affected by heteroskedasticity.

Consequently, we can take the empirical results of this study as indicating that either neglected non-linearity or ARCH effects may be present in the Austrian unemployment time series. Thus, further investigation is required. The results achieved are only a first step on the way to analysing inference procedures able to unambiguously detect non-linearity in real world small sized time series.

# Notes

[1]The results presented in this paper are part of a larger Ph.D. project of the first author, supervised by the second.

[2]Instead of using the residuals from a linear representation, the raw data can be examined through the use of the $k$ autocorrelation functions.

[3]The ARCH test of Engle [1982], which is asymptotically equivalent to the McLeod-Li test, shares the same shortcoming, assuming a correct conditional mean specification, including potential additive non-linearity (see Lumsdaine and Ng [1999] for more details and robustification of the ARCH test statistic).

[4]The test is also known as Teräsvirta's neural network test due to the fact it has been motivated by White's neural network test (see White [1989] and Lee, White, and Granger [1993]), the test statistic, however, is not based on neural network concepts such as neural network approximation theory.

# References


Barnett, W. A., Gallant, A. R., Hinich, M. J., Jungeilges, J. A., Kaplan, D. T., and Jensen, M. J. (1997). A single-blind controlled competition among tests for nonlinearity and chaos. *Journal of Econometrics*, **82**, 157–192.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**(3), 307–327.

Box, G. E. and Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Brock, W. A., Dechert, W. D., and Scheinkman, J. A. (1987). *A Test for Independence Based on the Correlation Dimension*. Work. pap., Department of Economics, University of Wisconsin at Madison, University of Houston, and University of Chicago.

Brock, W. A., Dechert, W. D., Scheinkman, J. A., and LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, **15**(3), 197–235.

Brock, W. A., Hsieh, D. A., and LeBaron, B. (1991). *Nonlinear dynamics, chaos, and instability: Statistical theory and economic evidence*. Cambridge, Mass. and London: MIT Press.

Brock, W. A. and Sayers, C. L. (1988). Is the business cycle characterized by deterministic chaos? *Journal of Monetary Economics*, **22**(1), 71–90.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. New York: Springer.

Burgess, S. M. (1992). Asymmetric employment cycles in Britain: Evidence and an explanation. *Economic Journal*, **102**, 279–290.

Craig, S. G., Kohlhase, J. E., and Papell, D. H. (1991). Chaos theory and microeconomics: An application to model specification and hedonic estimation. *Review of Economics and Statistics*, **73**(2), 208–215.

Cromwell, J. B., Labys, W. C., and Terraza, M. (1994). *Univariate Tests for Time Series Models*. Thousand Oaks, CA: SAGE Publications.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, **50**(4), 987–1007.

Engle, R. F., Lilien, D. M., and Robins, R. P. (1987). Estimating time varying risk premia in the term structure: The ARCH-M model. *Econometrica*, **55**(2), 391–407.

Granger, C. W. (1991). Developments in the nonlinear analysis of economic series. *Scandinavian Journal of Economics*, **93**(2), 263–276.

Granger, C. W. and Andersen, A. (1978). *An Introduction to Bilinear Time Series Models*. Göttingen: Vandenhoeck & Ruprecht.

Grassberger, P. and Procaccia, I. (1983). Measuring the strangeness of strange attractors. *Physica D*, **9**, 189–208.

Hsieh, D. A. (1989). Testing for nonlinear dependence in daily foreign exchange rates. *Journal of Business*, **62**(3), 339–368.

Lee, T.-H., White, H., and Granger, C. W. (1993). Testing for neglected nonlinearity in time series models. *Journal of Econometrics*, **56**, 269–290.

Lumsdaine, R. L. and Ng, S. (1999). Testing for ARCH in the presence of a possibly misspecified mean. *Journal of Econometrics*, **93**(2), 257–279.

McLeod, A. and Li, W. (1983). Diagnostic checking of ARMA time series models using squared residuals autocorrelations. *Journal of Time Series Analysis*, **4**, 269–273.

Ozaki, T. (1980). Non-linear time series models for non-linear random vibrations. *Journal of Applied Probability*, **17**, 84–93.

Peel, D. A. and Speight, A. E. H. (1998). The non-linear time series properties of unemployment rates. *Applied Economics*, **30**, 287–294.

Sentana, E. (1995). Quadratic ARCH models. *Review of Economic Studies*, **62**(4), 639–661.

Teräsvirta, T., Lin, C.-F., and Granger, C. (1993). Power of the neural network linearity test. *Journal of Time Series Analysis*, **14**(2), 209–220.

Tong, H. and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society*, **B42**, 245–292.

White, H. (1989). An additional hidden unit test for neglected non-linearity in multilayer feedforward networks. In *Proceedings of the International Joint Conference on Neural Networks, Washington, DC*, vol. I, pp. 451–455, San Diego, CA. SOS Printing.