Testing for non-linearity in daily sterling exchange rates

Article

Accepted Version

Testing for Nonlinearity in Daily Sterling Exchange Rates

by

Chris Brooks

Department of Economics, Faculty of Letters, PO Box 218,
Whiteknights, Reading, RG6 6AA. England

April 1995

Abstract

A number of tests for nonlinear dependence in time series are presented and implemented on a set of ten daily Sterling exchange rates covering the entire post-Bretton Woods era until the present day. Irrefutable evidence of nonlinearity is shown in many of the series, but most of this dependence can apparently be explained by reference to the GARCH family of models. It is suggested that the literature in this area has reached an impasse, with the presence of ARCH - effects clearly demonstrated in a large number of papers, but with the tests for nonlinearity which are currently available being unable to classify any additional nonlinear structure.

Keywords

exchange rates, financial time series, GARCH, nonlinearity tests

J.E.L. Classification

C22, C51, F31

1. Introduction

Testing for nonlinear dependence has become an important area of research in financial econometrics because of its profound implications for model adequacy, market efficiency, and predictability. If we find evidence of nonlinearity in financial time series, this suggests that, at least in the short term, forecasts may be improved by switching from a linear to a nonlinear modelling strategy, and furthermore, the tests may be viewed as general tests of model adequacy for linear models in the sense that if there is still dependence in the residuals of a linear model (of an albeit more complex form), the original linear models can no longer be viewed as an accurate representation of the data (Hinich and Patterson, 1989). Evidence of nonlinear dependence in market returns may also cast doubts on the informational efficiency of financial markets, since in theory (at least), it may be possible to devise a trading strategy which systematically generates positive returns with a probability considerably in excess of one half. It is also of interest to consider whether there is any difference in general between the extent of nonlinearity in widely traded currencies, such as the U.S. Dollar and Deutschmark relative to those which are infrequently traded or traded in considerably smaller volumes, such as the Austrian Schilling or Danish Krone. One may expect that, *a priori*, those currencies which are frequently traded should more closely follow the random walk model of weakly efficient markets (Fama, 1970), since their closer and more detailed inspection by traders and analysts should ensure that autoregressive dependencies of any form (i.e. linear or nonlinear), are quickly arbitraged away.

The structure of the remainder of the paper is as follows. Section 2 gives a brief survey of the findings of other studies which have investigated the use of the tests discussed here. Section 3 introduces and explores the data to be used, and section 4 gives a theoretical treatment of a number of tests for nonlinear dependence in time series. The

results generated thereof are described in section 5. Section 6 offers an analysis of the results and concludes. Finally, there follows an appendix of tabulated results.

2. A Summary of Previous Research in the Area

Since the techniques explored in this paper were popularised by Brock (1986), Hsieh (1989b) and co-workers, there has been an explosion of interest in what appears to be the science of finding structure where there is none, and forecasting the unforecastable. Series of financial returns often appear completely random to standard linear and spectral tests (Brock *et al.*, 1991), but there is a rapidly developing literature which argues that if a different approach, using more powerful techniques is used, it may be possible to uncover a more complex form of dependence in these series. Numerous studies have examined the possibility of nonlinear dependence over the last eight years or so, and these can be conveniently categorised into tests on macroeconomic and financial data. Papers by Brock and Sayers (1988), DeCoster and Mitchell (1991), Frank, Gencay and Stengos (1988), and Lee, White and Granger (1993) reside in the former category. Studies examining macroeconomic series have not been particularly successful, often beset with problems of insufficient data and excessive noise and measurement error (Ramsey *et al.*, 1990a, 1990b). The literature employing financial data has been far more voluminous, and some of the best examples, are Abhyankar *et al*. (1995), Hsieh (1989b), Scheinkman and LeBaron (1989a, 1989b), and Mayfield and Mizrach, (1992). With the notable exception of Willey (1992), the common finding is that there is overwhelming evidence of nonlinear structure across many types of financial data (stocks, bonds, futures, foreign exchange), which can reasonably be described as a generalised autoregressive conditionally heteroscedastic (GARCH) process (Ballie and Bollerslev, 1989). However, the question of whether this is the only, or indeed the best way to represent the nonlinear dependence which is found in these series still remains unresolved (Sola and Timmerman, 1994).

3. The Data and Preliminaries

The analysis presented here is based on just over twenty years of daily mid-price spot exchange rate data, denominated in Sterling. The sample period taken covers the whole of the post-Bretton Woods era until the present day, specifically from 2 January 1974 until 1 July 1994 inclusive. A set of ten currencies are analysed, namely the Austrian Schilling/Pound (hereafter denoted A), the Canadian Dollar/Pound (C), the Danish Krone/Pound (D), the French Franc/Pound (F), the German Mark/Pound (G), the Hong Kong Dollar/Pound (H), the Italian Lira/Pound (I), the Japanese Yen/Pound (J), the Swiss Franc/Pound (S), and the U.S. Dollar/Pound (U). The raw exchange rates were transformed into log-returns with all subsequent analysis being performed on these, which can be interpreted as a series of continuously compounded daily returns (Brock *et al.*, 1991), and constitutes a series of 5191 observations. This transformation has become standard in the finance literature, and one possible justification for using returns rather than raw data is that the raw data is likely to be nonstationary (see, for example, Corbae and Ouliaris, 1988 or Bleaney and Mizen, 1993)[1]. Nonstationarity may cause a spurious rejection of linearity (Hinich and Patterson, 1985) and furthermore, Guillaume *et al.* (1994, p6) argue that "...the changes of prices rather than the prices themselves are the variable of interest for traders maximising short-term investment returns".

Summary statistics for the returns are presented in table 1 of the appendix. All series show strong signs of non-normality, with significant excess kurtosis. This is clearly a salient feature of most financial data (Hsieh, 1988, 1989a). The Ljung Box (1978)

[1]The data were tested for the presence of unit root nonstationarity using the Dickey Fuller (Dickey and Fuller, 1979;Fuller, 1976), Phillips Perron (Phillips, 1987; Perron and Phillips, 1987; Phillips and Perron, 1988) and Sargan Bhargava (Sargan and Bhargava, 1983, Bhargava, 1986). The levels data and the log-levels data were found in all cases to be strongly *I(1)*, but there was no evidence of nonstationarity in the returns series.

portmanteau test for the joint significance of the first ten lags, given in the last column of the table, are generally significant for most of the series.

3.1 Fitting a Linear Model

A further issue to be explored at this stage is the question of whether there exists any simple linear dependence or daily seasonality in the data. This may need to be filtered out before testing for nonlinearity since some of the tests, such as the BDS test, have power against linear as well as nonlinear alternatives, and many of the tests employed (Tsay's test, White's neural network test, Engle's ARCH test, and Ramsey's RESET test), are calculated from the estimated residuals of a linear fit. The class of ARIMA, or Box Jenkins (Box and Jenkins, 1976) models can be used to fit a linear model to a time series. In this case, the unit root in the series has already been removed by the process of log-differencing, and since any moving average model can also be represented by an infinite order autoregression, the class of possible linear specifications is restricted to those of an autoregressive form. ARMA and MA models were also tested, but in general, the MA component adds little to the model, and can be restricted to zero on likelihood ratio grounds. Furthermore, the effects of MA and ARMA filtering on the null distribution of the nonlinearity test statistics is not well documented as it is with pure autoregressive models[2]. In each case, a dummy variable for the number of successive holiday days between trading (excluding weekends) was also included in the estimated regressions, together with daily dummies for Monday, Tuesday, Wednesday, and Thursday[3]. These dummies are incorporated in order to remove any linear seasonality in the data which may arise from the market microstructure or the non-homogeneous arrival of information throughout the week. This class of models has been widely used as a starting point for filtering linear dependence from time series data prior to testing for

[2]Granger *et al.*(1989), for example, test only AR(p) models "for simplicity".

nonlinearity (Hsieh, 1989a, 1989b, 1991; LeBaron, 1988, for example). Failure to remove seasonality from the data may lead to a spurious rejection of whiteness. The choice of autoregressive lag length can be made on a number of grounds, including the use of the Box Jenkins interactive procedure, or the use of various information criteria. Among these are the minimisation of Schwarz's Bayesian Information Criterion (Schwarz, 1978) and Akaike's Information Criterion (Akaike, 1974), or the maximisation of $\overline{R}^2$. An alternative, suggested by Hsieh (1989b), is to use the minimum number of lags that ensures that the Ljung Box Q(50) statistic is insignificant at the 10% level although this proposal seems somewhat arbitrary. Only results using the number of lags selected by SBIC are shown here, although all the nonlinearity tests were conducted using the numbers of lags selected using Hsieh's criterion and using an arbitrary ten lags for all the series, and although these are not shown, the results are virtually indistinguishable from those of shorter lag lengths.

It should be emphasised at the outset that the objective of this study is not to build a statistically adequate empirical model of exchange rate returns, but rather to choose an acceptable specification which will remove linear and linear seasonal components from the series, and yet will provide a parsimonious representation of the data. Indeed the linear regressions (not displayed) all perform rather poorly as stand-alone models, in common with many other studies (e.g. Scheinkman and LeBaron, 1989; Willey, 1992). Few lags are significant beyond the first lag, even at the 10% level, and those that are seem randomly placed. The holiday dummy is found to be significant (at the 10% level or better) only for the Canadian Dollar and Hong Kong Dollar. Many of the series exhibit no daily seasonality (A, D, F, G, S), while in the rest there appears to be a significant, positive Wednesday effect.

[3]A fifth dummy variable for Friday was not included (as it is, for example, in Ellis (1992)) to avoid the occurrence of perfect mulitcollinearity between the five daily dummies and the constant term.

3.2 Fitting a GARCH Model

The ARCH (AutoRegressive Conditional Heteroscedasticity) family of models was first suggested by Engle (1982), and has led to a subsequent explosion of articles on the subject. Instead of being conditionally fixed over time, the variance of the process is modelled as being dependent on lags of past squared residuals[4]. These models have found to accurately fit the data from a diversity of financial disciplines, although as Hall, Miles and Taylor (1989) note, the ARCH parameterisation of the conditional variance does not have any solid grounding in economic theory, but represents "a convenient and parsimonious representation of the data." A generalisation of the ARCH model, known as GARCH, was proposed by Bollerslev (1986), which can be viewed as an infinite order ARCH model. It was considered preferable to estimate a GARCH-(1,1) model, which embodies a more compact representation, for the purpose of filtering. The GARCH models estimated in this study were of the form

$$x_t = \beta_0 + \sum_{i=1}^{p} \beta_i x_{t-i} + \gamma_1 Mo + \gamma_2 Tu + \gamma_3 We + \gamma_4 Th + \gamma_5 Hol + \varepsilon_t \,,\ \varepsilon_t \sim N(0\,,\, h_t) \quad (1)$$

$$h_t = \alpha_0 + \alpha_1 h_{t-1} + \alpha_2 \varepsilon_{t-1}^{\,2} \quad (2)$$

where x_t represents the first difference of the log of the exchange rate, *Mo*, *Tu*, *We*, and *Th* are daily dummy variables, and *Hol* is a dummy variable representing the number of successive days between trading (excluding weekends), as described above. It is highly unlikely that a GARCH model of order greater than one in the autoregressive and moving average components would be required since, by definition, a GARCH-(1,1) model implies an infinitely long memory with respect to past innovations. Akgiray (1989) finds that all higher orders of GARCH from (5,3) to (1,1) can be restricted to zero under a likelihood ratio test. The models were estimated using the maximum likelihood algorithm

of Broyden, Fletcher, Golfarb and Shanno (BFGS) with mis-specification robust estimation of standard errors.

4. A Theoretical Derivation of the Tests

In this section, a brief theoretical description of each test for nonlinearity used is given together with the null distribution with which the test statistic is to be compared.

4.1 Tsay's Test

Tsay's Test (1986) is a more powerful (Tsay, 1986, p461) generalisation of Keenan's (1985) test, which includes cross-product terms at different lag lengths in the auxiliary regression, such as $x_{t-j}x_{t-k}$. The first step of the test is in common with Keenan's, that is to regress x_t on a constant and p own lags, saving the residuals, $\hat{u}_t$. Next, define a vector Z_t as

$$Z_t' = vech(Y_t' Y_t), \tag{3}$$

where *vech* denotes the vector half-stacking operation, and $Y_t = (x_{t-1},...,x_{t-p})$. That is, all possible combinations of the lagged terms (with $x_{t-j}x_{t-k}$ and $x_{t-k}x_{t-j}$ counting only once) are stacked in a single vector, which hence contains *(½p(p-1))* terms. The second step entails running a multivariate regression of Z_t on a constant and p lags, saving the residuals, $\hat{v}_t$. This step may be viewed as an estimation of a system of *½p(p+1)* separate regressions, and hence $\hat{v}_t$ will also have dimension *½p(p+1)* X *1*. Finally regress $\hat{u}_t$ on $\hat{v}_t$ and denote the residuals $\hat{e}_t$. The test statistic is given by:

$$\hat{F} = \frac{(\sum \hat{v}_t \hat{u}_t)(\sum \hat{v}_t' \hat{v}_t)^{-1}(\sum \hat{v}_t \hat{u}_t)/(\frac{1}{2}p(p-1))}{\sum \hat{e}_t^2/(T-p-(\frac{1}{2}p(p-1))-1} \tag{4}$$

[4]See Bollerslev *et al.* (1990) for an excellent survey of ARCH models and their derivatives.

where the summations are from $t=p+1$ to T. In the limit, $\hat{F}$ is distributed as a χ^2 variable with $\frac{1}{2}p(p+1)$ degrees of freedom, but Tsay states that "in practice [he] prefer[s] to use the approximate F-distribution" (p463).

4.2 The McLeod and Li Test

The McLeod and Li (1983) test is similar to the Ljung-Box (1978) test in the sense that it is a test based on the autocorrelation of the (in this case squared) residuals. In practice, the test is computed as follows. First, fit the "best" ARMA model to the time series, and denote the residuals by $\hat{u}_t$. The autocorrelation function at lag k is denoted by

$$\hat{r}_{uu}(k) = \frac{\sum_{t=K+1}^{T}(\hat{u}_t^2 - \hat{\sigma}^2)(\hat{u}_{t-k}^2 - \hat{\sigma}^2)}{\sum_{t=1}^{T}(\hat{u}_t^2 - \hat{\sigma}^2)^2} \qquad (5)$$

where $\hat{\sigma}^2 = \frac{1}{T}\sum u_t^2$

McLeod and Li show that for fixed p, $\sqrt{T}\,\hat{r}_{uu} = [\hat{r}_{uu}(1),\ldots,\hat{r}_{uu}(p)]$ is distributed asymptotically normally as $T\rightarrow\infty$. A significance test can then be undertaken using the portmanteau statistic

$$Q_{uu}^{*} = T(T+2)\sum_{i=1}^{p}\frac{\hat{r}_{uu}^2(i)}{(T-i)} \qquad (6)$$

which is asymptotically $\chi^2(p)$ if the $\hat{u}_t$ are independent. It is apparent from the formulation of the test statistic that the test is a squared residual analogue of the Ljung Box Q_a* statistic. Thus the test will be relatively powerful in detecting departures from linearity which may be attributed to the presence of ARCH effects, but as Lee *et al.* (1993) show, the test has some limited power against other nonlinear alternatives.

4.3 The RESET Test

Ramsey's RESET (Regression Error Specification Test), (Ramsey, 1969) is a general test for mis-specification of functional form which could again be viewed as a more general formulation of Keenan's test, which allows for functions of the fitted values of a higher order than two. The statistic is formulated via the following steps. First, the multivariate regression (with exogenous variables and/or with *p* own lags) with a constant is run, saving the residuals, $\hat{u}_t$, and fitted values, $\hat{y}_t$. Next, a regression which constitutes the alternative model of the form

$$y_t = \tilde{x}_t'\alpha + \beta_2 \hat{y}_t^2 + \ldots + \beta_k \hat{y}_t^k + v_t \qquad (7)$$

is run for k≥2, where $\tilde{x}_t'$ is a vector containing the original regressors, and k is the highest order of the fitted values from the first regression included in the second (auxiliary) equation. In practice, collinearity between the fitted values of different orders may be a problem, so the usual procedure is to form the principal components of the fitted values $(\hat{y}_t^2, \ldots, \hat{y}_t^k)$, and choosing the p^* largest, regressing z_t on these and $\tilde{x}_t'$ to give the residual $\hat{e}_t$. The test statistic is given by

$$RESET = \frac{(\hat{u}'\hat{u} - \hat{e}'\hat{e})/p^*}{\hat{e}'\hat{e}/(\hat{t}-k)} \qquad (8)$$

Under the null, $RESET \sim F(p^*, T-k)$. A Lagrange Multiplier version of the test is obtained by taking the squared multiple correlation coefficient (i.e. R^2) from the last regression and multiplying it by the sample size to form

$RESET2 = TR^2 \sim \chi^2(p^*)$ for large T under the null. RESET2 was also calculated, but the *p*-values in each case were almost identical, and hence only those for *RESET* are shown. The sample size is large here, so one would expect the asymptotic bounds of the tests to be closely approximated.

4.4 Engle's Test for ARCH

A test for the presence of ARCH in the residuals of a linear model is calculated as follows. First, run a linear regression of the form described above (that is, using the dummy variables and *m* lags of the dependent variable), saving the residuals, $\hat{\varepsilon}_t$. Then square the residuals, and regress them on *p* own lags to test for ARCH of order *p*. The LM test statistic is defined as TR^2 (the number of observations multiplied by the coefficient of multiple correlation) from the last regression, and is distributed as a $\chi^2(p)$ under the null of no order *p* ARCH effects.

4.5 White's Neural Network Test

White's (1989, 1990) neural network test for neglected nonlinearity uses a single hidden layer feed-forward neural network with additional direct connections from inputs to outputs. Network output, *o*, is given by

$$o = \tilde{x}'\alpha + \sum_{j=1}^{q} \beta_j \psi(\tilde{x}'\gamma_j) \qquad (9)$$

where $\tilde{x}$ is the input vector, γ_j are the connection strengths or weights, β_j are hidden to output weights, ψ is an "activation" or "squashing" function, and α represent the direct connections between input and hidden layers. In this case, the logistic, $\psi(\lambda) = (1-e^{-\lambda})^{-1}$ is used. If the null hypothesis of linearity is true, then the $\tilde{x}'\alpha$ (i.e. the direct connections between the inputs and outputs) will describe the structure and hence the optimal β_j will all be zero. The rather complex test statistic is given by

$$M_T = (T^{-1/2}\sum_{t=1}^{T}\varphi_t \hat{e}_t)\hat{W}_T^{-1}(T^{-1/2}\sum_{t=1}^{T}\varphi_t \hat{e}_t) \qquad (10)$$

where $\hat{e}_t$ are the estimated residuals of the linear model, $\varphi_t = (\psi(\tilde{x}_t'\Gamma_1),\ldots,\psi(\tilde{x}_t'\Gamma_q))'$ and Γ, the hidden unit activations, are chosen at random. $\hat{W}_T$ is a consistent estimator of $W^* = var(T^{-1/2}\sum_{t=1}^{T}\varphi_t e_t^*)$. The null hypothesis

tested in a Lagrange Multiplier framework, is given by H_0: $E(\varphi_t e_t^*) = 0$, against a two sided alternative that the equality does not hold. M_T is asymptotically $\chi^2(q)$ under the null as $T\rightarrow\infty$. Bonferroni bounds provide an upper limit on the *p*-value. If $p_1,...,p_k$ denote the ascending-ordered *p*-values corresponding to *k* draws of Γ, then the simple Bonferroni implies rejection of a linear null at the $100\alpha\%$ level if $p_1 \leq \alpha/k$, so that, in the limit, the simple Bonferroni *p*-value is given by $\alpha = kp_1$. Hochberg (1988) suggests an amendment to the Bonferroni method, which allows consideration of all the *p*-values rather than just the largest, which may have led to a loss of power. The modified Hochberg Bonferroni limit is given by $\alpha = \min_{i=1,...,k} (m-i+1)\, p_i$, so that H_0 is rejected if there exists an *i* such that $p_i \leq \alpha/(m-i+1)$, i=1,...,k.

4.6 The Bispectrum Test

The bispectrum is the double Fourier transform of the third order cumulant function, and this forms the basis of the test for Gaussianity and linearity suggested by Hinich (1982), following a previous paper by Subba Rao and Gabr (1980). The general third order moment (cumulant)[5] function is defined as

$$C_{xxx}(r,s) = E[x(t+r)x(t+s)x(t)] \qquad (11)$$

Since the estimates of the third order cumulants are hard to interpret (Hinich and Patterson, 1985), the double Fourier transform is calculated, leading to the bispectrum:

$$B_{xxx}(f_1,f_2) = \sum_{r=-\infty}^{\infty}\sum_{s=-\infty}^{\infty} C_{xxx}(r,s)\, exp[-2\pi i(f_1 r + f_2 s)] \qquad (12)$$

This follows directly as a third order analogue to the normal power spectrum of x_t, which is given by

$$S_{xx}(f) = \sum_{s=-\infty}^{\infty} C_{xx}(s)\, exp\,[-2\pi i f s] \qquad (13)$$

where $C_{xx}(s) = E[x(t+s)x(t)]$ is the second order moment function.

It has been proved (Hinich, 1982), that the skewness function $X(f_1,f_2)$ is given by

$$X^2(f_1,f_2) = \frac{\left|B_{xxx}(f_1,f_2)\right|^2}{S_{xx}(f_1)S_{xx}(f_2)S_{xx}(f_1+f_2)} = \frac{\mu_3^2}{\sigma_\varepsilon^6} \qquad (14)$$

where $\sigma_\varepsilon^2 = E(\varepsilon_t^2)$ and $\mu_3 = E(\varepsilon_t^3)$. The modulus of the bispectrum, B_{xxx}, is taken since in general it is a complex number. Ashley, Patterson and Hinich (1986, p174), prove that the bispectral linearity test is invariant to linear filtering. Linearity and Gaussianity of x_t are tested via the null hypotheses that $X(f_1,f_2)$ is constant over all frequencies and that $X(f_1,f_2)$ is zero over all frequencies respectively. The actual test statistic used in both cases reduces to

$$S = 2\sum_m \sum_n \left|X(f_m,f_n)\right|^2 \qquad (15)$$

Under the null hypothesis of Gaussianity, the test statistic is distributed asymptotically as a standard normal. When the null is of linearity, the test statistic is distributed approximately as a χ^2 random variable with two degrees of freedom. The 80% quantile of the empirical distribution, scaled by a function of the variance of the series is asymptotically distributed as a standard normal variable under the null that the underlying data generating process is linear, and constitutes the test statistic shown here, as suggested by Hinich and Patterson (1989). One parameter to be set by the user is the frame size, denoted *M*. Hinich and Patterson (1985) use a frame size equal to the square root of the number of observations. In this case, that is calculated as 71, but in a personal communication, Hinich recommends a reduction in the frame size to 60 to improve the power of the test. The relevant test statistics are estimated here using both frame sizes.

4.7 The BDS Test

[5]As Barnett *et. al.* (1994) correctly point out, the cumulants are the coefficients of the power series expansion of the logarithm of a distribution, rather than of the level, as is the case for the moments.

The test of Brock, Dechert and Scheinkman (BDS, 1987) takes the concept of the correlation integral and transforms it into a formal test statistic which is asymptotically distributed as a standard normal variable under the null hypothesis of independent and identical distribution (IID) against an unspecified alternative. The formulation of the test statistic is as follows. First, the "m-histories" of the series, $x_t^m = (x_t, x_{t+\tau}, .., x_{t+\tau(m-1)})$ are computed for time $t = 1, ..., T-m$, for embedding dimension m, and for time delay τ. The correlation integral is defined as

$$C_m(\varepsilon) = \frac{1}{(T-m+1)(T-m)} \sum_{\forall t,s} I_\varepsilon(x_t^m, x_s^m) \tag{16}$$

where I_ε is an indicator function that equals one if $\left\| x_t^m - x_s^m \right\| < \varepsilon$, and zero otherwise. $\|.\|$ denotes the supremum norm, which is the most widely used distance measure. Although the usual Euclidean norm is equivalent, it is computationally more intensive and hence is rarely used in practical applications. The correlation integral thus counts the number of points out of a possible total of $\frac{1}{2}(T(T-1))$ that are within a distance ε of each other in m-dimensional space. Next calculate the log of the correlation integral divided by the log of the distance, ε, and take the limit as ε is made progressively smaller. The BDS test statistic is defined as follows

$$W_{m,T}(\varepsilon) = T^{1/2} \frac{[C_T(\varepsilon) - C_1(\varepsilon)^m]}{\sigma_{m,\tau}(\varepsilon)} \tag{17}$$

where $\sigma_m(\varepsilon) = 4\,[K^m + 2(\sum_{j=1}^{m-1} K^{m-j} C(\varepsilon)^{2j}) + (m-1)^2 C(\varepsilon)^{2m} - m^2 K C(\varepsilon)^{2m-2}]$

and $K(\varepsilon)$ is estimated by $K(\varepsilon) = \frac{6\sum_{t,s,r} h_\varepsilon(x_t^m, x_s^m, x_r^m)}{[Tm(Tm-1)(Tm-2)]}$

and $h_\varepsilon(i,j,k) = [I_\varepsilon(i,j)\, I_\varepsilon(j,k) + I_\varepsilon(i,k)\, I_\varepsilon(k,j) + I_\varepsilon(j,i)\, I_\varepsilon(i,k)]/3$

Since the test has power against many forms of deviation from IID, the BDS test is usually carried out on the residuals of a linear and/or GARCH-type filter. This step is

known as pre-whitening or bleaching, and makes it possible to see if further dependence beyond that described by a linear or GARCH process is present in the data. From the results of extensive Monte Carlo trials, Brock *et al.* (1991) recommend the use of ε equal to between one and three halves the standard deviation of the data in order to maximise the power of the test. The results given here are for $\varepsilon / \sigma = 1$ only, although an implementation with $\varepsilon / \sigma = 0.5$, 0.75, 1.25, and 1.5 gave results which, whilst qualitatively similar, were less robust to changes in the embedding dimension. The results derived from the residuals of a linear model are also not shown, since these are very similar to those of the original returns data[6]. The BDS test is a two sided test so that, in general, rejection of the null of IID occurs when the estimated value of the W-statistic is more extreme (in either tail) than the corresponding statistic from the Normal tables. Hence, as a rule of thumb, there exists some degree of dependence in the data if the absolute value of the estimated test statistic takes on a value greater than two, although the null distribution is changed when filtering through a GARCH model is undertaken[7]. On the whole, one would expect to reject in the upper tail of the distribution, since more clustering of points in *m*-dimensional space than would be expected under randomness appears more likely than less.

5. Results

A full set of results for all the tests examined is given in tables 2 to 6 of an appendix. The result of each test is considered below, and the main features are drawn together and analysed in section 6. On the whole, there is a noteworthy degree of agreement between the tests, which are generally very different in formulation. The autoregressive lag length specified in the estimation procedure makes very little difference to the conclusion

[6]This is quite expected, since there is very little linear structure in the returns; these results are available on request.

reached. Tsay's test rejects the linear null for the German Mark, Italian Lira, Japanese Yen and U.S. Dollar. The McLeod and Li test always leads to rejection of linearity at all significance levels, and it appears that the test is of little empirical value for the analysis of financial data, since, as detailed above, the test always rejects as a consequence of ARCH effects, and is unable to discern any further or alternative structure. The RESET test leads to rejection for the German Mark, Hong Kong Dollar, Italian Lira and Swiss Franc, at the 0.1% level. As expected, there was extremely strong evidence for the presence of ARCH effects in the residuals of the linear models for all the series. The *p-values* for the test statistics are not reported since they are always zero. However, two further points are noteworthy for consideration with regard to the other nonlinearity tests. First, the magnitude of ARCH effects varies considerably between the series: it is highest for the Austrian Schilling, the Hong Kong Dollar, the Italian Lira, and the Swiss Franc, but especially for the latter. Second, as the lag length of the autoregressive filter in the mean term is increased, the value of the ARCH test statistic is reduced. This may be a result of either small lag lengths failing to fully remove linear components and hence biasing the test in favour of rejection of ARCH, or specification of an excessive lag length relative to that of the true data generating process, may remove some genuine nonlinear dependence from the data.

The neural network test may be more powerful than the tests shown above, and this has indeed been found to be the case (Lee *et al., op cit.*). Rejection of the linear null hypothesis certainly occurs more frequently than for all tests except McLeod & Li and Engle's ARCH. Only the Canadian Dollar, and to a lesser extent, the Hong Kong Dollar and Japanese Yen, show no evidence of nonlinearity, but once again, the test has some power against ARCH, and thus any interpretation of the results may be somewhat

[7]New critical values for tests on the residuals of a GARCH model are given in Hsieh (1991), table 13, although they do not differ markedly from those corresponding to the standard normal

ambiguous, although rejection of the linear null cannot be entirely attributed to ARCH, since rejection only occurs for some ofthe series, and yet all are characterised by ARCH effects.

Using the bispectrum test, the null hypothesis of Gaussianity is rejected in all cases for all the series (raw, linearly, and GARCH filtered data), although to a lesser extent for the Canadian Dollar than for the others. This result confirms the non-normality of the currencies suggested by the simple Bera Jarque (1981) test displayed in table 1[8]. The null hypothesis of linearity is very rarely rejected, even at the 10% level, indicating a lack of nonlinear dependence in any of the series, at least of the form that can be detected by the bispectrum test. This result is similar to that of Barnett *et al*. (1994) and is unsurprising given that the bispectrum test has low power against ARCH alternatives, and that the power of the test is adversely affected by strong kurtosis in the underlying series.

When conducting the BDS test, one usually pre-filters the data through linear and ARCH-type filters to obtain new series. This has been viewed as both an advantage (since a stronger picture of the precise nature of the nonlinearity present in the data, if any, can be observed), and a disadvantage (because filtering through a possibly mis-specified linear, or worse nonlinear, model may alter the null distribution of the test statistic The BDS test has been extensively applied in the literature, and on the whole, there appears to be less activity in these Pound-denominated series than in those denominated in U.S. Dollars or in stock market data (c.f., for example, Hsieh, 1989 and Abhyankar *et al.*, 1995 respectively). Linear filtering with up to ten lags of the dependent variable and the removal of daily seasonality and anomalies associated with holidays

---

distribution.

[8]Although, as Nachane and Ray (1993) note, a frequency domain approach to Gaussianity testing may be preferable to the time domain approach of Bera and Jarque, since the latter assumes that

makes very little difference to the estimated values of the test statistics, indicating that there exists further dependence in the data, which must be of a nonlinear form. Furthermore, contrary to the results of Hsieh (1989), but in agreement with those of Yadiv *et al.* (1993), there appears to be in some cases, particularly the Swiss Franc, considerable further nonlinearity once the effects of ARCH have been removed from the data, although the estimated values of the BDS statistics are, for all the other series, much reduced.

6. Analysis, Conclusions, and Suggestions for Further Research

On the whole, although the null hypothesis of linearity is rejected by almost all the tests used in this study, and across almost all of the series, these results should be put into perspective. It is possible, and perhaps useful, to categorise observed nonlinear dependence in financial time series into two groups: volatility clustering effects, and all other forms of nonlinearity. It has been shown that foreign exchange markets are characterised by what can be represented by ARCH or GARCH processes, and it appears that ARCH dominates many of the nonlinearity tests used here (Granger, 1993), to the extent that other forms of nonlinearity are masked and difficult to uncover. It is therefore important to consider which of the tests used have power against which particular nonlinear alternatives, and specifically, which do or do not have power against ARCH (such as the bispectrum test). A rejection of a linear null in the case where the test has little power against ARCH would clearly be an interesting result, since it would indicate that nonlinear dependence of a form in addition to, or instead of (but perhaps having the same properties as) ARCH is present in the data.

---

each observation is drawn from an identical distribution, while the former makes no such assumption and allows each realisation to be drawn from a distinct distribution.

The usual method for assessing the power of specific tests is to undertake a Monte Carlo study, simulating data (with given stochastic properties) from a pre-specified nonlinear model under the alternative, and calculating the number of times the test correctly rejects the linear or I.I.D. null. This procedure has been followed by a number of authors, including Barnett *et al.* (1994)[9], Hsieh (1991), Liu *et al.*(1992), Ashley, Patterson and Hinich (1986), Ashley and Patterson (1989), and Lee *et al.*(1993). The most important common conclusion is that no test is uniformly most powerful (UMP) against all nonlinear alternatives. Perhaps the results of Hinich's bispectrum test and the BDS test are most telling. Since the bispectrum test results in virtually no rejections of linearity, it appears that there is little evidence that the data could accurately be described as a bilinear, single state nonlinear moving average or single state nonlinear autoregressive model alone.

To conclude, it is evident from the numerous tests undertaken above, that nonlinear dependence is a salient feature of financial time series data. This is hardly surprising, given the huge volume of literature indicating the presence of ARCH effects in many financial markets. What becomes clear, however, from tests such as BDS, which allow ARCH-nonlinearity to be filtered from the data is that there still exists a degree small degree of nonlinearity of an unknown form, once volatility clustering is removed. In this study, it is true only of the Swiss Franc, although a number of other studies declare a similar result (Yadiv *et al.*, 1993) in spite of the fact that the GARCH-(1,1) model appears to fit the data extremely well. It may be either that the GARCH models used in the various studies are mis-specified and hence do not capture all of the dependence in the squared residuals, or, more likely, it may be that further patterns are present in the data. It is possible that this additional structure may be described by a switching or more

[9]Their work cannot be considered a full Monte Carlo simulation, since they use only one set of data for each model, which is generated cleanly with no added noise.

complex threshold model. Interestingly, the hypothesis suggested in the introduction to this chapter that one may anticipate more action in the less frequently traded currencies is not borne out by the data, with large volume currencies such as the U.S. Dollar, Japansese Yen, and German Mark exhibiting considerably more (possibly potentially forecastable) dependence than the Austrian Schilling, Canadian Dollar and Danish Krone.

It appears that an *impasse* has been reached in the literature on nonlinearity testing. GARCH effects have been found to dominate all other forms of nonlinearity, but few rival or additional models have come to light. Three new approaches seem worthy of comment. First, the "windowing" approach and the new test of Hinich (1995)[10] may prove useful, in that the short-term structure of financial series has often been somewhat neglected given the tendency in the literature towards ever longer sample periods with associated assumptions of parameter constancy in order to satisfy the data requirements of the more complex tests. These tests may be more powerful against some nonlinear alternatives, but this is at the expense of masking a myriad of structures which may be present for shorter period of time. Second, the test developed by Kaplan (1994), based on continuity in phase space, may be worthwhile as a portmanteau equivalent to the BDS test, since in principle at least, it allows the user to test for whiteness against any specified linear or nonlinear alternative. Third, as suggested by Peel and Speight (1994), it may be possible that GARCH models in combination with, for example, the self exciting threshold autoregressive (SETAR) model of Tong (1983, 1990), or the bilinear model of Granger and Anderson (1978), may more fully describe the data.

References

[10]The test is a time domain analogue of the bispectrum test which requires only relatively small samples.

Abhyankar, A.H., Copeland, L.S. and Wong, W. (1995) Nonlinear Dynamics in Real-Time Equity Market Indices: Evidence form the U.K. *Economic Journal*, forthcoming

Akaike, H.(1974) A New Look at the Statistical Model Identification *IEEE Transactions on Automatic Control*, **AC-19**(6), 716-723

Akgiray, V. (1989) Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts *Journal of Business*, **62**(1), 55-80

Ashley, R.A. and Patterson, D.M. (1989) Linear Versus Nonlinear Macroeconomies: A Statistical Test *International Economic Review*, **30**(3), 685-704

Ashley, R.A., Patterson, D.M. and Hinich, M.J. (1986) A Diagnostic Test for Nonlinear Serial Dependence in Time Series Fitting Errors *Journal of Time Series Analysis*, **7**(3), 165-178

Ballie, R.T. and Bollerslev, T. (1989) The Message in Daily Exchange Rates: A Conditional-Variance Tale *Journal of Business and Economic Statistics*, **7**(3), 297-305

Barnett, W.A., Gallant, A.R., Hinich, M.J., Jungeilges, J.A., Kaplan, D.T., and Jensen, M.J. (1994) A Single-Blind Controlled Competition Between Tests for Nonlinearity and Chaos *Washington University in St. Louis, Department of Economics Working Paper Series no. 167*

Bera, A.K. and Jarque, C.M. (1981) An Efficient Large-Sample Test for Normality of Observations and Regression Residuals *Australian National University Working Papers in Econometrics* 40, Canberra

Beveridge, S. and Oickle, C. (1994) A Comparison of Box-Jenkins and Objective Methods for Determining the Order of a Non-seasonal ARMA Model *Journal of Forecasting*, **13**, 419-434

Bhargava, A. (1986) On the Theory of Testing for Unit Roots in Observed Time Series *Review of Economic Studies*, **53**, 369-384

Bleaney, M. and Mizen, P. (1993) Do Real Exchange Rates Really Follow a Random Walk? *University of Nottingham Discussion Papers in Economics* No. 93/9

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

Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1990) ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence, working paper, North Western University, Georgia Tech., and University of Arizona

Brock, W.A. (1986) Distinguishing Random and Deterministic Systems: Abridged Version *Journal of Economic Theory*, **40**, 168-195

Brock, W.A., Dechert, W.D. and Scheinkman, J.A. (1991) A Test for Independence Based on the Correlation Dimension *Department of Economics, University of Wisconsin at Madison, University of Houston, University of Chicago, and University of Madison*

Brock, W.A., Hsieh, D.A and LeBaron, B. (1991) *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence* M.I.T. Press, Reading, Mass.

Brock, W.A. and Sayers, C.L. (1988) Is the Business Cycle Characterised by Deterministic Chaos? *Journal of Monetary Economics*, **22**, 71-90

Corbae, D. and Ouliaris, S. (1988) Cointegration Tests of Purchasing Power Parity *The Review of Economics and Statistics*, **70**, 508-511

DeCoster, G.P. and Mitchell, D.W. (1991) Nonlinear Monetary Dynamics *Journal of Business and Economic Statistics* **9**(4), 455-461

Dickey, D.A. and Fuller, W.A. (1979) Distribution of Estimators for Time Series Regressions with a Unit Root *Journal of the American Statistical Association*, **74**, 427-431

Ellis, J. (1992) An Investigation of Nonlinearities and Chaos in Exchange Rates *Unpublished PhD Thesis, Penn State University*

Engle, R.F. (1982) Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation *Econometrica*, **50**(4) 987-1007

Fama, E. (1970) Efficient Capital Markets: A Review of theory and Empirical Work *Journal of Finance*, **25**, 383-417

Frank, M.Z., Gencay, R. and Stengos, T. (1988) International Chaos? *European Economic Review*, **32**, 1569-1584

Fuller, W.A. (1976) *Introduction to Statistical Time Series* Wiley, N.Y.

Granger, C.W.J. (1993) Strategies for Modelling Nonlinear Time-Series Relationships *The Economic Record*, **69**(206), 233-238

Granger, C.W.J. and Anderson, A.P. (1978) *An Introduction to Bilinear Time Series Models* Vandenhoecht und Ruprecht, Göttingen

Granger, C.W.J., White, H. and Kamstra, M. (1989) Interval Forecasting: An Analysis Based Upon ARCH-Quantile Estimators *Journal of Econometrics*, **40**, 87-96

Guillaume, D.M., Dacorogna, M.M., Davé, R.R., Müller, U.A., Olsen, R.B., and Pictet, O.V. (1994) From the Bird's Eye to the Microscope: A Survey of the New Stylised Facts of the Intra-Daily Foreign Exchange Markets *Mimeo Olsen & Associates Research Group*

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

Hinich, M.J. (1995) Testing for Dependence in the Input to a Linear Time Series Model *Journal of Nonparametric Statistics*, forthcoming

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

Hinich, M.J. and Patterson, D.M. (1995) Detecting Epochs of Transient Dependence in White Noise *Mimeo. University of Texas at Austin*

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

Hinich, M.J. and Patterson, D.M. (1985) Evidence of Nonlinearity in Daily Stock Returns *Journal of Business and Economic Statistics*, **3**(1), 69-77

Hochberg, Y. (1988) A Sharper Bonferroni Procedure for Multiple Tests of Significance *Biometrika*, **75**, 800-802

Hsieh, D.A. (1991) Chaos and Nonlinear Dynamics: Application to Financial Markets *The Journal of Finance*, **46**(5), 1839-1877

Hsieh, D.A. (1989a) Modelling Heteroscedasticity in Daily Foreign Exchange Rates *Journal of Business and Economic Statistics*, **7**(3), 307-317

Hsieh, D.A. (1989b) Testing For Nonlinear Dependence in Daily Foreign Exchange Rates *Journal of Business*, **62**(3), 339-368

Hsieh, D.A. (1988) The Statistical Properties of Daily Foreign Exchange Rates: 1974-1983 *Journal of International Economics*, **24**, 129-145

Keenan, D.M. (1985) A Tukey Nonadditivity-Type Test for Time Series Nonlinearity *Biometrika*, **72**(1), 39-44

LeBaron, B. (1988) Nonlinear Puzzles in Stock Returns *Mimeo., University of Chicago*

Lee, T., White, H. and Granger, C.W.J. (1993) Testing for Neglected Nonlinearity in Time Series Models: A Comparison of Neural Network Methods and Alternative Tests, *Journal of Econometrics*, **56**, 269-290

Liu, T., Granger, C.W.J. and Heller, W.P. (1992) Using the Correlation Exponent to Decide Whether an Economic Series is Chaotic *Journal of Applied Econometrics*, **7 Supplement**, S25-S39

Ljung, G.M. and Box, G.E.P. (1978) On a Measure of Lack of Fit in Time Series Models, *Biometrika* **65**(2), 297-303

Mayfield, E.S. and Mizrach, B. (1992) On Determining the Dimension of Real-Time Stock-Price Data *Journal of Business and Economic Statistics*

McLeod, A.I. and Li, W.K. (1983) Diagnostic Checking ARMA Time Series Models Using Squared-Residual Autocorrelations *Journal of Time Series Analysis*, **4**(4), 269-273

Nachane, D.M. and Ray, D. (1993) Modelling Exchange Rate Dynamics: New Perspectives From the Frequency Domain *Journal of Forecasting*, **12**(5), 379-394

Peel, D.A. and Speight, A.E.H. (1994) Testing for Nonlinear Dependence in Inter-War Exchange Rates *Weltwirtschaftliches Archiv: Review of World Economics*,**130**(2), 391-417

Perron, P. and Phillips, P.C.B. (1987) Does GNP have a Unit Root?: A Re-evaluation *Economics Letters*, **23**, 139-145

Phillips, P.C.B. (1987) Time Series Regression with a Unit Root *Econometrica*, **55**, 277-301

Phillips, P.C.B. and Perron, P. (1988) Testing for a Unit Root in Time Series Regression *Biometrika*, **75**(2), 335-346

Ramsey, J.B. (1969) Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis *Journal of the Royal Statistical Society B*, **31**(2), 350-371

Ramsey, J.B., Sayers, C.L. and Rothman, P. (1990) The Statistical Properties of Dimension Calculations Using Small Datasets: Some Economic Applications *International Economic Review*, **31**(4), 991-1020

Ramsey, J.B. and Yuan, H-J. (1990) The Statistical Properties of Dimension Calculations Using Small Datasets *Nonlinearity*, **3**, 155-176

Sargan, J.D. and Bhargava, A. (1983) Testing Residuals from Least Squares Regression for Being Generated by the Gaussian Random Walk *Econometrica*, **51**, 153-174

Scheinkman, J.A. and LeBaron, B. (1989) Nonlinear Dynamics and GNP Data in *Economic Complexity: Chaos, Sunspots, Bubbles and Nonlinearity - International Symposium in Economic Theory and Econometrics* (eds.) W.A. Barnett, J. Geweke, and K. Shell, Cambridge University Press, 213-227

Scheinkman, J.A. and LeBaron, B. (1989) Nonlinear Dynamics and Stock Returns *Journal of Business*, **62**(3), 311-337

Schwarz, G. (1978) Estimating the Dimension of a Model *Annals of Statistics*, **6**, 461-464

Sola, M. and Timmermann, A. (1994) Fitting the Moments: A Comparison of ARCH and Regime Switching Models for Daily Stock Returns *Birkbeck College, University of London, Department of Economics Discussion Paper*

Subba Rao and Gabr (1980) A Test for Linearity of Stationary Time Series *Journal of Time Series Analysis,* **1**, 145-148

Tong, H. (1990) *Nonlinear Time Series: A Dynamical Systems Approach*, Oxford University Press

Tong, H. (1983) *Threshold Models in Nonlinear Time Series Analysis* Springer-Verlag, N.Y.

White, H. (1990) Connectionist Nonparametric Regression: Multilayered Feedforward Networks Can Learn Arbitrary Mappings *Neural Networks,* **3**, 535-550

White, H. (1989) Some Asymptotic Results for Learning in Single Hidden Layer Feedforward Network Models *Journal of the American Statistical Association,* **84**, 1003-1013

Tsay, R.S. (1986) Nonlinearity Tests for Time Series *Biometrika,* **73**(2), 461-466

Willey, T. (1992) Testing For Nonlinear Dependence in Daily Stock Indices *Journal of Economics and Business,* **44**, 63-74

Yadav, P.K., Paudyal, K. and Pope, P.F. (1993) Nonlinear Dependence in Individual Stock Returns: Does Size, Trading Frequency or Market Microstructure Matter? *The Management School, Lancaster University Working Papers in Accounting and Finance, Working Paper No. 94/004*

Acknowledgements

I am grateful to Dee Dechert, Mel Hinich and Tae Lee for generously providing their software. I would also like to thank Olan Henry and Seth Greenblatt for comments on an earlier version of this paper and Simon Burke and Mel Hinich for helpful discussions, whilst absolving them from any remaining errors.

Appendix of Results

Table 1: Summary Statistics for Returns Data

Series	Mean	Variance	Skewness	Kurtosis	Normality Test statistic[11]	LB(10)[12]
A	-0.0188	0.2816	-0.5591	6.8080	10295.43	51.3
C	-0.0016	0.4203	-0.1188	2.9541	1899.69	28.6
D	-0.0080	0.2705	-0.2934	5.0801	5656.58	32.1
F	-0.0058	0.2465	-0.0346	7.7472	12982.55	37.0
G	-0.0181	0.2444	-0.6079	6.1814	8584.09	38.1
H	0.0001	0.4918	-0.0599	5.3209	6126.73	26.5
I	0.0103	0.3018	0.0549	10.3139	23011.01	13.8
J	-0.0280	0.4232	-0.4918	4.0117	3690.20	63.8
S	-0.0250	0.3532	-0.3014	5.2854	6120.96	19.9
U	-0.0079	0.4190	-0.0693	3.5162	2678.38	46.6

[11]Bera and Jarque (1981). The test statistic is distributed as a χ^2 (2) = 5.99 at 5% under the null.
[12]Where LB(10) is the Ljung Box (10) statistic. The test statistic is distributed as a χ^2 (10) = 18.3 at 5% under the null.

Table 2: Nonlinearity Tests on Residuals of Linear Models on Returns Data[13]

		Test			
Series	Lags	Tsay	ARCH(4)	McLeod and Li	RESET
A	1	1.10 (0.29)	290.38 (0.00)	604.96 (0.00)	0.39 (0.76)
C	1	0.80 (0.37)	146.84 (0.00)	472.34 (0.00)	0.57 (0.63)
D	0	2.17 (0.14)	123.16 (0.00)	462.60 (0.00)	1.83 (0.14)
F	1	0.01 (0.94)	168.98 (0.00)	438.73 (0.00)	2.12 (0.10)
G	1	7.41 (0.01)	210.34 (0.00)	638.78 (0.00)	8.95 (0.00)
H	0	0.26 (0.61)	244.11 (0.00)	794.86 (0.00)	7.46 (0.00)
I	1	20.46 (0.00)	306.20 (0.00)	806.04 (0.00)	7.22 (0.00)
J	0	4.75 (0.03)	211.85 (0.00)	547.89 (0.00)	1.63 (0.18)
S	0	2.70 (0.10)	492.76 (0.00)	1181.39 (0.00)	16.93 (0.00)
U	1	5.26 (0.02)	195.96 (0.00)	913.6 (0.00)	1.83 (0.14)

[13] *p*-values are given in parentheses.

Table 3: Results of White's Neural Network Test on Returns Data

		P-values (not ordered)						
Series	Number of Inputs[14]	1	2	3	4	5	Simple Bonferroni	Hochberg Bonferroni
A	6	0.0005	0.0008	0.0004	0.0011	0.0035	0.0020	0.0019
C	6	0.3734	0.2480	0.7209	0.5541	0.6741	0.9240	0.7209
D	5	0.0745	0.0404	0.0622	0.0333	0.0249	0.1243	0.0745
F	6	0.0296	0.0204	0.0176	0.0232	0.0202	0.0881	0.0296
G	6	0.0000	0.0000	0.0000	0.0001	0.0001	0.0002	0.0001
H	5	0.0523	0.0235	0.0694	0.0849	0.1043	0.2615	0.2086
I	6	0.0002	0.0005	0.0371	0.0020	0.0005	0.0008	0.0008
J	5	0.0378	0.0477	0.0286	0.0382	0.0400	0.1428	0.0477
S	5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
U	6	0.0058	0.0700	0.5465	0.0118	0.0643	0.0291	0.0291

Table 4: Results of Bispectrum Tests on Returns Data

Series	framesize = 71			framesize = 60		
	Gaussianity H	Linearity, 0.80 Fractile	% Significant bispectral estimates	Gaussianity H	Linearity, 0.80 Fractile	% Significant bispectral estimates
A	24.68	-1.386	9.88	27.62	-1.659	11.56
C	11.02	-1.958	5.86	9.71	-3.786	7.11
D	20.43	1.383	6.17	21.96	-0.733	8.89
F	20.03	-1.639	8.33	22.10	-0.023	9.33
G	25.37	-0.067	8.33	28.01	-0.525	9.78
H	17.74	-0.138	7.41	21.03	2.626*	10.67
I	51.79	-2.475	11.42	52.85	-3.181	12.00
J	21.20	-3.204	7.72	16.30	-2.313	10.22
S	33.89	-2.011	11.73	27.39	0.464	12.44
U	17.60	-2.402	8.64	20.23	-1.468	8.44

[14]The inputs to the neural network are the five dummy variables outlined in the text, together with the number of lags of the dependent variable chosen by SBIC.

Table 5: BDS Test Results for Returns Data

	Embedding Dimension								
Series	2	3	4	5	6	7	8	9	10
A	20.108	17.611	14.984	12.810	11.249	10.231	9.375	8.519	7.794
C	2.232	2.960	2.191	6.972	8.739	9.847	11.491	12.243	12.492
D	2.194	2.223	2.061	2.064	2.124	2.186	2.236	2.262	2.274
F	0.096	0.115	0.096	0.088	0.047	0.031	0.015	0.012	0.015
G	0.913	1.321	1.343	1.381	1.382	1.586	1.762	1.871	1.971
H	1.366	1.509	1.999	3.030	3.751	6.069	7.615	8.559	9.100
I	1.328	1.466	1.452	1.458	1.385	1.371	1.356	1.330	1.306
J	0.596	1.751	1.898	2.585	2.720	3.476	3.853	4.024	4.197
S	4.991	13.019	14.385	14.670	14.402	13.950	13.490	13.044	12.619
U	1.732	2.269	2.906	6.048	8.273	9.684	10.667	11.091	11.192

Table 6: BDS Test Results for Standardised Residuals of a GARCH-(1,1) Model

	Embedding Dimension								
Series	2	3	4	5	6	7	8	9	10
A	3.587	3.050	2.768	2.472	2.207	1.975	1.772	1.591	1.430
C	-2.836	-3.732	-0.323	0.598	0.667	0.393	-0.026	-0.499	-0.997
D	-0.196	-0.245	0.094	0.231	0.348	0.387	0.405	0.435	0.440
F	-0.048	-0.090	0.202	0.917	0.770	1.200	1.415	1.517	1.551
G	0.314	0.397	0.745	1.236	1.426	1.578	1.672	1.719	1.728
H	-0.720	0.297	0.648	0.963	1.106	1.427	1.350	1.668	2.242
I	0.051	0.275	0.407	1.169	1.548	0.910	0.731	1.161	1.793
J	-0.055	-0.167	-0.200	-0.177	-0.167	-0.164	-0.147	-0.152	-0.158
S	2.137	6.446	7.036	7.787	7.719	7.367	7.900	8.432	8.958
U	-1.594	-2.161	1.566	2.473	3.596	3.935	3.916	3.721	3.429