Tests of non-linearity using LIFFE futures transactions price data

Article

Accepted Version

Tests of non-linearity using
LIFFE futures transactions price data

by

O. ap Gwilym, C. Brooks, A. Clare and S. Thomas[*]

Abstract

This paper presents and implements a number of tests for non-linear dependence and a test for chaos using transactions prices on three LIFFE futures contracts: the Short Sterling interest rate contract, the Long Gilt government bond contract, and the FTSE-100 stock index futures contract. While previous studies of high frequency futures market data use only those transactions which involve a price change, we use all of the transaction prices on these contracts whether they involve a price change or not. Our results indicate irrefutable evidence of non-linearity in two of the three contracts, although we find no evidence of a chaotic process in any of the series. We are also able to provide some indications of the effect of the duration of the trading day on the degree of non-linearity of the underlying contract. The trading day for the Long Gilt contract was extended in August 1994, and prior to this date there is evidence of only a linear structure in the return series. However, after the extension of the trading day we do find evidence of a non-linear return structure.

13 February, 2014

Keywords: LIFFE; GARCH; non-linearity tests; chaos; futures contracts; transactions data.

JEL #: C12, C52, G13

1. Introduction

A number of recent studies have examined the possibility that economic data are non-linear. These studies question whether linear models are an adequate characterisation of the data generating process, and attempt to determine whether non-linear time series models might be worthy of further investigation as forecasting tools. The discovery of a non-linear process which is chaotic however, has even more important implications for the modelling and prediction of the series under investigation[1]. The presence of chaos would mean that while long-term predictability is impossible, short-term forecasting using a non-linear model might be possible if we could obtain a reasonable approximation to the data generating process.

Previous literature in this area can be conveniently categorised into tests using (a) macro-economic data[2] and (b) financial data[3]. Because of a lack of data points and measurement error, studies which have used macroeconomic time series have not been particularly successful in identifying non-linearities (see Ramsey *et al.* (1990)). The literature employing financial data has been far more voluminous where overwhelming evidence exists of a non-linear, GARCH-type structure across many types of financial time series, *e.g.* stocks, bonds, futures, foreign exchange. However, in this paper we investigate whether this is the most appropriate way to represent the non-linear dependence which is found in financial time-series. To address this question we apply a number of tests for non-linearity using the intra-day *transactions data* on three futures contracts traded on LIFFE - the FTSE-100 stock index, the Short Sterling 3-month interest rate and the Long Gilt government bond contracts.

---

[1] See the collection by Dechert (1996) for a comprehensive treatment of the issues involved in testing for and modelling chaotic processes.
[2] For example see Brock and Sayers (1988), DeCoster and Mitchell (1991), Frank *et al.* (1988), or Lee *et al.* (1993).

The data set which we employ in this paper contrasts with the data used in intra-day studies of US futures contracts, for example see Ekman (1992), and Ederington and Lee (1993, 1995). These studies use *price change data* because price observers at markets such as the Chicago Mercantile Exchange (CME) record only those transactions which involve a change in contract price. A major problem with price change data is that a price change sequence will be either a reversal or a continuation (see Gosnell (1995) for a detailed discussion). As reversals are generally more common (between 60% and 90% in the Ederington and Lee (1995) data set), such data will exhibit a high degree of negative serial correlation. This could plausibly lead to spurious rejections of *iid* in econometric tests. Ederington and Lee (1995) mitigate this problem by smoothing the data (averaging the last two prices), which is successful in removing the negative serial correlation, but creates a new artificial structure in the data which may cause misleading inferences. In our transactions data set, reversals account for a much smaller proportion of price sequences, and consequently there is no significant negative serial correlation in the data. The significant difference in the nature of the data series used in this study therefore provides a motivation for re-considering the application of tests for non-linearity and chaos to high frequency financial return series.

The rest of this paper is organised as follows. In Section 2 we present preliminary statistics and describe out data set. In Section 3 we describe the tests for non-linear dependence and chaos which we employ here. We present the results in Section 4, and finally we provide a summary of our findings and some concluding remarks in Section 5.

---

[3] For example see Abhyankar *et al.* (1995), Hsieh (1989b), Scheinkman and LeBaron (1989a, 1989b), Mayfield and Mizrach, (1992), or Willey (1992)).

2. Data and Descriptive Statistics

We use transactions price data on three futures contracts traded at the London International Financial Futures and Options Exchange (LIFFE) - the FTSE-100 stock index, the Short Sterling 3-month interest rate and the Long Gilt government bond contracts - for the period 24 January 1992 to 30 June 1995. Pit trading in the Short Sterling 3-month interest rate futures contract is from 08:05-16:05GMT while the FTSE-100 stock index futures market opens from 08:35-16:10GMT. Between the beginning of our sample (24 January 1992) and 31 July 1994, the Long Gilt contract traded between the hours of 08:30-16:15GMT, but owing to its increasing popularity the market open was changed to 08:00GMT after that date. Therefore, in the empirical results which are presented in Section 3, we examine this series for the whole sample and for the two sub-samples which relate to the different trading hours. We are thus able to provide some indications of the effect of the duration of the trading day on the degree of non-linearity of the underlying series.

All three contracts are heavily traded in an open outcry system similar to that commonly used in US financial futures markets such as the Chicago Mercantile Exchange (CME). Client orders are time stamped by floor brokers and passed on to pit traders. On completion of an order, details are filled in on the trader's card, which is then time stamped. All trades are matched at the end of the day. The source of our data is the **LIFFE Time and Sales Data**, which contains information on the time to the nearest second, contract type, commodity code, delivery month, price, transaction code (bid, ask or trade), and traded volume. These are transaction prices captured by physical pit observation. The observers record every price, and hence the data set contains every price reported and not just price changes.

The transactions data which we employ in this paper contrasts with the price change data employed by Ederington and Lee (1995). Reversals account for between 60 and 90% of the observations in Ederington and Lee's (1995) data set, therefore such data will exhibit a high degree of negative serial correlation. In contrast to these figures the transactions price data which we use in this paper accounts for a much smaller proportion of the observations. For the FTSE-100 contract, percentage reversals, continuations and no price changes for the period are 24.5%, 13.3% and 62.2% respectively; for the Short Sterling contract the proportions are 17.3%, 3.7% and 79.0% respectively; and for the Long Gilt contract the proportions are 23.4%, 8.8% and 67.8% respectively.

To create a continuous series for the analysis, we use prices for the nearby contract but switch to the second nearby contract when its trading volume exceeds that of the nearby (see Ederington and Lee (1993, 1995)). For the Short Sterling and Long Gilt contracts the switch occurs, on average, 20 trading days before expiration of the nearby contract. For the FTSE-100 contract the switch occurs on the last trading day of the nearby contract. Hinich and Patterson (1985) show that non-stationarity in time-series data may cause a spurious rejection of linearity, therefore to avoid this problem in the tests which follow we transform the raw transactions price data into returns. The trading day is split into five-minute intervals, where the last observation in each interval is used to calculate the return on each of the series

$$R_i = \ln (P_i / P_{i-1}) \tag{1}$$

where R_i is the log return of the ith time observation. Summary statistics for the returns are presented in Table 1. All of the series show strong signs of non-normality; this departure from normality is primarily due to leptokurtosis rather than skewness. This is clearly a salient feature of most financial data (see for example Hsieh (1988, 1989a)).

3. A Description of the Tests

Before discussing the formulation of the tests for non-linearity in detail, it is worth giving a broad definition of what we mean by "non-linearity". The definition of nonlinearity employed here is due to De Lima (1997, p253), that a series with no nonlinearity has innovations under a standard Wold decomposition that are either *iid* or a martingale difference sequence.

In this section we provide a brief description of each test for non-linearity used in this paper, together with the null distribution with which the test statistic is to be compared. In addition to the tests described here, we also use Ramsey's (1969) RESET test and Engle's (1982) LM test for ARCH. These tests are not described here since they are in common usage and should therefore be familiar to the majority of readers.

3.1. Tsay's Test

Tsay's (1986) test is a more powerful 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 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. The second step entails running a multivariate regression of Z_t on a constant and p lags, saving the residuals, $\hat{v}_t$. The test statistic is given by

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

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

3.2. The McLeod and Li Test

The McLeod and Li (1983) test is effectively a Ljung-Box test for autocorrelation of the squared residuals. In practice, the test is computed as follows. In undertaking the test one must 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

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

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

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

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

which is asymptotically $\chi^2(p)$ if the $\hat{u}_t$ are independent. 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 non-linear alternatives.

3.3. White's Neural Network Test

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

$$o = \tilde{x}'\alpha + \sum_{j=1}^{q} \beta_j \psi(\tilde{x}'\gamma_j) \qquad (7)$$

where $\tilde{x}$ is the input vector, γ_j are the connection strengths or weights, β_j are hidden to output weights, and ψ is an "activation" or "squashing" function, (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 test statistic is given by

$$M_T = (T^{-1/2} \sum_{t=1}^{T} \varphi_t \hat{e}_t) \hat{W}_T^{-1} (T^{-1/2} \sum_{t=1}^{T} \varphi_t \hat{e}_t) \qquad (8)$$

where $\hat{e}_t$ are the estimated residuals of the linear model, $\varphi_t = (\psi(\tilde{x}_t'\Gamma_1),...,\psi(\tilde{x}_t'\Gamma_q))'$ 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 $T \to \infty$. 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 $\alpha = \min_{i=1,\dots,k} (m-i+1)\, p_i$, so that H_0 is rejected if there exists an *i* such that $p_i \leq \alpha/(m-i+1)$, i=1,...,k.

3.4. 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) function is defined as [4]

$$C_{xxx}(r,s) = \mathrm{E}[x(t+r)x(t+s)x(t)] \tag{9}$$

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

$$B_{xxx}(f_1,f_2) = \sum_{r=-\infty}^{\infty}\sum_{s=-\infty}^{\infty} C_{xxx}(r,s)\, \exp[-2\pi i(f_1 r + f_2 s)] \tag{10}$$

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

$$S_{xx}(f) = \sum_{s=-\infty}^{\infty} C_{xx}(s)\, \exp[-2\pi i f s] \tag{11}$$

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

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

$$X^2(f_1,f_2) = \frac{\left| B_{xxx}(f_1,f_2) \right|^2}{S_{xx}(f_1)S_{xx}(f_2)S_{xx}(f_1+f_2)} = \frac{\mu_3^2}{\sigma_\varepsilon^6} \tag{12}$$

where $\sigma_\varepsilon^2 = \mathrm{E}(\varepsilon_t^2)$ and $\mu_3 = \mathrm{E}(\varepsilon_t^3)$. The modulus of the bispectrum, B_{xxx}, is taken since in general it is a complex number. Ashley *et al.* (1986, p.174), 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

$$S = 2\sum_m\sum_n \left| X(f_m, f_n) \right|^2 \tag{13}$$

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. Hinich and Patterson (1985) use a frame size equal to the square root of the number of observations, and this is the value used in this study.

3.5. The BDS Test

The test of Brock, Dechert and Scheinkman (1987, now revised and published as Brock, Dechert, Scheinkman and LeBaron, 1996) 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,

---

[4] Barnett *et. al.* (1997) point out that the cumulants are the coefficients of the power series expansion of the

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

$$C_m(\varepsilon) = \frac{1}{(T-m+1)(T-m)}\sum_{\forall t,s} I_\varepsilon(x_t^m x_s^m)$$ (14)

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. 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

$$W_{m,T}(\varepsilon) = T^{1/2}\frac{[C_T(\varepsilon) - C_1(\varepsilon)^m]}{\sigma_{m,\tau}(\varepsilon)}$$ (15)

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

$K(\varepsilon)$ is estimated by $\quad K(\varepsilon) = \dfrac{6\sum_{t<s<r} h_\varepsilon(x_t^m, x_s^m, x_r^m)}{[(T-m+1)(T-m)(T-m-1)]}\quad$ 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 determinism 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 BDS test is a two sided test so that, in general, rejection of

---

logarithm of a distribution, rather than of the <u>level</u>, as is the case for the moments.

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. The null distribution is changed when filtering through a GARCH model is undertaken[5]. 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 would appear to be more likely than less. In the results which we present below, AR(10) and EGARCH models are used as pre-filters in this study following Hsieh (1991).

3.6. Testing for Chaos using the Largest Lyapunov Exponent

Varying definitions of chaos can be found in the literature, but the definition which will be used here is that a system is chaotic if it exhibits *sensitive dependence on initial conditions* (SDIC). This is a definition which is frequently used, although many others are possible (see Brock *et al.* (1991)). The concept of SDIC embodies the fundamental characteristic of chaotic systems that if an infinitesimal change, *δx(0)* is made to the initial conditions, then *δx(t)*, the corresponding change iterated through the system until time *t*, will grow exponentially with *t* (Ruelle (1990)). We can write

$$|\,\delta x(t)\,| \approx |\,\delta x(0)\,|\,e^{\lambda t}$$

The largest Lyapunov exponent (LE) measures the rate at which information is lost from a system, and is usually given in units of base 2, so that the measure can be interpreted as the information loss in bits per iteration. A positive largest Lyapunov exponent implies sensitive dependence, and therefore that we have evidence of chaos. This has important implications for the predictability of the underlying system, since the fact that all initial conditions are in practice estimated with some error (either measurement error or exogenous noise), will imply

---

[5] New critical values for tests on the residuals of an EGARCH model are given in Hsieh (1991, Table 13), although they do not differ markedly from those corresponding to the standard normal distribution.

that long term forecasting of the system is impossible as all useful information is likely to be lost in just a few iterations.

Arguably the only test explicitly formulated for chaos is the computation of the largest Lyapunov exponent. The spectrum of Lyapunov exponents can be defined as follows (see Wolf *et al.* (1985)). Consider an infinitesimally small hypersphere of radius ε. If we monitor the evolution of the sphere, it will become deformed into an ellipsoid as the system evolves over time. The Lyapunov exponent is then measured by the extent of the deformation, and is given by

$$\lambda_i = \lim_{t \to \infty} \lim_{\varepsilon(0) \to 0} \left[\frac{1}{T} \log_2 \left(\frac{\varepsilon_i(t)}{\varepsilon_i(0)} \right) \right] \tag{16}$$

where $\varepsilon_i(t)$ is the length of the ith principal axis of the ellipsoid at time t. In this paper, two separate algorithms for the calculation of Lyapunov exponents are employed. The first is an implementation of the Wolf *et al.* (1985) algorithm, and the second is a more recent technique due to Dechert and Gencay (1990, 1992). The Wolf *et al.* (1985) algorithm was the first method proposed for estimating Lyapunov exponents in time series data. However, only the largest exponent is calculated, and a number of authors (for example, Brock and Sayers, 1988) have found that the results of the estimation are highly sensitive to noise, which is particularly problematic in economic data where noise is more prevalent and data series are typically much shorter than in the physical sciences. The results given are for a delay time of one and are given in base 2; a largest Lyapunov exponent in base 2 can be interpreted as the loss of information in bits per iteration.

The new technique of Gencay and Dechert (1990, 1992) seems more promising in that, potentially, the whole spectrum of Lyapunov exponents can be estimated. Furthermore,

according to simulations on known chaotic data sets by the authors of the test (Dechert and Gencay (1992)), the algorithm is more powerful in the presence of noise than the earlier technique. The method uses a technique similar to that of Ellner *et al.* (1991), in that, unlike the Wolf *et al.* (1985) method which directly finds similar pairs of state vectors within the series and estimates how the subsequent trajectories diverge, the new procedures use Jacobian methods. These estimate the exponents through the intermediate step of estimating the individual Jacobian matrices. Using the terminology of Nychka *et al.* (1992), let $\hat{J}_t$ be the estimate of the Jacobian and $\hat{T}_m = \hat{J}_m ... \hat{J}_1$. The estimate of the Lyapunov exponents is given by

$$\lambda = \frac{1}{2m} \log \left| \hat{\upsilon}_1(m) \right| \qquad (17)$$

where $\hat{\upsilon}_1(m)$ is the largest eigenvalue of $(\hat{T}_m{}' \hat{T}_m)^{m/2}$. In practice, the method of Gencay and Dechert uses a single hidden layer feed-forward neural network to model the dynamics of the series and the spectrum of Lyapunov exponents are then calculated from the derivative matrices of the network models. With this technique, the user does not have to choose a value for the delay time (usually denoted τ), but one does have to select the number of inputs to the network (equivalent to the embedding dimension) and the number of hidden units in the intermediate layer, N. The inputs were selected as own lagged values of the series from *t-1* to *t-m*, where *m* is the number of inputs. The network is given by

$$\hat{x}_{N,m}(X; \beta, w, b) = \sum_{j=1}^{N} \beta_j \phi \left(\sum_{i=1}^{m} w_{ij} Z_i + b_j \right) \qquad (18)$$

where $\hat{x}$ is a vector of fitted values, Z is the input, β represents the hidden to output weights, and w and b represent the input to hidden weights. Let

$$x_t^m = (x_{t+m-1}, x_{t+m-2}, ..., x_t) \qquad (19)$$

The multivariate non-linear least squares minimisation problem is then given by

$$\min_{\beta,w,b} \sum_{t=0}^{T-m-1} [x_{t+m} - \hat{x}_{N,m}(x_t^m; \beta, w, b)]^2$$

and the activation function for the hidden layer is the sigmoid

$$\phi(p) = \frac{1}{1 + \exp(-p)} \qquad\qquad (20)$$

In our estimation we set the number of inputs at 1[6] and the number of hidden layers were allowed to vary from 1 to 10. These values were severely constrained by available CPU time, since the estimation is extremely data intensive[7].

4. Results

In this section we present the results of the non-linearity tests outlined in Section 3 using the transactions prices on three LIFFE futures contracts. These results are given in Tables 2 to 10. However, given the change in trading arrangements which occurs in the Long Gilt futures contract we discuss the results for the FTSE-100 and Short Sterling contracts in section 4a and the results for the Long Gilt futures contract in Section 4b.

4a. The LIFFE FTSE-100 and Short Sterling Futures Contracts

For the FTSE-100 and Short Sterling futures contracts, there is a general degree of agreement between the tests, which are very different in formulation. We report the results of the Tsay, McLeod and Li, RESET and Engle's ARCH tests in Table 2. The *Tsay test* leads to the rejection of the null of linearity for both the FTSE-100 and Short Sterling contracts. The *McLeod and Li test* always leads to the rejection of linearity at all significance levels for both

---

[6] We varied the number of inputs, *m* (equivalent to the embedding dimension) from 1 to 6, and the results are qualitatively identical to those presented here.
[7] To calculate the spectrum of Lyapunov exponents with just 1 input and 1-10 hidden layers took over 24 hours of C.P.U. time per series on a Sun SparcCentre 2000 with 8 processors. Thus it is impractical to use bootstrapping in this case to construct confidence intervals or to undertake tests of significance.

series. It also 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 hence is unable to discern any further or alternative structure so that *Engle's LM test for ARCH* may be used as a simpler alternative.

Strong rejections of linearity are also observed using the *Bispectrum test* shown in Table 3, although the cause of rejection cannot be attributed to ARCH effects since, by definition, the bispectrum test cannot detect dependence of this kind. This is an interesting result since it shows that non-linearity in the returns of the Short Sterling futures contract cannot be entirely attributed to ARCH effects, and that there must be some additional or alternative structure present in the data. The results of *White's neural network test* in Table 4 indicates that we should strongly reject the linear null for both of these contracts.

The *BDS test* applied to the unfiltered returns series, shown in Table 5 for the FTSE-100 contract and Table 6 for the Short Sterling contract, again leads us to reject the *iid* null hypothesis. The test statistics are still well in excess of their critical values when linear dependence has been removed, although they are somewhat reduced indicating that much of the "action" in the series must be of a non-linear form. Furthermore, the *BDS* test statistics based on the residuals of the E-GARCH filter are still in excess of their critical values at the 1% level, indicating that not all of the non-linearity in the series can be attributed to volatility pooling effects. In general then there seems to be extremely strong evidence of non-linearity in the returns of both of these contracts

We now turn to the test for a chaotic process. Frank *et al.* (1988), Dechert and Gencay (1992) and Blank (1991) compute the largest Lyapunov exponent of various financial and

17

macroeconomic time series. Of these, only Blank finds a positive exponent in agricultural futures price series which is robust to changes in the value of the user-adjustable parameters, such as the embedding dimension and delay time. This conjecture is, however, based on an average of only 336 observations, and uses the direct and less stable algorithm of Wolf *et al.* (1985). Our results in Table 10 (using the more robust indirect measure) show that all of the estimated Lyapunov exponents are negative, indicating no evidence of chaos in the returns of the FTSE-100 and Short Sterling contracts.

4b. The LIFFE Long Gilt Futures Contract

Although the *Tsay test* leads to the rejection of the null of linearity for both the FTSE-100 and Short Sterling contracts, this null is not rejected for the returns on the Long Gilt contract for the full sample (see Table 2). There is however some evidence of non-linearity after the trading day is extended on this contract after 1st August 1994. The *McLeod and Li* and Engle ARCH(4) tests lead to the rejection of the null of linearity for all three of the sample periods for this contract. The *RESET test* provides evidence of non-linearity both for the full sample and for the post July 1994 sample, but not for the pre-July 1994 sample. Thus Table 2 provides some evidence to suggest that the extension of the trading day for the Long Gilt contract lead to the return series for this contract becoming (more) non-linear.

However, when we employ the *Bispectrum test* we obtain a strong rejection of linearity for all three samples (Table 3). *White's neural network test* (Table 4), gives a slightly different picture: the linear null is not rejected for the Long Gilt series as a whole, or for the pre-July 1994 sample, but there does again appear to be evidence of non-linear dependence after the extension of the Long Gilt futures contract trading day. The BDS test for the Long Gilt Series indicates that for the full sample and both the sub-samples, there is strong evidence against

the iid null hypothesis. The value of the test statistic falls as the data are filtered through an autoregressive and then an EGARCH filter, but there remains further dependence in the data that cannot be explained with reference to these models. The results for the full sample (Table 7) compared with the two sub-samples (Tables 8 and 9) show considerably greater average values of the BDS statistic for the former than the latter, which may be attributed to the structural break in the former which is a further source of "non-linearity" causing the test to reject.

Table 10 shows that for the Gilt contract, both before and after the extension of the trading day, there is no evidence of chaos since the largest Lyapunov exponent is, in all cases, estimated to be negative.

Thus, to summarise, there is evidence of non-linearity in the Gilt series as a whole, and for the series after the extension of the trading day in August 1994, but not for the Gilt sample prior to the extension, except when we consider the bispectrum test which rejects linearity for all the samples.

5. Conclusions

On the whole, although the null hypothesis of linearity is rejected by almost all the tests used in this study for two of the three series, these results should be put into perspective. It is possible, and perhaps useful, to categorise observed non-linear dependence in financial time series into two groups: volatility clustering effects, and all other forms of non-linearity. It has been shown that many asset return series are characterised by what can be represented by ARCH or GARCH processes, and it appears that ARCH dominates many of the non-linearity tests used here (Granger (1993)), to the extent that other forms of non-linearity are masked

and difficult to uncover. The very nature of portmanteau non-linearity testing makes the tests prone to mis-classify nonstationarities, seasonalitites, and other departures from iid as non-linearities. De Lima (1997) has also shown that the valid application of some of the tests (in particular the McLeod and Li and bispectrum tests) is dependent upon the existence of finite higher order moments (8 for McLeod and Li, 6 for the bispectrum test). This condition is not fulfilled for many data series in economics and finance (see Loretan and Phillips, 1993). Fortunately, De Lima (1997) also demonstrates that the BDS test is not sensitive to moment condition failure, and the results of our BDS tests are highly indicative of non-linearity.

It is therefore important to consider which of the tests have power against which particular non-linear alternatives, and specifically, which do or do not have power against ARCH (such as the bispectrum test). A rejection of the linear null using the bispectrum test, as is the case here, is an important result, since it indicates that non-linear dependence of a form in addition to, or instead of (but perhaps having the same properties as) ARCH is present in the data. As Barnett *et al*. (1997) argue, rejection of the iid null by the Hinich test provides very strong evidence against the null hypothesis since the test is very conservative and can often fail to reject even if the data generating process is truly non-linear. It is also possible that the ARCH-type non-linearity is caused by volatility pooling around the open and close of the market as demonstrated by Brock and Kleidon (1992). The additional structure may be described by a switching or more complex threshold model.

However, an important caveat should be mentioned here, since the GARCH model has been "imposed" on the data and the residuals of the GARCH fit are tested for non-linearity. A finding of no further non-linearity is often taken to imply that the GARCH model is "accepted", although a large number of other parametric models fitted to data of this type

(with little structure in mean but slowly decaying autocorrelation in the conditional second moment) could plausibly lead to the same conclusion[8]. This procedure might also have implications for predictability. Brock *et al*. (1992), for example, use bootstrap specification tests to show that a number of models in the GARCH family incorrectly predict volatility following the buy / sell signals of certain technical trading rules.

To conclude, it is evident from the numerous tests undertaken above, that non-linear dependence is a salient feature of the financial time series data used in this paper. The results of applying the tests for non-linearity to the long Gilt series may also serve as a warning against using the tests on series which experience fundamental changes in their properties during the sample period. The extension of the trading day has altered the properties of the series, and has, more specifically, led to an increase in the extent of non-linearity in the data. However, the behaviour of the tests in the face of such a structural change is not known. Thus the result of applying the test to the entire sample may give an unpredictable and complex combination of the results from the two (or more) sub-samples and could considerably reduce the accuracy of conditional forecasts. In particular, for the BDS test, it is likely that a structural break would lead to a stronger rejection of the null. Brock *et al*. (1991) show that even two separate sets of *iid* observations with different means or variances will cause the test to reject if the two data series are spliced together. The result of some of the tests is to reject for one of the sub-samples and for the series as a whole; while for other tests, it appears that the effect of adding a sub-sample which does not display evidence of non-linearity to one that does, is to dilute the non-linear structure so that overall no evidence against the linear null is apparent. More research employing Monte Carlo simulations is required to quantify this effect for the individual tests.

---

[8] We are grateful to an anonymous referee for pointing this out.

References

Abhyankar, A.H., Copeland, L.S. and Wong, W. (1995) Non-linear Dynamics in Real-Time Equity Market Indices: Evidence form the U.K. *Economic Journal* 105, 864-888

Ashley, R.A., Patterson, D.M. and Hinich, M.J. (1986) A Diagnostic Test for Non-linear 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, R.A., Hinich, M.J., Jungeilges, J.A., Kaplan, D.T., and Jensen, M.J. (1997) A Single-Blind Controlled Competition Between Tests for Non-linearity and Chaos *forthcoming in Journal of Econometrics*

Blank, S.C. (1991) "Chaos" in Futures Markets ? A Non-linear Dynamical Analysis *The Journal of Futures Markets,* **11**(6), 711-728

Booth, G.G., Hatem, J.J. and Mustafa, C. (1990) Are German Stock Returns Chaotic? *Presented to the Institute for Quantitative Investment Research, 7-9 October, 1990, Cambridge, U.K.*

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

Brock, W.A., Dechert, W.D., Scheinkman, J.A. and LeBaron (1996). A Test for Independence Based on the Correlation Dimension *Econometric Reviews* 15, 197-235

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

Brock, W.A., and Kleidon , A.W. (1992) Periodic Market Closure and Trading Volume - A Model of Intra-day Bids and Asks *Journal of Economic Dynamics and Control* 16, 451-489.

Brock, W.A., Lakonishok, J. and LeBaron (1992) Simple Technical Trading Rules and the Stochastic Properties of Stock Returns *Journal of Finance* 47(5), 1731-1764

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

Dechert, W.D. (1996) *Chaos Theory in Economics: Methods, Models and Evidence*, Volume 66 of International Library of Critical Writings in Economics, Edward Elgar, cheltenham, UK.

Dechert, W.D. and Gencay, R. (1992) Lyapunov Exponents as a Nonparametric Diagnostic for Stable Analysis *Journal of Applied Econometrics,* **7** Supplement, S41-S60

Dechert, W.D. and Gencay, R. (1990) Estimating Lyapunov Exponents with Multilayer Feedforward Network Learning *mimeo, Dept. of Economics, University of Houston*

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

De Lima, P.J.F. (1997) On the Robustness of Non-linearity Tests to Moment Condition Failure *Journal of Econometrics* 76, 251-280

Ederington, L. H., and Lee, J.H. (1995) How Markets Process Information: News Releases and Volatility, *Journal of Finance*, **48,** 1161-1191.

Ederington, L. H., and Lee, J.H. (1993) The Short Run Dynamics of the Price Adjustment to New Information, *Journal of Financial and Quantitative Analysis*, **30**, 117-134.

Ekman, P. D. (1992) Intraday Patterns in the S&P500 Index Futures Market, *Journal of Futures Markets*, **12**, 365-382.

Ellner, S., Gallant, A.R., McCaffrey, D. and Nychka, D. (1991) Convergence Rates and Data Requirements for Jacobian-Based Estimates of Lyapunov Exponents From Data *Physics Letters A,* **153**(6,7) 357-363

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

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

Frank, M.Z. and Stengos, T. (1988) Some Evidence Concerning Macroeconomic Chaos *Journal of Monetary Economics,* **22**, 423-438

Gosnell, T. F. (1995) The Distribution of Reversals and Continuations and Tests for Intraday Market Efficiency, *Journal of Business Finance and Accounting*, **22**, 225-243.

Granger, C.W.J. and Teräsvirta, T. (1993) *Modelling Non-linear Economic Relationships* Oxford University Press

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. (1989) Evidence of Non-linearity in the Trade-by-Trade Stock Market Return Generating Process, in *Economic Complexity: Chaos, Sunspots, Bubbles and Non-linearity - 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 Non-linearity 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 Non-linear 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 Non-linear 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 Non-linearity *Biometrika,* **72**(1), 39-44.

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

Loretan, M. and Phillips, P.C.B. (1993) Testing the covariance Stationarity of Heavy-tailed Time Series: An Overview of the Theory with Applications to Several Financial Datasets *Journal of Empirical finance* 1, 211-248

Mayfield, E.S. and Mizrach, B. (1992) On Determining the Dimension of Real-Time Stock-Price Data *Journal of Business and Economic Statistics* **10**(3), 367-374

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.

Nychka, D., Ellner, S., Gallant, A.R. and McCaffrey, D. (1992) Finding Chaos in Noisy Systems *Journal of the Royal Statistical Society B,* **54**(2), 427-449.

Ott, E., Grebogi, C. and Yorke, J. (1990) Controlling Chaos *Physical Review Letters,* **64**(11), 1196-1199.

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.

Ruelle, D. (1990) Deterministic Chaos: The Science and the Fiction *Proceedings of the Royal Society of London A,* **427**, 241-248

Scheinkman, J.A. and LeBaron, B. (1989a) Non-linear Dynamics and GNP Data, in *Economic Complexity: Chaos, Sunspots, Bubbles and Non-linearity - 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. (1989b) Non-linear Dynamics and Stock Returns, *Journal of Business,* **62** (3), 311-337.

Shinbrot, T., Grebogi, C., Ott, E. and Yorke, J. (1993) Using Small Perturbations to Control Chaos *Nature,* **36**, 411-417.

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

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

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.

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

Wolf, A., Swift, J.B., Swinney, H.L. and Vastano, J.A. (1985) Determining Lyapunov Exponents from a Time Series *Physica D,* **16**, 285-317.

Table 1: Summary Statistics

	FTSE-100	Short-Sterling	Gilt (full sample)	Gilt 01/92-07/94	Gilt 08/94-06/95
Number of obs.	76944	81204	80077	57153	22924
Mean (X 10^{-6})	-117.3291	-8.6941	92.4367	-0.2116	313.4995
Variance	0.0078	0.0001	0.0031	0.0033	0.0026
Skewness	-0.3485	2.2418	0.1973	0.0960	0.5550
Kurtosis	22.7704	404.0876	9.7014	8.7753	12.7368
Maximum	1.6461	0.6401	0.9000	0.9000	0.9000
Minimum	-2.6567	-0.6032	-0.5661	-0.5661	-0.4526
Normality	1658726**	552592716**	31700**	184430**	157500**

Note: * and ** denote significance at the 5% and 1% levels respectively.

Table 2: Results for Tsay Test, McLeod and Li's Test, Engle's ARCH test, and Ramsey's RESET Test

	FTSE-100	Short-Sterling	Gilt (full sample)	Gilt 01/92-07/94	Gilt 08/94-06/95
Tsay	11.517**	69.930**	0.8243	0.422	10.924**
McLeod and Li	1771.599**	2878.188**	6459.190**	5336.435**	1399.927**
RESET	11.488**	293.804**	9.522**	0.457	12.968**
ARCH(4)	1427.554**	2216.172**	4355.774**	360.970**	1044.843**

Note: * and ** denote significance at the 5% and 1% levels respectively.

Table 3: Results of Bispectrum Test

	Gaussianity	Linearity	% sig. bispectral estimates
FTSE-100	140.22**	-1.582	10.18
Short-Sterling	5079.69**	37.094**	24.32
Gilt	207.76**	2.707**	12.60
Gilt 01/92-07/94	148.24**	5.567**	11.08
Gilt B 08/94-06/95	146.04**	-2.125**	12.53

Note: ** denotes significance at the 1% level.

Table 4: Results of White's Neural Network Test

	p-values (not ordered)						
						Simple Bonferroni	Hochberg Bonferroni
	1	2	3	4	5		
FTSE-100	0.00000	0.62610	0.01480	0.63913	0.10648	0.00000	0.00000
Short-Sterling	0.00105	0.00079	0.00067	0.00074	0.00062	0.00310	0.00105
Gilt (full sample)	0.1340	0.1403	0.1468	0.1318	0.1920	0.6700	0.19205
Gilt 01/92-07/94	0.35608	0.37473	0.36566	0.46202	0.43982	0.78038	0.46202
Gilt 08/94-06/95	0.00079	0.00083	0.00069	0.00061	0.00070	0.00306	0.00083

**Table 5: Results of BDS tests on Returns and Residuals of a Linear
and EGARCH Fit for FTSE-100 Series**

m \ Series	2	3	4	5	6	7	8	9	10
Raw Returns	52.27	67.56	78.81	89.90	101.90	115.42	130.88	149.36	171.76
Linear filter	51.72	66.81	77.69	88.60	100.42	113.69	128.81	146.85	168.70
EGARCH filter	38.42	49.79	58.88	68.24	78.26	90.01	103.89	121.11	142.66

**Table 6: Results of BDS tests on Returns and Residuals of a Linear
and EGARCH Fit for Short-Sterling Series**

m \ Series	2	3	4	5	6	7	8	9	10
Raw Returns	105.78	124.01	134.40	143.00	151.70	161.14	172.03	184.55	198.76
Linear filter	49.79	63.32	74.89	84.13	92.71	101.28	111.41	122.52	135.01
EGARCH filter	7.88	10.43	12.43	13.41	14.85	16.51	18.65	21.07	24.09

**Table 7: Results of BDS tests on Returns and Residuals of a Linear
and EGARCH Fit for Gilt Series (full sample)**

m \\ Series	2	3	4	5	6	7	8	9	10
Raw Returns	78.54	100.67	119.16	138.59	160.87	188.27	222.93	266.15	321.70
Linear filter	38.05	60.90	78.17	94.44	110.9	129.69	152.07	178.37	210.72
EGARCH filter	242.69	217.08	194.60	177.26	163.66	152.69	143.62	135.98	129.44

**Table 8: Results of BDS tests on Returns and Residuals of a Linear
and EGARCH Fit for Gilt Series (01/92-07/94)**

m \\ Series	2	3	4	5	6	7	8	9	10
Raw Returns	70.65	90.57	101.55	125.54	146.35	172.07	204.60	245.37	297.94
Linear filter	36.16	55.90	71.02	85.31	99.96	116.76	136.73	160.24	189.10
EGARCH filter	12.02	15.41	17.42	19.46	20.92	22.88	25.20	27.60	30.42

**Table 9: Results of BDS tests on Returns and Residuals of a Linear
and EGARCH Fit for Long Gilt Series (08/94-06/95)**

m \ Series	2	3	4	5	6	7	8	9	10
Raw Returns	35.26	44.73	52.19	59.82	68.00	77.64	89.52	103.72	121.31
Linear filter	8.67	18.70	27.12	35.68	44.03	53.11	63.80	76.27	92.14
EGARCH filter	8.81	10.50	11.86	13.06	13.77	14.40	15.22	15.99	17.02

Table 10: Results of Lyapunov Exponent Estimation

# hidden units \ Series	1	2	3	4	5	6	7	8	9	10
FTSE-100	-2.68	-2.72	-2.60	-2.56	-2.55	-2.57	-2.59	-3.06	-2.52	-2.55
Short-Sterling	-6.90	-5.32	-4.61	-5.04	-4.91	-3.18	-4.08	-4.19	-4.83	-4.51
Gilt (full sample)	-2.58	-2.60	-2.61	-2.61	-2.58	-2.58	-2.57	-2.56	-2.61	-2.69
Gilt 01/92-07/94	-10.14	-6.13	-5.42	-4.86	-3.96	-3.51	-3.42	-4.25	-4.10	-3.90
Gilt 08/94-06/95	-17.67	-4.68	-4.73	-5.13	-5.72	-5.46	-4.78	-5.31	-4.30	-4.07