Department of Economics

Testing for Neglected Nonlinearity in Long Memory Models

Richard T. Baillie and George Kapetanios

Working Paper No. 528 April 2005 ISSN 1473-0278

Queen Mary
University of London

Testing for Neglected Nonlinearity in Long Memory Models

Richard T. Baillie
Michigan State University and Queen Mary, University of London
and
George Kapetanios*
Queen Mary, University of London.

September 2004

Abstract

This paper constructs tests for the presence of nonlinearity of unknown form in addition to a fractionally integrated, long memory component in a time series process. The tests are based on artificial neural network structures and do not restrict the parametric form of the nonlinearity. The tests only require a consistent estimate of the long memory parameter. Some theoretical results for the new tests are obtained and detailed simulation evidence is also presented on the power of the tests. The new methodology is then applied to a wide variety of economic and financial time series.

JEL Classification: C22, C12, F31.
Key Words: Long Memory, Non-linearity, Artificial Neural Networks, Realized Volatility, Absolute Returns, Real Exchange Rates, Unemployment.

*The authors gratefully acknowledge the assistance of Yoryos Chortareas, Florence Hubert and the Federal Reserve Bank of Cleveland for supplying some of the data used in this paper.

1 Introduction

A considerable amount of recent work in time series econometrics has focused on alternative representations compared with the conventional I(0) and I(1) paradigms. In particular, there have been substantial developments in the modeling of long memory processes, and also in the mainly unrelated area of modeling non-linearity. However, there has been relatively little consideration of the issue of combining, or distinguishing between these types of processes. Notable exceptions are Diebold and Inoue (2001) who show how a process with Markov switching regime changes can be mistaken for a long memory process. Also, Kapetanios and Shin (2002) suggested a formal test for distinguishing between non-stationary long memory and nonlinear geometrically ergodic processes in small samples; while van Dijk, Frances, and Paap (2002) considered a long memory and Exponential Smooth Transition Autoregressive ($ESTAR$) model to represent the US unemployment rate. While the first two articles are concerned about the possibility of confusing non-linearity and long memory, the third paper addresses the possibility that a process may exhibit both long memory dynamics and non-linearity in the short memory dynamics.

This paper focuses on the issue of providing a general, formal testing framework for non-linearity in a time series process which may include a long memory, fractionally integrated component. One motivation for the study is to provide a basis for determining whether an apparent long memory model requires the addition of nonlinear terms. An attractive feature of our procedure is that it does not require specification of the exact parametric form of non-linearity, since a neural network approximation is used which is combined with the long memory component. Two classes of tests are considered; the first is based on artificial neural network approximations, while the second uses a Taylor series approximation. The power performance of the test statistics are shown to depend on the order of the neural network approximations and the number of lagged terms being included. The performance of the various test statistics are documented by means of an extensive simulation study with a variety of nonlinear data generating mechanisms. Some of the test statistics perform quite well and give rise to optimism that nonlinear effects can be distinguished *within* a long memory process. Our findings indicate the desirability of jointly modeling the nonlinear and long memory components of a time series. As noted by Granger and Teräsvirta (1993) the allowance for non-linearity can provide superior forecasts and improved economic intuition for short memory processes, and our results indicate how these effects can be tested and possibly incorporated into long memory processes.

As previously indicated, the main emphasis in our paper is to provide a workable testing strategy for testing against non-linearity in a long memory time series process. The power performance of the tests can be significantly affected in small samples by the use of relatively inefficient initial estimators of the long memory parameter. Hence, we consider the use of several estimators of the long memory parameter in both the time and frequency domains. Our simulation evidence is generally favorable to the Local Whittle estimator and also to a time domain approximate MLE where the long memory parameter is estimated jointly with terms from an artificial neural network expansion. These estimators are generally found to be preferable to using the Fox-Taqqu estimator, although we must note that the Fox-Taqqu estimator is estimating an $ARFIMA$ model unlike the proposed artificial neural network time domain estimators. Overall, the analysis shows the desirability of taking non-linearity into account when estimating long memory components. In particular, we document the extent to which the Local Whittle and other estimators of the long memory parameter is adversely affected by certain types of non-linearity.

The paper also includes an extensive application of the above methodology to various economic and financial time series. In general, the results indicate the widespread presence of both nonlinear and long memory components in many macroeconomic time series, including unemployment, monthly inflation rates and also in various definitions of real exchange rates. However, the application to financial market data is less clear. The daily absolute returns on seven major industrialized countries exchange rates against the US dollar are found to be well represented by pure long memory for only three series. A series of fifteen years of the daily logged Realized Volatility for the DM-$ appears to only possess marginal non-linearity in addition to long memory. However, the corresponding series for the Yen-$ and the Yen-DM is found to exhibit significant non-linearity. The daily logged Realized Volatility for five commodity futures contracts reveals almost pure long memory with no discernible non-linearity.

The structure of the rest of the paper is as follows. Section 2 presents the theoretical framework, with the details of the main theorem concerning the validity of the test statistics when a consistent estimator of the long memory parameter is used, is placed in an appendix. Section 3 discusses the various tests, and section 4 their implementation to the problem of testing for neglected non-linearity. Section 5 presents some detailed simulation evidence concerning the performance of the tests, while the next section discusses many different empirical examples. There is also a short conclusions section.

2 Nonlinear long memory models

Long memory, fractionally integrated processes were originally introduced by Granger and Joyeux (1980), Granger (1980) and Hosking (1981) to represent the slow hyperbolic rates of decay associated with the impulse response weights and autocorrelations of a series. See Beran (1994) and Baillie (1996) for detailed surveys of these models and the latter for discussion of the application to economics and finance. A univariate process with fractional integration in its conditional mean can be represented as

$$(1-L)^d y_t = u_t,\ t = 1, \dots, T \tag{1}$$

where L is the lag operator and where u_t is a short memory, $I(0)$ process; then y_t is said to be fractionally integrated of order d, or $I(d)$. In this study an I(0) process is defined according to de Jong and Davidson (2000), as a process whose partial sums converge weakly to Brownian motion. Hence, the parameter d represents the degree of "long memory" behavior for the series. For $-0.5 < d < 0.5$ the process is stationary and invertible; while for $0.5 < d < 1$, the process does not have a finite variance, but for $d < 1$ the impulse response weights are finite, which implies that shocks to the level of the series are mean reverting. If the short memory component can be represented by an $ARMA(p, q)$ process, then equation (1) becomes the $ARFIMA(p, d, q)$ model,

$$\phi(L)(1-L)^d y_t = \theta(L)\epsilon_t \tag{2}$$

where $E(\epsilon_t) = 0$, $E(\epsilon_t^2) = \sigma^2$, $E(\epsilon_t \epsilon_s) = 0$, $s \neq t$, and where $\phi(L)$ and $\theta(L)$ are polynomials in the lag operator of orders p and q respectively. The Wold decomposition, or infinite order moving average representation of this process is given by

$$y_t = \sum_{i=0}^{\infty} \psi_i(d)\epsilon_{t-i} \tag{3}$$

while the infinite order autoregressive representation is given by

$$y_t = \sum_{i=1}^{\infty} \pi_i(d) y_{t-i} + \epsilon_t \tag{4}$$

For high lag i, these coefficients decay at very slow hyperbolic rates of $\psi_i(d) \sim c_1 i^{d-1}$ and $\pi_i(d) \sim c_2 i^{-d-1}$, where c_1 and c_2 are constants. The hyperbolic decay that is generated by such a process is known as the 'Hurst effect', after Hurst (1951), who first discovered the phenomenon in hydrological time series data. This paper considers situations where the short memory process u_t maybe a nonlinear process rather than a conventional pure $ARMA$

process. For example, the long memory model in (1) can be combined with a short memory $ESTAR$ process,

$$u_t = \alpha_0 + \sum_{i=1}^{p_l} \alpha_i u_{t-i} + \sum_{i=1}^{p_n} \beta_i \left[1 - exp(-\gamma_1(u_{t-D} - \gamma_0)^2)\right] u_{t-i} + \epsilon_t \tag{5}$$

where D is the delay parameter. This model has been applied to the investigation of some macroeconomic series; see van Dijk, Frances, and Paap (2002), Michael, Nobay, and Peel (1996) and Sarantis (1999). However, there is no requirement to restrict attention to this particular form of non-linearity and u_t can be modelled in terms of other nonlinear structures such as, threshold autoregressions or bilinear models.

It should be noted that while the theoretical properties of long memory models were originally derived for the conditional mean, recent work has found strong empirical evidence for the presence of long memory in transformations of absolute returns in equity and currency markets and realized volatility series associated with financial markets in general; see, Ding, Granger, and Engle (1993), Andersen, Bollerslev, Diebold, and Labys (2001), Andersen, Bollerslev, Diebold, and Labys (2003) and others. There has also been a corresponding literature on the development of long memory $ARCH$ models, see Baillie, Bollerslev, and Mikkelsen (1996) and Bollerslev and Mikkelsen (1996), and long memory stochastic volatility models, see Breidt, Crato, and de Lima (1998). The methods developed in this paper can be directly applied to either the levels of an economic or financial time series data, or alternatively can be directly applied to absolute returns, or Realized Volatility, or any other metric of financial market data.

In practice, the long memory parameter is generally unknown. Hence, it is important to at least have a consistent estimate of the long memory parameter d, prior to the application of the non-linearity test. One approach considered in this study, is to apply a test for non-linearity to the series u_t, which is obtained by fractionally filtering the original series y_t using a non-parametric estimate of d. An alternative method is to jointly estimate d with the parameters of the nonlinear structure. This method is also considered in this paper. It should be noted that the application of standard $ARFIMA$ model estimation is inappropriate due to the possible neglected non-linearity, and will generally result in an inconsistent estimate of d, if non-linearity exists. Under the alternative hypothesis of neglected non-linearity, the construction of a test for non-linearity, using an $ARFIMA$ estimate of d, is likely to be less powerful than one based on the true value of d. It should be made clear that this issue is related primarily to the power of the test. Under the null hypothesis, d will be estimated consistently through an $ARFIMA$ model and, therefore, the test will be correctly sized.

The initial estimate of the long memory parameter may be based on approximate MLE in the time domain, or alternatively a non-parametric approach in the frequency domain by local Whittle or related techniques. The different effects of these estimators are analyzed in the Monte Carlo study in section 5 of this paper. The tests developed in the paper may be viewed as a first step to a parametric analysis of the neglected non-linearity through the use of a model belonging to a class of nonlinear models used to investigate weakly dependent stationary processes such as threshold autoregressive[1] (TAR) or smooth transition autoregressive ($STAR$) models.

Our proposed solution for estimating d is to consider a neural network type model for u_t. Once d is estimated, an estimate of u_t is obtained from fractionally filtering y_t. This estimate of u_t is then tested for non-linearity using standard neural network tests described in the next section.

3 Neural network models and tests

This section considers two different, but related tests for neglected non-linearity within the maintained hypothesis of long memory. In general, the conditional mean of u_t is allowed to be

$$u_t = F(u_{t-1} \ldots u_{t-p}) + \epsilon_t \tag{6}$$

which represents a possibly nonlinear autoregression involving the last p lags of the dependent variable. There are two methods of dealing with this that are now considered.[2]

3.1 The artificial neural network test

The null hypothesis of this test is that the conditional mean of u_t given lags of u_t is a linear function of the past information set, so that

$$P\left\{ E(u_t|u_{t-1} \ldots u_{t-p}) = \delta_0 + \sum_{i=1}^{p} \delta_i u_{t-p} \right\} = 1 \tag{7}$$

[1]Note that neural network specifications have been used to test for the presence of threshold type non-linearity, see, e.g., Lee, White, and Granger (1993).

[2]It should be noted that other tests for non-linearity have been proposed in previous literature. For example, Keenan (1985) and Tsay (1986) have suggested alternative tests based on Volterra expansions and are a different approach to the framework considered in our study. See Li (2004) and Granger and Teräsvirta (1993) for a review of alternative tests.

The implementation of the test in our case requires estimation of d from an auxiliary equation and the fractionally filtered series to be obtained from

$$\hat{u}_t = (1-L)^{\hat{d}} y_t \approx \sum_{i=1}^{t-1} \pi_i(\hat{d}) y_{t-i} \tag{8}$$

The various methods for the estimation of d will be spelled out in the next section. The implementation of the test requires use of the Lee, White, and Granger (1993) artificial neural network (henceforth ANN) testing framework, which specifies that the nonlinear part of $F(.)$ in (6) is given by $\sum_{j=1}^{q} \phi(\sum_{i=1}^{p} \gamma_{ij} \hat{u}_{t-i})$ where $\phi(\lambda)$ is the logistic function, given by $[1+\exp(-\lambda)]^{-1}$. As noted by Lee, White, and Granger (1993), this functional form can approximate arbitrarily well any continuous function.

The coefficients γ_{ij} are randomly generated from a uniform distribution over $[\gamma_l,\ \gamma_h]$. It should be noted that the use of random γ_{ij} has two purposes. First, it bypasses the need for computationally expensive estimation techniques and second, and most importantly, solves the identification problem for γ_{ij} since these parameters are not identified under the null hypothesis of linearity. For a given q, the constructed regressors $\phi(\sum_{i=1}^{p} \gamma_{ij} \hat{u}_{t-i})$, $j = 1, \ldots, q$ may suffer from multicollinearity. Following the suggestion of Lee, White, and Granger (1993), we also take the $\tilde{q}$ largest principle components of the constructed regressors excluding the largest one be used as regressors in

$$\hat{u}_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \hat{u}_{t-i} + \sum_{j=1}^{\tilde{q}} \beta_j \tilde{\phi}_{j,t} + \epsilon_t \tag{9}$$

where $\tilde{\phi}_{j,t}$ denotes the $(j+1)$-th principal component. A standard LM test is then be performed and Lee, White, and Granger (1993) suggest constructing the test statistic as TR^2, where R^2 is the uncentred squared multiple correlation coefficient of a regression of $\hat{\varepsilon}_t$ on a constant, $\hat{u}_{t-i}$, $i = 1 \ldots, p$, $\tilde{\phi}_{j,t}$, $j = 1, \ldots, \tilde{q}$, where $\hat{\varepsilon}_t$ is the residual of the regression of $\hat{u}_t$ on a constant and $\hat{u}_{t-i}$, $i = 1 \ldots, p$. Under the null hypothesis, this test statistic has an asymptotic $\chi^2_{\tilde{q}}$ distribution. Under the alternative hypothesis, this test is consistent as discussed in Stinchcombe and White (1998).

3.2 The Taylor expansion test

An alternative approach is motivated by the logistic neural network test proposed by Teräsvirta, Lin, and Granger (1993) and has also been used by Blake and Kapetanios (2003). That test approximates the logistic neural network by a Taylor expansion and tests for the significance of these additional terms when they are subsequently substituted into the model. In particular, Teräsvirta, Lin, and Granger (1993) suggest the use of the third order Taylor expansion.

In our framework, the model for $\hat{u}_t$ then takes the form

$$\hat{u}_t = \beta_0 + \sum_{i=1}^{p} \beta_i \hat{u}_{t-i} + \sum_{j=2}^{3} \sum_{i=1}^{p} \gamma_{0,i,j} \hat{u}_{t-i}^{j} + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{1,i,j} \hat{u}_{t-i} \hat{u}_{t-j} + \sum_{s=0}^{1} \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{2,s,i,j} \hat{u}_{t-i}^{2-s} \hat{u}_{t-j}^{s+1} + \epsilon_t \tag{10}$$

Clearly, this is just one particular form of Taylor series expansion that is being used to approximate the unknown function. Highly nonlinear data generating processes may well require higher order terms and for this reason it is desirable to also consider the second order expansion,

$$\hat{u}_t = \beta_0 + \sum_{i=1}^{p} \beta_i \hat{u}_{t-i} + \sum_{i=1}^{p} \gamma_{0,i,2} \hat{u}_{t-i}^{2} + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{1,i,j} \hat{u}_{t-i} \hat{u}_{t-j} + \epsilon_t \tag{11}$$

Similarly we may also wish to consider the fourth order Taylor series expansion of

$$\hat{u}_t = \beta_0 + \sum_{i=1}^{p} \beta_i \hat{u}_{t-i} + \sum_{j=2}^{4} \sum_{i=1}^{p} \gamma_{0,i,j} \hat{u}_{t-i}^{j} + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{1,i,j} \hat{u}_{t-i} \hat{u}_{t-j} + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{2,i,j} \hat{u}_{t-i}^{2} \hat{u}_{t-j}^{2} +$$

$$\sum_{s=0}^{1} \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{3,s,i,j} \hat{u}_{t-i}^{2-s} \hat{u}_{t-j}^{s+1} + \sum_{s=0}^{1} \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{4,s,i,j} \hat{u}_{t-i}^{3-2s} \hat{u}_{t-j}^{2s+1} + \epsilon_t \tag{12}$$

Clearly these are very general approximations with considerable numbers of terms and interactions. In order to restrict the number of parameters in the third and fourth order Taylor series expansions, it was decided to only consider cross products and powers of up to two lags. The restriction that the γ coefficients are all zero is tested using a Wald test. In what follows the models underlying these tests are denoted as the TLG_i models, $i = 2, 3, 4$.

4 Implementation of the tests

The first step in the implementation of the test is the estimation of d, which is needed to construct the filtered series $\hat{u}_t$ on which to apply the tests for non-linearity. We use three of the standard methods for the estimation of d. First, for many models an approximate MLE in the time domain is numerically straightforward. This is sometimes known as the conditional sum of squares (CSS) method and has been successfully applied to models such as $ARFIMA$ with $GARCH$; see Baillie, Chung, and Tieslau (1996). Although the method does not take into account starting values as considered by Sowell (1992a), it has been shown in several studies to perform well in sample sizes of 100 observations or more: see Cheung

(1993), Cheung and Diebold (1994) and Taqqu and Teverovsky (1998). The second technique is by MLE in the frequency domain using the method of Fox and Taqqu (1986). Both these methods assume Gaussianity of the disturbances. The third method is by one of the well known semi-parametric estimation procedures in the frequency domain. While there are now many possible semi-parametric estimator available, the method used in this paper is the Local Whittle estimator since it probably is the most widely used semi-parametric estimator. See Taqqu and Teverovsky (1997) for further discussion of its properties. Also, it generally has superior properties to other semi-parametric frequency domain estimators such as that proposed by Geweke and Porter-Hudak (1983). It should be noted that estimation in the time domain by assuming a pure $ARFIMA$ model is problematic under the alternative hypothesis of non-linearity, and it is expected that the tests would accordingly suffer from low power. Clearly, it seems desirable to consider a model which approximates the nonlinear structure of the series when estimating d. Following the discussion in the previous section, the TLG_i models for $i = 2, 3, 4$ are motivated as approximations to the nonlinear component of the processes and furthermore are straightforward to estimate by approximate MLE in the time domain.

We therefore use these models to estimate d by means of approximate MLE in the time domain, by assuming Gaussianity of the white noise process ϵ_t. The equations to be estimated are then

$$u_t(d) = \beta_0 + \sum_{i=1}^{p} \beta_i u_{t-i}(d) + \sum_{i=1}^{p} \gamma_{0,i,2} u_{t-i}^2(d) + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{1,i,j} u_{t-i}(d) u_{t-j}(d) + \epsilon_t \tag{13}$$

$$u_t(d) = \beta_0 + \sum_{i=1}^{p} \beta_i u_{t-i}(d) + \sum_{j=2}^{3} \sum_{i=1}^{p} \gamma_{0,i,j} u_{t-i}^j(d) + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{1,i,j} u_{t-i}(d) u_{t-j}(d) +$$
$$\sum_{s=0}^{1} \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{2,s,i,j} u_{t-i}^{2-s}(d) u_{t-j}^{s+1}(d) + \epsilon_t \tag{14}$$

$$u_t(d) = \beta_0 + \sum_{i=1}^{p} \beta_i u_{t-i}^j(d) + \sum_{j=2}^{4} \sum_{i=1}^{p} \gamma_{0,i,j} u_{t-j}^j(d) + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{1,i,j} u_{t-i}(d) u_{t-j}(d) +$$
$$\sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{2,i,j} u_{t-i}^2(d) u_{t-j}^2(d) + \sum_{s=0}^{1} \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{3,s,i,j} u_{t-i}^{2-s}(d) u_{t-j}^{s+1}(d) +$$

$$\sum_{s=0}^{1}\sum_{i=1}^{p-1}\sum_{j=i+1}^{p}\gamma_{4,s,i,j}u_{t-i}^{3-2s}(d)u_{t-j}^{2s+1}(d)+\epsilon_t \tag{15}$$

where $u_t(d) = y_t - \sum_{l=0}^{t-p}\pi(d)_l y_{t-l} \approx y_t - \sum_{l=0}^{\infty}\pi(d)_l y_{t-l} = (1-L)^d y_t$. The notation $u_t(d)$ is used to denote that the series is a function of d. This is in contrast to the final filtered series obtained from the approximate time domain MLE and which is denoted by $\hat{u}_t$. It is worth emphasizing that the TLG approximations are used twice in the testing procedure; first to estimate d and obtain $\hat{u}_t$, and then again to test $\hat{u}_t$ for non-linearity.

Both estimation and testing could be combined in a single step. However, the main interest of this study is to obtain a feasible test statistic for non-linearity. Hence it is convenient to distinguish the estimation of d from testing for neglected non-linearity in $\hat{u}_t$. The lag order of the models, p, may be determined by an information criterion or chosen a priori. Once d has been determined, the two non-linearity tests of the previous section can be applied by using the infinite AR representation of y_t to obtain u_t. Given the higher moments used in implementing the tests in equations (14) and (15), it is necessary to assume that $E(u_t^s) < \infty$ for $s = 16$.

Under the null hypothesis of no nonlinear structures and conditional upon knowing, or consistently estimating the lag order p, the following theorem is useful for subsequent testing.

Theorem 1 *Under the null hypothesis of linearity given by (7), given lag order p and for $d < 1/2$ the asymptotic distribution of the tests does not change when the tests are based on $\hat{u}_t = y_t - \sum_{l=0}^{t-p}\pi_l(\hat{d})y_{t-l}$ rather than $u_t = y_t - \sum_{l=0}^{t-p}\pi_l(d^0)y_{t-l}$ where $\hat{d}$ is obtained using (13)-(15) and d^0 is the true value of d.*

The proof is given in the Appendix. Since the asymptotic distribution of the test statistic for known d is simply a χ^2, the proposed tests, based on estimates of d are also asymptotically χ^2 distributed.

In order to clarify the nature of the tests being proposed, it is useful to consider the $p = 1$ case for illustrative purposes. As previously explained, the Monte Carlo analysis considers both parametric and a semiparametric frequency domain estimator for d. The approximate time domain MLE for d are obtained from the models

$$u_t(d) = \beta_0 + \beta_1 u_{t-1}(d) + \epsilon_t$$

$$u_t(d) = \beta_0 + \beta_1 u_{t-1}(d) + \gamma_{0,1,2}u_{t-1}^2(d) + \epsilon_t$$

$$u_t(d) = \beta_0 + \beta_1 u_{t-1}(d) + \gamma_{0,1,2}u_{t-1}^2(d)\gamma_{0,1,3}u_{t-1}^3(d) + \epsilon_t$$

$$u_t(d) = \beta_0 + \beta_1 u_{t-1}(d) + \gamma_{0,1,2} u_{t-1}^2(d) + \gamma_{0,1,3} u_{t-1}^3(d) + \gamma_{0,1,4} u_{t-1}^4(d) + \epsilon_t$$

The filtered series $\hat{u}_t$ is then obtained from the estimate of d, and the subsequent ANN non-linearity test consists of testing that $\beta_i = 0,\ i = 1, \ldots, \tilde{q}$ in the models

$$\hat{u}_t = \alpha_0 + \alpha_1 \hat{u}_{t-1} + \sum_{j=1}^{\tilde{q}} \beta_j \tilde{\phi}_{j,t} + \epsilon_t$$

and the TLG test consists of testing that $\gamma_{0,1,j} = 0,\ j = 2, 3, 4$ in the following regressions

$$\hat{u}_t = \beta_0 + \beta_1 \hat{u}_{t-1} + \gamma_{0,1,2} \hat{u}_{t-1}^2 + \epsilon_t$$

$$\hat{u}_t = \beta_0 + \beta_1 \hat{u}_{t-1} + \gamma_{0,1,2} \hat{u}_{t-1}^2 \gamma_{0,1,3} \hat{u}_{t-1}^3 + \epsilon_t$$

$$\hat{u}_t = \beta_0 + \beta_1 \hat{u}_{t-1} + \gamma_{0,1,2} \hat{u}_{t-1}^2 + \gamma_{0,1,3} \hat{u}_{t-1}^3 + \gamma_{0,1,4} \hat{u}_{t-1}^4 + \epsilon_t$$

The above approach is then implemented in the following simulation study to investigate the small sample properties of the procedures.

5 Simulation Study

This section reports the results obtained from a Monte Carlo study to investigate the size and power properties of the proposed new tests. The simulation experiment considers neglected non-linearity of the $ESTAR$ form and is consistent with the type of non-linearity investigated by van Dijk, Frances, and Paap (2002) in their analysis of US unemployment data. Tables 1 through 6 examine four experiments concerning the size of the test, where the model generating the data are $ARFIMA(1, d, 0)$ processes with the long memory parameter being either 0.4 or 0.6 and with the autoregressive coefficient being either 0 (Experiment 1) or 0.8 (Experiment 2). Our study then considers 8 power experiments where the alternative nonlinear hypothesis is a fractionally integrated model with $d = 0.4, 0.6$ and u_t following an $ESTAR$ model. The precise specification of the $ESTAR$ models is given below for experiments 3 through 10:

- Exp. 3 $\alpha_0 = 0$, $\gamma_0 = 0$, $\alpha_1 = 0.8$, $\beta_1 = -1.5$ $\gamma_1 = 0.01$
- Exp. 4 $\alpha_0 = 0$, $\gamma_0 = 0$, $\alpha_1 = 0.8$, $\beta_1 = -1$ $\gamma_1 = 0.01$
- Exp. 5 $\alpha_0 = 0$, $\gamma_0 = 0$, $\alpha_1 = 0.8$, $\beta_1 = -1.5$ $\gamma_1 = 0.05$
- Exp. 6 $\alpha_0 = 0$, $\gamma_0 = 0$, $\alpha_1 = 0.8$, $\beta_1 = -1$ $\gamma_1 = 0.05$
- Exp. 7 $\alpha_0 = 0$, $\gamma_0 = 0$, $\alpha_1 = 1.3$, $\beta_1 = -1.5$ $\gamma_1 = 0.01$

- Exp. 8 $\alpha_0 = 0$, $\gamma_0 = 0$, $\alpha_1 = 1.3$, $\beta_1 = -1$ $\gamma_1 = 0.01$
- Exp. 9 $\alpha_0 = 0$, $\gamma_0 = 0$, $\alpha_1 = 1.3$, $\beta_1 = -1.5$ $\gamma_1 = 0.05$
- Exp. 10 $\alpha_0 = 0$, $\gamma_0 = 0$, $\alpha_1 = 1.3$, $\beta_1 = -1$ $\gamma_1 = 0.05$

All the experiments, or designs, represent geometrically ergodic processes for u_t. The last four experiments allow for the corridor regime of the nonlinear process, (i.e. the regime closer to the mean of the process), to be locally explosive as the polynomial of the autoregressive part of the specification at the corridor regime has a root which is inside the unit circle. Such processes have been found to be of particular use for modeling certain macroeconomic series, such as US GDP by Kapetanios (2003). It is worth emphasizing that such processes are still geometrically ergodic. This result has been proven for $STAR$ models by Kapetanios, Shin, and Snell (2003) using the drift condition by Tweedie (1975). Since these processes are geometrically ergodic, they are also β-mixing and hence α-mixing by Davidson (1994, Ch. 14), with sufficiently rapidly decaying mixing coefficients. Hence, these processes are I(0) processes.

The ANN test and the TLG tests are applied to the process $\hat{u}_t = u_t(\hat{d})$ where $\hat{d}$ has been obtained from estimation of one of the four possible TLG_i, models, where $i = 1, 2, 3, 4$. The symbol TLG_1 refers to a linear $ARFIMA(p, d, 0)$ model; and where the linear model is estimated in both the time and frequency domains[3].

The Local Whittle semi-parametric estimator for d is obtained by minimizing the objective function

$$\log\left[\frac{1}{m}\sum_{j=1}^{m}\omega_j^{2d} I(\omega_j)\right] - \frac{2d}{m}\sum_{j=1}^{m}\log(\omega_j) \tag{16}$$

with respect to d, where $I(\omega_j)$ is the periodogram given by

$$I(\omega_j) = \frac{1}{2\pi T}\left|\sum_{j=1}^{T} y_t e^{i\omega_j t}\right|^2$$

and $m = [T^{0.5}]$. The ANN test is denoted by ANN_i^l, $l = t, f, s$, $i = 0, 1, 2, 3, 4$, where the subscript i refers to the TLG model used to estimate d and the superscript l refers to estimation in either the time domain t, or the frequency domain f, or by means of the semiparametric local Whittle estimator, s. The remaining notation includes the value $i = 0$, which indicates that the true value of d has been used as a benchmark for comparison of the various tests.

[3]For details on estimation in the frequency domain see Harvey (1989).

The TLG test is denoted by $TLG^l_{j,i}$, $l = t, f, s$, $i = 0, 1, 2, 3, 4$, $j = 3$, where the subscript i refers to the TLG model being used to estimate the d parameter. The superscript l refers to estimation in the time or frequency domain or the semiparametric local Whittle estimator, and finally the subscript j refers to the order of the Taylor expansion used to test the null hypothesis of linearity. Following suggestions by Teräsvirta, Lin, and Granger (1993) this study uses the value $j = 3$. Of course, different j could be envisaged following, e.g., Blake and Kapetanios (2003) but no significant difference in performance was observed and so for simplicity this study concentrates on the value of $j = 3$.

The error term ϵ_t is generated from a $NID(0, 1)$ process for all replications; and the results are presented for samples of size $T = 100$ and $T = 400$. Both rejection probabilities and the average estimates of d are reported in the Tables.

Following the advice in Lee, White, and Granger (1993) the parameter setting is for $q = 10$ in the application of the ANN test. The simulation also imposes $\tilde{q} = 2$, $\gamma_h = 2$ and $\gamma_l = -2$. The results in Tables 1 and 2 are for a data generating process with $d = 0.4$ and with $p = 1$. Tables 3 through 6 are for a similar design, but with $d = 0.6$ and covers both the cases of p of 1 and 2.

The results are quite revealing. First, it can be seen from Tables 2, 4 and 6, that all the time domain based estimates of d, perform relatively well in terms of bias and RMSE for most experiments. Note that these findings appear robust to both the stationary and invertible $ARFIMA(0, 0.4, 0)$ data generating process in Table 2 and the non-stationary $ARFIMA(0, 0.6, 0)$ process in Tables 4 and 6. However, the frequency domain Fox-Taqqu estimator can be seen to have substantial upward biases in the non stationary settings of Tables 4 and 6. The Fox-Taqqu estimator performs well in Table 2 for a sample size of $T = 400$, but the substantial upward bias is again evident for a sample size of $T = 100$ in Table 2. The Local Whittle estimator is superior to others for the majority of experiments and the stationary data generating process in Table 2, but tends to significantly over estimate d in the non-stationary environments in Tables 4 and 6.

A further interesting feature is the behaviour of the semi-parametric Local Whittle estimator, which is expected to work despite the presence of non-linearity in the short run dynamics. It is seen that for experiments 1 through 6 this estimator works well. However, in experiments 7 through 10, where the nonlinear process is still geometrically ergodic and also I(0) but highly persistent this estimator is severely biased upwards. Fortunately, the time domain approximate MLE based on the neural network specifications work quite well.

However, the most interesting results concern the rejection frequencies of the tests. The tests based on estimates of d using either the linear models or the semiparametric estimator are uniformly less powerful that the tests based on estimates of d from models approximating the nonlinear models. In some cases the advantage can reach 20%. More interestingly the proposed tests have comparable power to tests based on the true value of d. It is intuitive to expect that tests using the true value of d possess the property of reaching the upper bound in terms of power given specific linearity tests. Therefore, we can see that tests based on an estimate of d obtained from a neural network type model have clearly better power properties than tests based on an estimate of d obtained from $ARFIMA$ models.

6 Empirical Applications

There have been mixed findings for the presence of long memory in economic and financial time series. For example many studies have reported evidence of long memory in real GNP, see Diebold and Rudebusch (1989) and Sowell (1992b). Similarly there is a wealth of evidence on both long memory and non-linearity in inflation rates, but more mixed evidence in other macroeconomic time series, with rather strong evidence in measures of volatility in financial markets. This section considers some relevant applications.

Throughout the applications section we estimate the long memory parameter, d, using both a linear autoregressive fractionally integrated ($ARFIMA(p, d, 0)$) model and a model using a neural network approximation where we use a third order Taylor expansion. The model is estimated by minimising the conditional sum of squares. The use of this algorithm enables straightforward estimation for the neural network approximation model. In both cases the lag order, p, is chosen using the Bayesian information criterion with maximum lag order 4. The same maximum lag order is used throughout the empirical section for quarterly data whereas 12 is used for monthly data and 4 for daily data. We then apply both the ANN and TLG tests[4] using the estimate of d obtained both from the linear and nonlinear models. For the ANN test we follow Lee, White, and Granger (1993) and set $q = 10$, $\tilde{q} = 2$, $\gamma_h = 2$ and $\gamma_l = -2$.

[4]A third order Taylor expansion is used for the TLG test.

6.1 Real Exchange Rates

The first application considers real exchange rates, which have attracted a lot of attention in the literature; see, e.g., Papell (1997) and Diebold, Husted, and Rush (1991). Some of the main issues have been whether the series exhibit mean reversion, the duration of shocks, etc. The evidence has generally been mixed with less evidence of stationarity given data in the post Bretton Woods regime. For the sake of comparison we take a similar data set as Cheung and Lai (2001), who investigated the presence of long memory in Yen real exchange rates. One of the motivations of their paper was to aim to explain the puzzle of the inability to reject the null hypothesis of unit root nonstationarity using standard unit root tests. However, in order to augment our investigation, we consider other datasets, and construct bilateral real exchange rates against the *i-th* currency at time t $(q_{i,t})$ as $q_{i,t} = s_{i,t} + p_{j,t} - p^*_{i,t}$, where $j = JAPAN, US, GERMANY$ and $s_{i,t}$ is the corresponding nominal exchange rate (*i-th* currency per yen), $p_{j,t}$ the price level in the *j-th* numeraire country, and $p^*_{i,t}$ the price level of the *i-th* country. Thus, a rise in $q_{i,t}$ implies a real numeraire country currency appreciation against the *i-th* currency. The price levels are consumer price indices and all variables are in logs. All data are from the International Monetary Fund's *International Financial Statistics* in CD-ROM and are not seasonally adjusted. All the data are quarterly, spanning from 1960Q1 to 2000Q4 for the Yen and 1957Q1-1998Q4 for the US Dollar and German Mark. We consider a very large sample of countries in an attempt to make the empirical analysis more comprehensive. The countries we consider vary from dataset to dataset depending on data availability. Details on the countries for every dataset appear in the result Tables 7-9. We see that evidence for non-linearity is widespread in the datasets we consider. For the Yen real exchange rates, there are 16 countries (out of 33) for which both the ANN and TLG tests reject the null hypothesis of no neglected non-linearity at the 10% significance level when $\hat{d}$ from the neural network approximation model is used. When $\hat{d}$, estimated from the $ARFIMA$ model, is used both tests reject for 8 countries. The ANN and TLG tests seem in general to reach similar conclusions. There is a number of instances where the estimates of d from the linear and nonlinear model differ substantially. Looking at the countries for which rejection of the null hypothesis is obtained some interesting results arise. There seems to be little evidence for non-linearity in the series relating to major European countries. On the other hand there is evidence for non-linearity in the Asia/Pacific area real exchange rates.

However, the DM denominated exchange rates exhibit considerable evidence of non-linearity with 11 out of 18 countries rejecting the null of no non-linearity for the ANN test hen $\hat{d}$ from the neural network approximation model is used. When $\hat{d}$, estimated from the $ARFIMA$ model, is used tests reject at most for 6 countries. In this dataset there is evidence for non-linearity for European countries. For the US, we find again evidence for non-linearity. At least 6 countries out of 20 reject the null hypothesis.

6.2 European Unemployment Rates

This analysis examines a smaller dataset of unemployment rates for Finland, Germany, Greece, Ireland and Netherlands for the period 1965Q1-2003Q3. The data were obtained from the NiGem Database. The statistics are the same as those presented for the real exchange rate subsection. Results are presented in Table 10 and it is found that the null hypothesis of no neglected non-linearity is rejected for three of the five countries.

6.3 Monthly Inflation

The properties of monthly rates of inflation has been one of the most widely investigated series in the economics and econometrics literature. A central issue in much of this research has been the degree of persistence of the shocks, and is related to the controversy concerning the possible existence of a unit root in inflation. In particular, Ball and Cecchetti (1990), Brunner and Hess (1998)), Barsky (1987) and Nelson and Schwert (1977), have argued that US inflation contains a unit root so that shocks to inflation are completely persistent. Alternatively, Hassler and Wolters (1995), Baillie, Chung, and Tieslau (1996), Baum, Barkoulas, and Caglayan (1999), and others, have all found evidence that inflation is fractionally integrated. The above articles provide quite consistent evidence across countries and time periods that inflation is fractionally integrated with a differencing parameter which is significantly different from zero and unity. However, Brunner and Hess (1998) use non-linear methods, particularly switching regime models to represent inflation. The time series properties of inflation are important from a finance perspective, since as originally noted by Rose (1988), most asset pricing models require ex post real rates of interest to be stationary. For a long memory inflation process, this implies the existence of a non-standard form of cointegration between inflation and nominal interest rates. The type and properties of non-linearities are also important in terms of the real rate of interest.

The inflation series analyzed in this paper are monthly inflation defined as $y_t = \Delta ln(CPI_t)$, where CPI_t is monthly Consumer Price Index. The countries considered are US, Canada, France, Germany, Italy, Japan, UK, Argentina and Brazil. The series extend from January 1957 to April 1990. We use this dataset as it was previously analysed by Baillie, Chung, and Tieslau (1996). Results are presented in Table 11. The results are again very clear. There is overwhelming evidence for neglected non-linearity especially when d is estimated using a TLG approximation. A notable result is that obtained for France where we see that a negative estimate for d obtained using an $ARFIMA(p, d.0)$ is reversed when a neural network model is used.

6.4 Absolute Returns of Exchange Rates

Following Ding, Granger, and Engle (1993) there has been the widespread finding that absolute returns in speculative auction markets have pronounced long memory features, resulting in the very slow hyperbolic decay of their autocorrelations. Table 12 reports the results of our tests to the absolute returns from daily exchange rate returns for the seven freely floating exchange rates of Belgium, Canada, France, Germany, Italy, Japan and the UK vis a vis the US dollar, from March 1980 through June 1998; a total of 4,950 observations. Four of the absolute returns are found to have substantial non-linearity in addition to the long memory component.

6.5 Realized Volatility of Exchange Rates

Following the work of Andersen, Bollerslev, Diebold, and Labys (2001) and Andersen, Bollerslev, Diebold, and Labys (2003), there has been considerable recent interest in the computation and properties of Realized Volatility (RV) from high frequency financial market data. The RV series have the attraction of being a pure, model free measure of volatility and does not depend on the assumption of an $ARCH$, or stochastic volatility, or other model formulation. Andersen, Bollerslev, Diebold, and Labys (2003) and Andersen, Bollerslev, Diebold, and Labys (2001) show that under the assumption that logarithmic asset prices follow a univariate diffusion, the volatility is naturally measured by the associated quadratic variation process. The observed realized volatility is then calculated at the daily level by using the high frequency squared returns aggregated over the day. Table 13 of this study reports the results of the tests for non-linearity applied to the logarithm of the RV series of the DM-$, Yen-$ and DM-Yen; where the RV series are computed from 3,045 days from December 1, 1986 through June 30, 1999. Our estimates of d from the time domain MLE are extremely close to those of Andersen, Bollerslev, Diebold, and Labys (2001) and Andersen, Bollerslev, Diebold, and Labys (2003) with the estimates of d being 0.385, 0.433 and 0.440 for the DM-$, Yen-$ and DM-Yen respectively, compared with the respective estimates of 0.387, 0.413 and 0.430 reported by Andersen, Bollerslev, Diebold, and Labys (2003) from semi-parametric estimation. They also use a multivariate semi-parametric estimator for the combined series and report an estimate of d of 0.401. The tests for non-linearity are interesting. While the RV series are often considered to be virtually pure long memory, our findings in Table 13 indicate a failure to reject linearity at the .05 level for the DM-$, but strong rejections at that level for the Yen-$ and Yen-DM. It may be that the non-linearity is due to jumps and continuous price adjustments as suggested by Maheu and McCurdy (2002) in this context. Hence the approach of estimating switching regime models and/or threshold, or $STAR$ type models may give rise to further improvements in modeling and forecasting RV in currency markets.

6.6 Realized Volatility of Commodity Futures Returns

Table 14 investigates the presence of non-linearity on analogous RV series from commodity futures market returns. In particular, 370 days of high-frequency data on the futures prices of corn, soybeans, cattle, gasoline and gold from the Chicago Mercantile Exchange from May 3, 1999 through September 21, 2000 have been used. Since the commodities futures markets are not so deep as equity or currency markets, the data were sampled every 15 minutes at the high frequency level in order to construct the RV series. The results in Table 14 indicate that the series can be quite well approximated as fractional white noise processes. In general there is no evidence of neglected non-linearity, which suggests there is no evidence for additional modeling strategies of RV as suggested by Maheu and McCurdy (2002).

7 Conclusion

This paper has suggested some new tests for non-linearity in a time series process with a fractionally integrated component. Our suggested procedure does not require specification of the exact parametric form of non-linearity and is based on artificial neural network and Taylor series approximations. We find that using a linear model to estimate the long memory parameter d prior to applying linearity tests leads to a significant loss of power and we therefore suggest estimation of d using an approximate neural network model which is capable of picking up arbitrary forms of non-linearity. We find that this strategy entails no loss of power compared to the case of known d and we therefore recommend this approach. The test statistics generally perform quite well and indicate that non-linear effects can be distinguished within a long memory process.

We document the performance of different estimators of the long memory parameter. In the application section, the results indicate widespread presence of both non-linear and long memory components in many macroeconomic time series, including unemployment, monthly inflation rates and also in various definitions of real exchange rates. There also seems to be evidence of non-linearity in addition to long memory for daily absolute returns on some exchange rates against the US dollar; while there is also some evidence of non-linearity in Realized Volatility for currencies, but none for five commodity futures returns.

Appendix: Proof of Theorem 1

Define the following:

$$\mathbf{u}(d) = (u_1(d), \ldots, u_T(d))'$$

$$\mathbf{v}_t(d) = (1, u_{t-1}(d), u_{t-1}(d), \ldots, u_{t-p}(d))'$$

$$\mathbf{v}(d) = (\mathbf{v}_1(d), \ldots, \mathbf{v}_T(d))'$$

Let $\mathbf{z}_t(d)$ be the set of cross product regressors used to test the null hypothesis of neglected non-linearity for the TLG test. A similar analysis can be applied to the ANN test. Then,

$$\mathbf{z}(d) = (\mathbf{z}_1(d), \ldots, \mathbf{z}_T(d))'$$

$$\mathbf{M}_v(d) = I - \mathbf{v}(d)(\mathbf{v}(d)'\mathbf{v}(d))^{-1}\mathbf{v}(d)'$$

Then, the Wald test of the null hypothesis is given by

$$W(d) = 1/\hat{\sigma}^2\mathbf{u}(d)'\mathbf{M}_v(d)\mathbf{z}(d)(\mathbf{z}(d)'\mathbf{M}_v(d)\mathbf{z}(d))^{-1}\mathbf{z}(d)'\mathbf{M}_v(d)\mathbf{u}(d) =$$

$$1/\hat{\sigma}^2(d)\left[1/\sqrt{T}(\mathbf{u}(d)'\mathbf{M}_v(d)\mathbf{z}(d))\right]\left[1/T(\mathbf{z}(d)'\mathbf{M}_v(d)\mathbf{z}(d))\right]^{-1}\left[1/\sqrt{T}(\mathbf{z}(d)'\mathbf{M}_v(d)\mathbf{u}(d))\right]$$

Denote the true value of d by d^0. Then, the theorem is proven if we show that

$$W(d^0) - W(\hat{d}) = o_p(1) \tag{17}$$

under the null hypothesis. This follows if we show that

$$\hat{\sigma}^2 - \sigma^2 = o_p(1) \tag{18}$$

$$1/\sqrt{T}(\mathbf{u}(d^0)'\mathbf{M}_v(d^0)\mathbf{z}(d^0)) - 1/\sqrt{T}(\mathbf{u}(\hat{d})'\mathbf{M}_v(\hat{d})\mathbf{z}(\hat{d})) = o_p(1) \tag{19}$$

and

$$1/T(\mathbf{z}(d^0)'\mathbf{M}_v(d^0)\mathbf{z}(d^0)) - 1/T(\mathbf{z}(\hat{d})'\mathbf{M}_v(\hat{d})\mathbf{z}(\hat{d})) = o_p(1) \tag{20}$$

and $1/T(\mathbf{z}(d^0)'\mathbf{M}_v(d^0)\mathbf{z}(d^0))$ has a positive definite probability limit. The last statement is assumed to hold by assumption. Estimation of any of the models TLG_i, $i = 2, 3, 4$ can be shown straightforwardly to lead to an $\sqrt{T}$-consistent estimator of d or $d^0 - \hat{d} = O_p(T^{-1/2})$ under the null hypothesis. This follows easily from the analysis following Theorem 1 of Li and McLeod (1986). Further, this implies that (18) holds. We show that (19) holds. (20)

can be shown to hold similarly. Here, we note that the statement of the Theorem is valid for any T^{ϕ}-consistent estimator of d, i.e., any estimator such that $\hat{d} - d^0 = O_p(T^{-\phi})$, $\phi > 0$. We have

$$1/\sqrt{T}(\mathbf{u}(d^0)'\mathbf{M}_v(d^0)\mathbf{z}(d^0)) - 1/\sqrt{T}(\mathbf{u}(\hat{d})'\mathbf{M}_v(\hat{d})\mathbf{z}(\hat{d})) = 1/\sqrt{T}\left[(\mathbf{u}(d^0)' - \mathbf{u}(\hat{d})')\mathbf{M}_v(d^0)\mathbf{z}(d^0)\right] \tag{21}$$

$$+1/\sqrt{T}\left[\mathbf{u}(\hat{d})'(\mathbf{M}_v(d^0) - \mathbf{M}_v(\hat{d}))\mathbf{z}(d^0)\right] + 1/\sqrt{T}\left[\mathbf{u}(\hat{d})'\mathbf{M}_v(\hat{d})(\mathbf{z}(d^0) - \mathbf{z}(\hat{d}))\right]$$

Now examine the first term of (21)

$$1/\sqrt{T}\left[(\mathbf{u}(d^0)' - \mathbf{u}(\hat{d})')\mathbf{M}_v(d^0)\mathbf{z}(d^0)\right] = ||1/\sqrt{T}\left[(\mathbf{u}(d^0)' - \mathbf{u}(\hat{d})')\mathbf{z}^*(d^0)\right]||$$

where $||A||$ denotes matrix norm ($\equiv tr(A'A)$) and $\mathbf{z}^*(d^0)$ denotes the vector of residuals from a regression of $\mathbf{z}(d^0)$ on $\mathbf{v}(d^0)$. But

$$1/\sqrt{T}\left[(\mathbf{u}(d^0)' - \mathbf{u}(\hat{d})')\mathbf{z}^*(d^0)\right] = 1/\sqrt{T}\sum_{t=1}^{T} z_t^{*\prime}(d^0)(u_t(d^0) - u_t(\hat{d}))$$

So we need to show that

$$||1/\sqrt{T}\sum_{t=1}^{T} z_t^{*\prime}(d^0)(u_t(d^0) - u_t(\hat{d}))|| = o_p(1)$$

This holds if

$$1/T\sum_{t=1}^{T} z_t^{*\prime}(d^0)(u_t(d^0) - u_t(\hat{d})) = o_p(1)$$

and

$$1/T\sum_{t=1}^{T}\left(z_t^{*\prime}(d^0)(u_t(d^0) - u_t(\hat{d}))\right)^2 = o_p(1)$$

But

$$1/T\sum_{t=1}^{T} z_t^{*\prime}(d^0)(u_t(d^0) - u_t(\hat{d})) \leq \left(1/T\sum_{t=1}^{T} z_t^{*\prime}(d^0)^2\right)^{1/2}\left(1/T\sum_{t=1}^{T}(u_t(d^0) - u_t(\hat{d}))^2\right)^{1/2}$$

and

$$1/T\sum_{t=1}^{T}\left(z_t^{*\prime}(d^0)(u_t(d^0) - u_t(\hat{d}))\right)^2 \leq \left(1/T\sum_{t=1}^{T} z_t^{*\prime}(d^0)^4\right)^{1/2}\left(1/T\sum_{t=1}^{T}(u_t(d^0) - u_t(\hat{d}))^4\right)^{1/2}$$

By the moment assumptions on u_t we have that

$$\left(1/T\sum_{t=1}^{T} z_t^{*\prime}(d^0)^2\right)^{1/2} = O_p(1)$$

and

$$\left(1/T\sum_{t=1}^{T} z_t^{*\prime}(d^0)^4\right)^{1/2} = O_p(1)$$

But

$$u_t(d) = y_t - \sum_{l=1}^{t} b_l(d)y_{t-l}$$

and

$$u_t(d^0) - u_t(\hat{d}) = \sum_{l=1}^{t}(\pi_l(d^0) - \pi_l(\hat{d}))y_{t-l}$$

Therefore, by (4.17) of Wright (1995), for a T^{ϕ}-consistent estimator of d

$$u_t(d^0) - u_t(\hat{d}) = O_p(T^{-\phi})$$

Hence,

$$1/T\sum_{t=1}^{T}(u_t(d^0) - u_t(\hat{d}))^2 = o_p(1)$$

and

$$1/T\sum_{t=1}^{T}(u_t(d^0) - u_t(\hat{d}))^4 = o_p(1)$$

Further, it is easy to see that

$$u_t^i(d^0) - u_t^i(\hat{d}) = O_p(T^{-\phi})$$

for $i = 2, 3, 4$. To see this note that for, say, $i = 2$,

$$u_t^2(d^0) - u_t^2(\hat{d}) = (u_t(d^0) - u_t(\hat{d}))(u_t(d^0) + u_t(\hat{d})) = O_p(T^{-\phi})$$

Similar treatments can be used for higher values of i. Using the above, similar analysis can be shown to hold for the other terms of (21). Thus, the result is proven.

References

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001): "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96(2), 42–55.

——— (2003): "Modelling and Forecasting Realised Volatility," *Econometrica*, 71(2), 579–625.

Baillie, R. T. (1996): "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 73, 5–59.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1996): "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 74, 3–30.

Baillie, R. T., C. F. Chung, and M. A. Tieslau (1996): "Analysing Inflation by the Fractionally Integrated ARFIMA-GARCH Model," *Journal of Applied Econometrics*, 11, 23–40.

Ball, L., and S. G. Cecchetti (1990): "Inflation and Uncertainty at Short and Long Horizons," *Brookings Papers on Economic Activity*, pp. 215–254.

Barsky, R. B. (1987): "The Fisher Hypothesis and the Forecastability and Persistence of Inflation," *Journal of Monetary Economics*, 19, 3–24.

Baum, C. F., J. T. Barkoulas, and M. Caglayan (1999): "Persistence in International Inflation Rates," *Southern Economics Journal*, 65, 900–913.

Beran, J. (1994): *Statistics for Long Memory Processes*. Chapman & Hall.

Blake, A. P., and G. Kapetanios (2003): "A Radial Basis Function Artificial Neural Network Test for Neglected Nonlinearity," *Forthcoming in Econometrics Journal*.

Bollerslev, T., and H. O. Mikkelsen (1996): "Modelling and Pricing Long Memory in Stock Price Volatility," *Journal of Econometrics*, 73, 151–184.

Breidt, J. F., M. Crato, and P. J. F. de Lima (1998): "On the Detection and Estimation of Long Memory in Stochastic Volatility," *Journal of Econometrics*, 83, 325–348.

Brunner, A. D., and G. D. Hess (1998): "Are Higher Levels of Inflation Less Predictable? A State-Dependent Conditional Heteroskedasticity Approach," *Journal of Business and Economic Statistics*, 11, 187–197.

Cheung, Y. W. (1993): "Tests for Fractional Integration: A Monte Carlo Investigation," *Journal of Time Series Analysis*, 14, 331–345.

Cheung, Y. W., and F. X. Diebold (1994): "On Maximum Likelihood Estimation of the Differencing Parameter of Fractionally Integrated Noise with Unknown Mean," *Journal of Econometrics*, 62, 301–316.

Cheung, Y. W., and K. S. Lai (2001): "Long Memory and Nonlinear Mean Reversion in Japanese Yen-Based Real Exchange Rates," *Journal of International Money and Finance*, 20, 115–132.

Davidson, J. (1994): *Stochastic Limit Theory*, Advanced Tests in Econometrics. Oxford University Press.

de Jong, R. M., and J. Davidson (2000): "The Functional Central Limit Theorem and Convergence to Stochastic Integrals I: The Weakly Dependent Process," *Econometric Theory*, 16, 621–642.

Diebold, F. X., S. Husted, and M. Rush (1991): "Real Exchange Rates Under the Gold Standard," *Journal of Political Economy*, 99, 1252–1271.

Diebold, F. X., and A. Inoue (2001): "Long Memory and Regime Switching," *Journal of Econometrics*, 105, 131–159.

Diebold, F. X., and G. D. Rudebusch (1989): "Long Memory and Persistence in Aggregate Output," *Journal of Monetary Economics*, 24, 189–209.

Ding, Z., C. W. J. Granger, and R. F. Engle (1993): "A Long Memory Property of Stock Returns and a New Model," *Journal of Empirical Finance*, 1, 83–106.

Fox, R., and M. S. Taqqu (1986): "Large Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Processes," *Annals of Statistics*, 14, 517–532.

Geweke, J., and S. Porter-Hudak (1983): "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221–238.

Granger, C. W. J. (1980): "Long Memory Relationships and the Aggregation of Dynamic Models," *Journal of Econometrics*, 14, 227–238.

Granger, C. W. J., and R. Joyeux (1980): "An Introduction to Long Memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis*, 1, 15–39.

Granger, C. W. J., and T. Teräsvirta (1993): *Modelling Nonlinear Economic Relationships*. Oxford University Press.

Harvey, A. C. (1989): *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press.

Hassler, U., and J. Wolters (1995): "Long Memory in Inflation Rates: International Evidence," *Journal of Business and Economic Statistics*, 13, 37–45.

Hosking, J. R. M. (1981): "Fractional Differencing," *Biometrika*, 65, 165–176.

Hurst, H. E. (1951): "Long Term Storage Capacity of Reservoirs," *Transactions of the American Society of Civil Engineers*, 116, 770–799.

Kapetanios, G. (2003): "Threshold Models for Trended Time Series," *Empirical Economics*, 28, 687–708.

Kapetanios, G., and Y. Shin (2002): "Testing for Nonstationary Long Memory Against Nonlinear Ergodic Models," Unpublished Manuscript, University of Edinburgh.

Kapetanios, G., Y. Shin, and A. Snell (2003): "Testing for a Unit Root in the Nonlinear STAR Framework," *Journal of Econometrics*, 112(2), 359–379.

Keenan, D. M. (1985): "A Tukey Non Additivity Type test for Time Series Nonlinearity," *Biometrika*, 72, 39–44.

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

Li, W. K. (2004): *Diagnostic Testing.* Chapman and Hall.

Li, W. K., and A. I. McLeod (1986): "Fractional Time Series Modelling," *Biometrika*, 73, 217–221.

Maheu, J. M., and T. H. McCurdy (2002): "Nonlinear features of realized FX volatility," *Review of Economics and Statistics*, 84, 668–681.

Michael, P., R. A. Nobay, and D. A. Peel (1996): "Transactions Costs and Nonlinear Adjustment in Real Exchange Rates: an Empirical Investigation," *Journal of Political Economy*, 105, 862–879.

Nelson, C. R., and G. W. Schwert (1977): "Short-Term Interest Rates as Predictors of Inflation: On Testing the Hypothesis that the Real Rate of Interest is Constant," *American Economic Review*, 67, 478–486.

Papell, D. H. (1997): "Searching for Stationarity: Purchasing Power Parity under the Current Float," *Journal of International Economics*, 43, 313–332.

Rose, A. K. (1988): "Is the Real Interest Rate Stable?," *Journal of Finance*, 43, 1095–1112.

Sarantis, N. (1999): "Modelling Nonlinearities in Real Effective Exchange Rates," *Journal of International Money and Finance*, 18, 27–45.

Sowell, F. (1992a): "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models," *Journal of Econometrics*, 53, 165–188.

Sowell, F. B. (1992b): "Modelling Long Run Behaviour with the Fractional ARIMA Model," *Journal of Monetary Economics*, 29, 277–302.

Stinchcombe, M. B., and H. White (1998): "Consistent Specification Testing with Nuisance Parameters Present Only Under the Alternative," *Econometric Theory*, 14, 295–325.

Taqqu, M. S., and V. Teverovsky (1997): "Robustness of Whittle Type Estimators for Time Series with Long Range Dependence," *Stochastic Models*, 13, 723–757.

——— (1998): "On estimating the Intensity of Long Range Dependence in Finite and Infinite Variance Time Series," in *A Practical Guide to Long Memory Processes*, pp. 177–217. Santa Barbara, CA.

Teräsvirta, T., C. F. Lin, and C. W. J. Granger (1993): "Power of the Neural Network Linearity Test," *Journal of Time Series Analysis*, 14, 209–220.

Tsay, R. S. (1986): "Nonlinearity Tests for Time Series," *Biometrika*, 72, 461–466.

Tweedie, R. L. (1975): "Sufficient Conditions for Ergodicity and Recurrence of Markov Chains on a General State Space," *Stochastic Processes Appl.*, 3, 385–403.

van Dijk, D., P. H. Frances, and R. Paap (2002): "A Nonlinear Long Memory Model with Application to US Unemployment," *Journal of Econometrics*, 110(2), 135–165.

Wright, J. H. (1995): "Stochastic Orders of Magnitude Associated with Two-Stage Estimators of Fractional ARIMA Systems," *Journal of Time Series Analysis*, 16, 119–125.

Table 1. Test Rejection Probabilities for Models with $p = 1$ and d=0.4

ANN tests														
	T=100							T=400						
Exp	ANN_1^f	ANN_0^t	ANN_1^t	ANN_2^t	ANN_3^t	ANN_4^t	ANN^s	ANN_1^f	ANN_0^t	ANN_1^t	ANN_2^t	ANN_3^t	ANN_4^t	ANN^s
Exp 1	0.038	0.036	0.034	0.050	0.084	0.078	0.046	0.050	0.062	0.062	0.060	0.060	0.066	0.056
Exp 2	0.028	0.062	0.060	0.082	0.106	0.098	0.046	0.048	0.044	0.042	0.060	0.072	0.072	0.036
Exp 3	0.068	0.108	0.096	0.114	0.168	0.172	0.082	0.256	0.306	0.244	0.272	0.354	0.344	0.244
Exp 4	0.056	0.090	0.066	0.080	0.126	0.112	0.042	0.204	0.232	0.204	0.218	0.278	0.262	0.192
Exp 5	0.320	0.400	0.280	0.322	0.472	0.462	0.312	0.920	0.956	0.850	0.854	0.950	0.946	0.898
Exp 6	0.222	0.306	0.192	0.236	0.364	0.356	0.224	0.756	0.838	0.706	0.708	0.830	0.822	0.702
Exp 7	0.164	0.590	0.438	0.584	0.622	0.618	0.160	0.374	0.998	0.814	0.878	0.986	0.986	0.360
Exp 8	0.088	0.294	0.216	0.300	0.338	0.358	0.088	0.262	0.912	0.828	0.874	0.890	0.892	0.270
Exp 9	0.284	0.962	0.724	0.824	0.962	0.960	0.588	0.438	1.000	0.940	0.958	1.000	1.000	0.976
Exp 10	0.126	0.930	0.688	0.806	0.920	0.906	0.232	0.240	0.998	0.944	0.960	0.998	0.998	0.770

TLG tests														
	T=100							T=400						
Exp	$TLG_{3,1}^f$	$TLG_{3,0}^t$	$TLG_{3,1}^t$	$TLG_{3,2}^t$	$TLG_{3,3}^t$	$TLG_{3,4}^f$	TLG^s	$TLG_{3,1}^f$	$TLG_{3,0}^t$	$TLG_{3,1}^t$	$TLG_{3,2}^t$	$TLG_{3,3}^t$	$TLG_{3,4}^f$	TLG^s
Exp 1	0.038	0.036	0.032	0.052	0.082	0.076	0.050	0.058	0.054	0.054	0.058	0.058	0.064	0.052
Exp 2	0.030	0.058	0.058	0.086	0.104	0.100	0.046	0.042	0.046	0.042	0.060	0.080	0.076	0.034
Exp 3	0.068	0.110	0.098	0.112	0.180	0.172	0.082	0.258	0.312	0.256	0.278	0.368	0.354	0.244
Exp 4	0.054	0.086	0.066	0.072	0.122	0.112	0.044	0.204	0.236	0.198	0.226	0.276	0.274	0.192
Exp 5	0.322	0.412	0.292	0.328	0.478	0.464	0.314	0.924	0.960	0.856	0.864	0.958	0.952	0.898
Exp 6	0.218	0.308	0.202	0.238	0.368	0.360	0.224	0.756	0.840	0.712	0.732	0.840	0.834	0.720
Exp 7	0.170	0.588	0.436	0.588	0.630	0.620	0.156	0.376	0.998	0.822	0.874	0.986	0.990	0.356
Exp 8	0.088	0.294	0.212	0.300	0.342	0.350	0.088	0.268	0.916	0.828	0.868	0.890	0.892	0.264
Exp 9	0.290	0.960	0.726	0.834	0.968	0.962	0.598	0.438	1.000	0.948	0.964	1.000	1.000	0.976
Exp 10	0.126	0.928	0.700	0.808	0.926	0.916	0.236	0.242	1.000	0.948	0.962	1.000	1.000	0.778

Notes: Numbers reported are estimated rejection probabilities in 1000 replications.
For ANN_0^t the true value of d has been used.
For ANN_1^t, ANN_2^t, ANN_3^t, ANN_4^t, d has been estimated using an $ARFIMA(p, d, 0)$ model and (13), (14) and (15) respectively.
For ANN_1^f and ANN^s, d has been estimated using Fox-Taqqu and Local Whittle respectively.
For all ANN tests testing was carried out using (9).
For all TLG tests testing was carried out using (10).

Table 2. Properties of the Estimated Long Memory Parameter for $p = 1$ and d=0.4

	Average estimated d over replications											
	T=100						T=400					
Exp	FT	t_1	t_2	t_3	t_4	LW	FT	t_1	t_2	t_3	t_4	LW
Exp 1	0.409	0.201	0.211	0.223	0.248	0.388	0.409	0.361	0.358	0.356	0.356	0.399
Exp 2	0.744	0.383	0.390	0.381	0.402	0.544	0.480	0.375	0.385	0.385	0.387	0.433
Exp 3	0.603	0.327	0.346	0.347	0.363	0.475	0.436	0.347	0.350	0.354	0.362	0.413
Exp 4	0.635	0.324	0.336	0.349	0.359	0.488	0.443	0.362	0.367	0.370	0.372	0.426
Exp 5	0.454	0.203	0.234	0.290	0.308	0.415	0.408	0.292	0.301	0.364	0.366	0.392
Exp 6	0.507	0.262	0.286	0.303	0.322	0.436	0.404	0.298	0.314	0.357	0.357	0.397
Exp 7	1.239	0.331	0.382	0.398	0.398	1.197	1.119	0.258	0.311	0.423	0.420	1.201
Exp 8	1.249	0.385	0.413	0.402	0.402	1.168	1.103	0.393	0.414	0.435	0.434	1.127
Exp 9	1.062	0.259	0.334	0.388	0.395	0.777	1.008	0.231	0.262	0.388	0.390	0.530
Exp 10	1.174	0.294	0.353	0.392	0.399	0.987	1.122	0.237	0.268	0.386	0.390	0.728
	RMSE of estimate of d											
	T=100						T=400					
Exp	FT	t_1	t_2	t_3	t_4	LW	FT	t_1	t_2	t_3	t_4	LW
Exp 1	0.248	0.487	0.477	0.464	0.433	0.363	0.200	0.255	0.260	0.263	0.265	0.279
Exp 2	0.320	0.299	0.295	0.308	0.299	0.302	0.219	0.253	0.247	0.247	0.246	0.252
Exp 3	0.259	0.341	0.327	0.326	0.318	0.326	0.220	0.287	0.284	0.274	0.268	0.267
Exp 4	0.270	0.333	0.329	0.319	0.316	0.323	0.214	0.269	0.266	0.260	0.259	0.262
Exp 5	0.271	0.454	0.426	0.362	0.350	0.347	0.234	0.357	0.349	0.253	0.250	0.278
Exp 6	0.259	0.403	0.386	0.359	0.348	0.347	0.244	0.340	0.326	0.264	0.265	0.268
Exp 7	0.642	0.357	0.296	0.275	0.276	0.614	0.522	0.416	0.362	0.212	0.209	0.614
Exp 8	0.650	0.308	0.270	0.286	0.287	0.574	0.505	0.293	0.254	0.215	0.212	0.535
Exp 9	0.498	0.409	0.350	0.232	0.226	0.342	0.433	0.419	0.389	0.215	0.212	0.203
Exp 10	0.584	0.379	0.321	0.236	0.234	0.464	0.526	0.401	0.368	0.216	0.213	0.231

Notes: The estimators FT, t_1, t_2, t_3, t_4 and LW have been obtained using Fox-Taqqu, $ARFIMA(p, d, 0)$, (13), (14), (15) and Local Whittle respectively.

Table 3. Test Rejection Probabilities for Models with $p = 1$ and d=0.6

ANN tests														
	T=100							T=400						
Exp	ANN_1^f	ANN_0^t	ANN_1^t	ANN_2^t	ANN_3^t	ANN_4^t	ANN^s	ANN_1^f	ANN_0^t	ANN_1^t	ANN_2^t	ANN_3^t	ANN_4^t	ANN^s
Exp 1	0.048	0.032	0.042	0.058	0.082	0.078	0.060	0.034	0.036	0.044	0.048	0.048	0.052	0.046
Exp 2	0.038	0.056	0.054	0.068	0.116	0.112	0.042	0.024	0.042	0.034	0.054	0.074	0.070	0.026
Exp 3	0.080	0.084	0.088	0.110	0.156	0.144	0.064	0.246	0.316	0.272	0.302	0.358	0.354	0.260
Exp 4	0.044	0.076	0.072	0.090	0.136	0.124	0.054	0.152	0.202	0.170	0.198	0.272	0.270	0.176
Exp 5	0.304	0.410	0.314	0.360	0.488	0.484	0.324	0.884	0.950	0.860	0.868	0.942	0.954	0.874
Exp 6	0.182	0.260	0.210	0.246	0.314	0.320	0.196	0.688	0.804	0.720	0.744	0.812	0.810	0.686
Exp 7	0.130	0.626	0.460	0.564	0.626	0.618	0.308	0.294	0.992	0.610	0.750	0.932	0.936	0.458
Exp 8	0.058	0.306	0.228	0.296	0.322	0.320	0.128	0.216	0.930	0.626	0.696	0.752	0.752	0.520
Exp 9	0.262	0.974	0.774	0.888	0.978	0.970	0.568	0.508	0.998	0.964	0.976	1.000	0.996	0.956
Exp 10	0.182	0.900	0.690	0.790	0.912	0.896	0.334	0.292	0.998	0.944	0.958	0.996	0.996	0.774

TLG tests														
	T=100							T=400						
Exp	$TLG_{3,1}^f$	$TLG_{3,0}^t$	$TLG_{3,1}^t$	$TLG_{3,2}^t$	$TLG_{3,3}^t$	$TLG_{3,4}^f$	TLG^s	$TLG_{3,1}^f$	$TLG_{3,0}^t$	$TLG_{3,1}^t$	$TLG_{3,2}^t$	$TLG_{3,3}^t$	$TLG_{3,4}^f$	TLG^s
Exp 1	0.052	0.036	0.042	0.058	0.082	0.078	0.054	0.038	0.034	0.036	0.044	0.048	0.048	0.040
Exp 2	0.036	0.060	0.056	0.078	0.120	0.114	0.046	0.024	0.048	0.034	0.046	0.068	0.068	0.026
Exp 3	0.082	0.080	0.088	0.108	0.166	0.148	0.066	0.254	0.330	0.286	0.310	0.358	0.358	0.260
Exp 4	0.048	0.074	0.072	0.088	0.138	0.130	0.058	0.154	0.194	0.176	0.208	0.280	0.268	0.174
Exp 5	0.300	0.416	0.310	0.358	0.496	0.482	0.326	0.892	0.958	0.868	0.878	0.960	0.960	0.882
Exp 6	0.190	0.258	0.218	0.256	0.326	0.318	0.194	0.698	0.810	0.726	0.744	0.822	0.822	0.684
Exp 7	0.130	0.622	0.458	0.568	0.626	0.624	0.304	0.290	0.994	0.606	0.750	0.934	0.936	0.462
Exp 8	0.058	0.308	0.224	0.288	0.326	0.306	0.126	0.218	0.928	0.628	0.698	0.750	0.752	0.524
Exp 9	0.266	0.978	0.784	0.902	0.980	0.974	0.570	0.502	1.000	0.968	0.980	1.000	0.998	0.958
Exp 10	0.178	0.908	0.692	0.796	0.910	0.906	0.334	0.294	1.000	0.946	0.962	1.000	1.000	0.772

See notes in Table 1

Table 4. Properties of the Estimated Long Memory Parameter for $p = 1$ and d=0.6

	Average estimated d over replications											
	T=100						T=400					
Exp	FT	t_1	t_2	t_3	t_4	LW	FT	t_1	t_2	t_3	t_4	LW
Exp 1	0.659	0.444	0.466	0.465	0.479	0.602	0.642	0.568	0.572	0.569	0.568	0.641
Exp 2	0.990	0.643	0.669	0.658	0.676	0.738	0.785	0.613	0.621	0.621	0.622	0.658
Exp 3	0.859	0.607	0.622	0.609	0.630	0.674	0.704	0.577	0.590	0.584	0.593	0.633
Exp 4	0.909	0.613	0.634	0.625	0.649	0.701	0.717	0.591	0.604	0.609	0.615	0.660
Exp 5	0.753	0.550	0.573	0.556	0.567	0.643	0.671	0.535	0.540	0.582	0.581	0.632
Exp 6	0.771	0.544	0.568	0.557	0.572	0.641	0.666	0.554	0.566	0.579	0.576	0.628
Exp 7	0.538	0.558	0.581	0.594	0.594	1.133	0.857	0.422	0.448	0.568	0.572	1.200
Exp 8	0.237	0.540	0.551	0.549	0.565	1.080	0.454	0.482	0.490	0.517	0.519	1.086
Exp 9	1.212	0.516	0.577	0.596	0.602	0.920	1.182	0.421	0.454	0.591	0.594	0.750
Exp 10	1.260	0.533	0.592	0.609	0.616	1.087	1.245	0.422	0.460	0.590	0.591	0.918
	RMSE of estimate of d											
	T=100						T=400					
Exp	FT	t_1	t_2	t_3	t_4	LW	FT	t_1	t_2	t_3	t_4	LW
Exp 1	0.185	0.341	0.313	0.325	0.312	0.288	0.082	0.120	0.091	0.105	0.113	0.187
Exp 2	0.474	0.164	0.193	0.194	0.207	0.304	0.302	0.110	0.120	0.123	0.120	0.197
Exp 3	0.372	0.174	0.186	0.190	0.207	0.294	0.213	0.125	0.127	0.114	0.120	0.188
Exp 4	0.405	0.173	0.187	0.199	0.216	0.300	0.230	0.121	0.126	0.123	0.126	0.190
Exp 5	0.297	0.231	0.234	0.192	0.198	0.291	0.168	0.178	0.174	0.090	0.090	0.187
Exp 6	0.308	0.204	0.197	0.188	0.188	0.296	0.180	0.148	0.143	0.098	0.103	0.189
Exp 7	0.634	0.255	0.221	0.213	0.233	0.555	0.592	0.279	0.235	0.128	0.132	0.622
Exp 8	0.612	0.267	0.265	0.255	0.269	0.487	0.577	0.264	0.245	0.225	0.227	0.500
Exp 9	0.637	0.245	0.214	0.092	0.094	0.430	0.590	0.241	0.213	0.027	0.039	0.245
Exp 10	0.675	0.229	0.214	0.128	0.128	0.543	0.652	0.217	0.204	0.027	0.027	0.367
See notes in Table 2.												

Table 5. Test Rejection Probabilities for Models with $p = 2$ and d=0.6

ANN tests														
	T=100							T=400						
Exp	ANN_1^f	ANN_0^t	ANN_1^t	ANN_2^t	ANN_3^t	ANN_4^t	ANN^s	ANN_1^f	ANN_0^t	ANN_1^t	ANN_2^t	ANN_3^t	ANN_4^t	ANN^s
Exp 1	0.032	0.028	0.046	0.082	0.070	0.050	0.050	0.040	0.034	0.052	0.050	0.044	0.054	0.048
Exp 2	0.046	0.036	0.054	0.084	0.104	0.078	0.032	0.036	0.048	0.026	0.084	0.098	0.078	0.038
Exp 3	0.038	0.042	0.048	0.076	0.102	0.080	0.038	0.068	0.084	0.074	0.096	0.148	0.120	0.070
Exp 4	0.046	0.048	0.042	0.080	0.110	0.074	0.034	0.066	0.082	0.066	0.088	0.140	0.118	0.074
Exp 5	0.096	0.094	0.108	0.150	0.156	0.140	0.112	0.240	0.204	0.212	0.228	0.260	0.250	0.214
Exp 6	0.070	0.096	0.092	0.122	0.148	0.110	0.070	0.166	0.158	0.134	0.198	0.202	0.178	0.142
Exp 7	0.124	0.394	0.274	0.396	0.392	0.384	0.174	0.282	0.894	0.548	0.800	0.876	0.872	0.408
Exp 8	0.102	0.208	0.202	0.240	0.246	0.226	0.122	0.230	0.810	0.652	0.732	0.788	0.774	0.478
Exp 9	0.196	0.402	0.280	0.374	0.390	0.416	0.236	0.298	0.596	0.504	0.544	0.624	0.620	0.486
Exp 10	0.142	0.464	0.322	0.440	0.434	0.438	0.174	0.172	0.670	0.550	0.608	0.694	0.720	0.362

TLG tests														
	T=100							T=400						
Exp	$TLG_{3,1}^f$	$TLG_{3,0}^t$	$TLG_{3,1}^t$	$TLG_{3,2}^t$	$TLG_{3,3}^t$	$TLG_{3,4}^f$	TLG^s	$TLG_{3,1}^f$	$TLG_{3,0}^t$	$TLG_{3,1}^t$	$TLG_{3,2}^t$	$TLG_{3,3}^t$	$TLG_{3,4}^f$	TLG^s
Exp 1	0.038	0.036	0.042	0.062	0.106	0.090	0.036	0.042	0.044	0.046	0.042	0.058	0.056	0.044
Exp 2	0.054	0.050	0.056	0.088	0.158	0.120	0.050	0.040	0.042	0.042	0.068	0.132	0.130	0.038
Exp 3	0.068	0.054	0.064	0.082	0.156	0.120	0.070	0.166	0.168	0.188	0.212	0.306	0.278	0.166
Exp 4	0.046	0.062	0.056	0.080	0.140	0.104	0.058	0.128	0.124	0.118	0.150	0.226	0.192	0.118
Exp 5	0.212	0.268	0.232	0.266	0.390	0.340	0.224	0.832	0.888	0.848	0.858	0.896	0.892	0.824
Exp 6	0.166	0.172	0.164	0.200	0.304	0.246	0.158	0.572	0.646	0.568	0.604	0.690	0.672	0.556
Exp 7	0.134	0.398	0.342	0.416	0.480	0.434	0.186	0.546	0.996	0.836	0.944	0.996	0.996	0.474
Exp 8	0.068	0.172	0.188	0.216	0.260	0.228	0.090	0.466	0.848	0.744	0.826	0.860	0.856	0.442
Exp 9	0.416	0.860	0.694	0.774	0.896	0.858	0.494	0.748	1.000	0.988	0.994	1.000	1.000	0.952
Exp 10	0.222	0.768	0.560	0.664	0.804	0.748	0.230	0.304	1.000	0.960	0.968	1.000	1.000	0.706

See notes in Table 1

Table 6. Properties of the Estimated Long Memory Parameter for $p = 2$ and d=0.6

	Average estimated d over replications											
	T=100						T=400					
Exp	FT	t_1	t_2	t_3	t_4	LW	FT	t_1	t_2	t_3	t_4	LW
Exp 1	0.675	0.405	0.441	0.448	0.470	0.623	0.657	0.565	0.567	0.562	0.554	0.625
Exp 2	0.803	0.493	0.572	0.585	0.619	0.739	0.701	0.491	0.522	0.519	0.554	0.658
Exp 3	0.750	0.461	0.548	0.556	0.554	0.687	0.669	0.450	0.480	0.503	0.534	0.643
Exp 4	0.742	0.454	0.530	0.552	0.601	0.690	0.685	0.452	0.486	0.514	0.539	0.640
Exp 5	0.695	0.440	0.508	0.523	0.561	0.644	0.656	0.494	0.518	0.556	0.569	0.625
Exp 6	0.713	0.455	0.533	0.526	0.566	0.681	0.651	0.463	0.491	0.533	0.545	0.639
Exp 7	1.190	0.584	0.626	0.665	0.685	1.115	0.879	0.447	0.527	0.603	0.598	1.164
Exp 8	1.196	0.608	0.646	0.694	0.717	1.078	0.531	0.554	0.574	0.616	0.615	1.090
Exp 9	0.880	0.585	0.641	0.574	0.589	0.891	0.938	0.495	0.561	0.597	0.596	0.754
Exp 10	0.986	0.587	0.643	0.580	0.611	1.048	1.160	0.486	0.532	0.596	0.598	0.908
	RMSE of estimate of d											
	T=100						T=400					
Exp	FT	t_1	t_2	t_3	t_4	LW	FT	t_1	t_2	t_3	t_4	LW
Exp 1	0.236	0.430	0.364	0.362	0.355	0.280	0.105	0.150	0.132	0.158	0.194	0.192
Exp 2	0.424	0.434	0.359	0.358	0.331	0.314	0.253	0.318	0.300	0.330	0.275	0.199
Exp 3	0.385	0.441	0.350	0.333	0.335	0.303	0.218	0.367	0.322	0.294	0.267	0.195
Exp 4	0.391	0.446	0.369	0.340	0.332	0.303	0.233	0.369	0.323	0.293	0.248	0.196
Exp 5	0.359	0.430	0.352	0.311	0.300	0.303	0.189	0.302	0.260	0.197	0.165	0.195
Exp 6	0.357	0.439	0.346	0.333	0.318	0.320	0.198	0.340	0.286	0.220	0.203	0.193
Exp 7	0.629	0.311	0.251	0.153	0.188	0.531	0.582	0.292	0.195	0.074	0.043	0.580
Exp 8	0.616	0.297	0.265	0.173	0.200	0.483	0.639	0.222	0.173	0.073	0.073	0.503
Exp 9	0.504	0.339	0.281	0.213	0.208	0.399	0.434	0.261	0.187	0.033	0.035	0.239
Exp 10	0.579	0.293	0.245	0.215	0.197	0.497	0.591	0.242	0.183	0.038	0.039	0.357

See notes in Table 2.

Table 7. Tests for Non-Linearity on Yen real exchange rates.

Country	ANN_3^t	$TLG_{3,3}^t$	$d^{(1)}$	ANN_1^t	$TLG_{3,1}^t$	$d^{(2)}$
US	0.004	0.003	0.380	0.011	0.013	0.276
Germany	0.882	0.904	0.235	0.893	0.897	0.223
France	0.698	0.706	0.275	0.713	0.711	0.271
Italy	0.068	0.039	0.307	0.129	0.080	0.421
UK	0.189	0.204	0.224	0.239	0.256	0.317
Canada	0.170	0.175	0.415	0.230	0.233	0.327
Australia	0.004	0.004	0.187	0.006	0.005	0.238
Austria	0.002	0.001	0.435	0.003	0.003	0.381
Belgium	0.728	0.740	0.184	0.716	0.725	0.176
Denmark	0.540	0.521	0.425	0.585	0.552	0.475
Finland	0.095	0.099	0.844	0.407	0.424	0.482
Greece	0.021	0.018	0.160	0.045	0.045	0.219
Hungary	0.491	0.520	0.282	0.822	0.831	0.098
Iceland	0.003	0.002	0.206	0.003	0.002	0.204
Korea	0.012	0.012	0.116	0.011	0.011	0.168
Mexico	0.819	0.837	0.063	0.809	0.805	0.128
Netherlands	0.897	0.907	0.290	0.901	0.905	0.285
New Zealand	0.012	0.012	0.786	0.111	0.105	0.272
Norway	0.574	0.626	0.361	0.714	0.708	0.414
Portugal	0.017	0.014	0.251	0.047	0.047	0.368
Spain	0.816	0.848	0.330	0.880	0.870	0.362
Sweden	0.075	0.082	0.130	0.564	0.621	0.384
Switzerland	0.893	0.908	0.577	0.913	0.912	0.552
Turkey	0.059	0.056	0.433	0.108	0.080	0.311
Singapore	0.442	0.445	0.292	0.521	0.536	0.347
Malaysia	0.001	0.000	0.276	0.695	0.708	0.522
Indonesia	0.000	0.000	0.793	0.130	0.146	0.281
Thailand	0.000	0.000	0.282	0.000	0.000	0.327
Philippines	0.024	0.027	0.091	0.704	0.708	0.285
Sri Lanka	0.110	0.108	0.160	0.321	0.344	0.246
Chile	0.701	0.727	0.309	0.689	0.737	0.096
Colombia	0.303	0.315	0.249	0.314	0.322	0.276
Venezuela	0.286	0.300	0.379	0.923	0.936	0.074
No. of Rejections	16	16		8	10	

Notes: Probability values of neglected nonlinearity tests and estimated long memory parameters. The third column presents an estimate of d, denoted $d^{(1)}$, using the TLG approximation whereas the sixth column presents an estimate of d, denoted $d^{(2)}$ using an $ARFIMA(p, d, 0)$ model. No. of rejections reported at 10% significance level. W_3^t, $T_{3,3}^t$, W_1^t and $T_{3,1}^t$ are defined in page 12.

Table 8. Tests for Non-Linearity on DM real exchange rates.

Country	ANN_3^t	$TLG_{3,3}^t$	d	ANN_1^t	$TLG_{3,1}^t$	d
Australia	0.141	0.158	0.383	0.142	0.159	0.364
Austria	0.001	0.001	0.037	0.049	0.051	-0.171
Belgium	0.016	0.016	0.790	0.133	0.135	0.420
Canada	0.076	0.086	0.555	0.127	0.137	0.473
Finland	0.256	0.258	0.417	0.349	0.329	0.494
France	0.004	0.004	0.293	0.004	0.005	0.259
Italy	0.002	0.004	0.497	0.008	0.006	0.405
Luxemburg	0.001	0.004	0.772	0.128	0.134	0.351
Malta	0.691	0.706	0.402	0.762	0.751	0.435
Netherlands	0.313	0.340	0.112	0.327	0.345	0.133
New Zealand	0.948	0.962	0.202	0.958	0.962	0.200
Norway	0.520	0.559	0.399	0.929	0.913	0.554
Potrugal	0.037	0.037	0.586	0.322	0.334	0.419
S. Africa	0.000	0.000	0.202	0.001	0.000	0.275
Spain	0.086	0.092	0.593	0.271	0.262	0.389
Sweden	0.008	0.009	0.429	0.027	0.028	0.395
Switzerland	0.094	0.110	0.316	0.096	0.105	0.327
UK	0.440	0.458	0.490	0.750	0.769	0.350
No. of Rejections	11	10		6	5	

See notes in Table 7.

Table 9. Tests for Non-Linearity on US real exchange rates.

Country	ANN_3^t	$TLG_{3,3}^t$	d	ANN_1^t	$TLG_{3,1}^t$	d
Australia	0.037	0.038	0.173	0.040	0.040	0.239
Austria	0.070	0.074	0.337	0.082	0.085	0.406
Belgium	0.109	0.111	0.367	0.103	0.106	0.378
Canada	0.998	0.997	0.260	0.995	0.997	0.279
Finland	0.077	0.100	0.354	0.106	0.102	0.366
France	0.148	0.150	0.397	0.144	0.153	0.389
Germany	0.129	0.134	0.418	0.158	0.164	0.380
Greece	0.416	0.425	0.129	0.471	0.480	0.087
Italy	0.033	0.036	0.352	0.044	0.046	0.412
Japan	0.008	0.008	0.406	0.010	0.011	0.453
Luxemburg	0.146	0.155	0.407	0.182	0.210	0.365
Netherlands	0.933	0.933	0.222	0.917	0.932	0.202
New Zealand	0.453	0.482	0.387	0.443	0.483	0.393
Norway	0.052	0.062	0.213	0.060	0.081	0.317
Portugal	0.131	0.135	0.449	0.151	0.149	0.377
S. Africa	0.001	0.001	0.221	0.001	0.001	0.199
Spain	0.229	0.228	0.380	0.228	0.246	0.407
Sweden	0.315	0.330	0.382	0.411	0.425	0.466
Switzerland	0.660	0.691	0.336	0.668	0.684	0.398
UK	0.069	0.101	0.169	0.152	0.137	0.270
No. of Rejections	8	6		6	6	

See notes in Table 7.

Table 10. Tests for Non-Linearity on European Unemployment rates.

Country	ANN_3^t	$TLG_{3,3}^t$	d	ANN_1^t	$TLG_{3,1}^t$	d
Finland	0.109	0.115	1.461	0.113	0.141	1.255
Germany	0.000	0.000	1.231	0.017	0.016	0.490
Greece	0.000	0.000	0.708	0.607	0.596	1.318
Ireland	0.219	0.294	1.170	0.284	0.349	1.237
Netherlands	0.023	0.023	0.569	0.037	0.036	0.385
No. of Rejections	3	3		2	2	
See notes in Table 7.						

Table 11. Tests for Non-Linearity on Inflation rates.

Country	ANN_3^t	$TLG_{3,3}^t$	d	ANN_1^t	$TLG_{3,1}^t$	d
US	0.009	0.009	0.508	0.012	0.011	0.487
Canada	0.072	0.070	0.368	0.781	0.165	0.550
France	0.007	0.007	0.610	0.617	0.006	-0.217
Germany	0.433	0.438	0.292	0.441	0.438	0.284
Italy	0.025	0.031	0.314	0.145	0.158	0.422
Japan	0.002	0.003	0.170	0.967	0.077	0.367
UK	0.416	0.014	0.225	0.519	0.172	0.353
Argentina	0.000	0.000	0.224	0.005	0.000	0.270
Brazil	0.040	0.000	0.142	0.001	0.001	0.373
No. of Rejections	7	8		3	5	
See notes in Table 7.						

Table 12. Tests for Non-Linearity on Exchange Rate Daily Absolute Returns.

Country	ANN_3^t	$TLG_{3,3}^t$	d	ANN_1^t	$TLG_{3,1}^t$	d
Belgium	0.117	0.119	0.158	0.000	0.000	0.223
Canada	0.151	0.152	0.212	0.203	0.204	0.223
France	0.030	0.029	0.176	0.035	0.001	0.228
Germany	0.005	0.005	0.167	0.000	0.000	0.227
Italy	0.000	0.000	0.223	0.108	0.079	0.224
Japan	0.000	0.000	0.196	0.000	0.000	0.215
UK	0.216	0.226	0.208	0.263	0.231	0.211
No. of Rejections	4	4		4	5	
See notes in Table 7.						

Table 13. Tests for Non-Linearity on Exchange Rate Log Realized Volatility.

Exch. Rate	ANN_3^t	$TLG_{3,3}^t$	d	ANN_1^t	$TLG_{3,1}^t$	d
DM-$	0.084	0.084	0.385	0.877	0.111	0.458
Yen-$	0.039	0.039	0.433	0.043	0.043	0.442
Yen-DM	0.012	0.012	0.440	0.507	0.161	0.250
No. of Rejections	3	3		1	1	
See notes in Table 7.						

Table 14. Tests for Non-Linearity on Commodity Futures Log Realized Volatility.

Commodity	ANN_3^t	$TLG_{3,3}^t$	d	ANN_1^t	$TLG_{3,1}^t$	d
Corn	0.547	0.527	0.303	0.520	0.529	0.308
Soybeans	0.745	0.741	0.182	0.764	0.746	0.174
Cattle	0.202	0.193	0.290	0.199	0.235	0.244
Gasoline	0.239	0.244	0.176	0.243	0.245	0.169
Gold	0.805	0.772	0.210	0.932	0.775	0.207
No. of Rejections	0	0		0	0	
See notes in Table 7.						

Queen Mary
University of London

This working paper has been produced by
the Department of Economics at
Queen Mary, University of London

Department of Economics
Queen Mary, University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 5096
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/papers/wp.htm