Working Paper nº 06/06

Testing the Martingale Difference Hypothesis Using Integrated Regression Functions

J. Carlos Escanciano and Carlos Velasco

Facultad de Ciencias Económicas y Empresariales
Universidad de Navarra

Testing the Martingale Difference Hypothesis Using Integrated Regression Functions
J. Carlos Escanciano and Carlos Velasco
Working Paper No.06/06
May 2006
JEL No. C12.

ABSTRACT

This paper proposes an omnibus test for testing a generalized version of the martingale difference hypothesis (MDH). This generalized hypothesis includes the usual MDH, testing for conditional moments constancy such as conditional homoscedasticity (ARCH effects) or testing for directional predictability. Here we propose a unified approach for dealing with all of these testing problems. These hypotheses are long standing problems in econometric time series analysis, and typically have been tested using the sample autocorrelations or in the spectral domain using the periodogram. Since these hypotheses cover also nonlinear predictability, tests based on those second order statistics are inconsistent against uncorrelated processes in the alternative hypothesis. To circumvent this problem we introduce the pairwise integrated regression functions as measures of linear and nonlinear dependence. With our test there is no need to choose a lag order depending on sample size, to smooth the data or to formulate a parametric alternative model. Moreover, our test is robust to higher order dependence, in particular to conditional heteroskedasticity. Under general dependence the asymptotic null distribution depends on the data generating process, so a bootstrap procedure is considered and a Monte Carlo study examines its finite sample performance. Then we investigate the martingale and conditional heteroskedasticity properties of the Pound/Dollar exchange rate.

Juan Carlos Escanciano Reyero
Universidad de Navarra
Departamento de Métodos Cuantitativos
Campus Universitario
31080 Pamplona
jescanci@unav.es

TESTING THE MARTINGALE DIFFERENCE HYPOTHESIS USING INTEGRATED REGRESSION FUNCTIONS*

J. Carlos Escanciano[†]

Universidad de Navarra

Carlos Velasco

Universidad Carlos III de Madrid

April 28, 2006

Abstract

This paper proposes an omnibus test for testing a generalized version of the martingale difference hypothesis (MDH). This generalized hypothesis includes the usual MDH, testing for conditional moments constancy such as conditional homoscedasticity (ARCH effects) or testing for directional predictability. Here we propose a unified approach for dealing with all of these testing problems. These hypotheses are long standing problems in econometric time series analysis, and typically have been tested using the sample autocorrelations or in the spectral domain using the periodogram. Since these hypotheses cover also nonlinear predictability, tests based on those second order statistics are inconsistent against uncorrelated processes in the alternative hypothesis. To circumvent this problem we introduce the pairwise integrated regression functions as measures of linear and nonlinear dependence. With our test there is no need to choose a lag order depending on sample size, to smooth the data or to formulate a parametric alternative model. Moreover, our test is robust to higher order dependence, in particular to conditional heteroskedasticity. Under general dependence the asymptotic null distribution depends on the data generating process, so a bootstrap procedure is considered and a Monte Carlo study examines its finite sample performance. Then we investigate the martingale and conditional heteroskedasticity properties of the Pound/Dollar exchange rate.

Keywords: Nonlinear time series; Martingale difference hypothesis; empirical processes, exchange rates.

1

1. INTRODUCTION

The concept of martingale or martingale difference sequence (MDS) is central in many areas of economics, finance and other sciences. The martingale difference hypothesis (MDH) states that the best predictor, in the sense of least mean square error, of a real time series given some information set is just the unconditional mean. In this paper we are concerned with the case in which the time series to be predicted is a measurable real-valued transformation of a stationary vector time series $X_t \in \mathbb{R}^d$, $d \in \mathbb{N}$, $t = 0, \pm 1, \pm 2, ...$, and the information set is just the past values of the time series. That is, we would like to test

$$H_0 : E[Y_t \mid X_{t-1}, X_{t-2}, ...] = \mu, \qquad \mu \in \mathbb{R}. \tag{1}$$

where Y_t is a measurable real-valued transformation of X_t and $\mu = E[Y_t]$.

This hypothesis, referred to as the generalized MDH or simply the MDH, contains interesting testing problems as special cases. For instance, when Y_t is a power transformation of X_t, we are testing for conditional moments constancy; the usual MDH (when Y_t coincides with X_t) or testing for ARCH effects of a MDS are examples of this case. When $Y_t = I(X_t > c)$, $c \in \mathbb{R}^d$, where the indicator $I(A)$ is one when the event A occurs and zero otherwise, we are testing for no directional predictability, see Hong and Chung (2003) or Linton and Whang (2004). If $X_t = (X_{1t}, X_{2t}, X_{3t})$, X_{1t} and X_{2t} are random variables, $Y_t = X_{1t} - X_{2t}$ and $\mu = 0$, we are testing for equality of regression curves, see Ferreira and Stute (2004) for a recent reference. All these hypotheses are important in economics and finance.

Usually, these hypotheses have been typically tested using the autocorrelations or autocovariances or in the spectral domain using the periodogram. For instance, Cochrane (1988) proposed a variance ratio test for uncorrelatedness that has been widely used in finance. Durlauf (1991) proposed a spectral distribution based test, using the fact that under the MDH, the standardized spectral distribution function is a straight line. Recently, Deo (2000) has robustified Durlauf's (1991) test against certain forms of conditional heteroskedasticity.

However, all these tests are suitable for testing lack of serial correlation but not the MDH. In fact, they are not consistent against non-martingale difference sequences with zero autocorrelations, that is, when only nonlinear dependence is present, as commonly happens with economic and financial data, see e.g. our application to exchange rates. These tests are inconsistent because they only employ information contained in the second sample moments of the process.

Roughly speaking, there have been two main approaches in the literature for designing consistent tests for (1). The first approach is based on checking an infinite number of orthogonality conditions, see for instance, Bierens (1984), Stute (1997), Bierens and Ploberger (1997), Koul and Stute (1999), Whang (2000), Dominguez and Lobato (2003), Escanciano and Velasco (2003) or Escanciano (2006a),

among many others. The second line of research employs nonparametric estimates of the conditional expectation function, see for instance Härdle and Mammen (1993), Zheng (1996), Li (1999), Guerre and Lavergne (2005) and references therein.

Tests based on the second methodology have standard asymptotic null distributions, but they usually require strong assumptions on the data generating process (DGP), see e.g. Li (1999). More importantly, they require subjective choices of smoothing parameters and kernel functions, and usually statistical inferences can be sensitive to these selections.

On the other hand, tests statistics based on the first methodology do not demand in general the selection of any user-chosen parameters. They are consistent against Pitman's alternatives converging at the parametric rate and do not require strong conditions about the DGP, and although they have non-standard asymptotic null distributions, these can be well approximated by bootstrap methods, see Section 5. We propose a MDH test based on this methodology that maintains such desirable properties. If $I_t = (X_t, X_{t-1}, ...)$ is the information set at time t and $\mathcal{F}_t$ is the σ-field generated by I_t, the first methodology exploits the following equivalence principle

$$E[Y_t \mid I_{t-1}] = \mu \ a.s., \ \mu \ \in \mathbb{R} \Longleftrightarrow E[(Y_t - \mu)f(I_{t-1})] = 0, \tag{2}$$

for all bounded $\mathcal{F}_{t-1}$-measurable weighting function $f(\cdot)$. Tests are usually based on the discrepancy of the sample analog of $E[(Y_t - \mu)f(I_{t-1})]$ to zero.

The problem of testing over all possible weighting functions can be reduced to testing the orthogonality condition over a parametric family of functions. That is

$$E[(Y_t - \mu)f(I_{t-1})] = 0 \ \forall f(\cdot) \Longleftrightarrow E[(Y_t - \mu)W(I_{t-1}, x)] = 0 \ \forall x \in \Lambda,$$

for a suitable family $W(I_{t-1}, x)$ and set Λ, see Stinchcombe and White (1998). The most used weighting functions have been either exponential functions, see e.g. Bierens (1984), or indicators functions, as in Stute (1997).

Most of the above references test the MDH conditioning on a finite-dimensional information set, and therefore, they test for a particular Markov property and not the MDH, which involves an infinite number of lags. This solution is unsatisfactory because there could be structure in the conditional mean at omitted lags. Often, the maximum power could be achieved by using the correct lag order of the alternative. However, prior information on the conditional mean structure is usually not available. In addition, when a large number of conditioned variables is considered, the empirical power of those tests could be seriously affected by the curse of dimensionality.

There have been some proposals considering infinite-dimensional information sets. de Jong (1996) generalized Bierens' test to time series, and although this test has the appealing property of considering an increasing number of lags as the sample size increases, it requires numerical integration with dimension equal to the sample size, which makes this test unfeasible in applications where

3

the sample size is usually large, e.g. financial applications. Using a different methodology based on spectral methods, Hong (1999) proposed a MDH test, but his test is not robust to conditional heteroskedasticity and higher order dependence because it is based on an independence assumption. This drawback has been recently overcome in Hong and Lee (2005). Hong's (1999) approach involves the choice of a kernel, a bandwidth parameter and an integrating measure and, in general, statistical inferences are not robust to these choices. Escanciano and Velasco (2003) avoid bandwidth choices by means of a generalized spectral distribution function. In this paper we propose a MDH test which preserves the good properties of Escanciano and Velasco's (2003) test and that complements it because it leads to different power properties. The main methodology is assisted with the individual use of new nonlinear dependence measures defined in the paper, which can play an important role in explaining well-known stylized facts of economic and financial data, see Section 6 for examples.

Summarizing, the aims of this paper are two: first, to develop a methodology for testing the MDH that (i) does not involve the choice of any lag order or kernel function; (ii) avoid the curse of dimensionality; (iii) is consistent against a large class of nonlinear and linear deviations from (1); (iv) is robust to higher order dependence, such as conditional heteroskedasticity; (v) is simple to compute; and (vi) performs quite well in finite samples as will be shown below, and secondly, to propose new measures of nonlinear dependence which can provide new useful tools for studying nonlinear serial dependence in applied work, as it is illustrated in the exchange rate application of Section 6.

The layout of the article is as follows, in Section 2 we define the Integrated Pairwise Regression Functions and the Integrated Pairwise Autoregression Functions as our measures of dependence. In Section 3 we use these new dependence measures to set a general methodology to test (1). In Section 4 we study the asymptotic distribution of our test under the null. In Section 5 we propose a bootstrap approach and we present a simulation exercise comparing with competing tests. In Section 6 we present an empirical application of our test and dependence measures to exchange rates. In the Appendix we collect some asymptotic results on the new dependence measures. In the sequel, C is a generic constant that may change from one expression to another. Let $\wedge$ denotes the minimum, i.e, $a \wedge b = \min\{a, b\}$. Unless indicated, all convergences are taken as the sample size $n \longrightarrow \infty$.

2. GENERALIZED DEPENDENCE MEASURES

In this section we define a generalization of the usual autocovariances and crosscovariances to a nonlinear framework. It is well known that in the presence of nonlinearity (or non-Gaussianity) these measures do not characterize the dependence in the conditional mean and the practitioner needs more reliable measures such as the pairwise regression functions $E[Y_t \mid X_{t-j} = x]$. In general, estimation

of these functions involves nonparametric smoothing estimation with subjective bandwidth choices. Robinson (1983) has studied the large sample properties of kernel estimators of lagged conditional means $E[X_t \mid X_{t-j}]$ for various lags j, see also Auestad and Tjøstheim (1990). A natural way to overcome the smoothing approach is to consider cumulative measures. Assume that $X \in \mathbb{R}^d$, and the random variable Y is integrable, so that the regression function

$$m(x) = E[Y - \mu \mid X = x], \qquad x \in \mathbb{R}^d,$$

is well defined (up to a null set). By a measure-theoretic argument, the regression function $m(\cdot)$ can be characterized by the integrated regression function $\gamma(x)$ given by

$$\gamma(x) = E[(Y - \mu)I(X \le x)] = \int_{(-\infty, x]^d} E[Y - \mu \mid X = z]F(dz),$$

where the second equality follows by the law of iterated expectations and $F(\cdot)$ is the cumulative distribution function of X. In a time series context is particularly interesting the case when $Y = Y_t$ and $X = Y_{t-j}$, in that case, the measures

$$\gamma_j(x) = E[(Y_t - \mu)I(Y_{t-j} \le x)]$$

can be called the Integrated Pairwise Autoregression Functions (IPAF). In the general case in which Y_t is different from X_t, the measures

$$\gamma_j(x) = E[(Y_t - \mu)I(X_{t-j} \le x)] \tag{3}$$

are the Integrated Pairwise Regression Functions (IPRF). These measures are useful for testing interesting hypotheses in a nonlinear time series framework as in (1). They are able to pick out both linear and nonlinear dependence in the conditional mean. The natural estimator of $\gamma_j(x)$ based on a sample $\{Y_t, X_t\}_{t=1}^n$ is

$$\widehat{\gamma}_j(x) = \frac{1}{n-j} \sum_{t=1+j}^n (Y_t - \overline{Y}_{n-j})I(X_{t-j} \le x), \tag{4}$$

with

$$\overline{Y}_{n-j} = \frac{1}{n-j} \sum_{t=1+j}^n Y_t.$$

In the Appendix we develop the asymptotic theory for $\widehat{\gamma}_j(x)$ under the null hypothesis (1). As a consequence of Proposition 1 in the Appendix and the Continuous Mapping Theorem we have that under H_0,

$$KS_{Y|X}(j) := \sup_{x \in [-\infty, \infty]^d} \left| (n-j)^{\frac{1}{2}} \widehat{\gamma}_j(x) \right| = \max_{1 \le t \le n} \left| (n-j)^{\frac{1}{2}} \widehat{\gamma}_j(X_t) \right|$$

converges to the supremum of a Gaussian process, where the subscript $Y \mid X$ in $KS_{Y|X}(j)$ indicates that Y is the dependent variable and X the conditioning variable at lag j. In particular, under

homoscedasticity and $d = 1$, a standardized version of $\widehat{\gamma}_j(x)$ has a standard Brownian bridge as the limiting distribution and asymptotic inference is possible because the asymptotic quantiles are readily available, see Shorack and Wellner (1986). In the general case, the quantile can be approximated via a bootstrap approach, see Section 5. With the bootstrap critical values we can calculate uniform confidence bands for $\widehat{\gamma}_j(x)$ and the significance of $\gamma_j(x)$ can be tested, see Section 6 below. Note that $\widehat{\gamma}_j(x)$ can be useful for detecting nonlinearities graphically (see Tong (1990), p.12). In contrast with kernel estimators, $\widehat{\gamma}_j(x)$ does not depend on kernel and bandwidth choices, so tests are straightforward to implement. Our test statistic for the MDH uses all the measures $\widehat{\gamma}_j(x)$ simultaneously to test a pairwise version of (1). The plot of a standardization of $KS_{Y|X}(j)$ against the lag parameter $j \geq 1$ can be viewed as generalization of the usual autocovariance plot in linear dependence to nonlinear conditional mean dependence. We call this plot for the indicator case, the IPRF plot. We illustrate its use with an example.

Example 1 *Consider the following nonlinear moving average model (NLMA)*

$$Y_t = \varepsilon_{t-1}\varepsilon_{t-2}(\varepsilon_{t-2} + \varepsilon_t + 1) \qquad \varepsilon_t \sim i.i.d\ N(0,1).$$

It is easy to show that this DGP is uncorrelated but not a MDS with respect its past values since $E[Y_tY_{t-1}^2] \neq 0$, so the usual correlogram is useless. However, the IPRF PLOT can display some useful information. In Figure 1 we compare both plots for a simulated series with $n = 500$ of the NLMA process. Considering heteroskedasticity robust correlations as in Deo (2000), the confidence bands under the null of MDH are, respectively, $\pm 1/\sqrt{n}$ for Deo's linear correlogram and those derived from the bootstrap approach for the IPRF PLOT, see Section 5 below.

<center>*Please insert Figure 1 about here.*</center>

<center>**3. PAIRWISE CONSISTENT HYPOTHESIS TESTING**</center>

Our test is founded on a pairwise approach that leads to computationally feasible tests which are consistent against a broad class of alternatives, cf. Escanciano and Velasco (2003). We test that all the IPRF are identically zero, i.e.

$$\gamma_j(x) = 0 \qquad \forall j \geq 1, \forall x \in \mathbb{R}^d,\ a.s.. \tag{5}$$

If we define $\gamma_{-j}(\cdot) = \gamma_j(\cdot)$ for $j \geq 1$, we can consider the Fourier transform of the IPRF $\gamma_j(\cdot)$,

$$f(w,x) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma_j(x)e^{-ijw} \qquad \forall w \in [-\pi,\pi], x \in \mathbb{R}^d, \tag{6}$$

which contains all the information about (5). Because we are taking into account linear and nonlinear dependence, $f(w,x)$ is a generalization of the spectral (or cross-spectral) density function, cf. Hong

<center>6</center>

(1999). We can also consider the generalized spectral distribution function as the integral of $f(w, x)$,

$$H(\lambda, x) = 2 \int_0^{\lambda \pi} f(w, x) dw \qquad \forall \lambda \in [0, 1], x \in \mathbb{R}^d,$$

that is,

$$H(\lambda, x) = \gamma_0(x)\lambda + 2 \sum_{j=1}^{\infty} \gamma_j(x) \frac{\sin j\pi\lambda}{j\pi}. \tag{7}$$

Note that, both $f(w, x)$ and $H(\lambda, x)$ exist as functions in an appropriate Hilbert space. The generalized spectral distribution function contains all the information about the pairwise regressions functions and could be viewed as a generalization of the test statistic used in Durlauf (1991) and Deo (2000) because they considered the autocorrelations, instead of the pairwise regression functions, and only took into account the second moment implications of the MDH. By using $H(\lambda, x)$, we consider all the pairwise implications of the MDH, including both linear and non-linear conditional dependencies. Notice that a more flexible weighting scheme on the measure $\gamma_j(x)$ is possible via a kernel function and a lag-bandwidth parameter, but this approach would introduce some arbitrariness in the test via the kernel and bandwidth choices.

Our test is based on the sample analogue of (7),

$$\widehat{H}(\lambda, x) = \widehat{\gamma}_0(x)\lambda + 2 \sum_{j=1}^{n-1} (1 - \frac{j}{n})^{\frac{1}{2}} \widehat{\gamma}_j(x) \frac{\sin j\pi\lambda}{j\pi},$$

with $(1 - j/n)^{1/2}$ a finite sample correction factor which delivers a better finite sample performance and where $\{\widehat{\gamma}_j(x)\}_{j=1}^{n-1}$ are given by (4). Notice that all the $n - 1$ lags in the sample are used, and therefore, there is no need to choose a lag order. Therefore, all of the information in the conditional moment restriction is eventually accounted for. Because (5) is equivalent to $H(\lambda, x) = \gamma_0(x)\lambda$, the test is based on the discrepancy between $\widehat{H}(\lambda, x)$ and $\widehat{H}_0(\lambda, x) := \widehat{\gamma}_0(x)\lambda$. That is, we consider the process

$$S_n(\lambda, x) = (\frac{n}{2})^{\frac{1}{2}} \{\widehat{H}(\lambda, x) - \widehat{H}_0(\lambda, x)\} = \sum_{j=1}^{n-1} (n - j)^{\frac{1}{2}} \widehat{\gamma}_j(x) \frac{\sqrt{2} \sin j\pi\lambda}{j\pi}$$

to test H_0.

Informally speaking, under the null (1) the generalized sample spectral distribution $\widehat{H}(\lambda, x)$ will be approximately equal to $\widehat{H}_0(\lambda, x)$ and then, the process $S_n(\lambda, x)$ will converge in distribution as n increases, while under the alternative it is expected that $\widehat{H}(\lambda, x)$ will differ from $\widehat{H}_0(\lambda, x)$ and hence, $S_n(\lambda, x)$ will diverge to infinity when n increases.

In order to evaluate the distance of $S_n(\lambda, x)$ to zero, a norm has to be chosen. In this context the natural norm is the Cramér-von Mises (CvM) norm. If $F_n(x)$ is the usual empirical distribution function based on $\{X_t\}_{t=1}^n$, the CvM norm is

$$D_n^2 := \int (S_n(\lambda, x))^2 F_n(dx) d\lambda = \sum_{j=1}^{n-1} \frac{(n - j)}{n(j\pi)^2} \sum_{t=1}^{n} \widehat{\gamma}_j^2(X_t). \tag{8}$$

7

Our test rejects the null hypothesis for large values of the test statistic D_n^2.

4. ASYMPTOTIC NULL DISTRIBUTION

In this section, we first establish the null limit distribution of the empirical process $S_n(\lambda, x)$ under (1). The null limit distribution of the new test is the limit distribution of a functional of $S_n(\lambda, x)$. To further elaborate the foregoing points we need some notation. Let ν be the product measure of the probability distribution F and the Lebesgue measure on $[0, 1]$. Also, denote $\Pi = [0, 1] \times [-\infty, \infty]^d$ and $\eta := (\lambda, x) \in \Pi$. We consider $S_n(\eta)$ as a random element on the Hilbert space $L_2(\Pi, \nu)$ of all square integrable functions (with respect to the measure ν) with the inner product

$$\langle f, g \rangle := \int_\Pi f(\eta) g(\eta) d\nu(\eta) = \int_\Pi f(\lambda, x) g(\lambda, x) F(dx) d\lambda.$$

For recent applications on convergence results on Hilbert spaces in the econometric literature see Politis and Romano (1994), Chen and White (1996, 1998), Chen and Fan (1999), Escanciano and Velasco (2003) and Escanciano (2006b). If Z is a $L_2(\Pi, \nu)-$valued random variable, we say that Z has mean $m \in L_2(\Pi, \nu)$ if $E[\langle Z, f \rangle] = \langle m, f \rangle$, $\forall f \in L_2(\Pi, \nu)$. If $E \|Z\|^2 < \infty$ and Z has zero mean, then the covariance operator of Z, $C_Z(\cdot)$ say, is a continuous, linear, symmetric, positive definite operator from $L_2(\Pi, \nu)$ to $L_2(\Pi, \nu)$ defined by $C_Z(h) = E[\langle Z, h \rangle Z]$. Let $\Longrightarrow$ denote weak convergence in the Hilbert space $L_2(\Pi, \nu)$ endowed with the norm metric. Define

$$w_t(x) := \{I(X_t \leq x) - F(x)\}, \qquad t \in \mathbb{Z}, x \in \mathbb{R}^d$$

and

$$\Psi_j(\lambda) = \frac{\sqrt{2} \sin j\pi\lambda}{j\pi}, \qquad j \geq 1.$$

Assumption A1:

A1(a): $\{X_t\}_{t=1}^\infty$ is a stationary ergodic vector process in $\mathbb{R}^d$ with probability distribution function $F(\cdot)$ absolutely continuous with respect to Lebesgue measure. Also, Y_t is a measurable real-valued function of X_t.

A1(b): $E |Y_1|^2 < \infty$.

Note that Assumption A1 is mild, in particular it allows us to consider conditional heteroskedastic processes. Assumption A1(b) is much weaker than the moment assumptions used in Durlauf (1991) and Deo (2000) for testing the MDH, who assumed finite eight moments. The next theorem shows the null limit distribution of the $L_2(\Pi, \nu)-$valued random element $S_n(\eta)$. The reader is referred to Escanciano (2006b) for its proof.

THEOREM 1: *Under Assumption A1 and (1), the process $S_n(\eta)$ converges weakly to $S(\eta)$ on $L_2(\Pi, \nu)$,*

where $S(\eta)$ is a Gaussian process with mean zero and covariance operator

$$C_S(f) = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} E[(Y_1 - \mu)^2 \int_{\Pi \times \Pi} f(\eta_1)f(\eta_2)w_{1-j}(x_1)\Psi_j(\lambda_1)w_{1-k}(x_2)\Psi_k(\lambda_2)d\nu(\eta_1)d\nu(\eta_2)]$$

where $f \in L_2(\Pi, \nu)$, $\eta_1 = (\lambda_1, x_1)$ and $\eta_2 = (\lambda_2, x_2)$.

Remark 1. It is easy to show that under independence and using the equality

$$\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\Psi_j(\lambda)\Psi_k(\lambda') = (\lambda \wedge \lambda' - \lambda\lambda'), \tag{9}$$

the covariance function of the process $S(\eta)$ is equal to $\sigma^2(F(x \wedge x') - F(x)F(x'))(\lambda \wedge \lambda' - \lambda\lambda')$ where $\sigma^2 = E[(Y_1 - \mu)^2]$. Thus, in this case the asymptotic null distribution is nuisance parameter-free after standardization by a consistent estimate of σ^{-1}, and the quantiles of the asymptotic distribution of norms of $S(\eta)$ can be tabulated.

Next corollary shows the asymptotic distribution of our test statistic, and follows from the Continuous Mapping Theorem (Billingsley's (1968) Theorem 5.1), Theorem 1 and Lemma 3.1 in Chang (1990).

COROLLARY 1: *Under (1) and Assumption A1*

$$D_n^2 \xrightarrow{d} D_\infty^2 := \int (S(\lambda, x))^2 F(dx)d\lambda.$$

It can be shown that the asymptotic distribution of D_n^2 can be expressed as a weighted sum of independent χ_1^2 random variables with weights depending on the DGP. This could be a basis to approximate the asymptotic distribution of D_n^2 and obtain critical values of tests. We do not pursue this possibility here, on the contrary we approximate the asymptotic critical values of the test with the assistance of a bootstrap procedure in the next section.

The consistency and local power properties for D_n^2 can be discussed following the lines of Section 3.2 in Escanciano and Velasco (2003). The reader is referred to the latter reference for the asymptotic power local properties of the CvM test based on D_n^2. In particular, under A1

$$\frac{1}{n}D_n^2 \xrightarrow{P} \sum_{j=1}^{\infty}\frac{1}{(j\pi)^2}\int_{\mathbb{R}^d} |\gamma_j(x)|^2 F(dx).$$

Thus the test is consistent provided there exists at least one $j \geq 1$ such that $\gamma_j(x) \neq 0$ for some subset of $\mathbb{R}^d$ with positive Lebesgue measure. Note however, that there could exist non-MDS such that $\gamma_j(x) = 0$, $\forall j \geq 1$, and our test will not be able to detect such alternatives. An example of such alternatives is $Y_t = \varepsilon_{t-1}\varepsilon_{t-3} + \varepsilon_t$, where $\varepsilon_t \sim i.i.d$ N(0,1). It is interesting to formally characterize the class of non-MDS that satisfies (5), but this is beyond the scope of this paper.

Moreover, it can be shown that under mild assumptions our test based on D_n^2 is asymptotically admissible, that is, there does not exist a test that is uniformly more powerful than our test for a general class of local alternatives to H_0, see Corollary 4 in Escanciano and Velasco (2003). Thus, the test proposed in this paper should be viewed not as competing with but as a complement to that considered in Escanciano and Velasco (2003) since they have different power properties and both are asymptotically admissible.

5. BOOTSTRAP APPROXIMATION AND FINITE SAMPLE PERFORMANCE

In this section we estimate the distribution of $S_n(\eta)$ by that of

$$S_n^*(\eta) = \sum_{j=1}^{n-1} (n-j)^{1/2} \widehat{\gamma}_j^*(x) \frac{\sqrt{2} \sin j\pi\lambda}{j\pi},$$

with

$$\widehat{\gamma}_j^*(x) = (n-j)^{-1} \sum_{t=1+j}^{n} (Y_t - \overline{Y}_{n-j})\{I(X_{t-j} \leq x) - F_{n-j}(x)\}W_t,$$

$$F_{n-j}(x) = (n-j)^{-1} \sum_{t=1+j}^{n} I(X_{t-j} \leq x)$$

and where $\{W_t\}$ is a sequence of independent random variables with zero mean, unit variance, bounded support and also independent of the sequence $\{X_t\}_{t=1}^n$. We construct the bootstrap CvM norm as

$$D_n^{*2} := \int (S_n^*(\lambda, x))^2 F_n(dx)d\lambda.$$

This procedure is similar to the wild bootstrap, see, e.g. Hardle and Mammen (1993). Its theoretical justification can be found in Escanciano and Velasco (2003). In particular, we can simulate the critical values for the tests statistics D_n^2 and computing uniform confidence bands for $\gamma_j(x)$ by the following algorithm:

1. Calculate the test statistic D_n^2 with the original sample.

2. Generate W_t, a sequence of independent random variables with zero mean, unit variance, bounded support and that is independent of the sample $\{X_t\}_{t=1}^n$.

3. Compute $\widehat{\gamma}_j^*(x)$, $S_n^*(\eta)$ and D_n^{*2}.

4. Repeat steps 2 and 3 B times and compute the empirical $(1-\alpha)th$ sample quantile of D_n^{*2} with the B values, $D_{n,\alpha}^{*2}$. The proposed test rejects the null hypothesis at the significance level α if $D_n^2 > D_{n,\alpha}^{*2}$.

In order to examine the finite sample performance of the proposed test we carry out a simulation experiment with several DGP's under the null and under the alternative. The emphasis in these simulations is on the comparison between indicator-based tests for testing the classical MDH, that is, when $Y_t = X_t$ and $d = 1$. The first block of models considered here have been used by Dominguez and Lobato (2003) and thus, they will be useful for comparing both tests. We also compare the tests with the conditional heteroskedasticity robust Durlauf's (1991) test proposed by Deo (2000).

We briefly describe our simulation setup. We denote D_n^2 the new CvM test statistic defined in (8). Let $\overline{Y} = n^{-1}\sum_{t=1}^n Y_t$ be the usual sample mean.

Dominguez and Lobato (2003) have considered a MDH test taking into account a fixed number of lags. We denote by CvM_P and KS_P the CvM and Kolmogorov-Smirnov statistics respectively, with P as the number of lags used. These statistics are based on the multivariate integrated regression function, i.e.

$$CvM_P = \frac{1}{n^2}\sum_{j=1}^n \left[\sum_{t=1}^n (Y_t - \overline{Y})I(\widetilde{z}_{t,P} \leq \widetilde{z}_{j,P})\right]^2,$$

$$KS_P = \max_{1 \leq i \leq n}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^n (Y_j - \overline{Y})I(\widetilde{z}_{j,P} \leq \widetilde{z}_{i,P})\right|,$$

where $\widetilde{z}_{t,P} = (Y_{t-1}, ..., Y_{t-P})$ is the P-lagged values of the series. To save space, only results for $P = 1$ and 2 are presented. Note that, CvM_1 and KS_1 are functionals of a process studied by Koul and Stute (1999) in a more general context.

Recently, Deo (2000) has proposed a correction of the Durlauf's (1991) test to take into account conditional heteroskedasticity. The corrected statistic is

$$DURC := \sum_{j=1}^{n-1} n\widehat{a}_j^2\left(\frac{1}{j\pi}\right)^2$$

where,

$$\widehat{a}_j = \widehat{\rho}_j\left[(n-j)^{-1}\sum_{t=1}^{n-j}(Y_t - \overline{Y})^2(Y_{t+j} - \overline{Y})^2\right]^{-1/2},$$

$$\widehat{\rho}_j = (n-j)^{-1}\sum_{t=1}^{n-j}(Y_t - \overline{Y})(Y_{t+j} - \overline{Y}).$$

We have considered the factor $(n-j)^{-1}$ in $\widehat{\rho}_j$ instead of n^{-1} as used in Deo (2000) because it gives a better finite sample performance. Note that we do not consider a kernel or weighting function. Under the null hypothesis of the MDS and some additional assumptions (see Deo (2000)),

$$DURC \xrightarrow{d} \int_0^1 B^2(t)dt \text{ as } n \longrightarrow \infty,$$

11

where $B(t)$ is the standard Brownian bridge on [0,1]. The 10%, 5% and 1% asymptotic critical values are obtained from Shorack and Wellner (1986, p.147) and are 0.347, 0.461 and 0.743 respectively, although we have also used in the simulations empirical critical values. In the sequel $\varepsilon_t \sim i.i.d \ N(0,1)$. The first block of models considered in the simulations are two martingale difference sequences:

1. A sequence of $i.i.d$ N(0,1) variates.

2. GARCH(1,1) processes:

$$Y_t = \varepsilon_t \sigma_t, \qquad \sigma_t^2 = w + \alpha Y_{t-1}^2 + \beta \sigma_{t-1}^2,$$

with $w = 0.001$ and the following combinations for (α, β): (0.01, 0.97), (0.09, 0.89) and (0.09, 0.90), we call these processes GARCH1, GARCH2 and GARCH3 respectively.

And the following non-martingale sequences:

3. Non-Linear Moving Average (NLMA) process: $Y_t = \varepsilon_{t-1}\varepsilon_{t-2}(\varepsilon_{t-2} + \varepsilon_t + 1)$.

4. Bilinear Processes:

$$Y_t = \varepsilon_t + b_1 \varepsilon_{t-1} Y_{t-1} + b_2 \varepsilon_{t-1} Y_{t-2},$$

with (b_1, b_2): $(0.15, 0.05)$ and $(0.25, 0.15)$, we call these process BIL-I and BIL-II respectively.

Note that the second and third GARCH models have unbounded eighth and sixth moment respectively. Also note that the NLMA process is uncorrelated and, therefore, tests based on autocorrelations, see for instance Box and Pierce (1970), Durlauf (1991) or Hong (1996), have no asymptotic power against this model.

We consider for the experiments under the null a sample size of $n = 100$ and under the alternative $n = 100, 200$ and 300. The number of Monte Carlo experiments is 1000 and the number of bootstrap replications is $B = 500$. In all the replications 200 pre-sample data values of the processes were generated and discarded. Random numbers were generated using IMSL ggnml subroutine. We employ a sequence $\{W_t\}$ of $i.i.d$ Bernoulli variates where $P(W_t = 0.5(1 - \sqrt{5})) = (1 + \sqrt{5})/2\sqrt{5}$ and $P(W_t = 0.5(1 + \sqrt{5})) = 1 - (1 + \sqrt{5})/2\sqrt{5}$. Note that the third moment of W_t is equal to 1 and hence the first three moments of the bootstrap series coincide with the three moments of the original series, see Stute, Gonzalez-Manteiga and Presedo-Quindimil (1998).

In Table I and Table II we show the empirical rejection probabilities (RP) associated with the three nominal levels 10%, 5% and 1%. The results for D_n^2, CvM_1, KS_1, CvM_2, KS_2 and $DURC$ show that all the tests have good size properties and are robust to thick tails given their behavior with GARCH models. For $DURC$ we also present in brackets the RP using the empirical critical values based on simulations with $i.i.d$ standard normal variables and 10.000 replications.

Please insert Tables I and II about here.

In Table III we report the empirical power against the NLMA process. It increases with the sample size n, as expected. Generally, the Kolmogorov-Smirnov test has more empirical power than CvM tests. There is no test which dominates uniformly the others, although the statistics CvM_P and KS_P for $P \geq 2$ have less power as P increases. Deo's (2000) statistic, $DURC$, has no power against this alternative, as expected, because this NLMA model is uncorrelated. Among the statistics D_n^2, CvM_1 and KS_1 the difference is not substantial, because the NLMA process has dependence in the conditional mean only for the first lags.

Please insert Table III about here.

In Tables IV and V we show the RP for the bilinear models. As in the case of the NLMA, there is no test which dominates uniformly the others and, again, the statistics CvM_P and KS_P for $P \geq 2$ have less power in almost all cases. In the BIL-II case, CvM_1 and KS_1 perform slightly better than D_n^2, but the difference is not substantial. The empirical power of $DURC$ is lower in all cases.

Please insert Tables IV and V about here.

In these examples the dependence is present only at the first lags, and therefore there is not too much practical difference between D_n^2, CvM_1 and KS_1. For models with higher order dependence structure is expected that D_n^2 will have more empirical power than CvM_1 and KS_1 because D_n^2 considers all lags. To illustrate the latter fact, we consider the following second block of models:

5. A linear ARMA(1,2) model:

$$(1 - 0.3L)Y_t = (1 - 0.5L^2)\varepsilon_t,$$

   where L is the lag shift operator, i.e. $LY_t = Y_{t-1}$.

6. A non-Gaussian moving average model (NGMA): $Y_t = \exp(\varepsilon_t) - 0.7\exp(\varepsilon_{t-3})$.

7. A threshold autoregressive model (TAR)

$$Y_t = 0.1Y_{t-3} - 0.5Y_{t-4} + \varepsilon_t \text{ if } Y_{t-3} \geq 1$$

$$Y_t = -0.5Y_{t-3} + 0.4Y_{t-4} + \varepsilon_t \text{ if } Y_{t-3} < 1.$$

The empirical power for a sample size of $n = 100$ and level 5% and the same design as Tables III-V is shown in Table VI. Now, D_n^2 has more empirical power than CvM_P and KS_P, $P = 1$ and 2, because of the higher lag dependence. $DURC$ has better empirical power properties than D_n^2 for the ARMA(1,2) alternative, as expected, since there is no gain in considering nonlinear dependence when

13

testing against this linear model. On the contrary, our test is superior to $DURC$ for the NGMA and TAR alternatives. The conclusion for these higher order dependence cases is that is more preferred to sum-up pairwise information at various lags than consider many of them simultaneously. Though our additive approach restricts the hypothesis to be tested (from (1) to (5)), it is able to break the curse of the dimensionality that affects the statistics CvM_P and KS_P for moderate P.

Please insert Table VI about here.

Summarizing, D_n^2 has omnibus good empirical power against all linear and nonlinear dependencies, independently if the dependence is at high or low lags. In many cases it presents better empirical power properties than competing tests while showing and overall power against all the alternatives considered for moderate sample sizes.

6. EXCHANGE RATES DYNAMICS

In this section we investigate by means of our generalized spectral distribution MDH-test and the IPAF measures the dynamics of the daily log price changes of the British Pound exchange rate in terms of the US Dollar exchange rate (BPUSD). This problem has been explored in, e.g., Hsieh (1989), Gallant, Hsieh and Tauchen (1991), Bera and Higgins (1997) and examined recently by Dominguez and Lobato (2003), among others. The data consist in two samples, first from January 2nd, 1974 to December 31st, 1983 (BPUSD1) and the second period from December 12th, 1985 to February 28th, 1991 (BPUSD2). For a better comparison, we discard the 10% final observations in BPUSD2 as in Bera and Higgins (1997). Then, the number of total observations for BPUSD1 and BPUSD2 are 2505 and 1210 respectively. The rates of change are calculated by taking the logarithmic differences between successive trading days, i.e., $Y_t = 100 \log(r_t/r_{t-1})$, where r_t denote the U.S. Dollar price of a Pound at time t. Table VII provides summary statistics of the data. The sample distribution of the data has heavy tails in the two periods, specially in BPUSD1, and the sample kurtosis coefficients are substantially larger than that of the standard normal distribution (which is 3). Earlier investigations focused on the linear predictability (or lack thereof) of the exchange rates. Usually, exchange rates data are serially uncorrelated. Here we use Deo's (2000) test statistic to check if the data are uncorrelated. Previous results have shown that BPUSD1 have little linear dependence, see e.g. Table 2 in Hsieh (1989). This is in agreement with the p-values of $DURC$ test in Table VIII. These results also show that BPUSD2 has not linear dependence, which is in agreement with the findings of Bera and Higgins (1997).

Please insert Table VII about here.

There has been some evidence in the literature supporting that exchange rate changes exhibit nonlinear dependence, see e.g. Hsieh (1989). We now investigate the type of the serial dependence presented in the BPUSD by means of the IPAF and the generalized spectral distribution test. An important problem is to distinguish whether the serial dependence affects the conditional mean or the conditional variance. The solution of this problem is crucial because it has important implications in economics. Hsieh (1989) proposed a test based on third order cumulants to discriminate between both types of dependence and found evidence in favor of the multiplicative dependence, that is, evidence that the serial dependence affects the conditional variance, and rejecting the dependence in the conditional mean. To test these hypotheses we use our generalized spectral test and present the results for the two periods in Table VIII. To facilitate interpretations we show the p-values for D_n^2, CvM_P and KS_P for $P = 1$ and 2. For BPUSD1, Dominguez and Lobato (2003) found evidence for $P = 1$ against the MDH, but for $P = 2$ they supported the MDH. Note that this is a contradictory result. The statistic D_n^2 shows strong evidence against the MDH, which disagrees with the findings of Hsieh (1989). These results can be explained by the fact that the data has third order cumulants equal to zero, but is not a MDS. For the second period we find evidence supporting the MDH, with all the tests. Then, a nonlinear model for the conditional mean can not explain the nonlinear dependence in the BPUSD2, in particular the bilinear model used in Bera and Higgins (1997).

Please insert Table VIII about here.

To gain insight in the serial dependence properties of the data we consider the IPAF for the two periods and also for the squares of the data. In Tables IX and X we show the Kolmogorov-Smirnov tests statistics for the IPAF at different lags and different combinations of variables, and the bootstrap 95% quantile to indicate the significance of the IPAF for the BPUSD1 and BPUSD2. For the definition of the corresponding Kolmogorov-Smirnov test statistics $KS_{Y|X}(j)$ see Section 2. Table IX reveals that the nonlinearity in the conditional mean is significative at lags 1 and 5 for the BPUSD1, confirming a weekly effect in this daily data set. The BPUSD1 also seems to be highly heteroskedastic at all lags and conditionally asymmetric at lag $j = 1$. The results in Table X display a different behavior for the BPUSD2, in which all the IPAF are not significant except for some integrated conditional variances.

Please insert Tables IX and X about here.

We also plot in Figures 2 and 3 the IPRF at lag $j = 1$ for the BPUSD1 data levels and squares respectively (Y_t and Y_t^2 as dependent variables and Y_{t-1} as the conditioning variable). We also plot the corresponding uniform confidence bands under the null of the MDH and under the alternative. The pattern of the IPAF for the data confirms some stylized facts about exchange rate changes. For instance, Figure 3 shows the well-known "Leverage effect" in the exchange rates, in which volatility

is higher when past rates changes are negative. The IPAF for Y_t^3 as dependent variable, which has not been shown for the sake of space, reveals that large changes in exchange rates are often negative.

On the other hand, Figure 2 shows the IPAF for the BPUSD1 data and confirms that the MDH is rejected at lag $j = 1$, although the autocorrelation at lag $j = 1$ is not significatively different from zero. This reveals that the conditional mean at this lag can not be linear and there is nonlinear dependence. Since the slope of the IPAF at each point is proportional to the regression function we can conclude that the regression function has the same sign as the lagged exchange rate change. This feature supports the well-known fact that the sample autocorrelation at lag $j = 1$ of exchange rates is usually positive.

Please insert Figures 2 and 3 about here.

Summarizing, our new test finds nonlinear dependence in the conditional mean of the BPUSD1 exchange rate changes, contrasting with some previous studies which assume that exchange rate changes are very nearly to be unpredictable given past prices, and in agreement with more recent findings by Escanciano and Velasco (2003) and Hong and Lee (2003). The nonlinearity in the conditional mean of BPUSD1 suggests that additional effort has to be dedicated to investigate the form of such nonlinearity in the conditional mean before modeling the conditional variance.

APPENDIX: ASYMPTOTIC RESULTS FOR THE IPRF

In this section we show the weak convergence of $\widehat{\gamma}_j(\cdot)$ in (4) under H_0 (and mild regularity conditions.) Throughout the Appendix $j \geq 1$ is fixed.

Assumption B1:

B1(a): $\{X_t\}_{t=1}^{\infty}$ is a stationary ergodic vector process in $\mathbb{R}^d$ with probability distribution function $F(\cdot)$ absolutely continuous with respect to Lebesgue measure. Also, Y_t is a measurable real-valued function of X_t.

B1(b): $E[Y_t^4 |X_{t-j}|^{1+\delta}] < \infty$, for some $\delta > 0$.

B1(c): The conditional density of Y_t given X_{t-j} is (uniformly) bounded and continuous.

The next proposition shows the null limit distribution of $\widehat{\gamma}_j(\cdot)$.

PROPOSITION 1: *Suppose that Assumption B1 holds. Then, under (1) the process $(n-j)^{\frac{1}{2}}\widehat{\gamma}_j(\cdot)$ converges weakly to $B_j(\cdot)$ on the Skorohod space $D[-\infty, \infty]^d$, where $B_j(\cdot)$ is a Gaussian process with zero mean and covariance function*

$$K_j(x, x') = E[(Y_1 - \mu)^2 w_{1-j}(x) w_{1-j}(x')]$$

where $w_t(x) := \{I(X_t \leq x) - F(x)\}$.

PROOF OF PROPOSITION 1: First, write

$$\widehat{r}_j(x) = \frac{1}{n-j} \sum_{t=1+j}^{n} (Y_t - \mu)\{I(X_{t-j} \leq x) - F(x)\}.$$

Note that

$$\widehat{\gamma}_j(x) = \widehat{r}_j(x) - \left[\frac{1}{n-j} \sum_{t=1+j}^{n} (Y_t - \mu)\right] \left[\frac{1}{n-j} \sum_{t=1+j}^{n} (I(X_{t-j} \leq x) - F(x))\right]. \qquad (10)$$

Now, it is straightforward to show that under (1) for fixed K, $1 < K < n$,

$$(n-j)^{-1/2} \sum_{t=1+j}^{n} (Y_t - \mu) = O_P(1), \quad \forall 1 \leq j < K,$$

whereas by stationarity, ergodicity and monotonicity of $F_{n-j}(x)$ we have that the Glivenko-Cantelli Theorem holds, i.e.

$$\sup_{x \in \mathbb{R}^d} |F_{n-j}(x) - F(x)| = o_p(1), \quad \forall 1 \leq j < K.$$

Hence, for each fixed j, $1 \leq j < K$,

$$\sup_{x \in \mathbb{R}^d} |(n-j)^{1/2}\{\widehat{\gamma}_j(x) - \widehat{r}_j(x)\}| = o_P(1) \ as \ n \to \infty.$$

Therefore, it is sufficient to show the weak convergence of the process $\widehat{r}_j(x)$ which follows from Lemma 2 in Dominguez and Lobato (2004).

REFERENCES

[1] Auestad, B. and Tjøstheim, D., (1990). Identification of nonlinear time series: First order characterization and order determination. Biometrika 77, 669-687.

[2] Bera, A. K and Higgins, M. L., (1997). ARCH and bilinearity as competing models for nonlinear dependence. Journal of Business and Economic Statistics 15, 43-50.

[3] Bierens, H. J., (1984). Model specification testing of time series regressions. Journal of Econometrics 26, 323-353.

[4] Bierens, H. and Ploberger, W., (1997). Asymptotic theory of integrated conditional moment tests. Econometrica 65, 1129-1151.

[5] Billingsley, P., (1968). Convergence of Probability Measures. Wiley, New York.

[6] Box, G. and Pierce, D., (1970). Distribution of residual autocorrelations in autorregressive integrated moving average time series models. Journal of the American Statistical Association 65, 1509-1527.

[7] Chang, N. M., (1990). Weak convergence of a self-consistent estimator of a survival function with doubly censored data. Annals of Statistics 18, 391-404.

[8] Chen, X. and White, H., (1996). Laws of large numbers for Hilbert space-valued mixingales with applications. Econometric Theory 12, 284-304.

[9] ――――― (1998). Central limit and functional central limit theorems for Hilbert-valued dependent heterogeneous arrays with applications. Econometric Theory 14, 260-284.

[10] Chen, X. and Fan, Y., (1999). Consistent hypothesis testing in semiparametric and nonparametric models for econometric time series. Journal of Econometrics 91, 373-401.

[11] Cochrane, J. H., (1988). How big is the random walk in GNP? Journal of Political Economy 96, 893-920.

[12] de Jong, R.M., (1996). The Bierens' test under data dependence. Journal of Econometrics 72, 1-32.

[13] Deo, R. S., (2000). Spectral tests of the martingale hypothesis under conditional heroscedasticity. Journal of Econometrics 99, 291-315.

[14] Dominguez, M.A and Lobato, I.N., (2003). A consistent test for the martingale difference hypothesis. Econometric Reviews 22, 351-377.

[15] ——— (2004). Consistent estimation of models defined by conditional moment restrictions. Econometrica 72, 1601–1615.

[16] Durlauf, S., (1991). Spectral based testing of the martingale hypothesis. Journal of Econometrics 50, 355-376.

[17] Escanciano, J. C., and Velasco, C., (2003). Generalized spectral tests for the martingale difference hypothesis. Forthcoming in Journal of Econometrics.

[18] Escanciano, J. C., (2006a). Model checks using residual marked empirical processes. Forthcoming in Statistica Sinica.

[19] ——— (2006b). Goodness-of-fit tests for linear and non-linear time series models. Forthcoming in Journal of the American Statistical Association.

[20] Ferreira, E. and Stute, W., (2004). Testing for differences between conditional means in a time series context. Journal of the American Statistical Association 99, 169-174.

[21] Gallant, A. R., Hsieh, D. A. and Tauchen, G., (1991). On Fitting a Recalcitrant Series: The Pound/Dollar Exchange Rate, 1974-1983. in Nonparametric and Semiparametric Methods in Econometrics and Statistics. Eds. W. A. Barnett, J. Powell and G. Tauchen, Cambridge, U.K. Cambridge University Press, pp. 199-240.

[22] Guerre, E. and Lavergne, P., (2005). Rate-optimal data-driven specification testing for regression models. Annals of Statistics 33, 840-870.

[23] Härdle, W. and Mammen, E., (1993). Comparing nonparametric versus parametric regression fits. Annals of Statistics 21, 1926-1974

[24] Hong, Y., (1996). Consistent testing for serial correlation of unknown form. Econometrica 64, 837-864.

[25] ——— (1999). Hypothesis testing in time series via the empirical characteristic function: A generalized spectral density approach. Journal of the American Statistical Association 84, 1201-1220.

[26] Hong, Y. and Chung, J., (2003). Are the directions of stock price changes predictable? Statistical theory and evidence. Working paper.

[27] Hong, Y. and Lee, T.H., (2003). Inference on predictability of foreign exchange rate changes via generalized spectrum and nonlinear time series models. Review of Economics and Statistics 85, 1048-1062.

[28] Hong, Y., and Lee, Y. J., (2005). Generalized spectral tests for conditional mean models in time series with conditional heteroskedasticity of unknown form. Review of Economic Studies 72, 499-541.

[29] Hsieh, D. A., (1989). Testing for nonlinear dependence in daily foreign exchange rates. Journal of Business 62, 339-368.

[30] Koul H.L. and Stute W., (1999). Nonparametric model checks for time series. Annals of Statistics 27, 204-236.

[31] Li, Q., (1999). Consistent model especification tests for time series econometric models. Journal of Econometrics 92, 101-147.

[32] Linton, O. and Whang, Y-J., (2004). The Quantilogram: with an application to evaluating directional predictability. Forthcoming in Journal of Econometrics.

[33] Politis, D. and J. Romano, (1994). Limit theorems for weakly dependent Hilbert space valued random variables with application to the stationary bootstrap. Statistica Sinica 4, 461-476.

[34] Robinson, P.M., (1983). Nonparametric estimators for time series. Journal of Time Series Analysis 4, 185-207.

[35] Shorack, G. and Wellner, J., (1986). Empirical Processes with Applications to Statistics. Wiley: New York.

[36] Stinchcombe, M. and White, H., (1998). Consistent specification testing with nuisance parameters present only under the alternative. Econometric Theory 14, 295-325.

[37] Stute, W., (1997). Nonparametic model checks for regression. Annals of Statistics 25, 613-641.

[38] Stute, W., Gonzalez-Manteiga, W. and Presedo-Quindimil, M., (1998). Bootstrap approximations in model checks for regression. Journal of the American Statistical Association 93, 141-149.

[39] Tong, H., (1990). Nonlinear Time Series: A Dynamic System Approach. Clarendon Press: Oxford.

[40] Whang, Y-J., (2000). Consistent bootstrap tests of parametric regression functions. Journal of Econometrics 98, 27-46

[41] Zheng, X., (1996). A consistent test of funtional form via nonparametric estimation technique. Journal of Econometrics 75, 263-289.
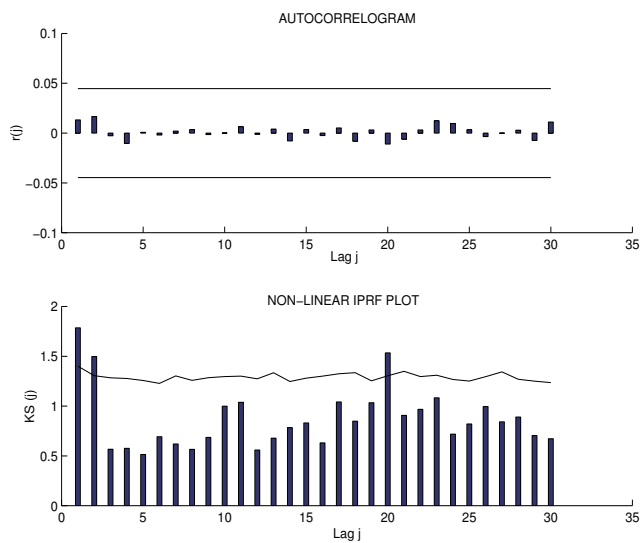
FIG. 1. Linear Correlogram and Nonlinear IPRF Plot for the NLMA data

TABLE I: SIZE OF TESTS.

	IID			GARCH1		
	10%	**5%**	**1%**	**10%**	**5%**	**1%**
D_n^2	9.9	5.0	0.8	10.2	4.7	0.9
CvM_1	9.3	4.7	0.8	9.7	5.1	0.7
KS_1	10.8	5.6	0.8	11.2	5.9	0.7
CvM_2	10.3	6.1	0.9	10.6	4.9	1.3
KS_2	11.5	6.5	1.9	10.4	6.2	1.2
$DURC$	10.8	5.0	1.1	10.8	4.8	0.9
	[9.6]	[4.2]	[1.0]	[9.1]	[4.1]	[0.9]

TABLE II: SIZE OF TESTS.

	GARCH2			GARCH3		
	10%	**5%**	**1%**	**10%**	**5%**	**1%**
D_n^2	9.3	5.2	1.1	10.4	4.9	1.2
CvM_1	9.2	4.6	1.1	10.4	4.9	1.1
KS_1	10.1	5.4	1.3	11.0	6.1	1.3
CvM_2	10.8	5.9	0.9	9.5	4.9	0.9
KS_2	10.3	5.5	1.1	10.3	6.1	0.8
$DURC$	12.0	5.0	1.0	10.2	4.1	0.7
	[8.8]	[3.9]	[0.6]	[8.4]	[4.1]	[0.6]

TABLE III

POWER OF TESTS

NLMA		n=100			n=200			n=300	
	10%	**5%**	**1%**	**10%**	**5%**	**1%**	**10%**	**5%**	**1%**
D_n^2	28.0	17.9	3.7	37.5	24.5	9.3	49.1	35.5	14.2
CvM_1	26.4	16.1	3.6	37.1	24.5	9.9	49.6	35.6	14.1
KS_1	26.9	16.3	3.5	41.4	26.7	10.5	52.9	40.1	17.4
CvM_2	20.6	12.5	2.2	28.5	16.9	4.3	35.0	23.0	8.4
KS_2	21.5	11.3	2.4	32.9	21.9	6.4	41.9	28.8	11.8
$DURC$	15.0	6.6	0.9	14.2	7.1	0.9	14.0	7.1	1.5
	[12.8]	[6.1]	[1.0]	[11.8]	[5.7]	[0.8]	[12.9]	[6.2]	[1.1]

TABLE IV

POWER OF TESTS

BIL-I		n=100			n=200			n=300	
	10%	**5%**	**1%**	**10%**	**5%**	**1%**	**10%**	**5%**	**1%**
D_n^2	18.6	9.9	1.9	30.2	18.5	4.4	45.8	26.3	6.7
CvM_1	19.2	9.8	1.8	29.2	17.2	4.0	46.9	26.5	7.1
KS_1	21.1	10.2	2.8	28.4	17.8	4.8	43.9	29.2	8.1
CvM_2	13.0	7.3	1.9	17.8	9.2	1.9	24.9	12.2	2.2
KS_2	14.4	7.3	1.5	19.5	10.7	2.6	26.5	15.9	5.1
$DURC$	13.7	6.1	1.7	11.8	6.1	1.1	15.8	9.0	2.4
	[12.2]	[5.3]	[1.6]	[13.5]	[6.5]	[1.8]	[14.1]	[8.5]	[2.2]

TABLE V

POWER OF TESTS

BIL-II		n=100			n=200			n=300	
	10%	**5%**	**1%**	**10%**	**5%**	**1%**	**10%**	**5%**	**1%**
D_n^2	39.3	23.6	7.1	75.1	56.1	21.4	90.0	78.9	37.4
CvM_1	41.4	23.9	7.5	76.3	58.5	21.1	92.9	79.5	33.9
KS_1	43.8	28.5	10.5	72.9	56.8	26.5	88.3	77.1	42.6
CvM_2	22.6	12.1	2.1	40.8	22.0	4.3	61.1	38.1	9.3
KS_2	27.6	17.1	5.5	48.1	33.8	14.2	67.2	53.7	26.9
$DURC$	21.4	11.9	2.8	23.1	14.2	4.9	35.3	24.0	9.8
	[19.7]	[11.2]	[2.5]	[25.4]	[17.0]	[5.7]	[33.6]	[22.2]	[9.3]

Table VI

Power of Tests

n=100	ARMA(1,2)	NGMA	TAR
D_n^2	74.9	23.0	16.6
CvM_1	35.2	8.3	7.9
KS_1	32.0	7.9	7.9
CvM_2	32.8	1.7	8.6
KS_2	36.7	3.7	11.9
$DURC$	83.0	14.1	14.2

Table VII

Summary Statistics of Log Price Changes

for $100\log(S_t/S_{t-1})$

	1974-1983	1985-1991
Mean	-0.0185	-0.0227
Median	0.0000	0.0000
SD	0.5572	0.4766
Skewness	-0.5145	0.0315
Kurtosis	8.4885	4.7584
Maximum	3.4344	1.9597
Minimun	-3.8427	-2.2520

Table VIII

P-values for the BPUSD

	1974-1983	1985-1991
n	2505	1210
D_n^2	0.002	0.943
CvM_1	0.004	0.896
KS_1	0.000	0.706
CvM_2	0.016	0.870
KS_2	0.000	0.673
$DURC$	0.230	0.860

Lag j	$KS_{Y|Y}(j)\ [q_{0.95}]$	$KS_{Y^2|Y}(j)\ [q_{0.95}]$	$KS_{Y^2|Y^2}(j)\ [q_{0.95}]$	$KS_{Y^3|Y}(j)\ [q_{0.95}]$
1	1.95* [1.31]	2.02* [1.29]	2.74* [1.24]	1.48* [1.21]
2	0.91 [1.39]	1.56* [1.31]	2.91* [1.35]	0.78 [1.24]
3	0.71 [1.39]	2.23* [1.26]	2.78* [1.34]	0.55 [1.18]
4	1.14 [1.34]	1.59* [1.26]	2.30* [1.29]	0.53 [1.25]
5	1.46* [1.37]	2.76* [1.38]	3.50* [1.41]	0.94 [1.32]
6	0.56 [1.41]	1.59* [1.32]	2.77* [1.31]	0.86 [1.17]
7	0.91 [1.31]	1.35* [1.29]	2.41* [1.28]	0.52 [1.21]
8	0.88 [1.32]	1.93* [1.26]	2.44* [1.26]	1.08 [1.20]
9	1.09 [1.33]	1.71* [1.30]	2.17* [1.25]	0.99 [1.21]
10	0.84 [1.38]	1.47* [1.25]	2.45* [1.26]	0.46 [1.18]
20	1.03 [1.34]	1.62* [1.30]	2.16* [1.27]	0.82 [1.16]
30	0.66 [1.33]	1.21 [1.26]	1.98* [1.37]	0.57 [1.16]
40	0.74 [1.38]	1.26 [1.31]	1.56* [1.26]	0.86 [1.19]
50	1.18 [1.31]	1.91* [1.38]	1.28* [1.26]	1.41* [1.20]

Note:* Significantly different from zero at the 5% level (bootstrap test)

Lag j	$KS_{Y|Y}(j)\ [q_{0.95}]$	$KS_{Y^2|Y}(j)\ [q_{0.95}]$	$KS_{Y^2|Y^2}(j)\ [q_{0.95}]$	$KS_{Y^3|Y}(j)\ [q_{0.95}]$
1	0.50 [1.37]	1.40* [1.33]	0.82 [1.27]	0.76 [1.33]
2	0.78 [1.31]	0.51 [1.23]	0.67 [1.31]	0.63 [1.18]
3	0.86 [1.36]	1.11 [1.27]	1.63* [1.29]	0.82 [1.25]
4	0.68 [1.35]	0.82 [1.27]	1.31 [1.33]	0.69 [1.26]
5	0.46 [1.31]	1.19 [1.32]	1.66* [1.30]	0.77 [1.21]
6	0.70 [1.33]	1.17 [1.38]	1.89* [1.37]	0.66 [1.26]
7	0.83 [1.34]	0.82 [1.31]	1.43* [1.34]	0.66 [1.24]
8	0.55 [1.30]	0.77 [1.33]	1.20 [1.32]	0.79 [1.27]
9	1.20 [1.28]	0.71 [1.30]	0.94 [1.25]	0.89 [1.20]
10	0.88 [1.33]	1.15 [1.31]	1.55* [1.30]	0.82 [1.29]
20	0.91 [1.31]	0.89 [1.33]	1.06 [1.34]	1.27 [1.30]
30	0.96 [1.29]	0.69 [1.31]	1.16 [1.33]	0.96 [1.30]
40	0.96 [1.30]	0.88 [1.31]	1.13 [1.34]	0.71 [1.28]
50	1.07 [1.34]	1.23 [1.37]	1.18 [1.30]	0.72 [1.30]

Note:* Significantly different from zero at the 5% level (bootstrap test)
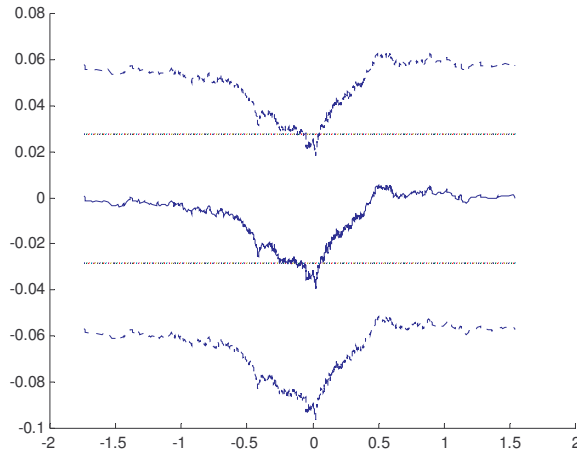
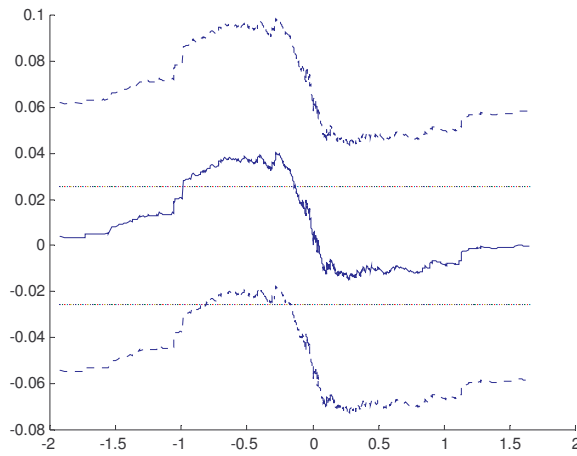FIG. 2. IPAF for Y_t BPUSD1 at lag $j = 1$ (conditioned on Y_{t-1}).

FIG. 3. IPAF for Y_t^2 BPUSD1 at lag $j = 1$ (conditioned on Y_{t-1}).