

TESTING THE MARTINGALE HYPOTHESIS*

J. CARLOS ESCANCIANO

Indiana University, Bloomington, USA

IGNACIO N. LOBATO

Instituto Tecnológico Autónomo de México, México D.F., Mexico

October 25, 2007

Abstract

This article examines testing the Martingale Difference Hypothesis (MDH) and related statistical inference issues. The earlier literature on testing the MDH was based on linear measures of dependence, such as sample autocorrelations, for instance the classical Box-Pierce Portman-teau test and the Variance Ratio test. In order to account for the existing nonlinearity in economic and financial data, two directions have been entertained. First, to modify these classical approaches by taking into account the possible nonlinear dependence. Second, to use more sophisticated statistical tools such as those based on empirical processes theory or the use of generalized spectral analysis. This paper discusses these developments and applies them to exchange rate data.

Keywords and Phrases: Martingale; Martingale Difference Hypothesis; Spectral Analysis; Generalized Spectral Analysis; Exchange Rates.

1. Introduction
2. Preliminaries
3. Tests based on linear measures of dependence
 - 3.1 Tests based on a finite-dimensional conditioning set
 - 3.2 Tests based on an infinite-dimensional conditioning set
4. Tests based on nonlinear measures of dependence
 - 4.1 Tests based on a finite-dimensional conditioning set
 - 4.2 Tests based on an infinite-dimensional conditioning set
5. Related Hypothesis

*Escanciano acknowledges financial support from the Spanish Ministerio de Educación y Ciencia, Reference numbers SEJ2004-04583/ECON and SEJ2005-07657/ECON. Lobato acknowledges financial support from the Mexican CONACYT, reference number 59028, and from Asociación Mexicana de Cultura.

1. INTRODUCTION

Martingale testing has historically received an enormous attention in econometrics. One of the main reasons is the efficient market hypothesis and the many ideas related to it. In addition, many economic theories in dynamic contexts in which expectations are assumed to be rational lead to such dependence restrictions on the underlying economic variables; see e.g. Hall (1978), Fama (1991), LeRoy (1989), Lo (1997) and Cochrane (2005). These have prompted a vast research in macro and financial economics which have stimulated a huge interest in developing suitable econometric techniques. This econometric research has grown around the theme of lack of predictability of macro or financial series, but this topic has flourished in different branches, emphasized different methodological aspects, and appeared under different subject names.

When looking at assets prices, the idea of lack of predictability has been commonly referred to as the random walk hypothesis. Unfortunately, the term random walk has been used in different contexts to mean different statistical objects. For instance, in Campbell, Lo and MacKinlay (1997) textbook, they distinguish three types of random walks according to the dependence structure of the increment series. Random walk 1 corresponds to independent increments, random walk 2 to mean-independent increments, and random walk 3 to uncorrelated increments. Of these three notions, the two relevant for financial econometrics are the second and the third. The notion of random walk 1 is clearly rejected in financial data for many reasons, the most important is the volatility, that is, the lack of constancy of the variance of current asset returns conditional on lagged asset returns. Within this terminology, this article will focus basically on the idea of random walk 2, but we will also discuss some aspects associated to random walk 3. A martingale would correspond to random walk 2, and it plainly means that the best forecast of tomorrow's asset price is today's. Then, the asset returns, which are unpredictable, are said to form a martingale difference sequence. Since asset prices are not stationary, from a technical point of view, it is simpler to handle asset returns, and instead of testing that prices follow a martingale, it is more common to test that returns follow a martingale difference sequence.

Given the huge literature that has developed, it is unavoidable that the present article reflects the authors' personal interests. It is important at the outset to stress what this article does not cover. We do not consider unit root tests, which is a topic covered in many references, see for instance, Laudrup and Jansson (2006). We do not address technical analysis, which assumes predictability, and focuses on the best ways of constructing a variety of charts, to forecast a series. We do not consider out of sample prediction tests because they assume particular models under the alternative, see Inoue and Kilian (2004) and Clark and West (2006). We do not examine chaos tests, which are

motivated by deterministic nonlinear models, see references in Barnett and Serletis (2000) and Chan and Tong (2002). What we address would be called conditional mean independence testing in the statistical literature.

The outline of the paper is the following. Section 2 contains the preliminary definitions and an overview of the data that we will employ to illustrate the different techniques. Section 3 studies martingale difference tests based on linear measures of dependence both in the time and frequency domains. Section 4 is devoted to tests based on nonlinear measures of dependence. Section 5 discusses briefly some hypotheses related to the martingale difference hypothesis and Section 6 concludes.

2. PRELIMINARIES

The Martingale Difference Hypothesis (MDH, hereinafter) plays a central role in economic models where expectations are assumed to be rational. The underlying statistical object of interest is the concept of a martingale or, alternatively, the concept of martingale difference sequence (mds). Mathematically speaking, we say that X_t forms a martingale with respect to its natural filtration, when $E[X_t | X_{t-1}, X_{t-2}, \dots] = X_{t-1}$ almost surely (*a.s.*). As commented in the introduction, from a technical point of view, it is simpler to work with the first differences, $Y_t = X_t - X_{t-1}$, and we say that Y_t follows a mds when $E[Y_t | Y_{t-1}, Y_{t-2}, \dots] = 0$ *a.s.* More generally, we state that the MDH holds when, for a real-valued stationary time series $\{Y_t\}_{t=-\infty}^{\infty}$, the following conditional moment restriction holds *a.s.*

$$E[Y_t | Y_{t-1}, Y_{t-2}, \dots] = \mu, \quad \mu \in \mathbb{R}. \quad (1)$$

The MDH slightly generalizes the notion of mds by allowing the unconditional mean of Y_t to be nonzero and unknown. The MDH states that the best predictor, in the sense of least mean square error, of the future values of a time series given the past and current information set is just the unconditional expectation. The MDH is called conditional mean independence in the statistical literature, and it means that past and current information are of no use to forecasting future values of a mds. In Section 5 we discuss extensions of this basic version of the MDH.

As commented in the introduction, there is a vast empirical and theoretical literature on the MDH. In order to systematize part of this literature we start by introducing the following definitions. Let $I_t = \{Y_t, Y_{t-1}, \dots\}$ be the information set at time t and let \mathcal{F}_t be the σ -field generated by I_t . Then, the following equivalence is fundamental because it formalizes the characteristic property of a mds, that is, the fact that Y_t is linearly unpredictable given any linear or nonlinear transformation of the past $w(I_{t-1})$. That is,

$$E[Y_t | I_{t-1}] = \mu \text{ a.s.}, \mu \in \mathbb{R} \iff E[(Y_t - \mu)w(I_{t-1})] = 0, \quad (2)$$

for all \mathcal{F}_{t-1} -measurable weighting function $w(\cdot)$ (such that the moment exists). Equation (2) is fundamental to understand the motivation and main features behind many tests for the MDH. There are two challenging features present in the definition of a mds: first, the information set at time t , I_t , will typically include the infinite past of the series, and second, the number of functions $w(\cdot)$ is also infinite. We will classify the extant theoretical literature on testing the MDH, according to what types of functions $w(\cdot)$ are employed. Section 3 analyses the case where linear $w(\cdot)$ are employed, that is, the use of tests based on linear measures of dependence. Section 4 analyses the case where an infinite number of nonlinear w 's are employed, that is, the use of tests based on nonlinear measures of dependence. In both sections, we divide the extensive literature according to whether the tests account for a finite number of lags or not, that is, whether they assume that $w(I_{t-1}) = w(Y_{t-1}, \dots, Y_{t-P})$ for some $P \geq 1$ or not.

We shall illustrate some of the available methods for testing the MDH by applying them to exchange rates returns. The martingale properties of exchange rates returns have been studied previously by many authors leading to mixed conclusions. For instance, Bekaert and Hodrick (1992), Escanciano and Velasco (2006a, 2006b), Fong and Ouliaris (1995), Hong and Lee (2003), Kuan and Lee (2004), LeBaron (1999), Levich and Thomas (1993), Liu and He (1991), McCurdy and Morgan (1988) and Sweeney (1986) find evidence against the MDH for nominal or real exchange rates at different frequencies, whereas Diebold and Nason (1990), Fong, Koh and Ouliaris (1997), Hsieh (1988, 1989, 1993), McCurdy and Morgan (1987) and Meese and Rogoff (1983a,b) find little evidence against the MDH. Here, we consider data that consists of four daily and weekly exchange rate returns on the Euro (Euro), Canadian Dollar (Can), the sterling Pound (£) and the Japanese Yen (¥) against the US dollar. The daily data is taken from January 2, 2004 to August 17, 2007, with a total of 908 observations. As for the weekly data, we consider the returns on Wednesdays from January 2, 2000 to August 17, 2007, with a total of 382 observations. The daily noon buying rates in New York City certified by the Federal Reserve Bank of New York for customs and cable transfers purposes are obtained from <http://www.federalreserve.gov/Releases/h10/hist>. In Figure 1 we have plotted the evolution of these four daily series for the whole period from January 2, 2000 to August 17, and, again similarly to previous analysis, the main two features of these plots are their unpredictability and their volatility. Table I provides summary statistics for the most relevant aspects of the marginal distribution of the data. Similarly to most financial series the main feature from Table 1 is the kurtosis that, in the line of previous studies, is larger for daily than for weekly data. Note that skewness is moderate and slightly negative for daily data. As it has been observed repeatedly before, the marginal distribution of weekly data is closer to the normal distribution than that of daily data.

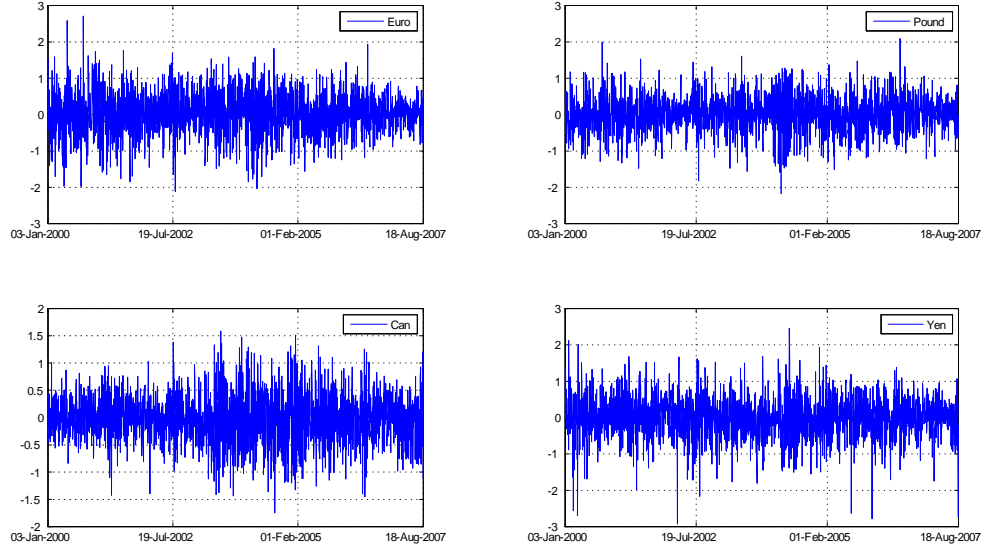


Figure 1. Daily returns of the Euro (Euro), Canadian Dollar (Can), the sterling Pound (Pound) and the Japanese Yen (Yen) against the US dollar. Data from January 2, 2000 to August 17, 2007.

TABLE I
SUMMARY STATISTICS OF EXCHANGE RATES RETURNS

	Daily				Weekly			
	Euro	£	Can	¥	Euro	£	Can	¥
n	908	908	908	908	382	382	382	382
Mean	0.0076	0.0113	-0.0213	0.0068	0.0738	0.0552	-0.0832	0.0352
Median	0.0000	0.0221	-0.0080	0.0279	0.0781	0.0763	-0.0864	0.0141
SD	0.5423	0.5332	0.5036	0.5670	1.3539	1.1407	0.9410	1.2525
Skewness	-0.1263	-0.0976	-0.0196	-0.3763	0.0540	0.0545	0.0846	-0.2945
Kurtosis	3.7602	3.4927	3.1345	5.0746	3.0555	2.9649	2.8875	3.0895
Maximum	1.9358	2.0930	1.5129	2.4519	4.4680	3.4830	2.8128	3.1835
Minimum	-2.0355	-2.1707	-1.7491	-2.7859	-3.1636	-3.2307	-2.7067	-4.3058

3. TESTS BASED ON LINEAR MEASURES OF DEPENDENCE

Recall the mds definition in equation (2) that should hold for any function $w(\cdot)$. The simplest approach is to consider linear functions $w(\cdot)$, such as $w(I_{t-1}) = Y_{t-j}$, for some $j \geq 1$. Hence, a necessary (but not sufficient, in general) condition for the MDH to hold is that the time series is uncorrelated, i.e.

$$\gamma_j = \text{Cov}(Y_t, Y_{t-j}) = E[(Y_t - \mu)Y_{t-j}] = 0 \quad \text{for all } j \geq 1, \quad (3)$$

where γ_j denotes the autocovariance of order j . In principle, one should test that all autocovariances or autocorrelations are zero. However, the most employed tests just consider that a finite number of autocorrelations are zero. As commented in the introduction, we will separately address these two cases.

Notice that the early literature, which includes some distinguished references such as Yule (1926), Bartlett (1955), Grenander and Rosenblatt (1957) or Durbin and Watson (1950), essentially assumed Gaussianity and, hence, identified three concepts: lack of serial correlation, mds and independence. In the time series literature the term white noise is commonly used to denote an uncorrelated series that can present some form of dependence. Obviously, a white noise series is neither necessarily independent nor mds since dependence can be reflected in other aspects of the joint distribution such as higher order moments. The distinction between these three concepts has been stressed recently in econometrics. In fact, during the last years a variety of models designed to reflect nonlinear dependence has been studied in the econometrics literature. For instance, in empirical finance, ARCH and bilinear models have been widely studied, see Bera and Higgins (1993, 1997) and Weiss (1986) for a comparison. These models are suitable to reflect the nonlinear dependence structure found in many financial series.

Tests for white noise have been proposed both in the time domain and in the frequency domain. The time domain has mainly, but not exclusively, focused on a finite number of lags, while the frequency domain has been more suitable to address the infinite dimensional case.

3.1 Tests based on a finite-dimensional conditioning set

In the time domain the most popular test (apart from the Durbin-Watson which is designed to test for lack of first order serial correlation using regression residuals) has been the Box-Pierce (Box and Pierce, 1970) Portmanteau Q_p test. The Q_p test is designed for testing that the first p autocorrelations of a series (possibly residuals) are zero. The number p can be considered to be fixed or to grow with the sample size n . In this section we will assume that p is fixed.

Suppose that we observe raw data $\{Y_t\}_{t=1}^n$. Then, γ_j can be consistently estimated by the sample

autocovariance

$$\hat{\gamma}_j = (n-j)^{-1} \sum_{t=1+j}^n (Y_t - \bar{Y})(Y_{t-j} - \bar{Y}),$$

where \bar{Y} is the sample mean, and we also introduce $\hat{\rho}_j = \hat{\gamma}_j / \hat{\gamma}_0$ to denote the j -th order autocorrelation. The Q_p statistic is just

$$Q_p = n \sum_{j=1}^p \hat{\rho}_j^2,$$

but it is commonly implemented via the Ljung and Box (1978) modification

$$LB_p = n(n+2) \sum_{j=1}^p (n-j)^{-1} \hat{\rho}_j^2.$$

Note that Q_p (or LB_p) only takes into account the linear dependence up to the lag p . When p is considered fixed, the Q_p test statistic applied to independent data follows a χ_p^2 asymptotic null distribution since the asymptotic covariance matrix of the first p autocorrelations of an independent series is the identity matrix. Hence, it is useful to write $Q_p = (\sqrt{n}\hat{\rho})' I^{-1} (\sqrt{n}\hat{\rho})$ where $\hat{\rho} = (\hat{\rho}_1, \dots, \hat{\rho}_p)'$.

Note, however, that when the series present some kind of nonlinear dependence, such as conditional heteroskedasticity, this asymptotic null covariance matrix is no longer the identity. In fact, denoting $\rho = (\rho_1, \dots, \rho_p)'$, for a general time series the asymptotic distribution of $\sqrt{n}(\hat{\rho} - \rho)$ is $N(0, T)$ where the (i,j) -th element of T is given by (see e.g. Romano and Thombs, 1996)

$$\gamma_0^{-2}(c_{ij} - \rho_i c_{0j} - \rho_j c_{0i} + \rho_i \rho_j c_{00})$$

where, for $i, j = 0, 1, \dots, p$,

$$c_{ij} = \sum_{d=-\infty}^{\infty} \{E[(Y_t - \mu)(Y_{t-i} - \mu)(Y_{t+d} - \mu)(Y_{t+d-j} - \mu)] - E[(Y_t - \mu)(Y_{t-i} - \mu)] E[(Y_{t+d} - \mu)(Y_{t+d-j} - \mu)]\}.$$

Under alternative assumptions the matrix T can be simplified and this will lead to several modified versions of the Box-Pierce statistic. When this matrix is still diagonal, as it happens under mds and additional moment restrictions, which, for instance, are satisfied by Gaussian GARCH models and many stochastic volatility models, the natural approach is to robustify the Q_p by standardizing it by a consistent estimation of its asymptotic variance, i.e.,

$$Q_p^* = n \sum_{j=1}^p \frac{\hat{\rho}_j^2}{\tau_j}$$

where

$$\tau_j = \frac{1}{\hat{\gamma}_0^2} \sum_{t=1+j}^n (Y_t - \bar{Y})^2 (Y_{t-j} - \bar{Y})^2.$$

We have followed Lobato, Nankervis and Savin (2001) notation and denoted the robustified Q_p by Q_p^* . This statistic has appeared in different versions, see for instance, Diebold (1986), Lo and MacKinlay

(1989), Robinson (1991), Cumby and Huizinga (1992), Bollerslev and Wooldridge (1992) and Bera and Higgins (1993). The Q_p^* statistic (or its Ljung-Box analog) should be routinely computed for financial data instead of the standard Q_p (or the LB_p). However, this is not typically the case, see Lobato, Nankervis and Savin (2001) for details.

For the general case, the asymptotic covariance matrix of the first p autocorrelations is not a diagonal matrix. Hence, for this general case both the Q_p and the Q_p^* tests are invalid. However, under mds the matrix T can be greatly simplified so that its $ij - th$ element takes the form $E[(Y_t - \mu)^2(Y_{t-i} - \mu)(Y_{t-j} - \mu)]$ that can be easily estimated using its sample analog. This is the approach followed by Guo and Phillips (2001). For the general case, that includes mds and non mds processes, the asymptotic covariance matrix of the first p autocorrelations is a complicated nondiagonal matrix. Hence, for this general case, the literature has proposed the following two modifications of the Q_p test. The first one is to modify the Q_p statistic by introducing a consistent estimator of the asymptotic null covariance matrix of the sample autocorrelations, \hat{T} , so that the modified Q_p statistic retains the χ_p^2 asymptotic null distribution. Lobato, Nankervis and Savin (2002) name this statistic $\tilde{Q}_p = (\sqrt{n}\hat{\rho})' \hat{T}^{-1} (\sqrt{n}\hat{\rho})$. The main drawback of this approach is that in order to construct \hat{T} a bandwidth number has to be introduced, see Lobato, Nankervis and Savin (2002) for details. This approach works for general dependence structures that allow for the asymptotic covariance matrix of the first p autocorrelations to take any form. The second modification has been studied by Horowitz, Lobato, Nankervis and Savin (2006) who employ a bootstrap procedure to estimate consistently the asymptotic null distribution of the Q_p test for the general case. They compare two bootstrap approaches, a single and a double blocks-of-blocks bootstrap, and the final recommendation is to employ a double blocks-of-blocks bootstrap after prewhitening the time series. This solution presents a similar problem, though, namely the researcher has to choose arbitrarily a block length number. The previous papers considered raw data, but Francq, Roy and Zakoïan (2005) have addressed the use of the Q_p statistic with residuals. They propose to estimate the asymptotic null distribution of the Q_p test statistic for the general weak dependent case. However, their approach still requires the selection of p , and of several additional arbitrary numbers necessary to estimate consistently the needed asymptotic critical values.

These previous references represent an effort to address the problem of testing for mds using the standard linear measures (autocorrelations) but allowing for nonlinear dependence. Lobato (2001) represents an alternative approach with a similar spirit. The target is to avoid the problem of introducing a user-chosen number and the idea is to construct an asymptotically distribution free statistic. Although this approach delivers tests that handle nonlinear dependence and control properly the type I error in finite samples, its main theoretical drawback is its inefficiency in terms of local power.

A related statistic, which has been commonly employed in the empirical finance literature, see Cochrane (1988) and Lo and MacKinlay (1989), is the variance ratio that takes the form

$$VR_p = 1 + 2 \sum_{j=1}^{p-1} \left(1 - \frac{j}{p}\right) \hat{\rho}_j.$$

Under independence, $\sqrt{np}(VR_p - 1)$ is asymptotically distributed as $N(0, 2(p-1))$. Although this test can also be robustified and it can be powerful in some occasions, it presents the serious theoretical limitation of being inconsistent. For instance, González and Lobato (2003) considered an MA(2) $y_t = e_t - 0.4597e_{t-1} + 0.10124e_{t-2}$. For this process $VR_3 = 0$ in spite that the first two autocorrelations are non zero. The problem with variance ratio statistics resides in the possible existence of compensations between autocorrelations with different signs, and this may affect power severely. Related to VR tests, Nankervis and Savin (2007) have proposed a robustified version of the Andrews and Ploberger's (1996) test that appear to have very good finite sample power with the common empirical finance models. Also related, Delgado and Velasco (2007) have recently considered a large class of directional tests based on linear combinations of autocorrelations. Their tests are shown to be optimal in certain known local alternative directions and are asymptotically equivalent to Lagrange Multiplier tests. Finally, we mention Kuan and Lee (2004) who propose a correlation-based test for the MDH that instead of using lagged values of Y_t as the function $w(\cdot)$, they employ some other arbitrary $w(\cdot)$. This test shares with all the tests analyzed in this section the problem of inconsistency derived from not using a whole family of functions $w(\cdot)$.

3.2 Tests based on a infinite-dimensional conditioning set

The approach presented in the previous subsection laid naturally in the time domain since a finite number of autocorrelations were tested. However, when the infinite past is considered, the natural framework for performing inference is the frequency domain. The advantage of the frequency domain is the existence of one object, namely, the spectral density, that contains the information contained in all the autocovariances. Hence, in the frequency domain, the role previously taken by autocorrelations is now carried by the spectral density function. Define the spectral density $f(\lambda)$ implicitly by

$$\gamma_k = \int_{\Pi} f(\lambda) \exp(ik\lambda) d\lambda \quad k = 0, 1, 2, \dots$$

where $\Pi = [-\pi, \pi]$. Define also the periodogram as $I(\lambda) = |w(\lambda)|^2$ where $w(\lambda) = n^{-1/2} \sum_{t=1}^n x_t \exp(it\lambda)$. Although the periodogram is an inconsistent estimator of the spectral density, it can be employed as a building block to construct a consistent estimator. The integral of the spectral density is called the spectral distribution, which, under the MDH, is linear in λ .

For this infinite lag case, the MDH implies as null hypothesis of interest that $\gamma_k = 0$ for all $k \neq 0$,

and equivalently, in terms of the spectral density, the null hypothesis states that $f(\lambda) = \gamma_0/2\pi$ for all $\lambda \in \Pi$.

The advantage of the frequency domain is that the problem of selecting p , which was present in the previous subsection, does not appear because the null hypothesis is stated in terms of *all* autocorrelations, as summarized by the spectral density or distribution. The classical approach in the frequency domain involves the standardized cumulative periodogram, that is,

$$Z_n(\lambda) = \sqrt{T} \left(\frac{\sum_{j=1}^{\lfloor \lambda T/\pi \rfloor} I(\lambda_j)}{\sum_{j=1}^T I(\lambda_j)} - \frac{\lambda}{\pi} \right),$$

where $\lambda_j = 2\pi j/n$, $j = 1, 2, \dots, n/2$, are called the Fourier frequencies. Based on $Z_n(\lambda)$, the two classical tests statistics are the Kolmogorov-Smirnov

$$\max_{j=1, \dots, T} |Z_n(\lambda_j)|,$$

and the Cramer von Mises

$$\frac{1}{T} \sum_{j=1}^T Z_n(\lambda_j)^2.$$

These tests statistics have been commonly employed (see Bartlett (1955) and Grenander and Rosenblatt(1957)) because when the series y_t is not only white noise but also independent (or mds with additional moment restrictions), it can be shown that the process $Z_n(\lambda)$ converges weakly in $D[0, \pi]$ (the space of cadlag functions in $D[0, \pi]$) to the Brownian bridge process, see Dahlhaus (1985). Hence, asymptotic critical values are readily available for the independent case. In fact, Durlauf (1991) has shown that the independence assumption can be relaxed to conditional homoskedastic mds. For the mds case with conditional heteroskedasticity (and some moment conditions), Deo (2000) slightly modified this statistic so that the standardized cumulative periodogram retained the convergence to the Brownian bridge. Deo's test can be interpreted as a continuous version of the robustified Box-Pierce statistic, Q_p^* . Notice that in Deo's setup there is no need of introducing any user-chosen number since under the stated assumptions (see condition A in Deo (2000, p. 293)) the autocorrelations are asymptotically independent. As Deo comments, his assumption (vii) is the main responsible for the diagonality of the asymptotic null covariance matrix of the sample autocorrelations. However, for many common models, such as GARCH models with asymmetric innovations, EGARCH models and bilinear models, Deo's condition (vii) does not hold and the autocorrelations are not asymptotically independent under the null hypothesis. Hence, for the general case, Deo's test is not asymptotically valid. Deo's Cramer-von Mises test statistic can also be written in the time domain as

$$DEO_n := \sum_{j=1}^{n-1} n \frac{\hat{\rho}_j^2}{\tau_j} \left(\frac{1}{j\pi} \right)^2.$$

More general weighting schemes for the sample autocovariances $\hat{\rho}_j$ than the ones considered here are possible. Under the null hypothesis of the mds and some additional assumptions (see Deo (2000)),

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

where $B(t)$ is the standard Brownian bridge on $[0,1]$. The 10%, 5% and 1% asymptotic critical values can be obtained from Shorack and Wellner (1986, p.147) and are 0.347, 0.461 and 0.743, respectively. For extensions of this basic approach see also Paparoditis (2000) and Delgado, Hidalgo and Velasco (2005), among others.

Under general weak dependent assumptions (see Dahlhaus (1985)) the asymptotic null distribution of the process $Z_n(\lambda)$ is no longer the Brownian bridge but, in fact, it converges weakly in $D[0, \pi]$ to a zero mean Gaussian process with covariance given by

$$\begin{aligned} & \frac{\pi G(\pi)}{F(\pi)^2} \left\{ \frac{G(\lambda \wedge \mu)}{G(\pi)} + \frac{F(\lambda)F(\mu)}{F(\pi)^2} - \frac{F(\lambda)G(\mu)}{F(\pi)G(\pi)} - \frac{F(\mu)G(\lambda)}{F(\pi)G(\pi)} \right. \\ & \left. + \frac{F_4(\lambda, \mu)}{G(\pi)} + \frac{F_4(\pi, \pi)}{G(\pi)} \frac{F(\lambda)F(\mu)}{F(\pi)^2} - \frac{F_4(\mu, \pi)}{G(\pi)} \frac{F(\lambda)}{F(\pi)} - \frac{F_4(\lambda, \pi)}{G(\pi)} \frac{F(\mu)}{F(\pi)} \right\}, \end{aligned}$$

where $F(\lambda)$ denotes the spectral distribution function, $F(\lambda) = \int_0^\lambda f(\omega) d\omega$, $G(\lambda) = \int_0^\lambda f(\omega)^2 d\omega$, and $F_4(\lambda, \mu) = \int_0^\lambda \int_0^\mu f_4(\omega, -\omega, -\theta) d\omega d\theta$, where $f_4(\boldsymbol{\lambda})$, with $\boldsymbol{\lambda} \in \Pi^{q-1}$, denotes the fourth order cumulant spectral density, see expression (2.6.2) in Brillinger (1981, p.25). The important message from the previous complicated covariance is that the asymptotic null distribution depends on the nature of the data generating process of y_t . Therefore, no asymptotic critical values are available. Chen and Romano (1999, p.628) propose to estimate the asymptotic distribution by means of either the block bootstrap or the subsampling technique. Unfortunately, these bootstrap procedures require the selection of some arbitrary number and in a general framework no theory is available about their optimal selection. Alternative bootstrap procedures which do not require the selection of a user-chosen number such as resampling the periodogram as in Franke and Hardle (1992) or in Dahlhaus and Janas (1996) will not estimate consistently the asymptotic null distribution because of the fourth order cumulant terms.

Lobato and Velasco (2004) considered the use of the statistic

$$M_n = \frac{T^{-1} \sum_{j=1}^T I(\lambda_j)^2}{\left(T^{-1} \sum_{j=1}^T I(\lambda_j) \right)^2} - 1$$

under general weak dependence conditions. This statistic was previously considered by Milhøj (1981) who employed M_n as a general goodness of fit test statistic for time series. Milhøj informally justified the use of this statistic for testing the adequacy of linear time series models, but since he identified white noise with i.i.d. (see his p. 177), his analysis does not automatically apply in general contexts.

Beran (1992) and Deo and Chen (2000) have also employed the M_n test statistic as a goodness-of-fit tests for Gaussian processes. Statistical inference is especially simple with M_n since its asymptotic null distribution is normal even after parametric estimation. We note that the continuous version of M_n can be expressed in the time domain as an statistic proportional to $\sum_{j=0}^{n-1} \hat{\rho}_j^2$, which shows the difficulty of deriving the asymptotic properties in the time domain since the $\hat{\rho}_j$ may not be asymptotically independent.

In the time domain, Hong (1996) has considered p as growing with n and hence, has been able to derive a consistent test in the time domain for the case of regression residuals. In this framework p can be interpreted as a bandwidth number that needs to grow with n so that his test can handle the fact that the null hypothesis implies an infinite number of autocovariances. Hong restricted to the independent case while Hong and Lee (2003) have extended Hong's procedure to allow for conditional heteroskedasticity. However, notice that their framework still restricts the sample autocorrelations to be asymptotically independent.

An alternative solution recently explored by Escanciano and Lobato (2007) consists on modifying the Box-Pierce statistic using an adaptive Neyman test that would take the form

$$N_n = Q_p^*$$

where

$$\tilde{p} = \min\{m : 1 \leq m \leq p_n; L_m \geq L_h, h = 1, 2, \dots, p_n\}, \quad (4)$$

where

$$L_p = Q_p^* - \pi(p, n, q),$$

and p_n is an upper bound that grows slowly to infinite with n , and

$$\pi(p, n, c) = \begin{cases} p \log n, & \text{if } \max_{1 \leq j \leq p_n} \left| \frac{\hat{\rho}_j^2}{\tau_j} \right| \leq \sqrt{q \log n} \\ 2p, & \text{if } \max_{1 \leq j \leq p_n} \left| \frac{\hat{\rho}_j^2}{\tau_j} \right| > \sqrt{q \log n}, \end{cases}$$

where q is some fixed positive number. Our choice of q is 2.4 and it is motivated from an extensive simulation study in Inglot and Ledwina (2006) and from simulations in Escanciano and Lobato (2007). Small values of q result in the Akaike's criterion choice, while large q 's lead to the choice of the Schwarz's criterion. Moderate values, such as 2.4, provide a "switching effect" in which one combines the advantages of the two selection rules, that is, when the alternative is of high frequency (i.e. when only the significant autocorrelations are at large lags j) Akaike is used whereas if the alternative is of low-frequency (i.e. if the first autocorrelations are different from zero) Schwarz is chosen. The previous adaptive test is an improvement with respect to the traditional Box-Pierce and Hong's approaches because the N_n test is more powerful and less sensitive to the selection of the bandwidth number p_n than theses approaches, and more importantly, because it avoids the

estimation of the complicated variance-covariance matrix T since its asymptotic distribution is χ_1^2 for general mds processes.

Summarizing, testing the MDH using linear measures of dependence presents two challenging features. The first aspect is that the null hypothesis implies that an infinite number of autocorrelations are zero. This feature has been addressed successfully in the frequency domain under severe restrictions on the dependence structure of the process. The second feature is that the null hypothesis allows the time series to present some form of dependence beyond the second moments. This dependence entails that the asymptotic null covariance matrix of the sample autocorrelations is not diagonal, so that it has n^2 non-zero terms (contrary to Durlauf (1991) and Deo (2000) who consider a diagonal matrix, and hence, it has only n non-zero elements). This aspect has been handled by introducing some arbitrary user-chosen numbers whose selection complicates statistical inference.

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

To circumvent this problem we could take into account higher order moments, as in Hinich and Patterson (1992). They proposed to use the bispectrum, i.e., the Fourier transform of the third order cumulants of the process, but again, this test is not consistent against non-martingale difference sequences with zero third order cumulants.

In Table II we report the robust version (to conditional heteroskedasticity) of the first five autocorrelations, the Ljung and Box's (1978) test, that is a corrected Q_p^* statistic, which we call LB_p^* , Deo's (2000) modification of Durlauf's test statistic and the Escanciano and Lobato's (2007) test based on N_n to check whether or not our exchange rates changes are uncorrelated. For further evidence of linear independence see Figures 2 to 9 in Section 4.2. This Table is in agreement with previous findings that have shown that exchange rates have no linear dependence, see for instance, Table 2 in Hsieh (1989), Bera and Higgins (1997), Hong and Lee (2003) and references therein.

TABLE II

LINEAR PREDICTABILITY OF EXCHANGE RATES RETURNS

	Daily				Weekly			
	Euro	£	Can	¥	Euro	£	Can	¥
$\hat{\rho}_1$	-0.047	0.001	-0.016	-0.020	0.018	0.046	-0.023	0.054
$\hat{\rho}_2$	0.003	0.007	-0.028	-0.015	-0.002	-0.008	0.031	-0.024
$\hat{\rho}_3$	-0.046	-0.055	-0.001	-0.016	0.049	-0.031	0.011	0.010
$\hat{\rho}_4$	-0.002	0.028	-0.060	0.013	0.024	-0.043	0.015	-0.041
$\hat{\rho}_5$	-0.002	0.003	-0.063	0.039	0.036	-0.024	0.052	-0.095
LB_5^*	4.071	3.586	8.045	2.452	1.795	2.191	1.781	4.900
LB_{15}^*	15.516	13.256	15.181	6.670	9.139	7.451	10.266	18.861
LB_{25}^*	28.552	26.568	19.756	13.155	18.746	32.584	21.786	24.519
LB_{50}^*	61.922	64.803**	49.887	37.428	42.559	59.107	41.140	58.756
$DURC_n$	0.253	0.055	0.095	0.063	0.043	0.114	0.050	0.167
N_n	1.889	0.021	0.253	0.380	0.151	0.827	0.208	1.105

Note: * and ** significantly different from zero at the 5% and 10% level, respectively.

4. TESTS BASED ON NONLINEAR MEASURES OF DEPENDENCE

Arguably, testing for the MDH is a challenging problem, since in order to verify it, we must check that a very large class of transformations of the past does not help to predict the current value of the series, see (2). An important step through the development of consistent tests was made when the econometricians realized that is not necessary to take a very *large* class of functions in (2) but just a convenient parametric class of functions, satisfying certain properties. Domínguez and Lobato (2003) called this methodology the global approach and Escanciano (2007a) called it the integrated approach. Most of this section will be devoted to a careful study of this approach.

Note, however, that there exists an alternative methodology that is based on the direct estimation of the conditional expectation $E[Y_t | \tilde{Y}_{t,P}]$ where $\tilde{Y}_{t,P} = (Y_{t-1}, \dots, Y_{t-P})'$ for some P finite. This approach can be called the smoothing approach (since smoothing numbers are required for this non-parametric estimation) or a local approach, see Domínguez and Lobato (2003). Tests within the local approach have been proposed by Wooldridge (1992), Yatchew (1992), Horowitz and Härdle (1994), Zheng (1996), Fan and Huang (2001), Horowitz and Spokoiny (2001) and Guerre and Lavergne (2005), to mention just a few; see Hart (1997) for a comprehensive review of the local approach when $P = 1$. Among these tests based on local methods, the test recently proposed by Guay and Guerre (2006) seems to be especially convenient for testing the MDH for two reasons. First, it has been justified for time series under conditional heteroskedasticity of unknown form. Second, it is an

adaptive data-driven test. Their test combines a chi-square statistic based on nonparametric Fourier series estimators for $E[Y_t | \tilde{Y}_{t,P}]$ coupled with a data-driven choice for the number of components in the estimator. To construct their test a nonparametric estimator of the unknown conditional variance is needed. Notice that a practically relevant problem of the local approach arises when P is large or even moderate. The problem is motivated by the sparseness of the data in high-dimensional spaces, which leads to most test statistics to suffer a considerable bias, even for large sample sizes. In the next subsection, we will consider an approach that helps to alleviate this problem.

This section focuses on integrated tests. We divide the extensive literature within this integrated approach according to whether the tests consider functions of a finite number of lags or not, that is, whether $w(I_{t-1}) = w(\tilde{Y}_{t,P})$ for some $P \geq 1$ or not. We stress at the outset that the main advantage of the tests considered in this section is that they are consistent for testing the MDH (at least when the information set has a finite number of variables), contrary to the tests considered in Section 3. The main disadvantage is that their asymptotic null distributions are, in general, not standard, what means that no critical values are ready available. In this situation, the typical solution is to employ the bootstrap to estimate this distribution.

4.1 Tests based on a finite-dimensional conditioning set

The problem of testing over all possible weighting functions can be reduced to testing the orthogonality condition over a parametric family of functions, see e.g. Stinchcombe and White (1998). Although still the parametric class has to include an infinite number of elements, the complexity of the class to be considered is substantially simplified and makes it possible to test for the MDH.

The methods that we review in this subsection use $w(I_{t-1}) = w_0(\tilde{Y}_{t,P}, x)$ in (2), where, as stated above, $\tilde{Y}_{t,P} = (Y_{t-1}, \dots, Y_{t-P})'$ and w_0 is a known function indexed by a parameter x . That is, these methods check for any form of predictability from the lagged P values of the series. The test statistics are based on a “distance” from the sample analogue of $E[(Y_t - \mu)w_0(\tilde{Y}_{t,P}, x)]$ to zero.

The exponential function $w_0(\tilde{Y}_{t,P}, x) = \exp(ix' \tilde{Y}_{t,P})$, $x \in \mathbb{R}$, was first considered in Bierens (1982, 1984, 1990); see also Bierens and Ploberger (1997). One version of the Cramér-von Mises (CvM) test of Bierens (1984) leads to the test statistic

$$CvM_{n,\exp,P} = n^{-1} \hat{\sigma}^{-2} \sum_{t=1}^n \sum_{s=1}^n (Y_t - \bar{Y})(Y_s - \bar{Y}) \exp\left(-\frac{1}{2} \left| \tilde{Y}_{t,P} - \tilde{Y}_{s,P} \right|^2\right),$$

where

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n (Y_t - \bar{Y})^2.$$

Indicator functions $w_0(\tilde{Y}_{t,P}, x) = 1(\tilde{Y}_{t,P} \leq x)$, $x \in \mathbb{R}$, were used in Stute (1997) and Koul and Stute (1999) for model checks of regressions and autoregressions, respectively, and in Domínguez

and Lobato (2003) for the MDH problem.

Domínguez and Lobato (2003), extending to the multivariate case the results of Koul and Stute (1999), considered the CvM and Kolmogorov-Smirnov (KS) statistics, respectively,

$$\begin{aligned} CvM_{n,P} &: = \frac{1}{\hat{\sigma}^2 n^2} \sum_{j=1}^n \left[\sum_{t=1}^n (Y_t - \bar{Y}) 1(\tilde{Y}_{t,P} \leq \tilde{Y}_{j,P}) \right]^2, \\ KS_{n,P} &: = \max_{1 \leq i \leq n} \left| \frac{1}{\hat{\sigma} \sqrt{n}} \sum_{t=1}^n (Y_t - \bar{Y}) 1(\tilde{Y}_{t,P} \leq \tilde{Y}_{i,P}) \right|. \end{aligned}$$

As mentioned above, an important problem of the local approach (also shared by other methods) arises in the case where P is large or even moderate. The sparseness of the data in high-dimensional spaces implies severe biases to most test statistics. This is an important practical limitation for most tests considered in the literature because these biases still persist in fairly large samples. Motivated by this problem, Escanciano (2007a) proposed the use of $w_0(\tilde{Y}_{t,P}, x) = 1(\beta' \tilde{Y}_{t,P} \leq u)$, where $x = (\beta, u) \in \mathbb{S}^d \times \mathbb{R}$, with $\mathbb{S}^d = \{\beta \in \mathbb{R}^d : |\beta| = 1\}$, and defined CvM tests based on this choice. We denote by $PCVM_{n,P}$ the resulting CvM test in Escanciano (2007a). Also recently, Lavergne and Patilea (2007) has proposed dimension-reduction bootstrap consistent test for regression models based on nonparametric kernel estimators of one-dimensional projections. Their proposal falls in the category of local-based methods, though.

As mentioned earlier, the asymptotic null distribution of integrated tests based on $w_0(\tilde{Y}_{t,P}, x)$ depends on the data generating process (DGP) in a complicated way. Therefore, critical values for the tests statistics can not be tabulated for general cases. One possibility, only explored in the literature for the case $P = 1$ by Koul and Stute (1997), consists of applying the so-called Khmaladze's transformation (Khmaladze, 1981) to get asymptotically distribution free tests. Extensions to $P > 1$ are not available yet. Alternatively, we can approximate the asymptotic null distributions by bootstrap methods. The most relevant bootstrap procedure for testing the MDH has been the wild bootstrap (WB) introduced in Wu (1986) and Liu (1988). For instance, this approach has been employed in Domínguez and Lobato (2003) and Escanciano and Velasco (2006a, 2006b) to approximate the asymptotic distribution of integrated MDH tests. The asymptotic distribution is approximated by replacing $(Y_t - \bar{Y})$ by $(Y_t - \bar{Y})(V_t - \bar{V})$, where $\{V_t\}_{t=1}^n$ is a sequence of independent random variables (r.v's) with zero mean, unit variance, bounded support and also independent of the sequence $\{Y_t\}_{t=1}^n$. Here, \bar{V} is the sample mean of $\{V_t\}_{t=1}^n$. The bootstrap samples are obtained resampling from the distribution of V_t . A popular choice for $\{V_t\}$ is a sequence of i.i.d. Bernoulli variates with $P(V_t = 0.5(1 - \sqrt{5})) = (1 + \sqrt{5})/2\sqrt{5}$, and $P(V_t = 0.5(1 + \sqrt{5})) = 1 - (1 + \sqrt{5})/2\sqrt{5}$.

We have applied several tests within the integrated methodology to our exchange rates data. In Table III we report the wild bootstrap empirical values. In our application we have considered the values $P = 1$ and $P = 3$ for the number of lags used in $CvM_{n,\text{exp},P}$, $CvM_{n,P}$, $KS_{n,P}$ and $PCVM_{n,P}$.

TABLE III
TESTING THE MDH OF EXCHANGE RATES RETURNS

EMPIRICAL P-VALUES								
	Daily				Weekly			
	Euro	£	Can	¥	Euro	£	Can	¥
$CvM_{n,\text{exp},1}$	0.028	0.322	0.744	0.842	0.453	0.086	0.876	0.488
$CvM_{n,\text{exp},3}$	0.164	0.320	0.898	0.666	0.743	0.250	0.076	0.258
$CvM_{n,1}$	0.020	0.354	0.628	0.822	0.610	0.146	0.863	0.388
$CvM_{n,3}$	0.192	0.424	0.798	0.588	0.916	0.893	0.720	0.500
$KS_{n,1}$	0.016	0.220	0.502	0.740	0.726	0.176	0.836	0.542
$KS_{n,3}$	0.036	0.280	0.734	0.526	0.986	0.810	0.224	0.654
$PCVM_{n,1}$	0.020	0.354	0.626	0.822	0.610	0.146	0.863	0.388
$PCVM_{n,3}$	0.248	0.438	0.790	0.664	0.746	0.443	0.566	0.414

Our results favor the MDH with all exchange rates at both frequencies, weekly and daily, with the exception of the daily Euro for $P = 1$. Surprisingly enough, we obtain contradictory results for this exchange rate when $P = 3$. These contradictory results have been previously documented in e.g. Escanciano and Velasco (2006a) and rather than to a true lack of evidence against the MDH, they may be due to a lack of power of the tests.

Although the consideration of an omnibus test, as those discussed in this section, is naturally the first idea when there is no a priori information about directions in the alternative hypothesis, it is worth noting that there is an important limitation of omnibus tests: despite their capability to detect deviations from the null in any direction, it is well-known that they only have reasonable nontrivial local power against very few orthogonal directions, see Janssen (2000) and Escanciano (2008) for theoretical explanations and bounds for the number of orthogonal directions.

A possible solution to overcome the “lack” of power of omnibus tests is provided by the so-called Neyman’s smooth tests. They were first proposed by Neyman (1937) in the context of goodness-of-fit of distributions, and since then, there has been a plethora of research documenting their theoretical and empirical properties. In the context of MDH testing, a recent data-driven smooth test has been proposed by Escanciano and Mayoral (2007). Their test is based on the principal components of the marked empirical processes resulting from the choice $w_0(\tilde{Y}_{t,1}, x) = 1(Y_{t-1} \leq x)$ with $x \in \mathbb{R}$. This test is an extension to nonlinear dependence of order one, i.e. for $P = 1$, of the test based on N_n . As shown by these authors, this test possesses excellent local power properties and compares favorably to omnibus tests and other competing tests. The test statistic is

$$T_{n,\tilde{p}} = \sum_{j=1}^{\tilde{p}} \hat{\epsilon}_{j,n}^2, \quad (5)$$

with \tilde{p} as in (4), but with $T_{n,p}$, defined by (5), replacing Q_p^* there, and where $\hat{\epsilon}_{j,n}$ are the sample principal components of a certain CvM test; the reader is referred to Escanciano and Mayoral (2007) for details. The asymptotic null distribution of $T_{n,\tilde{p}}$ is a χ_1^2 .

We have applied the adaptive data-driven test based on $T_{n,\tilde{p}}$ to our exchange rates data. The results are reported in Table IV and support our previous conclusions. Only the MDH for the daily Euro exchange rate is rejected at 1% with $T_{n,\tilde{p}}$.

TABLE IV
TESTING THE MDH OF EXCHANGE RATES RETURNS
BOOTSTRAP P-VALUES. DATA-DRIVEN TESTS

	Daily				Weekly			
	Euro	£	Can	¥	Euro	£	Can	¥
$T_{n,\tilde{p}}$	0.049	0.847	0.514	0.876	0.622	0.133	0.747	0.299

4.2 Tests based on an infinite-dimensional information set

The aforementioned references test the MDH conditioning on a finite-dimensional information set, and therefore, they may miss some dependence structure in the conditional mean at omitted lags. In principle, the maximum power could be achieved by using the correct lag order P of the alternative. However, prior information on the conditional mean structure is usually not available.

There have been some proposals considering infinite-dimensional information sets. First, de Jong (1996) generalized Bierens' test to time series, and although his test had the appealing property of considering an increasing number of lags as the sample size increases, it required numerical integration with dimension equal to the sample size, which makes this test unfeasible in applications where the sample size is usually large, e.g. financial applications. Second, Domínguez and Lobato (2003) suggest constructing a test statistic as a weighted average of all the tests statistics established for a fixed number of lags. However, Dominguez and Lobato did not further analyze the test neither the selection of the measure to weight the different statistics.

Using a different methodology based on the generalized spectral density approach of Hong (1999), Hong and Lee (2003) proposed a MDH bootstrap test; see also Hong and Lee (2005). Tests based on the generalized spectral density involve three choices: a kernel, a bandwidth parameter and an integrating measure, and, in general, statistical inferences are sensitive to these choices. This fact motivated Escanciano and Velasco (2006a, 2006b) to propose MDH by means of a generalized spectral distribution function.

The generalized spectral approach is based on the fact that the MDH implies that

$$H_0 : \gamma_{j,w}(x) = 0 \quad \forall j \geq 1, \text{ for all } x, \quad (6)$$

where $\gamma_{j,w}(x) = E[(Y_t - \mu)w_0(Y_{t-j}, x)]$ and where $w_0(Y_{t-j}, x)$ is any of the parametric functions of the previous section. The generalized spectral approach of Hong is based on the choice $w_0(Y_{t-j}, x) = \exp(ixY_{t-j})$. Escanciano and Velasco (2006a) considered the latter choice, and Escanciano and Velasco (2006b) used $w_0(Y_{t-j}, x) = 1(Y_{t-j} \leq x)$, and called the measures $\gamma_{j,ind}(x) = E[(Y_t - \mu)1(Y_{t-j} \leq x)]$ the Integrated Pairwise Autoregression Functions (IPAF). The name follows from the fact that

$$\gamma_{j,ind}(x) = E[(Y_t - \mu)1(Y_{t-j} \leq x)] = \int_{-\infty}^x E[Y - \mu \mid Y_{t-j} = z]F(dz),$$

where F is the cdf of Y_t . The measures $\gamma_{j,w}(x)$ can be viewed as a generalization of the usual autocovariances to measure the conditional mean dependence in a nonlinear time series framework. They can be easily estimated from a sample. For instance, the IPAF's can be estimated by

$$\hat{\gamma}_{j,ind}(x) = \frac{1}{n-j} \sum_{t=1+j}^n (Y_t - \bar{Y})1(Y_{t-j} \leq x). \quad (7)$$

Moreover, as proposed by Escanciano and Velasco (2006b), nonlinear correlograms can be used to formally assess the nonlinear dependence structure in the conditional mean of the series. These authors define the KS test statistic as

$$KS(j) := \sup_{x \in [-\infty, \infty]^d} \left| (n-j)^{\frac{1}{2}} \hat{\gamma}_{j,ind}(x) \right| = \max_{1+j \leq t \leq n} \left| (n-j)^{\frac{1}{2}} \hat{\gamma}_{j,ind}(Y_{t-j}) \right|.$$

The asymptotic quantile of $KS(j)$ under the MDH can be approximated via a wild bootstrap approach. With the bootstrap critical values we can calculate uniform confidence bands for $\hat{\gamma}_j(x)$ and the significance of $\gamma_j(x)$ can be tested. The plot of a standardization of $KS(j)$ 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. Escanciano and Velasco (2006b) called this plot the Integrated Pairwise Regression Functions (IPRF) plot.

In Figures 2 to 9 we plot the IPRF for our exchange rates returns. The common feature of these graphs is the lack of dependence in the exchange rates, both linear and nonlinear. Only a few isolated statistics seem to be significative, but the evidence is very weak. It seems that the IPRF support the MDH for these data sets.

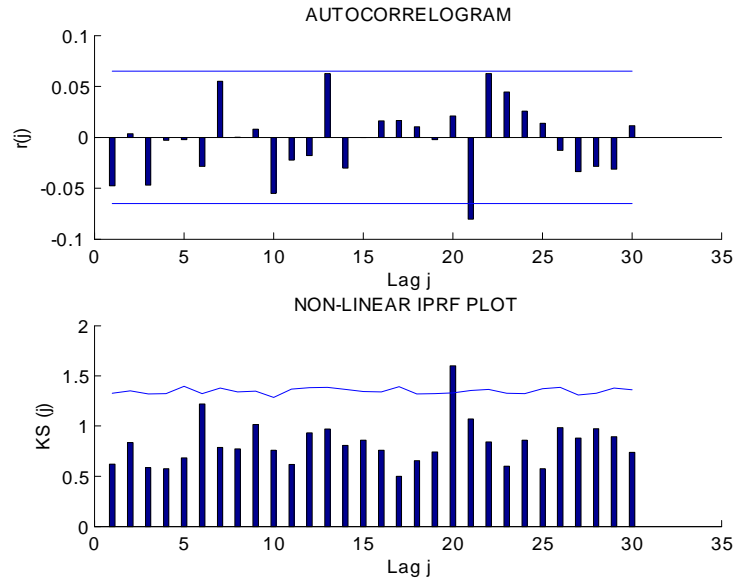


Figure 2. IPRF for the daily Euro. Top graph is the heteroskedasticity robust autocorrelation plot.
Bottom graph is the IPRF plot.

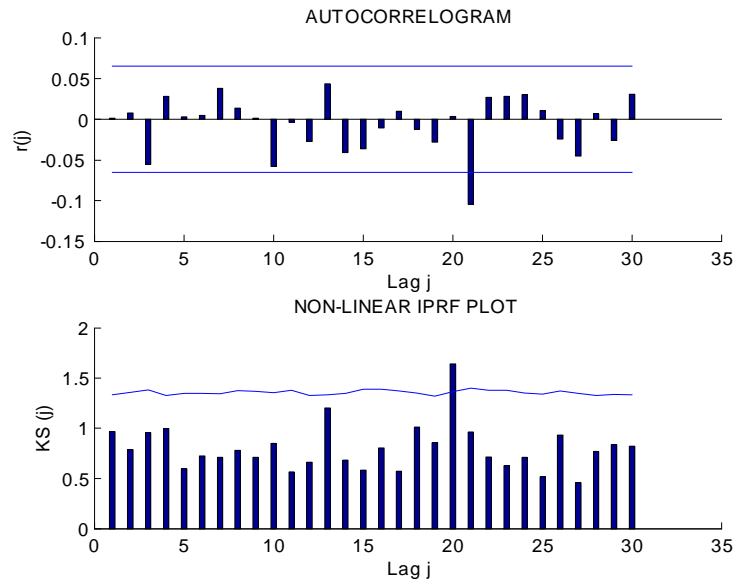


Figure 3. IPRF for the daily Pound. Top graph is the heteroskedasticity robust autocorrelation plot. Bottom graph is the IPRF plot.

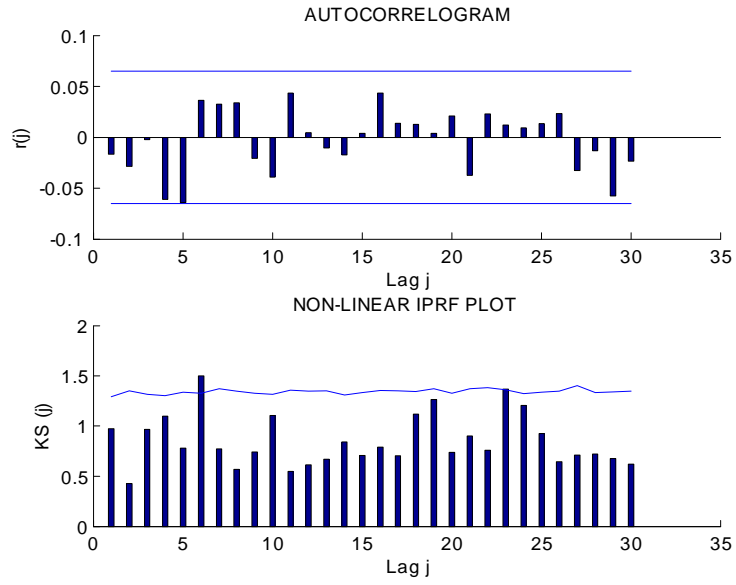


Figure 4. IPRF for the daily Can. Top graph is the heteroskedasticity robust autocorrelation plot.
 Bottom graph is the IPRF plot.

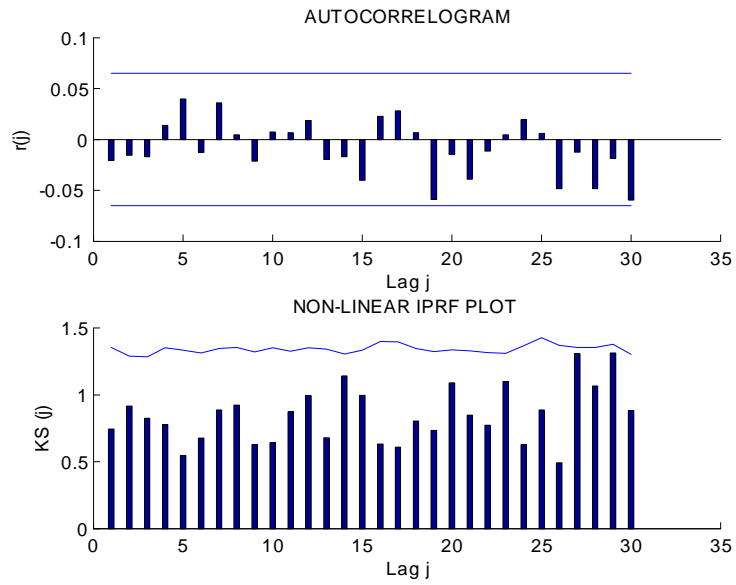


Figure 5. IPRF for the daily Yen. Top graph is the heteroskedasticity robust autocorrelation plot.
 Bottom graph is the IPRF plot.

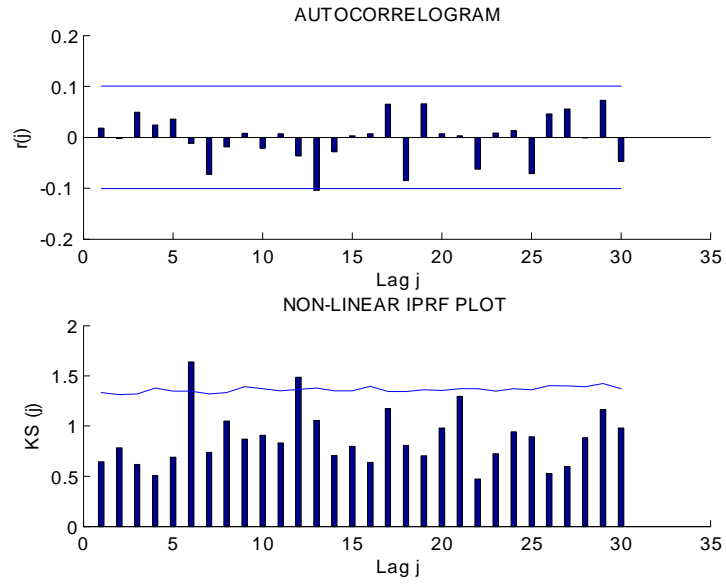


Figure 6. IPRF for the weekly Euro. Top graph is the heteroskedasticity robust autocorrelation plot. Bottom graph is the IPRF plot.

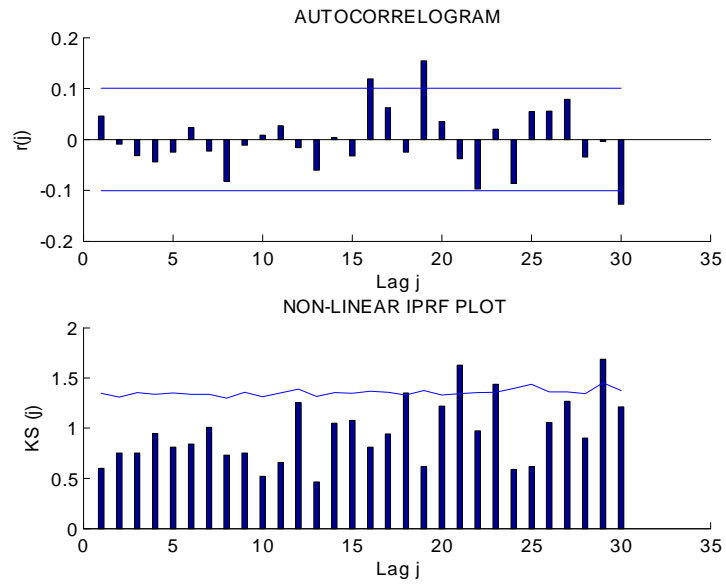


Figure 7. IPRF for the weekly Pound. Top graph is the heteroskedasticity robust autocorrelation plot. Bottom graph is the IPRF plot.

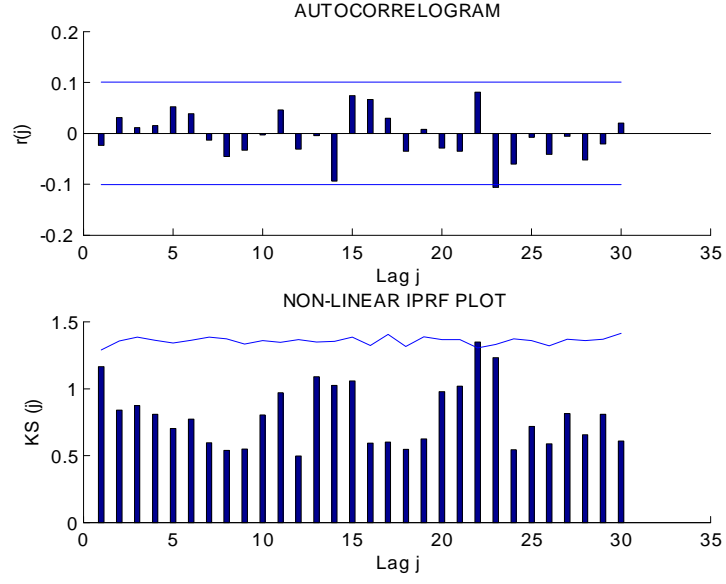


Figure 8. IPRF for the weekly Can. Top graph is the heteroskedasticity robust autocorrelation plot. Bottom graph is the IPRF plot.

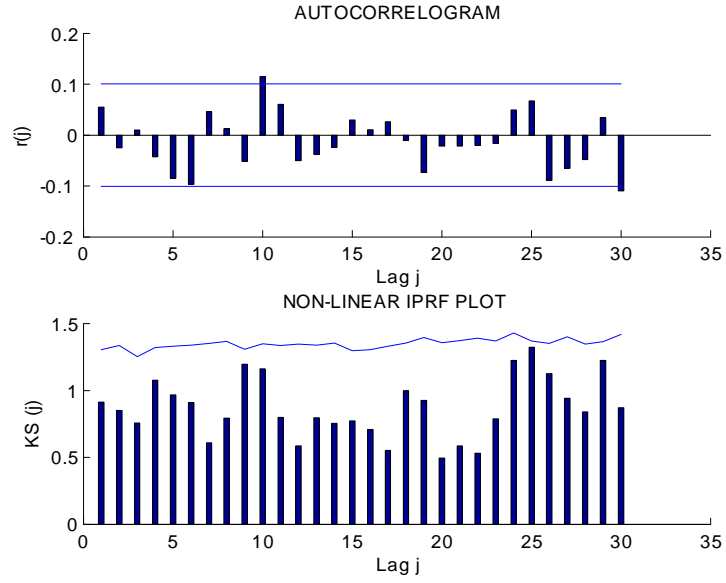


Figure 9. IPRF for the weekly Yen. Top graph is the heteroskedasticity robust autocorrelation plot. Bottom graph is the IPRF plot.

We now describe a generalized spectral approach to consider simultaneously all the nonlinear measures of dependence $\{\gamma_{j,w}(\cdot)\}$. Define $\gamma_{-j,w}(\cdot) = \gamma_{j,w}(\cdot)$ for $j \geq 1$, and consider the Fourier

transform of the functions $\gamma_{j,w}(x)$,

$$f_w(\varpi, x) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_{j,w}(x) e^{-ij\varpi} \quad \forall \varpi \in [-\pi, \pi]. \quad (8)$$

Note that $f_w(\varpi, x)$ is able to capture pairwise non-martingale difference alternatives with zero autocorrelations. Under the MDH, the condition $f_{0,w}(\varpi, x) = (2\pi)^{-1}\gamma_0(x)$ holds, that is, the generalized spectral density $f_w(\varpi, x)$ is flat in ϖ . Hong (1999) proposed the estimators

$$\hat{f}_w(\varpi, x) = \frac{1}{2\pi} \sum_{j=-n+1}^{n-1} \left(1 - \frac{|j|}{n}\right)^{\frac{1}{2}} k\left(\frac{j}{p}\right) \hat{\gamma}_{j,\text{exp}}(x) e^{-ij\varpi}$$

and

$$\hat{f}_{0,w}(\varpi, x) = \frac{1}{2\pi} \hat{\gamma}_{0,w}(x)$$

to test the MDH, where $k(\cdot)$ is a symmetric kernel and p a bandwidth parameter. He considered a standardization of an L_2 -distance using a weighting function $W(\cdot)$

$$\begin{aligned} L_{2,n}^2(p) &= \frac{\pi}{2} \int_{-\pi}^{\pi} \int_{\mathbb{R}} n \left| \hat{f}_w(\varpi, x) - \hat{f}_{0,w}(\varpi, x) \right|^2 W(dx) d\varpi \\ &= \sum_{j=1}^{n-1} (n-j) k^2\left(\frac{j}{p}\right) \int_{\mathbb{R}} |\hat{\gamma}_{j,w}(x)|^2 W(dx). \end{aligned} \quad (9)$$

Under the null of MDH and some additional assumptions, Hong and Lee (2005) showed that a convenient standardization of $L_{2,n}^2(p)$ converges to a standard normal random variable. The centering and scaling factors in this standardization depend on the higher dependence structure of the series.

Alternatively, the generalized spectral distribution function is

$$H_w(\lambda, x) = 2 \int_0^{\lambda\pi} f_w(\varpi, x) d\varpi \quad \lambda \in [0, 1],$$

which, after some algebra, boils down to

$$H_w(\lambda, x) = \gamma_{0,w}(x)\lambda + 2 \sum_{j=1}^{\infty} \gamma_{j,w}(x) \frac{\sin j\pi\lambda}{j\pi}. \quad (10)$$

Tests can be based on the sample analogue of (10), i.e.

$$\hat{H}_w(\lambda, x) = \hat{\gamma}_{0,w}(x)\lambda + 2 \sum_{j=1}^{n-1} \left(1 - \frac{j}{n}\right)^{\frac{1}{2}} \hat{\gamma}_{j,w}(x) \frac{\sin j\pi\lambda}{j\pi},$$

where $(1 - \frac{j}{n})^{\frac{1}{2}}$ is just a finite sample correction factor. The effect of this correction factor is to put less weight on very large lags, for which we have less sample information. Because under the MDH, $H_w(\lambda, x) = \gamma_0(x)\lambda$, tests can be based on the discrepancy between $\hat{H}_w(\lambda, x)$ and $\hat{H}_{0,w}(\lambda, x) := \hat{\gamma}_0(x)\lambda$. That is, we can consider the process

$$S_{n,w}(\lambda, x) = \left(\frac{n}{2}\right)^{\frac{1}{2}} \{\hat{H}_w(\lambda, x) - \hat{H}_{0,w}(\lambda, x)\} = \sum_{j=1}^{n-1} (n-j)^{\frac{1}{2}} \hat{\gamma}_{j,w}(x) \frac{\sqrt{2} \sin j\pi\lambda}{j\pi} \quad (11)$$

to test for the MDH.

In order to evaluate the distance of $S_n(\lambda, x)$ to zero, a norm has to be chosen. One norm considered in practice is the Cramér-von Mises norm

$$D_{n,w}^2 = \int_{\mathbb{R}} \int_0^1 |S_{n,w}(\lambda, x)|^2 W(dx) d\lambda = \sum_{j=1}^{n-1} (n-j) \frac{1}{(j\pi)^2} \int_{\mathbb{R}} |\hat{\gamma}_{j,w}(x)|^2 W(dx), \quad (12)$$

where $W(\cdot)$ is a weighting function satisfying some mild conditions. $D_{n,w}^2$ has the attractive convenience of being free of choosing any smoothing parameter or kernel, and it has been documented to deliver tests with good power properties (cf. Escanciano and Velasco, 2006a,b).

Among the members of this class of test statistics, the most common choices are

$$D_{n,\exp}^2 = \hat{\sigma}^{-2} \sum_{j=1}^{n-1} (n-j) \frac{1}{(j\pi)^2} \sum_{t=j+1}^n \sum_{s=j+1}^n (Y_t - \bar{Y}_{n-j})(Y_s - \bar{Y}_{n-j}) \exp(-0.5(Y_{t-j} - Y_{s-j})^2)$$

and

$$D_{n,ind}^2 = \hat{\sigma}^{-2} \sum_{j=1}^{n-1} \frac{(n-j)}{n(j\pi)^2} \sum_{t=1}^n \hat{\gamma}_{j,ind}^2(X_t),$$

where $\hat{\gamma}_{j,ind}$ is given in (7). The test statistic $D_{n,\exp}^2$ is based on $w_0(Y_{t-j}, x) = \exp(ixY_{t-j})$ and the standard normal cdf as the weighting function $W(\cdot)$, whereas $D_{n,ind}^2$ is based on $w_0(Y_{t-j}, x) = 1(Y_{t-j} \leq x)$ and the empirical cdf as the function W .

We have applied these two generalized spectral distribution based tests to our exchange rates data. The results are reported in Table V and support our previous conclusions. Only the MDH for the daily Euro exchange rate is rejected.

TABLE V
TESTING THE MDH OF EXCHANGE RATES RETURNS
BOOTSTRAP P-VALUES. GENERALIZED SPECTRAL TESTS

	Daily				Weekly			
	Euro	£	Can	¥	Euro	£	Can	¥
$D_{n,\exp}^2$	0.023	0.450	0.680	0.913	0.670	0.123	0.360	0.586
$D_{n,ind}^2$	0.016	0.343	0.640	0.923	0.800	0.253	0.526	0.524

5. RELATED HYPOTHESES

In this article we have considered testing the MDH that, in statistical terms, just implies that the mean of an economic time series is independent of its past. The procedures studied in this article can be straightforwardly applied for testing the following generalization of the MDH

$$H_0 : E[Y_t | X_{t-1}, X_{t-2}, \dots] = \mu, \quad \mu \in \mathbb{R}.$$

where Y_t is a measurable real-valued transformation of X_t and $\mu = E[Y_t]$. This null hypothesis, which is referred to as the generalized MDH, contains many interesting testing problems as special cases. For instance, when Y_t is a power transformation of X_t , this null hypothesis implies constancy of conditional moments. The leading case in financial applications is the case where $Y_t = X_t^2$, because when X_t follows an mds, this null hypothesis means that there is no volatility in the series X_t , that is, X_t is conditionally homoskedastic. The cases $Y_t = X_t^3$ or $Y_t = X_t^4$ would respectively test for no dynamic structure in the third (conditionally constant skewness) and fourth (conditionally constant kurtosis), see for instance, Bollerslev (1987) and Engle and González-Rivera (1991). Another relevant case is when $Y_t = 1(X_t > c)$, $c \in \mathbb{R}^d$. In this case, the null hypothesis tested represents no directional predictability, see e.g. Linton and Whang (2007). Other situation of interest occurs when the null hypothesis of interest is the equality of the regression curves of two random variables, X_{1t} and X_{2t} , say; in this case, $Y_t = X_{1t} - X_{2t}$, $\mu = 0$, see Ferreira and Stute (2004) for a recent reference.

Note also that most of the procedures considered in this article are also applicable for testing the null hypothesis that a general dynamic nonlinear model is correctly specified. In this situation, the null hypothesis of interest establishes that

$$\exists \theta_0 : E[\psi(Y_t, X_t, \theta_0) \mid X_t] = 0,$$

where ψ is a given function, Y_t is a vector of endogenous variables and X_t is a vector of exogenous variables. Test statistics can be constructed along the lines described in this article. The main theoretical challenge in this framework is the way of handling the estimation of the parameters. There are basically three alternative approaches. First, to estimate the asymptotic null distribution of the relevant test statistics by estimating its spectral decomposition (e.g. Horowitz (2006) or Carrasco, Florens and Renault (2007)). Second, to use the bootstrap to estimate this distribution, see Wu (1986) and Stute, W., González-Manteiga, W.G. and M. Presedo-Quindmil (1998). Third, to transform the test statistic via martingalization to yield an asymptotically distribution free test statistic.

Finally, in this article we have considered testing for mds instead of testing for martingale. Recall that X_t is a martingale with respect to its natural filtration, when $E[X_t \mid X_{t-1}, X_{t-2}, \dots] = X_{t-1}$ *a.s.* Testing for martingale presents the additional challenge of handling nonstationary variables. Park and Whang (2005) considered testing that a first-order Markovian process follows a martingale by testing that the first difference of the process conditionally on the last value has zero mean, that is,

$$E(X_t - X_{t-1} \mid X_{t-1}) = 0 \text{ a.s.} \quad (13)$$

Hence, they allow for a singular nonstationary conditioning variable. This restrictive Markovian framework has the advantage of leading to tests statistics which are asymptotically distribution

free, and hence, they do not need to transform their statistics or to use bootstrap procedures to obtain critical values. Similarly, note that many of the procedures described in Section 4 also lead to asymptotically distribution free tests in this restrictive framework. As shown in Escanciano (2007b), the extension to the multivariate conditioning case in (13) leads to non-pivotal tests and some resampling procedure is necessary.

6. CONCLUSIONS

This article has presented a general panoramic on the literature of testing for the MDH. This area started at the beginning of the last century by developing tests for serial correlation and experimented a renewed interest recently because of the nonlinear dependence present in economic and, specially, financial series. The initial statistical tools were based on linear dependence measures such as autocorrelations or the spectral density function. These tools were initially considered motivated by the observation that economic time series follow normal distributions. Since in the last twenty five years it has been stressed the non normal behavior of financial series, the statistical and econometrics literature followed two alternative approaches. The first one targeted to robustify the well-established linear measures to allow for non-linear dependence. This approach has the advantage of its simplicity since it leads typically to standard asymptotic null distributions. However, its main limitation is that it cannot detect nonlinear dependence. The second approach considered nonlinear measures of dependence. Its advantage is that it is more powerful, its disadvantage is that asymptotic null distributions are nonstandard. Nowadays, this feature is hardly a drawback because the increasing availability of computing resources has allowed the implementation of bootstrap procedures that can estimate the asymptotic null distributions with relative ease.

The definition of martingale involves the information set of the agent that typically contains the infinite past of the economic series. This feature implies that, in practice, it is practically impossible to construct a test which, although it may be consistent theoretically, has power for any possible violation of the null hypothesis. The *pairwise* approach, which admittedly does not deliver consistent tests, leads to tests with reasonable power for common alternatives. Another sensible possibility to reduce this dimensionality problem is to consider alternatives of a single-index structure, i.e. where the conditioning set is given by a univariate, possibly unknown, projection of the infinite-dimensional information set. More research is clearly needed in this direction.

In this article we have illustrated the different methodologies with exchange rate data that typically satisfy the MDH, as we have seen. Stock market data is not such a clear cut case. Rejecting the MDH leads to the challenge of selecting a proper model. In this respect, data-driven adaptive tests are informative, since they provide an alternative model in case of rejection. Notably, the principal component analysis provided in Escanciano and Mayoral (2007) represents a clear, theoretically well

motivated approach, that coupled with an effective choice for the number of components can help in this selection process.

References

- Andrews, D.W. K., and W. Ploberger (1996), "Testing for Serial Correlation Against an ARMA (1, 1) Process," *Journal of the American Statistical Association*, 91, 1331-1342.
- Barnett, W. and A. Serletis (2000), "Martingales, nonlinearity, and chaos", *Journal of Economics. Dynamics and Control*, 24, 703-724
- Bartlett, M.S. (1955), *An Introduction to Stochastic Processes with Special Reference to Methods and Applications*, Cambridge University Press, London.
- Bekaert, G. and R. Hodrick (1992), "Characterizing predictable components in excess returns on equity and foreign exchange markets", *Journal of Finance*, 467-509
- Bera, A.K. and M.L. Higgins (1993) "ARCH models: properties, estimation and testing", *Journal of Economic Surveys*, 7, 305-362.
- Bera, A.K. and M.L. Higgins (1997) "ARCH and bilinearity as competing models for nonlinear dependence", *Journal of Business and Economic Statistics*, 15, 43-51.
- Beran, J. (1992), "A Goodness-of-fit Test for Time Series with Long Range Dependence", *Journal of the Royal Statistical Society B*, vol. 54, pp.749-760.
- Bierens, H. J. (1982), "Consistent model specification tests", *Journal of Econometrics* 20, 105-134.
- Bierens, H. J. (1984), "Model specification testing of time series regressions", *Journal of Econometrics* 26, 323-353.
- Bierens, H. J. (1990), "A consistent conditional moment test of functional form", *Econometrica* 58, 1443-1458.
- Bierens, H. J. and W. Ploberger (1997), "Asymptotic theory of integrated conditional moment tests", *Econometrica* 65, 1129-1151.
- Bollerslev, T. (1987), "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return", *The Review of Economics and Statistics*, 69, 542-547.
- Bollerslev, T and J. M. Wooldridge, (1992) "Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances," *Econometric Reviews*, 11, 143-179.
- Box, G.E.P. and D. A. Pierce (1970) "Distribution of residual autocorrelations in autoregressive integrated moving average time series models," *Journal of the American Statistical Association* 65, 1509-1526.
- Brillinger, D.R. (1981), *Time Series: Data Analysis and Theory*, Holden Day, San Francisco.
- Campbell, J. Y., A.W. Lo and A.C. MacKinlay (1997), *The econometrics of financial markets*, Princeton University Press.

- Carrasco, M., J.P. Florens, and E. Renault (2007), “Linear Inverse Problems in Structural Econometrics”, *Handbook of Econometrics*, Vol. 6, eds. J.J. Heckman and E.E. Leamer.
- Chan, K.S. and H. Tong (2001), *Chaos: a statistical perspective*, Springer.
- Chen, H. and J.P. Romano (1999), “Bootstrap-assisted goodness-of-fit- tests in the frequency domain”, *Journal of Time Series Analysis*, 20, 619-654.
- Clark, T.E. and K.D. West (2006), “Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis”, *Journal of Econometrics*, 135, 155–186.
- Cochrane, J. H. (1988), “How big is the random walk in GNP?”, *Journal of Political Economy*, 96 893-920.
- Cochrane, J. H. (2005), *Asset Pricing*. Princeton University Press. Princeton, New Jersey.
- Cumby, R. E. and J. Huizinga (1992) “Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variable Regressions,” *Econometrica*, 60, 185-196.
- Dahlhaus, R. (1985), “On the asymptotic distribution of Bartlett’s U_p -statistic”, *Journal of Time Series Analysis*, 6, 213-227.
- Dahlhaus, R. and D. Janas (1996), “A frequency domain bootstrap for ratio statistics in time series analysis”, *The Annals of Statistics*, 24, 1934-1965.
- de Jong, R. M. (1996), “The Bierens’ tests under data dependence”, *Journal of Econometrics* 72, 1-32.
- Delgado, M. A., J. Hidalgo, and C. Velasco (2005), “Distribution free goodness-of-fit tests for linear models,” *The Annals of Statistics*, 33, 2568-2609.
- Delgado, M. A. and C. Velasco (2007), “A new class of distribution-free tests for time series models specification”, preprint.
- Deo, R.S. (2000), “Spectral tests of the martingale hypothesis under conditional heteroscedasticity”, *Journal of Econometrics*, 99, 291-315.
- Deo, R.S. and W. Chen (2000), “On the integral of the squared periodogram”, *Stochastic Processes and their Applications*, 85, 159-176.
- Diebold, F. X. (1986), “Testing for serial correlation in the presence of heteroskedasticity”, *Proceedings of the American Statistical Association, Business and Economics Statistics Section*, 323-328.
- Diebold, F. X and J.A. Nason (1990), “Nonparametric exchange rate predictions?”, *Journal of International Economics*, 28, 315-332.
- Dominguez, M. and I.N. Lobato (2003) “Testing the martingale difference hypothesis”, *Econometric Reviews*, 22, 351-377.
- Durbin, J. and G.S. Watson, (1950), “Testing for Serial Correlation in Least Squares Regression: I”,

- Biometrika, 37, 409-428.
- Durlauf, S. N. (1991), "Spectral based testing of the martingale hypothesis", *Journal of Econometrics*, 50, 355-376.
- Engle, R.F. and G. González-Rivera (1991), "Semiparametric ARCH Models", *Journal of Business & Economic Statistics*, 9, 345-359.
- Escanciano, J. C. (2007a), "Model Checks Using Residual Marked Empirical Processes," *Statistica Sinica*, 17, 115-138.
- Escanciano, J. C. (2007b): "Weak convergence of non-stationary multivariate marked processes with applications to martingale testing", *Journal of Multivariate Analysis*, 98, 1321-1336.
- Escanciano, J. C. (2008), "On the Lack of Power of Omnibus Specification Tests," forthcoming in *Econometric Theory*.
- Escanciano, J.C. and I.N. Lobato (2007), "Data-Driven Portmanteau Tests for Testing for Serial Correlation", preprint.
- Escanciano, J. C. and S. Mayoral (2007) "Data-Driven Smooth Tests for the Martingale Difference Hypothesis, " preprint.
- Escanciano, J. C., and C. Velasco (2006a), "Generalized Spectral Tests for the Martingale Difference Hypothesis," *Journal of Econometrics*, 134, 151-185.
- Escanciano, J. C. and C. Velasco (2006b), "Testing the Martingale Difference Hypothesis Using Integrated Regression Functions", *Computational Statistics & Data Analysis*, 51, 2278-2294.
- Fama, E. F. (1991), "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25, 383-417.
- Fan, J. and L. Huang (2001), "Goodness-of-Fit Tests for Parametric Regression Models," *Journal of the American Statistical Association*, 96, 640-652.
- Ferreira, E., and W. Stute (2004) "Testing for Differences Between Conditional Means in a Time Series Context," *Journal of the American Statistical Association*, 99, 169-174.
- Fong, W. M., S.K. Koh, and S. Ouliaris (1997), "Joint variance ratio test of the martingale hypothesis for exchange rates", *Journal of Business and Economics Statistics*, 15, 51-59.
- Fong, W. M. and S. Ouliaris (1995), "Spectral tests of the martingale hypothesis for exchange rates", *Journal of Applied Econometrics* 10, 255-271.
- Francq, C., R. Roy & J.-M. Zakoïan (2005), "Diagnostic checking in ARMA models with uncorrelated errors", *Journal of the American Statistical Association*, 13, 532-544.
- Franke, J. and W. Hardle (1992), "On bootstrapping kernel spectral estimates", *The Annals of Statistics*, 20, 121-145.

- González, M and I.N. Lobato (2003), “Contrastes de autocorrelación”, *Gaceta de Economía*, 14, pp. 41-57.
- Grenander, U. and M. Rosenblatt (1957), *Statistical Analysis of Stationary Time Series*, Chelsea Publishing Company, New York.
- Guay, A., and E. Guerre (2006), “A Data-Driven Nonparametric Specification Test for Dynamic Regression Models,” *Econometric Theory*, 22, 543-586.
- Guerre, E., and P. Lavergne (2005), “Rate-optimal data-driven specification testing for regression models”, *Annals of Statistics* 33, 840-870.
- Guo, B. and P. C.B. Phillips (2001), “Testing for Autocorrelation and Unit Roots in the Presence of Conditional Heteroskedasticity of unknown Form,” *Cowles Foundation for Research in Economics*, Yale University-
- Hall, R. E. (1978), “Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence”, *Journal of Political Economy* 86, 971-987.
- Hart, J.D. (1997). *Nonparametric Smoothing and Lack-of-Fit Tests*. Springer Verlag, New-York.
- Hinich, M. and D. Patterson (1992) “A new diagnostic test of model inadequacy which uses the martingale difference criterion”, *Journal of Time Series Analysis* 13, 233-252.
- Hong, Y. (1996), “Consistent Testing for Serial Correlation of Unknown Form”, *Econometrica*, 64, 837-864.
- Hong, Y., (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.
- Hong, Y. and T.H. Lee (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.
- Hong, Y. and Y.J. Lee (2005), “Generalized spectral tests for conditional mean models in time series with conditional heteroskedasticity of unknown form”, *Review of Economic Studies*, 43, 499-541.
- Horowitz, J. (2006) “Testing a Parametric Model Against a Nonparametric Alternative with Identification Through Instrumental Variables”, *Econometrica*, 74, 521-538.
- Horowitz, J. L. and W. Härdle (1994) “Testing a parametric model against a semiparametric alternative”, *Econometric Theory* 10, 821-848.
- Horowitz, J., I.N. Lobato, J.Nankervis and N.E. Savin, (2006) “Bootstrapping the Box-Pierce Q test: a robust test of uncorrelatedness”, *Journal of Econometrics*, 133, 841-862.
- Horowitz, J.L., and V.G. Spokoiny (2001) “An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative”, *Econometrica* 69, 599-632.
- Hsieh, D. A. (1988), “Statistical properties of daily foreign exchange rates”, *Journal of International*

- Economics, 24, 129-145.
- Hsieh, D. A. (1989), "Testing for nonlinear dependence in daily foreign exchange rates", *Journal of Business* 62, 339-368.
- Hsieh, D. A. (1993), "Implication of nonlinear dynamics for financial risk management", *Journal of Finance and Quantitative Analysis*, 28, 41-64.
- Inglot, T., and T. Ledwina (2006), "Data Driven Score Tests of Fit for Semiparametric Homoscedastic Linear Regression Model", Preprint 665, Institute of Mathematics Polish Academy of Sciences.
- Inoue, A. and L. Kilian (2004), "In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?", *Econometric Reviews*, 23(4), 371-402.
- Janssen, A. (2000), "Global power functions of goodness of fit tests", *Annals of Statistics*, 28, 239-253.
- Khmaladze, E. V. (1981), "Martingale Approach in the Theory of Goodness-of-Fit Tests", *Theory of Probability and its Applications* 26, 240-257.
- Koul, H. L. and W. Stute (1999), "Nonparametric model checks for time series", *The Annals of Statistics* 27, 204-236.
- Kuan, C.-M. and W. Lee (2004), "A new test of the martingale difference hypothesis", *Studies in Nonlinear Dynamics & Econometrics*, 8.
- Laudrup N. and M. Jansson (2006), "Improving size and power in unit root testing", *Palgrave Handbook of Econometrics*, Vol. I, chapter 7, pp 252-277.
- Lavergne, P. and V. Patilea (2007), "One for all and all for one: dimension reduction for regression checks", forthcoming in *Journal of Econometrics*.
- Lebaron, B. (1999), "Technical trading rule profitability and foreign exchange intervention", *Journal of International Economics*, 49, 125-143.
- LeRoy, S. F. (1989), "Efficient Capital Markets and Martingales," *Journal of Economic Literature*, 27, 39-44.
- Levich, R. and L. Thomas (1993), "The significance of technical trading rule profits in the foreign exchange market: a bootstrap approach", *Journal of International Money and Finance*, 12, 451-474.
- Linton, O. and Y.-J. Whang (2007), "The Quantilogram: with an application to evaluating directional predictability", *Journal of Econometrics*, 141, 250-282.
- Liu, R. Y. (1988), "Bootstrap procedures under some non-i.i.d models", *The Annals of Statistics* 16, 1696-1708.
- Liu, C. Y. and He, J. (1991), "A variance ratio test for random walks in foreign exchange rates", *Journal of Finance*, 46, 773-785.

- Ljung, G.M. and G. E. P. Box (1978), "On a Measure of Lack of Fit in Time Series Models", *Biometrika*, 65, 297-303.
- Lo, A. W. (1997), *Market efficiency: stock market behaviour in theory and practice*, Vol. I and II (Edward Elgar).
- Lo, A.W. and A.C. MacKinlay (1989), "The size and power of the variance ratio test in finite samples: a Monte Carlo investigation", *Journal of Econometrics*, 40, 203-238.
- Lobato, I.N. (2001), "Testing that a dependent process is uncorrelated", *Journal of the American Statistical Association*, vol.96, pp.1066-1076.
- Lobato, I.N., J. C. Nankervis, and N.E. Savin (2001) "Testing for autocorrelation using a modified Box-Pierce Q test," *International Economic Review* 42, 187-205.
- Lobato, I.N., J. C. Nankervis, and N.E. Savin (2002), "Testing for Zero Autocorrelation in the Presence of Statistical Dependence," *Econometric Theory*, 18, 730-743.
- Lobato, I.N. and C. Velasco (2004), "A General Test for White Noise", preprint.
- McCurdy, T. H. and I.G. Morgan (1987), "Testing of the martingale hypothesis for foreign currency futures with time-varying volatility", *International Journal of Forecasting*, 3, 131-148.
- McCurdy, T. H. and I.G. Morgan (1988), "Testing of the martingale hypothesis in Deutsche Mark futures with model specifying the form of heteroskedasticity", *Journal of Applied Econometrics*, 3, 187-202.
- Meese, R. A. and K. Rogoff (1983a), "Empirical exchange rates models of the seventies", *Journal of International Economics*, 14, 3-24.
- Meese, R. A. and K. Rogoff (1983b), "The out of sample failure of empirical exchange rates models: sampling error or misspecification?", in J. Frenkel (Ed.), *Exchange Rates and International Economics* (Chicago: University of Chicago Press).
- Milhøj, A. (1981), "A test of fit in time series models", *Biometrika*, 68, 177-187.
- Nankervis J.C., and N.E. Savin (2007) "Testing for serial correlation: Generalized Andrews-Ploberger tests", preprint.
- Neyman, J. (1937), "Smooth Test for Goodness of Fit," *Scandinavian Aktuarietidskr*, 20, 149-199.
- Paparoditis, E. (2000), "Spectral density based goodness-of-fit tests for time series models", *Scandinavian Journal of Statistics* 27, 143-176.
- Park, J. Y., and Y.J. Whang (2005), "Testing for the Martingale Hypothesis", *Studies in Nonlinear Dynamics and Econometrics*, Issue 2, Article 2.
- Robinson, P. M., (1991) "Testing for Strong Serial Correlation and Dynamic Conditional Heteroskedasticity in Multiple Regression," *Journal of Econometrics*, 47, 67-84.

- Romano, J. L., and L.A. Thombs (1996), "Inference for Autocorrelations Under Weak Assumptions," *Journal of the American Statistical Association*, 91, 590-600.
- Shorack, G. and J. Wellner (1986), *Empirical Processes with Applications to Statistics* (Wiley: New York).
- Stinchcombe, M. and H. White (1998), "Consistent specification testing with nuisance parameters present only under the alternative", *Econometric Theory* 14, 295-325.
- Stute, W. (1997), "Nonparametric model checks for regression", *The Annals of Statistics* 25, 613-641.
- Stute, W., W. G. González-Manteiga, and M. Presedo-Quindmil (1998), "Bootstrap approximations in model checks for regression", *Journal of the American Statistical Association*, 83, 141-149.
- Sweeny, R. J. (1986), "Beating the foreign exchange market", *Journal of Finance*, 41, 163-182.
- Weiss, A. (1986), "ARCH and bilinear time series models: comparison and combination", *Journal of Business and Economic Statistics*, 4, 59-70.
- Wooldridge, J. (1992), "A test for functional form against nonparametric alternatives", *Econometric Theory* 8, 452-475.
- Wu, C. F. J., (1986), "Jackknife, bootstrap and other resampling methods in regression analysis with discussion", *The Annals of Statistics* 14, 1261-1350.
- Yatchew, A. J. (1992) "Nonparametric regression tests based on least squares", *Econometric Theory* 8, 435-451.
- Yule, G.U. (1926), "Why do we sometimes get nonsense correlations between time-series? A study in sampling and the nature of time-series", *Journal of the Royal Society*, vol. 89, pp.2-9, 30-41.
- Zheng, X. (1996) "A consistent test of functional form via nonparametric estimation technique", *Journal of Econometrics* 75, 263-289.