Robust Analysis of the Martingale Hypothesis

Christian Gourieroux * and Joann Jasiak †

preliminary version
March 1, 2016

The authors thank E., Guerre, J., Kim, and J.M., Zakoian for helpful comments.

*University of Toronto and CREST, *e-mail*: gouriero@ensae.fr
†York University, Canada, *e-mail*: jasiakj@yorku.ca.

Robust Analysis of the Martingale Hypothesis

Abstract

The martingale hypothesis is commonly tested in various time series, including the financial and economic data with trends. In practice, there exists a variety of martingale processes and not all of them are nonstationary, like the random walks. In particular, some martingales are stationary processes with heavy-tailed marginal distributions. These martingales display local trends and bubbles, and can feature volatility induced "mean-reversion". The aim of our paper is to develop tests of the martingale hypothesis, which are robust to the type of martingale proces that generated the data, and are valid for nonstationary as well as stationary martingales.

Keywords: Martingale Hypothesis, Recurrence, Noncausal Process, Stationary Martingale, Nadaraya-Watson Estimator.

JEL number: C52, C13, C16.

1 Introduction

The class of martingale processes is very large and includes the well-known nonstationary processes, such as the random walks as well as some specific stationary processes. All these processes display trends, which can be either global, local or just short-lived bubbles. In practice, a number of economic and financial time series display this behavior. For example, the time series of retail price indexes may have explosive behavior due to nonstationarity, while the exchange rates series tend to have local trends without taking on extreme values. Loosely speaking, the martingale property is a kind of "mean reverting" behavior [1], although that "mean reversion" can occur very infrequently. This implies that trends in the trajectories of martingale processes can be perceived as global or local depending on the length of the sampling period and the type of observed martingale. Therefore, visual inspection of a time series and traditional methods of analysis may suggest nonstationarity in martingale processes that are stationary.

This paper explores nonparametric tests of the martingale hypothesis under the maintained hypothesis of homogenous (recurrent) Markov process. Let us consider discrete time observations $y_0, y_1, ...y_T$ on a scalar process (y_t). The two above hypotheses concern the form of the transition distribution of y_t given the past values $\underline{y_{t-1}} = \{y_{t-1}, ..., y_0\}$. The maintained hypothesis of the process being a homogenous Markov is written in terms of the transition density function (p.d.f) denoted $f(y_t|\underline{y_{t-1}})$ as:

$$H = \left\{ f(y_t|\underline{y_{t-1}}) = f(y_t|y_{t-1}), \ \ \forall \underline{y_t} \right\}. \tag{1.1}$$

Under H, the null hypothesis of martingale can be written either in terms of the process(y_t), or of its first difference, i.e. the martingale difference sequence (MDS): $\Delta y_t = (y_t - y_{t-1})$ as follows:

$$\begin{aligned} H_0 &= \{E(y_t|y_{t-1}) = y_{t-1}, \ \forall y_{t-1}\} \\ &= \{E(y_t - y_{t-1}|y_{t-1}) = 0, \ \forall y_{t-1}\}. \end{aligned} \tag{1.2}$$

with respect to the same conditioning set [2] $\underline{y_{t-1}}$.

[1] The mean reversion is rather meant as mode reversion as many martingales have no finite mean.

[2] The set $\underline{y_{t-1}}$ differs from the set $\underline{\Delta y_{t-1}}$ due to the effect of the initial value y_0.

The tests of the Markov and martingale hypotheses appear commonly in financial applications in the context of market efficiency [3]. The Markov hypothesis represents the "informational efficiency". More specifically, if a market is informationally efficient, then all available information on a future of a stock price is "fully reflected" in the current market price [Fama (1965), (1970), (1991)].

The martingale hypothesis corresponds to another type of market efficiency, which is the impossibility to beat the market. It can be summarized by considering market with two assets: a riskfree asset with zero interest rate and a risky asset with price y_t at date t. A well-known property of a martingale is the Doob's optional stopping time Theorem that implies:

$$E(y_{t+\nu}|y_t) = y_t, \quad \forall y_t, \tag{1.3}$$

for any upper bounded stopping time ν [4].

If an investor buys a unit of the risky asset at date t at price y_t, then, no matter how sophisticated is his/her strategy of chosing the best date to sell, his/her expected gain will be the same as if he/she had invested in the riskfree asset only [5].

Therefore it is common to test the martingale hypothesis in various price series (possibly discounted) such as stock prices, market prices, commodity prices, exchange rates, retail price indexes, or zero-coupon prices with given maturity. Some of these processes, such as the commodity prices or exchange rates can be stationary, while satisfying the martingale condition.

This paper introduces a test statistic for testing the null hypothesis of martingale that includes all possible martingales, which can be either stationary, or nonstationary. Our purpose is fill the gap in the literature on martingale hypothesis tests, as the existing tests do not include the stationary martingles under the null hypothesis.

More specifically, the null martingale hypothesis H_0, examined in this paper, concerns

[3] Other interpretations have been provided for macroeconomic applications [see e.g. Hall (1978)].

[4] A stopping time is a discrete variable such that $\{\nu = h\}$ is a function of $(\underline{y_{t+h-1}})$, for any $h \geq 1$.

[5] When the riskfree interest rate is not equal to 0 and varies stochastically in time, the same reasoning applies to the compoundly discounted stock price:

$$y_t = B(0,1)B(1,2), ..., B(t-1,t)p_t, \tag{1.4}$$

where $B(\tau, \tau+1)$ denotes the price at τ of the short-term zero coupon bond (with time to maturity 1).

the function $m(y) = E(y_t|y_{t-1} = y)$, which under the null is equal to y [6]. As $m(y)$ is a functional parameter, it has to be tested through a functional statistic that depends on state y. We follow Karlsen, Tjostheim (2001) and consider the nonparametric kernel estimator of the autoregression $y \to m(y) = E(Y_{t+1}|Y_t = y)$, that is the Nadaraya-Watson (NW) estimator. Karlsen, Tjostheim (2001) show that such kernel estimator with appropriately chosen kernel and bandwidth is consistent of function m and of the confidence interval of m. While these results are independent of the stationarity or nonstationarity of the process, they depend on the recurrence property, a condition that is often satisfied by the data (see Section 3, for the recurrence condition). We observe that the convergence of the NW estimator is not uniform in y, especially in the extremes. Therefore, any scalar statistic based on this estimator can lead to spurious results if the extremes are given too much weight. Therefore, out proposed scalar test statistic of the martingale hypothesis based on the NW estimator of m is adjusted for the extremes to avoid the effects of the lack of uniformity.

The paper is organized as follows.

Section 2 summarizes the main tests of the martingale hypothesis that exist in the literature, and distinguishes the implicit and alternative hypotheses considered.

Section 3 provides various examples of martingales, such as random walks, time discretized scalar diffusions, double autoregressive processes, and the noncausal Cauchy processes. We discuss their properties of stationarity and their properties of recurrence, such as the frequency at which those processes return in a neighborhood of a given value. As an illustration simulated paths of the processes are provided.

Section 4 considers the nonparametric kernel estimation of the autoregression $y \to m(y) = E(Y_{t+1}|Y_t = y)$, and illustrates the properties of the NW estimator in different types of martingales described in Section 2.

Section 5 introduces the test of the martingale hypothesis for stationary and nonstationary processes. As an example, the effect of the lack of uniformity on the nonconvergence of the OLS estimator of regression coefficient illustrated.

Section 6 provides the Monte Carlo results and illustrates the properties of the propsed test in finite sample. Section 6 concludes. Appendix 1 provides a summary of the literature on martingale tests. Appendix 2 outlines the assumptions for the asymptotic properties

[6] Under continuity conditions introduced in Section 3.

of the NW estimator. Appendix 3 discusses the properties of the tests based on the mean squared error (MSE) of $\hat{m}_T(y_t)$ proposed in [Gao, King, Lu, Tjostheim (2009)].

2 Martingale Hypothesis Tests in the Literature

There exists a large body of literature on the tests of the martingale hypothesis, which is summarized in Appendix 1. Table 1 distinguishes the main types of tests and provides the implicit null hypothesis $\bar{H}_0$ and the implicit alternative hypothesis $\bar{H}_A$. The implicit null hypothesis (resp. implicit alternative hypothesis) includes all the dynamics which are asymptotically (for T large) accepted by that test procedure (resp. rejected by the test procedure). An optimal test procedure is expected to be such that the implicit null hypothesis coincides with the null hypothesis of interest: $\bar{H}_0 = H_0$. Among the tests outlined in Table 1 are tests such that : $H_0 \cap \bar{H}_A \neq \emptyset$, so that the implicit null hypothesis can be rejected in some martingale processes. Table 1 also includes tests such that $H_0 \subset \bar{H}_0$, but $\bar{H}_0 - H_0 \neq \emptyset$. These tests accept the implicit null hypothesis in processes which are not martingales. In both cases, the testing procedures are not properly designed.

Given this observation, let us now comment on selected tests displayed in Table 1 [see Escanciano, Lobato (2009)b for another survey]. Historically, the first procedures are based on the analysis of the sequence of estimated correlations of the differenced series $\Delta y_t = y_t - y_{t-1}$. This sequence can be examined in terms of its correlations by the well-known Box-Pierce tests, the Ljung Box tests, the portmanteau statistics, or by means of its spectral density [see e.g. Durlauf (1991), and Deo (2000), Lobato, Nankervis, Savin (2001) for extensions]. These test procedures have the two drawbacks mentioned above and can accept non-martingale dynamics, as only the absence of linear serial dependence is accounted for, as well as they can reject martingales when the difference Δy_t has no second-order moment, due to fat tails. The same remark applies to all procedures based on autocovariances, such as the variance ratio test introduced in Lo, MacKinlay (1988) [see also Chen, Deo (2006)].

Some other procedures proposed in the literature rely on testing if the autoregressive coefficient in a dynamic model of the process (y_t) is equal to one [see e.g. White (1958), (1959), Dickey, Fuller (1979), Phillips (1987)a, Phillips, Perron (1988), Chan, Wei (1988)]. In general, these procedures test the null of nonstationarity of the process against the alternative of stationarity. Thus, the implicit null hypothesis can be rejected if the process is a stationary martingale.

The null martingale hypothesis H_0, examined in this paper, concerns the nonparametric autoregression function $m(y) = E(y_t|y_{t-1} = y)$. In this respect, four types of (functional) test procedures have been recently suggested in the literature:

i) Tests based on a functional estimator of function $a(w) = E[\Delta y_t exp(iw\Delta y_{t-1})]$ [Hong (1999), Escanciano, Velasio (2006)].

ii) Tests based on the kernel estimator of the autoregressive function: $g(u) = E(\Delta y_t|\Delta y_{t-1} = u)$. These tests involve an implicit Markov assumption of Δy_t instead of a Markov assumption on the process (y_t), the only Markov assumption with economic interpretation.

iii) Tests based on a nonparametric estimator of the function:

$$c(y) = E(\Delta y_t 1_{y_{t-1} \leq y}),$$

[see, Park, Whang (2005), Phillips, Jin (2014)].

iv) Tests based on a kernel estimator $\hat{m}_T$ of function m [Park, Phillips (1998), Gao, King, Lu, Tjostheim (2009), Myklebust, Karlsen, Tjosheim (2012)].

The two latter tests focus on the dependence between y_t and y_{t-1} under the maintained Markov hypothesis. The difference between them is as follows. The first approach is global, as function c is defined by an unconditional expectation, while the second test based on the nonlinear autoregression is local. This has an important consequence for the outcome of the test. While under the local approach one can find an asymptotically pivotal statistic for H_0, that is a statistic based on a nonparametric $\hat{m}_T$ with a fixed asymptotic distribution under H_0 [Karlsen, Tjostheim (2001)], this is not the case under the global approach, as the derivation of such an asymptotic pivotal statistic requires additional assumptions under the null,such as weak conditional heteroscedasticity (WCH, see Appendix 1) and the existence of second-order moment of the differenced series Δy_t. Thus, the implicit alternative hypothesis contains martingales whose first difference Δy_t has no second-order moment and is characterized by strong conditional heteroscedasticity.

3 The Martingales

This section defines the martingale process (y_t) that satisfies the mainatined Markov hypothesis. As an illustration, examples of nonstationary martingales, such as the random walk, and stationary martingales, such as the double autoregressive process and the noncausal Cauchy AR(1) are discussed.

3.1 Definition

Let $y = (y_t,\ t = 0, 1, 2, ...)$ denote a scalar process in discrete time. We assume that it is a Markov process of order one that takes values in the interval $(-\infty + \infty)$, and has a continuous transition density function denoted by:

$$f(y_{t+1}|y_t) = lim_{dy \to 0} \frac{1}{dy} P[Y_{t+1} \in (y_{t+1}, y_{t+1} + dy)|Y_t = y_t].$$

Let us introduce the following assumptions :

Assumption A.1

i) For any value y_{t+1}, the transition density is a continuous function of y_t.

ii) The transition density is strictly positive for any y_t, y_{t+1}.

The continuity assumption is needed to determine the density, without ambiguity. It is also needed for the local analysis of the transition density and its kernel estimation. The strict positivity (with the recurrence condition) ensures that the process takes its values from any open interval.

The Markov process is recurrent [7] if for any y and any interval $(a, b), b > a$, we have $P[S_{(a,b)} < \infty|Y_t = y] = 1$, where $S_{(a,b)} = inf\{t : Y_t \in (a, b)\}$ is the first time of entry in (a, b).

Assumption A.2 For any value of y_t, the integral $\int |y_{t+1}|^p f(y_{t+1}|y_t) dy_{t+1}$ exists and is a continuous function of y_t.

When $p = 1$, the conditional expectation is:

$$m(y) = E(Y_{t+1}|Y_t = y). \tag{3.1}$$

[7] This is a condition of Harris recurrence with the associated Lebesgue measure λ. The exact condition is : "if for any y, λ a.e.", but the a.e. can be omitted due to the assumption of continuity on the transition p.d.f.

When $p = 2$, the conditional variance is defined as:

$$\eta^2(y) = V(Y_{t+1}|Y_t = y). \tag{3.2}$$

Definition 1: Suppose that assumptions A.1 and A.2 hold with $p = 1$. The process Y is a martingale if and only if $E(Y_{t+1}|Y_t = y) = y, \; \forall y \iff m(y) = y, \; \forall y$.

The martingale has finite first-order conditional moments for h=1 and more generally for any fixed $h = 1, 2, ..$:

$$E[Y_{t+h}|Y_t = y_t] = E[E[Y_{t+h}|Y_{t+1}]|Y_t = y_t] = E[Y_{t+1}|Y_t = y_t] = y_t.$$

In general, a martingale does not have finite unconditional moments, that is it can be such that $E|Y_t| = \infty$.

An integrable martingale, which is not conditionally equal to a constant (a consequence of Assumption A.1.) is necessarily a nonstationary process. Indeed, by the convexity inequality, we have:

$$E(|Y_t||Y_{t-1}) > |E(Y_t|Y_{t-1})| = |Y_{t-1}|,$$

and by taking the expectation on both sides:

$$E(|Y_t|) > E(|Y_{t-1}|).$$

Thus, the nonstationarity results from the bahavior of $E|Y_t|$.

As a consequence, a stationary martingale cannot be integrable and features (unconditional) fat tails.

Let us now provide the examples of martingales, which are commonly examined in the literature.

3.2 Random Walk

A random walk is a process defined as:

$$Y_t = Y_{t-1} + u_t, \; t \geq 1, \tag{3.3}$$

where (u_t) is a sequence of independent, identically distributed (i.i.d.) variables, and the initial value Y_0 is assumed to be independent of this sequence.

A random walk satisfies the martingale definition if the first moment of u_t exists: $E|u_t| < \infty$, and is equal to $Eu_t = 0$. When the p.d.f. of u_t is strictly positive, a random walk satisfies the recurrence condition.

3.3 Time discretized diffusion process

Let us consider a diffusion process $\{Y_t, t \in (0, \infty)\}$ that satisfies:

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t,\ t \geq 0, \tag{3.4}$$

where (W_t) is a Brownian motion, μ and σ are the drift and volatility functions, respectively:

$$\mu(y_t) = \lim_{dt \to 0} \frac{1}{dt} E[Y_{t+dt} - Y_t | Y_t = y_t],\ \ \sigma^2(y_t) = \lim_{dt \to 0} \frac{1}{dt} V[Y_{t+dt} | Y_t = y_t],$$

and the initial value Y_0 of this process is independent of the evolution of the Brownian motion $W_t, t \geq 0$.

The standard local Lipschitz conditions are assumed to hold, which ensures the existence of this diffusion process and that $\sigma(y) > 0,\ \forall y \in (-\infty, +\infty)$. Due to this positivity condition the diffusion process takes any values in $(-\infty, +\infty)$ [see Revuz, Yor (1991), Section IX, paragraph 2].
We have:

$$E(dY_t | Y_t) = E(Y_{t+dt} | Y_t) - Y_t = \mu(Y_t)dt.$$

Hence, at an infinitesimal horizon dt, the diffusion process satisfies the martingale condition if and only if the drift is zero. From the iterated expectation theorem, it follows that the time discretized diffusion:

$$dY_t = \sigma(Y_t)dW_t, \tag{3.5}$$

is a martingale in discrete time.

The recurrence property of a scalar diffusion process depends on the asymptotic behavior of its scale function S which has to be such that [see e.g. Khasminskii (1980), Karatzas, Shreve (1991), Ex. 7.13.]:

$$lim_{y \to -\infty} S(y) = -\infty, \;\; lim_{y \to +\infty} S(y) = +\infty.$$

For a scalar diffusion, the scale function is:

$$S(y) = \int_0^y exp[\int_0^x -\frac{2\mu(x)}{\sigma^2(x)} dx] dy,$$

which becomes $S(y) = y$ for a martingale diffusion process [8]. Therefore all martingale diffusions taking values on R satisfy the recurrence property.

The martingale process (3.5) can be stationary as well as nonstationary. From [Kutoyants (2013), Th.1.16 and condition (2.1), Karatzas, Shreve (1991), p221], this martingale process is stationary, if $\int_{-\infty}^{+\infty} \frac{dy}{\sigma^2(y)} < \infty$, and is nonstationary, otherwise. If $\sigma(y)$ is equal to $|y|^\gamma$ for large y, the stationarity condition is satisfied for $\gamma > 0.5$. Thus, for a stationary martingale, the "mean reversion" effect is not due to the drift, but rather to the volatility function. To satisfy the stationarity condition, the volatility increases for large values of y, and allows the process to revert to its median. Conley et al. (1997) refer to this effect as the volatility induced mean-reversion, in the context of processes with finite mean.

However, this remark has to be considered with caution. The solution of a diffusion equation exists up to a random time, called the explosion time that is not necessarily in $[0, \infty]$. For a martingale diffusion, the explosion time is a.s. infinite (resp. a.s. finite) if and only if $k(-\infty) = k(+\infty) = \infty$, where:

$$k(y) = \int_0^y [2 \int_0^z \frac{1}{\sigma^2(z)} dz] dy,$$

(resp. otherwise) [see, Revuz, Yor (1991), p.357, ex.2.15]. It follows directly that a martingale diffusion always satisfies the condition for infinite explosion time.

3.4 The Double Autoregressive Process

An analogue of the martingale diffusion process (2.5) written in discrete time is:

$$Y_t = Y_{t-1} + \eta(Y_{t-1}) \epsilon_t, \tag{3.6}$$

[8] We say that this process is in its natural scale.

where (ϵ_t) is a sequence of i.i.d. variables, with mean 0 [9]. This process does not always satisfy the recurrence property. For example, when $\eta(Y_{t-1}) = Y_{t-1}$, we have:

$$|Y_t| = |Y_{t-1}|\,|1 + \epsilon_t|,$$

and, by taking the logarithm on both sides:

$$log|Y_t| = log|Y_{t-1}| + \log|1 + \epsilon_t| = log|Y_{t-1}| + Elog|1 + \epsilon_t| + u_t,$$

where the sequence $(u_t) = (log|1+\epsilon_t| - E(log|1+\epsilon_t|))$ is i.i.d. with mean 0. Thus $log|Y_t|$ is a random walk with drift, which is explosive [10] and non-recurrent [see, e.g. Bandi, Phillips (2009), Example 2, and references therein]. As the recurrence property is invariant with respect to scale transformations, $(|Y_t|)$ and (Y_t) do not satisfy the recurrence property either.

Depending on the pattern of the volatility function, the martingale (3.6) can be either nonstationary, or stationary. If $\eta(Y_{t-1}) = \eta$, than y_t is the random walk of Section 3.2. An example of a stationary martingale follows from the following model examined in Ling (2004):

$$Y_t = Y_{t-1} + \sqrt{c + aY_{t-1}^2}\,\epsilon_t, \quad c \geq 0, a > 0, \tag{3.7}$$

where ϵ_t is i.i.d. with a symmetric distribution and mean zero. This process is stationary if

$$Elog|1 + \epsilon_t\sqrt{a}| < 0,$$

[Borkovec, Kluppelberg (1998), Th. 3.3]. This type of martingale is another process with volatility induced "mean-reversion". This example shows that in the class of simple autoregressive processes, a unit root alone does not imply nonstationarity.

3.5 Noncausal Process

A martingale process (y_t) can be time reversible or time irreversible, depending on whether its dynamics in the calendar and the reversed time are identical or not.

[9] A time discretized diffusion process satisfying (3.5) does not generally satisfy condition (3.6) due to the effect of time aggregation.

[10] to $+\infty$, if $Elog|\epsilon_t| > 0$, to $-\infty$, if $Elog|\epsilon_t| < 0$.

Noncausal processes are processes defined in the reversed time [see Rosenblatt (2000) for noncausal linear processes.]. A Markov noncausal process is also Markov in the calendar time [Cambanis, Fakhre-Zakeri (1994)] [11], and its transition density in the calendar time can be inferred from its transition density in the reversed time. Among the noncausal linear processes, the noncausal autoregressive Cauchy process of order one is of special interest [see, Gourieroux, Zakoian(2015)] [12] and defined as:

$$Y_t = \rho^* Y_{t+1} + \sigma^* \epsilon_t^*, \tag{3.8}$$

where $0 < \rho^* < 1$ and (ϵ_t^*) is a sequence of i.i.d. Cauchy variables.

As the first-order moment of a Cauchy variable does not exist, it follows that both the unconditional first-order moment and the noncausal conditional first-order moment do not exist, that is :

$$E|Y_t| = \infty, \quad E(|Y_t|\,|Y_{t+1}) = \infty.$$

Let us now consider the causal transition density function. The marginal density of (Y_t) is Cauchy with scale $1/(1-\rho^*)$ and the causal conditional density is:

$$f(y_t|y_{t-1}) = \frac{1}{\pi\sigma^*} \frac{1}{1 + (y_{t-1} - \rho^* y_t)^2/\sigma^{*2}} \frac{\sigma^{*2} + (1-\rho^*)^2 y_{t-1}^2}{\sigma^{*2} + (1-\rho^*)^2 y_t^2}.$$

Thus, in the calendar time, the process has conditional moments up to any order p, $p < 4$. The first and second causal conditional moments are:

$$E(Y_t|Y_{t-1}) = Y_{t-1}, \quad E(Y_t^2|Y_{t-1}) = \frac{1}{\rho^*} Y_{t-1}^2 + \frac{\sigma^{*2}}{\rho^*(1-\rho^*)}.$$

Hence, the noncausal autoregressive Cauchy process is a stationary martingale in the calendar time although its first two marginal moments do not exist. The trajectory of the process features recurrent bubbles with a rate of explosion of about $1/\rho^*$.

The noncausal Cauchy AR(1) is Markov and recurrent. In particular, it is recurrent in reversed time and the number of visits in any given interval of the calendar and the reversed time are equal.

[11] see, Revuz, Yor (1991), Section III, paragraph 4, for a similar result for processes in continuous time.

[12] The existence and uniqueness of the strictly stationary solution are proven in Gourieroux, Zakoian (2015).

Let us now consider the associated martingale difference sequence:

$$\Delta Y_t = Y_t - Y_{t-1},$$

with the conditional variance:

$$\begin{aligned} V(\Delta Y_t | Y_{t-1}) &= V(Y_t | Y_{t-1}) \\ &= E(Y_t^2 | Y_{t-1}) - [E(Y_t | Y_{t-1})]^2 \\ &= (\frac{1}{\rho^*} - 1) Y_{t-1}^2 + \frac{\sigma^{*2}}{\rho^*(1-\rho^*)}. \end{aligned}$$

It follows directly that (ΔY_t) cannot satisfy the weak conditional heteroscedasticity (WCH) assumption (see Appendix 1), because the time averaged volatility:

$$\frac{1}{T} \sum_{t=1}^{T} V(\Delta Y_t | Y_{t-1}) = (\frac{1}{\rho^*} - 1) \frac{1}{T} \sum_{t=1}^{T} Y_{t-1}^2 + \frac{\sigma^{*2}}{\rho^*(1-\rho^*)},$$

does not converge when T tends to infinity, as the marginal second-order moment does not exist $E|Y_t|^2 = +\infty$.

3.6 Trajectories

As shown in the previous sections, the class of martingales includes stationary as well as nonstationary processes, with linear or nonlinear dynamics. In order to illustrate the behavior of the processes discussed in the previous sections, we show below the simulated trajectories of the following processes:

a) random walk with i) Gaussian errors $N(0,1)$, ii) t-Student errors with 3 degrees of freedom

b) time-discretized diffusion process with i) $\sigma(y) = \sqrt{1 + |y|^{0.6}}$, ii) $\sigma(y) = \sqrt{1 + |y|}$.

c) noncausal autoregressive Cauchy process with i) $\rho^* = 0.5$, ii) $\rho^* = 0.8$.

[Insert Figure 1 : Trajectories of Random Walks]

[Insert Figure 2 : Trajectories of Time Discretized Diffusions]

[Insert Figure 3 : Trajectories of NonCausal Cauchy AR(1)]

All the simulated trajectories have the same length $T = 200$.

These processes satisfy all the recurrence property although the time to recurr can be large in random walks and time discretized diffusions with volatility induced "mean reversion". Figures 1 and 2 show trajectories that take mostly positive values. The Noncausal AR(1) is recurring quite often as it displays short-lived bubbles and takes value 0 also quite often [13].

The difference between the trajectories of the two types of stationary martingales is in the duration of their departures form the most commonly observed value. In noncausal Cauchy AR(1) processes, the rate of growth of a local trend depends on the persistence of the series, and ends with a sudden drop. In the discretized diffusion processes, the local trends are longer lasting and end due to the volatility-induced reversion to the median.

4 Asymptotic Results

This section discusses the kernel-based estimation of a nonparametric autoregression from recurrent Markov processes. Karlsen, Tjostheim (2001)(see Appendix 2) show that the Nadaraya-Watson (NW) estimator with appropriately chosen kernel and bandwidth leads to consistent estimation of function m and its confidence interval. Although this results is independent of the stationarity or nonstationarity of the process, it depends on the recurrence condition being satisfied by the data.

We show how the asymptotic behavior of the kernel-based density estimator and of the NW autoregression estimator may change, depending on whether the data generating process is stationary or not.

4.1 Nonparametric estimators

(a) Mixing processes

Let us consider kernel estimators of the stationary density $f(y)$, of the autoregression function $m(y)$ and of the local volatility $\eta^2(y)$ for stationary mixing and ergodic processes, defined below:

[13] A noncausal Cauchy with a positive drift can be used to represent positive price processes, as the martingale condition is invariant with respect to the drift.

$$\hat{f}_T(y) = \frac{1}{Th_T} \sum_{t=1}^{T} K\left(\frac{y_t - y}{h_T}\right), \tag{4.1}$$

$$\hat{m}_T(y) = [\sum_{t=1}^{T} y_{t+1} K\left(\frac{y_t - y}{h_T}\right)]/[\sum_{t=1}^{T} K\left(\frac{y_t - y}{h_T}\right)], \tag{4.2}$$

$$\hat{\eta}_T^2(y) = \{\sum_{t=1}^{T} [y_{t+1} - \hat{m}_T(y)]^2 K\left(\frac{y_t - y}{h_T}\right)\}/\{\sum_{t=1}^{T} K\left(\frac{y_t - y}{h_T}\right)\}, \tag{4.3}$$

where K is a kernel such that

$$\int K(y)dy = 1, \int yK(y)dy = 0, \int K^2(y)dy \equiv k_2 > 0, \tag{4.4}$$

and h_T is a bandwith that tends to 0 when the number of observations T tends to infinity. If the bandwith tends to zero at an appropriate rate, both estimators $\hat{f}_T(y)$, $\hat{m}_T(y)$ tend to their theoretical counterparts:

$$\hat{f}_T(y) \to f(y), \text{ and } \hat{m}_T(y) \to m(y) \text{ and } \hat{\eta}^2(y) \to \eta^2(y) \tag{4.5}$$

that are assumed to exist. Their speed of convergence is is equal to $1/\sqrt{Th_T}$ and is smaller than the speed of convergence of the parametric estimators. Their limiting distributions are Gaussian. For example, we have:

$$\sqrt{Th_T}[\hat{m}_T(y) - m(y)] \xrightarrow{d} N[0, \frac{k_2 \eta^2(y)}{f(y)}]. \tag{4.6}$$

(b) Random Walk

As shown in Park, Phillips (1998), these results are modified for a random walk, which has no stationary density. The kernel estimator of the "density" has still some asymptotic convergence properties. More precisely:

$$lim_{T\to\infty} \hat{f}_T(y) = 0, \text{ but } lim_{T\to\infty} a.s. \sqrt{T} \hat{f}_T(y) = L, \tag{4.7}$$

where L is a stochastic limit function of the local time of the Brownian motion and on the assumption on the initial value Y_0. In fact, $\hat{f}_T(y)$ measures the frequency of sejourn in the interval $(y, y + dy)$. This frequency is asymptotically negligeable due to the long range explosions of the process.

The Nadaraya-Watson estimator of the autoregression is still consistent:

$$lim_{T\to\infty}\hat{m}_T(y) = m(y)(= y), \tag{4.8}$$

and, upon an appropriate standardization, we have:

$$\sqrt{Th_t}^4[\hat{m}_T(y) - m(y)] \xrightarrow{d} \sqrt{\frac{k_2\eta^2}{L}}U, \tag{4.9}$$

where U is a standard normal variable, independent of L, and we take into consideration that $\eta^2(y) = \eta^2$ is constant in a random walk. Thus, the limiting distribution of $\hat{m}_T$ is a mixture of Gaussian distributions.

From conditions (4.7) and (4.9), it follows that the asymptotic CI estimator for $m(y) = y$ at 95% is:

$$\left\{\hat{m}_T(y) \pm 1.96\left[\frac{k_2\hat{\eta}_T^2}{Th_T\hat{f}_T(y)}\right]^{0.5}\right\}, \tag{4.10}$$

where $\hat{\eta}_T^2$ is a consistent estimator of η^2. This is the standard CI estimator used in the analysis of stationary and mixing processes (with conditional homoscedasticity).

4.2 The consistency

By following a different approach, Karlsen, Tjostheim (2001) have shown that this result is valid for a large class of Markov processes, that satisfy a recurrence property [14]. Their findings can be summarized as follows:

i) Upon an appropriate standardization, the kernel estimator of the density tends to a nonzero limit, which can be either deterministic (in a stationary mixing process), or stochastic (in a random walk).

ii) The Nadaraya-Watson autoregression estimator is consistent, and, when standardized, it tends in distribution to a variable which is a mixture of Gaussian variables.

iii) The standard estimator of the asymptotic confidence interval (4.9) remains valid,

$$\left\{\hat{m}_T(y) \pm 1.96\left[\frac{k_2\hat{\eta}_T^2(y)}{Th_T\hat{f}_T(y)}\right]^{0.5}\right\}, \tag{4.11}$$

[14] Athreya, Atuncar (1998) were the first authors to use the property of recurrent Markov chain to derive the asymptotic behavior of kernel estimators.

taking into account a possible conditional heteroscedasticity.

The regularity conditions that ensure the validity of the above results are given in Appendix 2 and discussed in Section 5. These include the existence of the conditional first-order moment for the convergence of $\hat{m}_T$, the existence of the conditional second-order moments for the estimated confidence interval, and conditions on the speed of convergence of the bandwith to zero, explained in the next section. Neither the stationarity conditions, nor (strong) mixing conditions are needed.

4.3 The recurrence assumption

Let us consider the recurrence condition for Markov processes given in Section 3. AS the Markov chain with continuous state space can be well approximated by a chain with discrete state space, let us first consider the effect of the recurrence assumption on such a discrete state space chain $\tilde{Y}_t$ such that $P(\tilde{Y}_t = y|\tilde{Y}_{t+1} = x)$ denotes the elementary transition probability with $P(\tilde{Y}_t = x|\tilde{Y}_{t-1} = x) = 0,\ \forall x$. Let $S_{1,y}$ denote the first time of entry in $\{y\}$. Under the recurrence condition, $P(S_{1,y} < \infty|Y_0 = x) = 1, \forall x$, we can construct the sequence $S_{1,y} < S_{2,y} < \cdots S_{k,y} \cdots$ of successive entry times in y. Since the process is Markov with a time homogenous transition, it regenerates itself at each entry time. As a consequence, the variables:

$$[S_{k+1,y} - S_{k,y}, \tilde{Y}_{S_{k,y}+1}, \cdots, \tilde{Y}_{S_{k+1,y}+1}],\ k \geq 1, \tag{4.12}$$

are independent, identically distributed [Nummelin (1978),(1984), p. 76]. In this discrete state space framework, the estimated autoregression is:

$$\hat{m}_T(y) = \sum_{t=1}^{T} (\tilde{Y}_{t+1} 1_{\tilde{Y}_t = y}) / \sum_{t=1}^{T} 1_{\tilde{Y}_t = y}, \tag{4.13}$$

that is the crude discrete state counterpart of the Nadaraya-Watson estimator.
Let $K_T = sup\{k : S_{k,y} \leq T\} - 1$ denote the number of total regenerations over the sampling period. Up to the effect of the first and last observed values, we have:

$$\hat{m}_T(y) \approx \frac{1}{K_T} \sum_{k=1}^{K_T} \tilde{Y}_{S_{k,y}+1}. \tag{4.14}$$

As variables $\tilde{Y}_{S_{k,y}+1}$ are i.i.d., and K_T is large when T tends to infinity, we can apply the standard limit theorems for i.i.d. variables conditional on K_T. Thus, for T large, i.e. K_T large, we have approximately:

$$\begin{aligned}\hat{m}_T(y) &\overset{d}{\sim} N[E(\tilde{Y}_{S_{k,y}+1}),\ \frac{1}{K_T}V(\tilde{Y}_{S_{k,y}+1})] \\ &= N[E[\tilde{Y}_{t+1}|Y_t = y], \frac{1}{K_T}V(\tilde{Y}_{t+1}|Y_t = y)],\end{aligned}$$

or

$$\sqrt{K_T}[\hat{m}_T(y) - m(y)] \approx \eta(y)U, \tag{4.15}$$

where U is a standard normal variable independent of K_T. This result provides the asymptotic confidence interval estimator:

$$[\hat{m}_T(y) \pm 1.96\frac{\hat{\eta}_T(y)}{\sqrt{K_T}}], \tag{4.16}$$

which is valid regardless of the stationarity and the mixing conditions being satisfied.

The confidence interval estimator (4.16) is the analogue of confidence interval estimator (4.11) that includes a kernel because of the continuous state-space. This CI estimator is also equivalent to the one used commonly for stationary processes. The speed of convergence of the estimator $\hat{m}_T$ depends on how frequent are the visits in a vicinity of y. The asymptotic distribution of $\hat{m}_T$, upon a deterministic standardization, depends on the distribution of the standardized number of visits.

This speed of convergence is directly linked to the tail behavior of the distribution of $S_{k+1,y} - S_{k,y}$. Also, an appropriately chosen bandwith has to tend to zero at the same rate to ensure the validity of these asymptotic results(see Appendix 2).

4.4 Illustration

To illustrate the estimation of the autoregression function in a martingale process, we consider below the three types of martingales presented in Section 3.6. We use an Epanechnikov kernel to satisfy the condition of kernel with compact support (see Assumption a in Appendix 2) and a basic bandwidth of order $h_T =$ ct $T^{-0.22}$ (see the discussion on the choice of the bandwidth in Appendix 2). The nonparametric kernel estimates are rather sensitive to the choice of the kernel, especially to the remaining constant ct. For each series, we fix the bandwidth proportionally to the range computed from the difference between the 10th and 90th percentiles of the series, so that there are about $K_T = 10$ observations in the associated equiweighted kernel. This bandwidth is used in the estimation of $f(y)$ and $\eta^2(y)$. That bandwidth is divided by two when it is used for the estimation of $m(y)$ in order to avoid excess smoothing.

The theory suggests the following differences between the results for nonstationary and stationary martingales.

i) The speed of convergence of $\hat{m}(y)$ is higher in stationary processes, as they recurr more frequently.

ii) This effect can be compensated by its components that are eatimators of the conditional variance $\eta^2(y)$ and of the "invariant" density $\pi(y)$, respectively. For the stationary martingales (that are the time discretized diffusion and the noncausal Cauchy AR(1) process, $\eta^2(y)$ tends to infinity when y tends to $\pm\infty$, and this effect is cumulated with the fact that the stationary density $\pi(y) = f(y)$ tends to zero at infinity. This explains the poor accuracy of the functional estimator of $m(y) - y$ in the extremes.

For the Gaussian random walk both the theoretical volatility $\eta^2(y) = \eta^2$ and the theoretical stationary density $\pi(y) = 1$ (the Lebesgue measure is invariant) are constant [15].

[Insert Figure 4 : Estimation of $m(y) - y$; Gaussian Random Walk]

[Insert Figure 5 : Estimation of $m(y) - y$; Time Discretized Diffusion $\sigma^2(y) = 1 + |y|^{0.6}$]

[Insert Figure 6 : Estimation of $m(y) - y$; Noncausal AR Cauchy $\rho^* = 0.8$]

Figures 4 to 6 provide the kernel estimates of $m(y) - y$, along with their asymptotic confidence intervals. We observe that:

i) The set of values from which the nonparametric estimate is computed varies accross the series.

ii) Over the set of most commonly observed values of the series, the value of zero is inside the confidence interval. That means that locally, the test of the martingale will not reject the null hypothesis.

iii) The estimator of $m(y) - y$ is different from zero over the set of extreme values of

[15] The nonparametric bounds are computed locally without taking into account this information.

the series, where the confidence interval widens up. This is due to the lack of uniformity in y that slows down the convergence of $\hat{m}(y)$. This effect is noticeable in the noncausal Cauchy AR(1) process due to infrequently observed large bubbles.

5 Scalar Test Statistics of the Martingale Hypothesis

The weak convergence of the NW estimator derived in Karlsen, Tjostheim (2001) does not imply that the convergence is uniform in y. As noted in Section 4.4, the effect of the lack of uniformity is visible in the Figures, where the accuracy of the estimator in the extremes can be rather poor, especially for the stationary discretized diffusion or for the noncausal autoregressive Cauchy process. This is mainly a consequence of strong conditional heteroscedasticity in the extremes.[16] This explains why standard parametric test statistics can have unexpected asymptotic behavior under the null hypothesis H_0. First, we discuss this effect for the OLS estimator of the autoregressive coefficient. Next, we propose alternative estimators for the autoregressive coefficient assumed to be constant, and not necessarily equal to 1. Next, we introduce scalar test statistics of the martingale hypothesis based on the NW estimator and on its standardized version.

5.1 OLS estimator of the autoregressive coefficient

Let us consider the standard analysis of the unit root hypothesis based on the OLS estimator of the autoregressive coefficient:

$$\hat{\rho}_T = \sum_{t=1}^{T} Y_{t+1} Y_t / \sum_{t=1}^{T} Y_t^2. \tag{5.1}$$

This estimator is close to the quantity $\tilde{\rho}_T = \sum_{t=1}^{T} [\frac{(\hat{m}_T(Y_t)}{Y_t} Y_t^2) / \sum_{t=1}^{T} Y_t^2$, that is a weighted sum of the ratio $\hat{m}_T(y)/y$ evaluated at the observed values. If the process is stationary, we expect $\hat{\rho}_T$ to be close to $\int \frac{m(y)}{y} \pi(y) dy$, where π is the stationary distribution of the process, that is close to $\int 1\pi(y)dy = 1$.

The example of the noncausal Cauchy AR(1) process contradicts that expected result. That process is a martingale and satisfies the regularity conditions for the convergence

[16] This effect can be eliminated from the test of the martingale hypothesis by introducing additional assumptions of either i.i.d. errors, or conditional homoscedasticity, or weak conditional homoscedasticity (see Appendix 1).

of the Nadaraya-Watson estimator, that is $\hat{m}_T(y) \to y$ for any y. Davis, Resnick (1986) have shown that the OLS estimator $\hat{\rho}_T$ converges to the value ρ^*, $0 < \rho^* < 1$, of the autoregressive coefficient in the reversed time representation (4.7), rather than to 1.

The convergence of $\hat{m}_T(y)$ to $m(y)$ is not uniform in y and the standard parametric OLS estimator of the autoregressive coefficient is strongly influenced by large observations.

More precisely, in the noncausal AR Cauchy framework, the asymptotic variance of $\hat{m}_T(y)/y$ is proportional to $\frac{\eta^2(y)}{y^2 f(y)}$, up to a factor that depends on T (and h_T). The OLS estimator is a combination of the ratios $\hat{m}_T(y)/y$ with weights $\alpha(y) \propto y^2 f(y)$. Therefore the "variance" of this OLS estimator is proportional to :

$$\int \alpha^2(y) \frac{\eta^2(y)}{y^2 f(y)} dy = \int y^2 f(y) \eta^2(y) dy.$$

From Section 4.5, we know that $\eta^2(y)$ is of order y^2 and $f(y)$ of order $1/y^4$ at infinity. Therefore $y^2 f(y)\eta^2(y)$ is of order 1, and the "variance" of the OLS estimator does not exist. This explains the result of the convergence of the OLS estimator of ρ reported by Davis, Resnick (1986) [see also Figures 7-8].

5.2 Robust estimation of the autoregressive coefficient

Let us now consider the robust estimation of the autoregressive parameter. For this purpose, we introduce the parametric hypothesis $ARH_0 = \{m(y) = \rho y\}$ nested in the maintained hypothesis H of recurrent Markov process. The three hypotheses are nested as follows :

$$H_0 = \{m(y) = y\} \subset ARH_0 = \{m(y) = \rho y\} \subset H.$$

Next, we look for estimators of ρ, which are consistent for any process in H, stationary as well as nonstationary.

Consistent estimators can be obtained by fixing a grid of values of the state $y_j, j = 1, \ldots, J$, and using the asymptotic distributions :

$$\hat{m}_T(y_j)/y_j \simeq \rho + \left[\frac{k_2 \hat{\eta}_T^2(y_j)}{(y_j)^2 T h_T \hat{f}_T(y_j)}\right]^{1/2} U_j, j = 1, \ldots, J, \tag{5.2}$$

where $U_j, j = 1, \ldots, J$ are standard normal independent variables, and are independent of the $\hat{f}_T(y_j), j = 1, \ldots, J$. Expression (5.2) is a linear model in ρ with conditional heteroscedasticiy. Thus parameter ρ can be estimated by feasible Generalized Least Squares as :

$$\tilde{\rho}_T = \sum_{j=1}^{J} \left[\frac{\hat{m}_T(y_j)\hat{f}_T(y_j)}{\hat{\eta}_T^2(y_j)} \right] / \sum_{j=1}^{J} \left(\frac{y_j^2 \hat{f}_T(y_j)}{\hat{\eta}_T^2(y_j)} \right), \tag{5.3}$$

which corresponds to a weighted average of $\hat{m}_T(y_j)/y_j$, with weights proportional to $\alpha(y_j) = y_j^2 \hat{f}_T(y_j)/\hat{\eta}_T^2(y_j)$. This estimator has a complicated asymptotic distribution, and leads to a straightforward confidence interval estimator:

$$\tilde{\rho}_T \pm 1.96 \left[\sum_{j=1}^{J} \frac{y_j^2 \hat{f}_T(y_j)}{\hat{\eta}_T^2(y_j)} \right]^{0.5}. \tag{5.4}$$

This clarifies why the OLS estimator is the benchmark in the unit root literature. It is conceived for processes such that the variables $y_t - \rho y_{t-1}$ are i.i.d., that is, for the random walk with $\rho = 1$. The weights of the OLS estimator have been selected accordingly and are not appropriate for other types of AR recurrent Markov processes.

Other consistent estimators can also be constructed from a random grid based on observations $y_1, \ldots, y_T$. In the case of stationary martingales, we have to avoid too much weight being assigned to the extremes. This can be accomplished by using more appropriate weighting. For instance, we can consider estimator such as :

$$\rho_T^* = \sum_{t=1}^{T} \hat{m}_T(Y_t) w_T(Y_t), \tag{5.5}$$

where the weighting function $w_T(y) \geq 0, \int_{-\infty}^{+\infty} w_T(y)dy = 1$, decreases to zero when y tends to infinity at an appropriate rate in the spirit of what has been proposed by Collomb, Hardle (1986) to robustify a nonparametric regression. The aim of our paper is not to discuss such weighting. Let us just suggest to use a shrinkaged OLS estimator instead of an OLS estimator :

$$\hat{\rho}_T^s = \sum_{t=1}^{T} (Y_{t+1} Y_t \mathbb{1}_{Y_t \in A}) / (\sum_{t=1}^{T} (Y_t^2 \mathbb{1}_{Y_t \in A})), \tag{5.6}$$

where A is a given finite interval. Such a shrinkage might allow to recover the consistency, but clearly this approach will be subefficient since it has not estimated the right weights as in the feasible GLS approach proposed above. The analysis of the asymptotic property of such an estimator $\hat{\rho}_T^s$, with possibly an interval A_T increasing with T, will demand other tools than the functional Central Limit Theorem from Karlsen, Tjostheim (2001), typically limit theorems for U-statistics [see e.g. Dedecker, Prieur (2007), Elharfaoui, Harel (2008)]. This analysis is out of the scope of the present paper.

To illustrate the properties of the shrinkages OLS estimators, we show in Figures 7-8 below, the OLS estimators computed from observations that fall between the percentiles (100-a) and a. Each OLS estimator is computed from a regression without an intercept in Figure 7 and with an intercept in Figure 8. The three series used in the estimation: the random walk, the discretized diffusion and the noncausal Cauchy AR(1) are of length T=1000.

We observe that for all three series, the shrinkaged OLS estimates are closer to 1 in the regression without intercept. As expected, $\hat{\rho}$ is close to 1 in the random walk, for any value of a. In the discretized diffusion process, the estimates are close to 1 when computed from a sufficiently large range of obsrvation values. The results from the OLS estimation of the noncausal Cauchy AR(1) are twofold. As suggested by the result by Davis, Resnick (1986), the OLS estimator computed from all observations is close to 0.8. More surprisingly, when it is computed from observations close to the median is far from 1. This implies that the lack of uniform convergence of $\hat{m}(y)$ concerns not only the extremes, but also the the observations close to the center of the distribution.

[Insert Figure 7 : Shrinkaged OLS Estimate as a Function of a]

[Insert Figure 8 : Shrinkaged OLS Estimate as a Function of a with Intercept]

We observe that the OLS estimator is unreliable in stationary martingales, and so are any test statistics involving that estimator.

5.3 Scalar test statistics of the Martingale Hypothesis

Let us now introduce scalar test statistics that are suitable for testing the null hypothesis of martingale in nonstationary and stationary processes. A natural idea is to base the test on a measure of difference between the estimated autoregression $\hat{m}_T(y)$ and the identity function y. This is a robust approach involving the pivotal functional statistic :

$$\hat{\xi}_T(y) = [\hat{m}_T(y) - y] / \left[\frac{k_2 \hat{\eta}_T^2(y)}{Th_T \hat{f}_T(y)} \right]^{0.5}. \tag{5.7}$$

We consider scalar test statistics, that are the analogues of quantities of the type $\int_{-\infty}^{+\infty} [\hat{\xi}_T(y)]^2 w_j(y) dy$, where w is a weighting function. For a given grid $y_j, j = 1, \ldots, J$, with associated weights $w_j, j = 1, \ldots, J$, we consider the scalar statistic :

$$\hat{\xi}_T = \sum_{j=1}^{J} w_j [\hat{\xi}_T(y_j)]^2, \tag{5.8}$$

where $w_j > 0, \sum_{j=1}^{J} w_j = 1$.

If J is large and $\sum_{j=1}^{J} w_j^2$ close to zero, we have by the standard Central Limit Theorem :

$$\hat{\xi}_T \sim 1 + [\sum_{j=1}^{J} w_j^2]^{0.5} U, \tag{5.9}$$

where U is standard normal. We reject the null hypothesis of martingale if $\frac{[\hat{\xi}_T - 1]}{[\sum_{j=1}^{J} w_j^2]^{0.5}} > 2$, accept it, otherwise.

By construction, we get an asymptotically similar test, where the asymptotic size of the test (i.e. 5%) does not depend on the type of martingale in the null hypothesis.

To illustrate the application of the above test, we consider the three series introduced in Section 4.4 of length $T = 200$ and compute the test statistics evaluated at the percentiles given in Table 2 below:

Table 2 : Percentile Grid

percentile	30 %	40%	50 %	60 %	70 %
Random Walk	-4.94	-3.72	1.10	5.13	7.12
Discr. Diffusion	-0.20	0.49	1.27	2.02	3.58
Noncausal AR(1)	-2.35	-1.10	-0.10	2.59	4.77

From each series, we caclulate:

$$\hat{\xi}_3 = \hat{\xi}_T(40\%)^2 + \hat{\xi}_T(50\%)^2 + \hat{\xi}_T(60\%)^2$$

$$\hat{\xi}_5 = \hat{\xi}_T(30\%)^2 + \hat{\xi}_T(40\%)^2 + \hat{\xi}_T(50\%)^2 + \hat{\xi}_T(60\%)^2 + \hat{\xi}_T(70\%)^2.$$

These statistics take the following values:

Table 3 : Chi-Square Test Statistics

	Random Walk	Discr. Diffusion	Noncausal AR(1)
$\hat{\xi}_3$	0.461	0.690	0.021
$\hat{\xi}_5$	0.896	1.506	0.137

These values need to be compared with the critical values of $\chi^2(3)$ and $\chi^2(5)$ distributions at $\alpha = 0.95$, which are 7.81 and 11.07, respectively.
In order to examine the finite sample behavior of the test statistics, we simulate 100 trajectories of the noncausal Cauchy AR(1), Gaussian Random Walk and the diffusion process from Section 4.4. of lenght T=200 and T=400. In the test statistics computed from these three processes, we use the bandwith equal to the interpercentile range between the 10th and 90th precentiles divided by either 2 or 5.

The choice of the bandwith has a strong impact on the finite-sample distribution of the test statistic. Table 4 below and Figure 9 illustrate that problem.

Table 4 : Critical values of test statistics

		T=200		**T=400**	
		/5 (34)	/2 (80)	/5 (68)	/2 (160)
Cauchy AR(1)	ν=3	0.57	0.95	3.17	7.28
Cauchy AR(1)	ν=5	1.58	3.40	4.35	10.63
R.Walk	ν=3	2.12	3.88	4.04	8.53
R. Walk	ν=5	2.85	4.37	4.75	12.25
Diffus.	ν=3	1.35	1.74	2.41	5.65
Diffus.	ν=5	2.04	3.40	3.52	8.44

We observe that the quantiles of the test statistics at 95%differ from the quantiles of the limiting $chi(3)$ and $chi(5)$ distributions. Also, the critical values of the test statistics vary accross the processes.

Figure 9 below compares the densities of various test statistics.

[Insert Figure 9: Finite sample density of test statistics T=400]

Figure 9 provides the densities of $\hat{\xi}(3)$ and $\hat{\xi}(5)$. They are computed from the noncausal AR(1) and random walk processes with bandwith equal to the interpercentile range between the 10th and 90th precentiles divided by 5 and from the diffusion process with bandwith divided by 2. We observe that the densities of the test statistics depend on the process. Also, as suggested in Table 4, their quantiles do not overlap with the quantiles of the $chi(3)$ and $chi(5)$ distributions. A similar problem is encountered in the test Proposed in Gao, King, Lu, Tjostheim (2004). We illustrate that problem with their test procedure in Technical Appendix 2 at www.yorku.ca/ jasiakj.

6 Concluding Remark

The martingale processes can display trends that are long-lasting or short, such as the local trends due to "mean reversion" and bubbles. As various types of trends can be associated with either nonstationary or stationary martingales, it is important to develop tests statistics that are applicable to both types of these processes. We found that the traditional testing procedures based on global test statistics are unable to handle both types of these processes, while the local analysis provides promissing results. We proposed a kernel based test statistic for testing the martingale hypothesis and applied that test to various martingale processes. The questions for further research are i) how to solve the problem of the lack of uniformity in y of the kernel estimator in the extremes as well as in the centre of the distribution that is relevant for the noncaulal Cauchy AR(1) and ii) what are the properties of the test statistics calculated from a function of percentiles along the path of the process instead of being calculated at fixed points.

References

[1] Athreya, K., and G., Atuncar (1998): "Kernel Estimation for Real Valued Markov Chains", Sankhya, 60, 1-17.

[2] Bandi, F., and P., Phillips (2009) : "Nonstationary Continuous-Time Processes", in Handbook of Financial Econometrics, Y., Ait Sahalia, and L., Hansen eds., 140-199, Elsevier.

[3] Bhargava, A. (1986): "On the Theory of Testing for Unit Root in Observed Time Series", Review of Economic Studies, 53, 369-384.

[4] Borkovec, M., and C., Kluppelberg (1998): "The Tail of the Stationary Distribution of an Autoregressive Process with ARCH(1) Errors", Annals of Applied Probability, 11, 1220-1241.

[5] Box, G., and D., Pierce (1970): " Distribution of Residual Autocorrelation in Autoregressive Integrated Moving Average Time Series", JASA, 65, 1509-1527.

[6] Cambanis S., and I., Fakhre-Zakeri (1994): "Prediction of Heavy Tailed Autoregressive Sequences: Forward versus Reversed Time", Theory of Probability and its Applications, 39, 217-233.

[7] Chan, K., and H., Tong (1985) : "On the Use of the Deterministic Lyapunov Function for the Ergodicity of Stochastic Difference Equations", Advances in Applied Probability", 17, 666-678.

[8] Chan, N., and C., Wei (1988) : "Limiting Distribution of Least Squares Estimates of Unstable Autoregressive Processes", Annals of Statistics, 16, 367-401

[9] Chen, W., and R., Deo (2006): "The Variance Ratio Statistic at Large Horizon", Econometric Theory, 23, 206-234.

[10] Collomb, G., and W., Hardle (1986): "Strong Uniform Convergence Rates in Robust Nonparametric Time Series Analysis and Prediction: Kernel Regression Estimation from Dependent Observations", Stochastic Processes and Their Applications, 23, 77-89.

[11] Conley, T., Hansen, L., Luttmer, E., and J., Scheinkman (1997): "Short Term Interest Rates as Subordinated Diffusions", Review of Financial Studies, 10, 525-577.

[12] Davis, R., and S., Resnick (1986): "Limit Theory for the Sample Covariance and Correlation Functions of Moving Averages", The Annals of Statistics, 14, 533-558.

[13] Dedecker, J., and C., Prieur (2007) : "An Empirical Central Limit Theorem for Dependent Sequences", Stochastic Processes and Their Applications, 117, 121-142.

[14] Delattre, S., Hoffmann, M., and M., Kessler (2002) : "Dynamic Adaptive Estimation of a Scalar Diffusion", DP 762, LPMA, Univ. Paris 6.

[15] Deo, R. (2000): "Spectral Tests of the Martingale Hypothesis under Conditional Heteroscedasticity", Journal of Econometrics, 99, 291-315.

[16] Dickey, D., and W., Fuller (1979): "Distribution of the Estimators for Autoregressive Time Series with Unit Root", Econometrica, 49, 1057-1072.

[17] Durlauf, S. (1991): "Spectral Based Testing of the Martingale Hypothesis", Journal of Econometrics, 50, 355-376.

[18] Elharfaoui, E., and M., Harel (2008) : "Central Limit Theorem of the Smoothed Empirical Distribution Functions for Asymptotically Stationary Absolutely Regular Stochastic Processes", Journal of Applied Mathematics and Stochastic Analysis.

[19] Escanciano, J., and I., Lobato (2009a): "An Automatic Portmanteau Test for Serial Correlations", Journal of Econometrics, 151, 140-169.

[20] Escanciano, J., and I., Lobato (2009b): "Testing the Martingale Hypothesis", in Pelgrave Handbook of Econometrics, eds K., Patterson and T., Mills, New York, Palgrave MacMillan, 972-1003.

[21] Escanciano, J., and C., Velasio (2006): "Generalized Spectral Tests for the Martingale Difference Hypothesis", Journal of Econometrics, 134, 151-185.

[22] Fama, E. (1965): " The Behavior of Stock Market Prices", Journal of Business, 38, 34-105.

[23] Fama, E. (1970): ”Efficient Capital Markets: A Review of Theory and Empirical Work”, Journal of Finance, 25, 383-417.

[24] Fama, E. (1991): ”Efficient Capital Markets II”, Journal of Finance, 46, 1575-1618.

[25] Gao, J., and M., King (2012): ”An Improved Nonparametric Unit Root Test”, DP Monash University.

[26] Gao, J., King, M., Lu, D., and D., Tjostheim (2009): ”Specification Testing in Nonlinear and Nonstationary Time Series Autoregression”, Annals of Statistics, 37, 3893-3928.

[27] Gourieroux, C., and A., Hencic (2014): ”Noncausal Autoregressive Model in Application to Bitcoin/USD Exchange Rate”, in ”Econometrics of Risk”, Series: Studies in Computational Intelligence, Springer

[28] Gourieroux, C., and J., Jasiak (2014): ”Filtering, Prediction and Simulation Methods for Noncausal Processes” York University D.P.

[29] Gourieroux, C., and J.M., Zakoian (2015): ” Explosive Bubble Modelling by Noncausal Processes”, forthcoming JRSS.

[30] Guay, A., and E., Guerre (2006): ” A Data-Driven Nonparametric Specification Test for Dynamic Regression Models”, Econometric Theory, 22, 543-586.

[31] Guerre, E. (2004) : ”Design Adaptive Pointwise Nonparametric Regression Estimation for Recurrent Markov Time Series”, LSTA D.P., Univ. Paris 6.

[32] Hall, G. (1978): ”Stochastic Implications of the Life Cycle Permament Income Hypothesis: Theory and Evidence”, Journal of Political Economy, 86, 971-987.

[33] Hamilton, J. (1994): ”Time Series Analysis”, Princeton University Press, Princeton.

[34] Hong, Y. (1999): ”Hypothesis Testing in Time Series via the Empirical Characteristic Function: A Generalized Spectral Density Approach”, JASA, 94, 1201-1220.

[35] Hong, Y., and T., Lee (2005): ”Generalized Spectral Test for Conditional Mean Model in Time Series with Conditional Heteroscedasticity of Unknown Form”, Review of Economic Studies”, 72, 499-541.

[36] Kallianpur, G., and H., Robbins (1954): "The Sequence of Sums of Independent Random Variables", Duke Math, 21, 285-307.

[37] Karatzas, I., and S., Shreve (1991) : "Brownian Motion and Stochastic Calculus", Springer Verlag, New-York.

[38] Karlsen, M., and D., Tjostheim (2001): "Nonparametric Estimation in Null Recurrent Time Series", Annals of Statistics, 29, 372-416.

[39] Khasminskii, R. (1980):"Stochastic Stability of Differential Equations", Second edition (2011), Stochastic Modelling and Applied Probability, 66, Springer.

[40] Kuan, C., and W., Lee (2004): "A New Test of the Martingale Difference Hypothesis", Studies in Nonlinear Dynamics and Econometrics, 8, 1-24.

[41] Kutoyants, Y. (2013): "Statistical Inference for Ergodic Processes", Springer.

[42] LeRoy, S. (1989): "Efficient Capital Markets and Martingales", Journal of Economic Literature, 27, 1583-1621.

[43] Ling, S. (2004): "Estimation and Testing Stationarity for Double Autoregressive Models", JRSS B, 66, 63-78.

[44] Ljung, G., and G., Box (1978): "On a Measure of a Lack of Fit in Time Series Models", Biometrika, 65, 297-303

[45] Lo, A., and A., McKinlay (1988) : "Stock Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test", Review of Financial Studies, 1, 41-66.

[46] Lobato, J., Nankervis, J., and N., Savin (2001): "Testing for Zero Autocorrelations in the Presence of Statistical Dependence", Econometric Theory, 18, 730-743.

[47] Meyn, S., and R., Tweedie (1993): "Markov Chains and Stochastic Stability", Springer Verlag.

[48] Myklebust, T., Karlsen, H., and D., Tjostheim (2012): "Null Recurrent Unit Root Processes", Econometric Theory, 28, 1-41.

[49] Nummelin, E. (1978): " A Splitting Technique for Harris Recurrent Markov Chain", Zeitschrift fur Wahr. und Verwandte Gebiete, 43, 309-318.

[50] Nummelin, E. (1984): "General Irreducible Markov Chains and Nonnegative Operators", Cambridge University Press.

[51] Park, J., and P., Phillips (1998): "Nonstationary Density Estimation and Kernel Autoregression", Cowles Foundation DP 1181, Yale University.

[52] Park, J., and Y., Whang (2005): "Testing for the Martingale Hypothesis", Studies in Nonlinear Dynamics and Econometrics, 9, 1-30 .

[53] Phillips, P. (1987a): "Towards a Unified Asymptotic Theory for Autoregression", Biometrika, 74, 535-547.

[54] Phillips, P. (1987b): "Time Series Regression with a Unit Root", Econometrica, 55, 277-302.

[55] Phillips, P., and S., Jin (2014): "Testing the Martingale Hypothesis", DP Cowles Foundation 1443.

[56] Phillips, P., and P., Perron (1988): "Testing for a Unit Root in a Time Series Regression", Biometrika, 75, 335-345.

[57] Revuz, D., and M., Yor (1991): "Continuous Martingales and Brownian Motion", Springer Verlag.

[58] Rosenblatt, M. (2000): "Gaussian and Non-Gaussian Time Series and Random Fields", Springer Verlag, New York.

[59] Stock, J. (1994): "Unit Root, Structural Breaks and Trends", in Handbook of Econometrics, Vol 4, Chapter 46, 2749-2831, Elsevier, Amsterdam.

[60] White, J. (1958): "The Limiting Distribution of the Serial Correlation Coefficient in the Explosive Case", Annals of Math-Statistics, 29, 1188-1197.

[61] White, J. (1959): "The Limiting Distribution of the Serial Correlation Coefficient in the Explosive Case II", Annals of Math-Statistics, 30, 831-834.

APPENDIX 1

A Survey of the Literature on the Tests of the Martingale Hypothesis

This Appendix provides a survey of the literature on the tests of the martingale hypothesis and describes the null implicit and alternative hypotheses, including the possibly restrictive regularity conditions, such as the existence of moments, the independence and mixing conditions additionally introduced to derive the asymptotic distributions.

In Table 1, we focus on the tests where the process is a martingale under the null. Hence, we are not interested in tests where under the null there exists an additional effect such as a deterministic trend, other autoregressive effects, effects of other exogenous variables, seasonalities, or structural breaks. This extensive literature has been developed rather for macroeconomic applications than for Finance, and is typically the basis of the cointegration theory. These extensions are not suitable for testing the form of market efficiency, we're interested in. We refer to Stock (1994), chapter 3, Hamilton (1994) chapter 17 for surveys on such extensions.

Table 1 shows three types of procedures:

1) the so-called tests of the difference of martingale hypothesis which check white noise properties of the process $\Delta y_t = y_t - y_{t-1}$, and lead to portmanteau-type tests. They use implicitely as information set $\underline{\Delta y_{t-1}}$ instead of $\underline{y_{t-1}}$.

2) the test of the value of the autoregressive coefficient, often called test of unit root. They are based on parametric methods and usually require strong additional assumptions to derive the asymptotic distribution of the parametric estimator.

3) the test of the pattern of the autoregressive function, not assumed to be linear a priori.

Table 1: The Tests of the Martingale Hypothesis in the Literature

source	**the general hypothesis**		**the null hypothesis**		**stationarity**		**parametric (P)** or nonparametric (NP)
	model	error term	type of martingale	restriction	under H_0	under H_A	
Box, Pierce (70) Ljung, Box (78)	$y_t = y_{t-1} + u_t$	u_t strictly stationary under H_0 $Eu_t^4 < \infty$	not only martingale	$\gamma(h) = 0$	NS	NS	NP (Pm)
White (58, 59) Dickey, Fuller (79)	$y_t = \rho y_{t-1} + u_t$	$u_t\ IIN(0, \sigma^2)$	Gaussian R.W.	$\rho = 1$	NS	S/NS	P(OLS)
Lo, Mc Kinlay (88)	$y_t = y_{t-1} + u_t$	u_t strictly stationary $E\|u_t\|^4 < \infty$	RW	$\gamma(h) = 0$	NS	NS	NP (VR)
Phillips (87) Phillips, Perron (88)	$y_t = \rho y_{t-1} + u_t$	$\sup_t E\|u_t\|^\gamma < \infty$ with $\gamma > 2$ (u_t) strong mixing		$\rho = 1$	NS	S/NS	P(OLS)
Chan, Wei (88)	$y_t = \rho y_{t-1} + u_t$	(u_t) MDS, $E_{t-1}(u_t^2) = 1$ $\sup_t E_{t-1}(u_t^\gamma) < \infty, \gamma > 2$		$\rho = 1$	NS	NS	P(OLS)
Durlauf (91) Deo (00)	$y_t = y_{t-1} + u_t$	WCH, WCC $E\|u_t^8\| < 0$		$\gamma(h) = 0$ $\forall h \geq 1$	NS	NS	NP (SP)
Park, Phillips (98) Gao, King, Lu Tjostheim (09)	$y_t = m(y_{t-1}) + u_t$	(u_t) iid $(0, \sigma^2)$ $Eu_t^4 < \infty$	RW	$m(y) = y$	NS	NS	NP (localized Pm)
Park, Whang (05) Phillips, Jin (14)	$y_t = m(y_{t-1}) + u_t$	WCH		$m(y) = y$	NS	NS	NP
Gao, King (12)	$y_t = m(y_{t-1}) + u_t$	WCH		$m(y) = y$	NS	NS	
Hong (99) Escanciano, Velasio (06)	$y_t = y_{t-1} + u_t$	u_t strictly stationary $E(u_t)^{2+\delta} < \infty$		$E[\Delta y_t \| \Delta y_{t-1}] = 0$	NS	NS	NP (MSE and KS based on estimated $E(\Delta y_t 1_{y_{t-1} < y})$.
Hong, Lee (05)							(SP)
Guay, Guerre (06)	$y_t = y_{t-1} + u_t$	u_t strictly stationary		$E[\Delta y_t \| \Delta y_{t-1}] = 0$	NS	NS	NP (NW)

NW (Nadaraya, Watson), VR (Variance Ratio), SP (Spectral Density), Pm (Portmanteau Statistic), WCH : (Weakly Conditional Heteroscedastic), WCC (Weakly Conditionally Correlated)

A number of testing methods assume that the changes in the process $\Delta y_t = u_t$ satisfy an assumption of weak conditional heteroscedasticity (denoted WCH). A typical example is the assumption that:

$$E[u_t|\underline{y_{t-1}}] = 0, \quad E[u_t^2|\underline{y_{t-1}}] < \infty,$$

and $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E[u_t^2|\underline{y_{t-1}}] = \sigma^2 \ \ (a.1)$.

Similarly, we encounter in the literature a condition of weak conditional correlation (WCC), which is the analogue of (a.1) for correlations. Typically it is written as:

$$lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} u_{t-r}u_{t-s}E(u_t^2|\underline{u}_{t-1}) = \sigma^4\gamma(r,s). \quad (a.2)$$

This assumption is in particular implied by the conditional homoscedasticity

$E[u_t^2|\underline{y_{t-1}}] = \sigma^2, \forall t$, and by the condition of u_t i.i.d. $(0,\sigma^2)$.

Let us now explain why such additional assumptions are not innocuous, and in particular lead to focus on nonstationary martingales.

Under such a WCH condition, we have:

$$\begin{aligned} V(y_{t+H}|\underline{y_t}) &= V(y_{t+H} - y_t|\underline{y_t}) = V(\sum_{h=1}^{H} u_{t+h}|\underline{y_t}) \\ &= \sum_{h=1}^{H} E[E(u_{t+h}^2|\underline{y_{t+h-1}})|\underline{y_t}] \\ &= E[E(\sum_{h=1}^{H} u_{t+h}^2|\underline{y_{t+h-1}})|\underline{y_t}]. \end{aligned}$$

Therefore for H large, we get:

$$V(y_{t+H}|\underline{y_t}) \sim E(H\sigma^2|y_0) = H\sigma^2,$$

and the process (y_t) is nonstationary in variance with the same evolution of the variance as for a random walk.

APPENDIX 2

The Functional Central Limit Theorem for Kernel Estimators

Let us briefly describe the required assumptions for derivation of a functional central limit theorem (FCLT) for the process $[\hat{f}_T(y), \hat{m}_T(y)]$, where $\hat{f}_T$ (resp. $\hat{m}_T$) is the standard kernel estimator of the "stationary" density (resp. the Nadaraya-Watson estimator of the autoregression function). This functional limit theorem has been derived in Karlsen, Tjostheim (2001) [see also Guerre (2004), Bandi, Phillips (2008)]. It relies on the assumption of recurrence of the Markov process (y_t), that follows from the idea of Athreya, Atuncar (1998) that appeared in the literature on kernel estimation.

In this Section, we use an assumption of positive Harris recurrence associated with the Lebesgue measure [see Meyn, Tweedic (1993) for additional discussion on the notion of recurrence].

Assumption a.1 : The process (Y_t) is a Markov process and is positive λ-Harris recurrent, that is :

$$P_x[S_{(a.b)} < \infty] = 1, \forall x, \lambda.a.e., \forall a < b.$$

The advantage of the recurrence property is that it is satisfied by stationary Markov processes, as well as by a large class of nonstationary Markov processes.

Under Assumption a.1, the process will visit any interval (a, b) infinitely many times. We denote by $K_T(a, b)$ the number of visits in (a,b) included between 0 and T. Moreover the process admits an invariant measure π, say. This measure is defined up to a scale. If the invariant measure is finite, then $\pi^* = \pi/\pi(-\infty, +\infty)$ is a probability distribution, the process is stationary and π^* is its marginal distribution. Otherwise, π is $\sigma-$ finite, but not finite. This problem arises in nonstationary process, for which no time independent marginal distribution exist. The numbers of visits in two given intervals (a,b) and (c,d) are closely related as follows:

$$\lim as_{T\to\infty} \frac{K_T(a, b)}{K_T(c, d)} = \frac{\pi(a, b)}{\pi(c, d)}.$$

Intuitively, the rates of divergence of these numbers of visits are of the same order. The assumption below concerns that rate of divergence.

Assumption a.2 : $K_T(a,b)$ behaves approximately as T^β, with $\beta \in (0,1)$. More precisely, $T^{\beta-\varepsilon} \ll K_T(.,b) \ll T^{\beta+\varepsilon}$ a.s., for all ε, where $a_T \ll b_T$ means $a_T = o(b_T)$.

Under Assumptions a.1. and a.2., it is possible to derive a functional limit theorem for the number of visits. The notation $\overset{d}{\Rightarrow}$ means the weak convergence in the space of cadlag functions defined on $(0,\infty), \mathcal{D}(0,\infty)$, say.

Let us consider the process of standardized numbers of visits :

$\tilde{K}_T(a,b) = \{K_{[zT]}(a,b)/T^\beta, z \in (0,\infty)\}$, where $[.]$ denotes the integer part. This is a process indexed by $z \in (0,\infty)$. Then we have [see Karlsen, Tjustheim (2001), Th 3.2 and Lemma 3.6] :

Functional Limit Theorem for $\tilde{K}_T$.

Under Assumptions a.1-a.2, we have :

$$\tilde{K}_T(a,b)/\pi(a,b) \overset{d}{\Rightarrow} M_\beta,$$

where M_β is the Mittag-Leffler process with parameter β, that is the inverse $M_\beta = S_\beta^{-1}$ of the one-sided stable Levy process with marginal characteristic function : $E[\exp ivS_\beta(z)] = \exp(iv^\beta z)$.

In order to derive a FCLT that is suitable for kernel estimation, the following three steps need to be completed:

i) Check that the FLT for $\tilde{K}_T$ remains valid for a standardized local number of visits in a neighbourhood of y, that corresponds to the behaviour of the kernel estimator of the invariant density. That conditio will be satisfied under a set of assumptions concerning the kernel and the bandwidth. As before, the limiting process M_β will be the same, i.e. independent of the interval and/or of the state chosen for local analysis.

ii) Check that there is independence between these numbers of global and local visits and the average evolution of the process between the regeneration points. This independence property is expected to hold for stationary processes, where $\beta = 1$ and the limit is deterministic, and therefore independent of anything else.

iii) It would be nice to have a FCLT not only with respect to the "proportion" z of observations which is considered, but also with respect to the state y of localisation.

Below, we adapt Theorem 4.1 in KT. (2001) to our kernel estimation framework and write workable conditions to get the result.[17]

The main additional assumptions are as follows:

Assumption a.3 : There exists a continuous version of the conditional mean $m(y) = E(Y_t|Y_{t-1} = y)$ and of the conditional variance $\eta^2(y) = V(Y_t[Y_{t-1} = y)$.

In particular, functions m and η^2 are integrable on any bounded interval with respect to the invariant measure. However, in the case of stationary martingale where $E(Y_t) = +\infty$, they are not integrable on (semi) infinite intervals.

Assumption a.4 : The kernel K is positive, with $\int K(u)du = 1, \int uK(u)du = u$, $k_2 \equiv \int K^2(u)du < \infty$. The support of the kernel is contained in a compact interval.

Assumption a.5 : The bandwidth is such that $T^{\beta/2-\varepsilon} \gg \frac{1}{h_T} \gg T^{\beta/5+\varepsilon}$.

This condition cannot be satisfied for all β values in $(0, 1)$ and as noted in [Karlsen, Tjostheim (2001), Lemma 3.4 and 3.7], it might be necessary to estimate β, for instance by $\hat{\beta} = \log K_T(a, b)/\log T$, for a given $(a.b)$. However such an estimator should not be very accurate.

We follow another approach that is focused on stationary martingales ($\beta = 1$), and on nonstationary martingales with "explosive patterns" that are small or equal to those of a Gaussian random walk ($\beta = 0.5$).[18]. Then we find that:
$1/h_T = cT^{0.22}$, where c is a constant, satisfies Assumption a.5 for all $\beta \in (0.5, 1)$.

Note that the condition of β null-recurrence can be circumvented by introducing a data driven adaptive bandwidth [see e.g. Delattre, Hoffmann, Kessler (2002), Guerre (2004), Guay, Guerre (2006)].

---

[17]We only introduce the main regularity conditions.

[18]This is a consequence of the reflection principle for Brownian motion [see e.g. Revuz, Yor (1991), p100, Prop.3.7]. Loosely speaking let us consider a Brownian motion (B_t), and denote by L_t the local time. Then the distribution of (B_t) and L_t are the same. In particular the p.d.f. of $Z_t = L_t/T^{0.5}$ is $\frac{2}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right)$. This provides the value $\beta = 0.5$.

The result below follows from Theorem 4.1 in Karlsen, Tjostheim (2001) (see also Corollary 4.2 and Th. 5.4).

Asymptotic behaviour of $\hat{f}_T(y), \hat{m}_T(y)$

Let us consider J values $y_1 < \ldots < y_J$. Then, under Assumptions a.1-a.5, we have the following weak convergences :
$T^{1-\beta} \hat{f}_T(y_j)/\pi(y_j) \xrightarrow{d} M_\beta(1), \forall j = 1, \ldots, J$, (with an appropriately standardized invariant density),

$$\left([Th_T \hat{f}_T(y_j)]^{1/2}[\hat{m}_T(y_j) - m(y_j)]\right) \stackrel{d}{\Rightarrow} [k_2 diag(\eta^2(y_j))]^{1/2} U,$$

where U is a standard J-dimensional Gaussian vector independent of $M_\beta(1)$.

This result is not a FCLT with respect to state y. However this multidimensional version is sufficient in practice to develop interpretable scalar statistics for testing the martingale hypothesis.

The property above highlights the role of the quantities $\hat{\xi}_T(y) = [Th_T \hat{f}_T(y)/\eta_T^2(y)]^{1/2}(\hat{m}_T(y) - y)$ as statistics with a fixed distribution under the martingale hypothesis. This distribution is also independent of β.

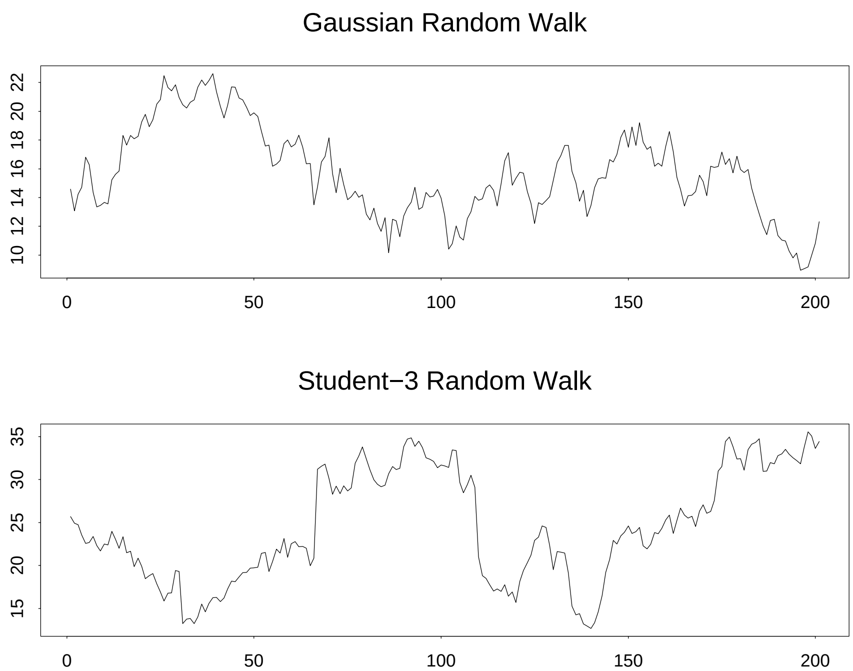

Figure 1: Random Walk

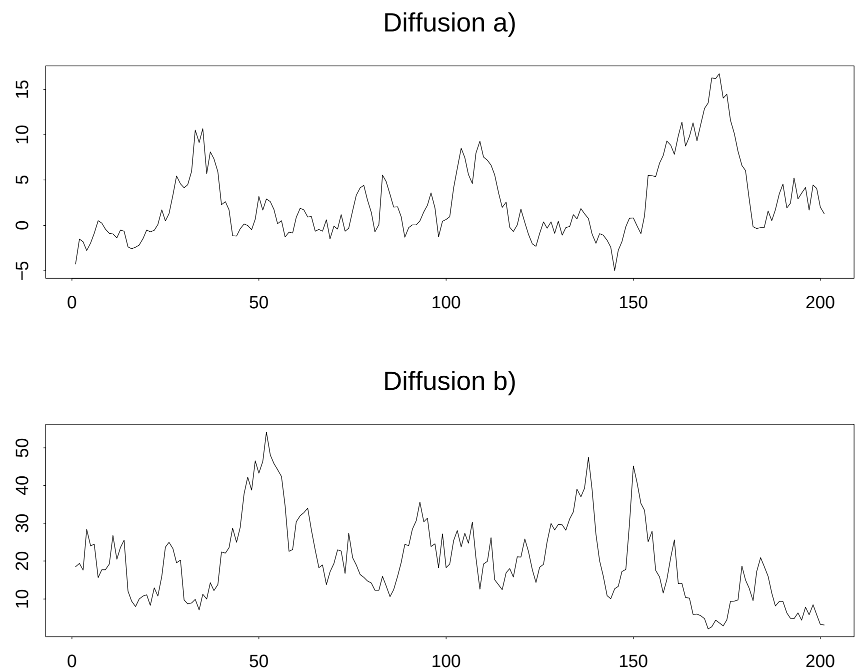

Figure 2: Time Discretized Diffusion

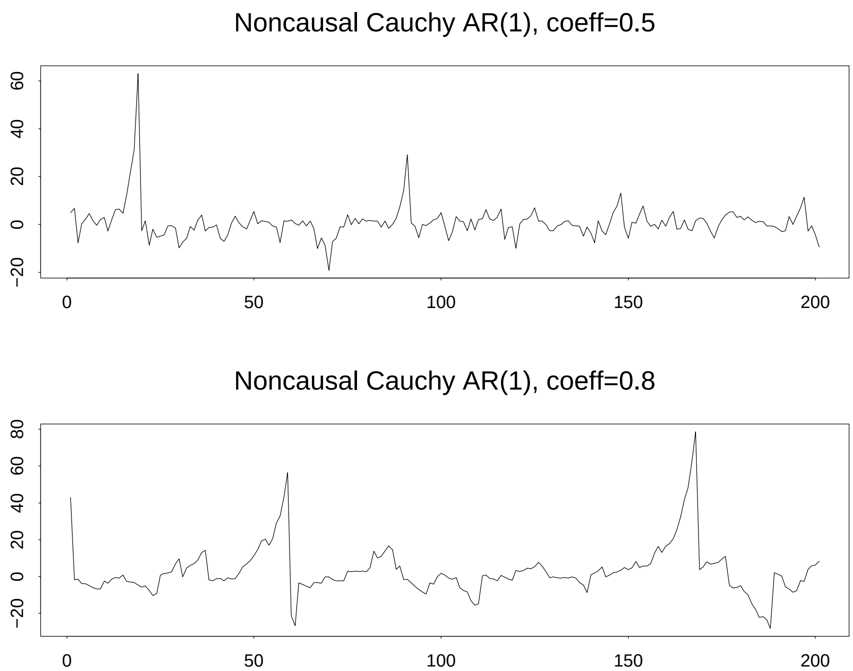

Figure 3: Noncausal Cauchy AR(1)

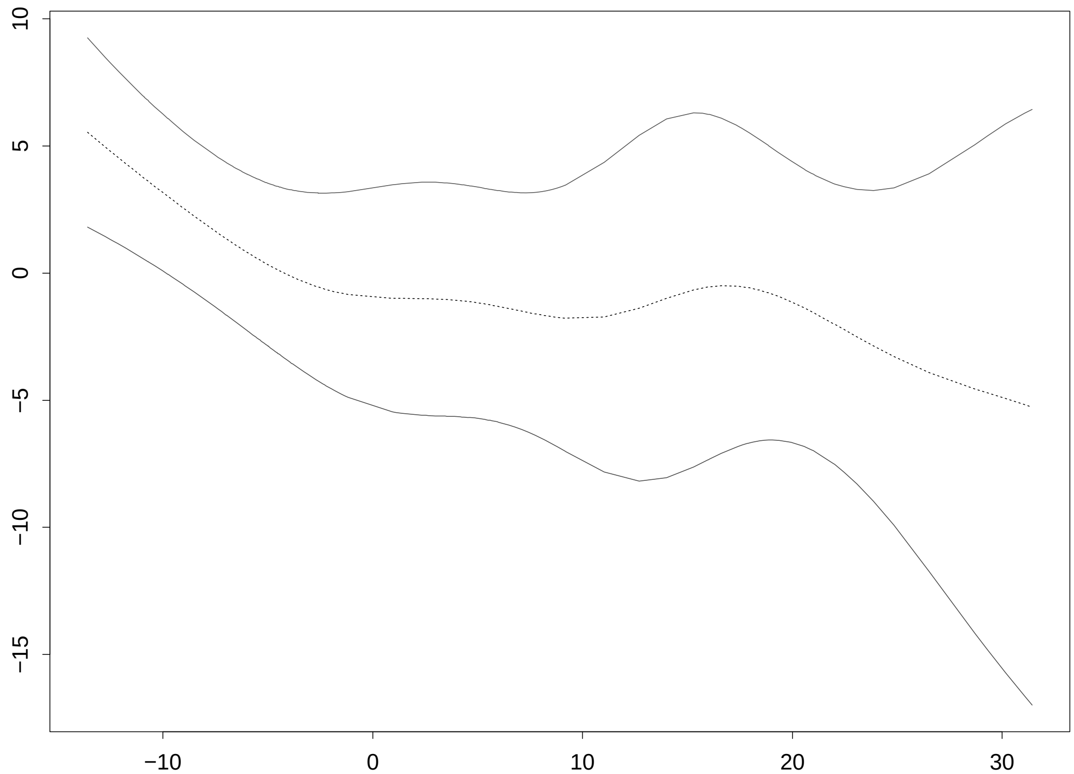

Figure 4: Estimated $m(y) - y$, Gaussian Random Walk

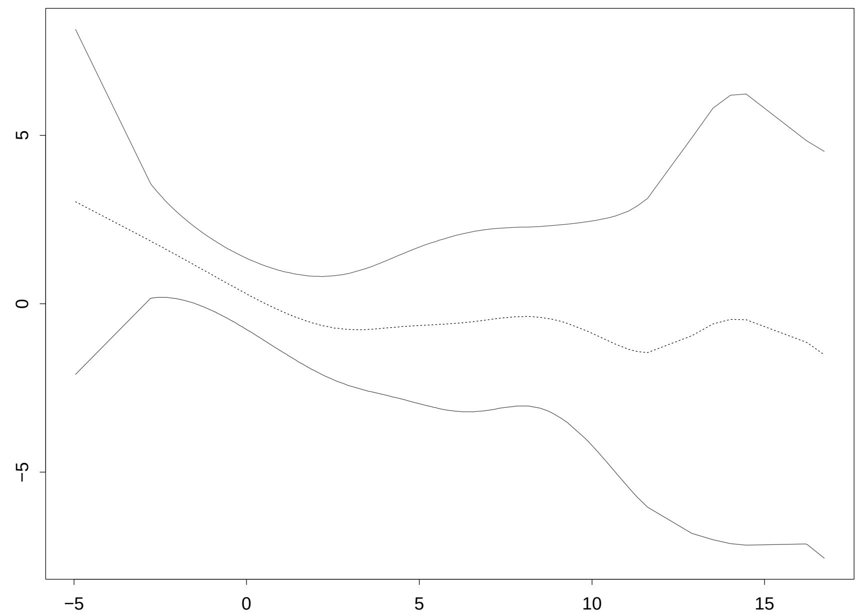

Figure 5: Estimated $m(y) - y$, Time Discretized Diffusion

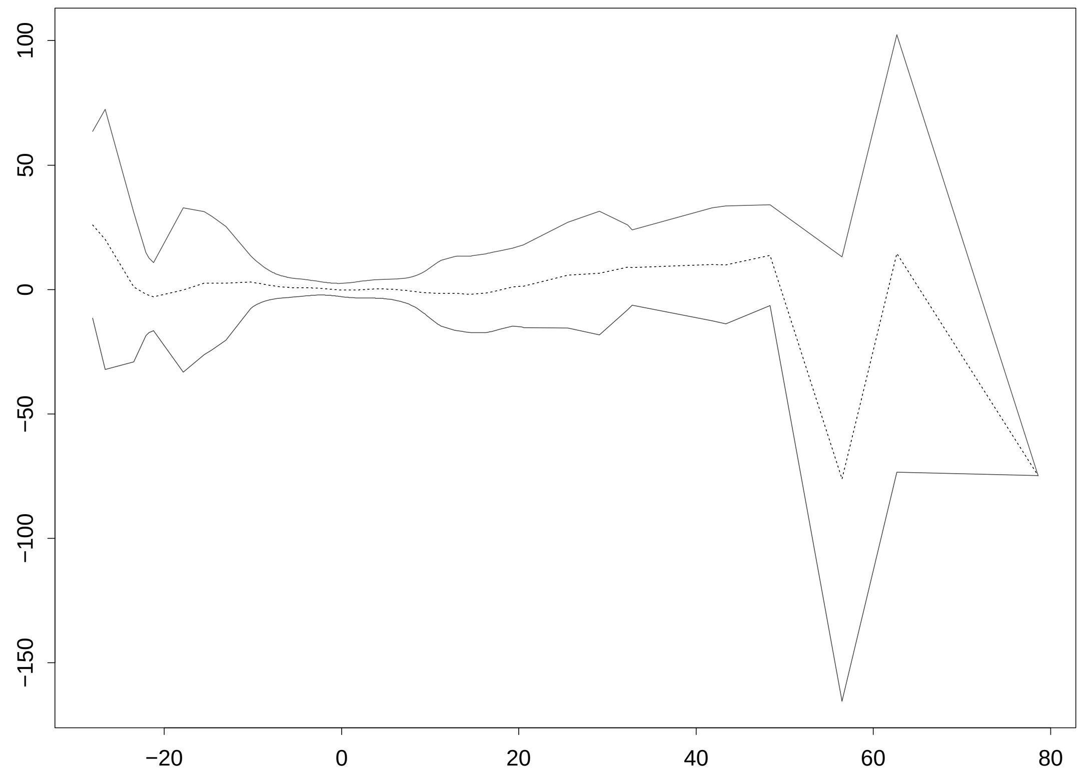

Figure 6: Estimated $m(y) - y$, Noncausal Cauchy AR(1)

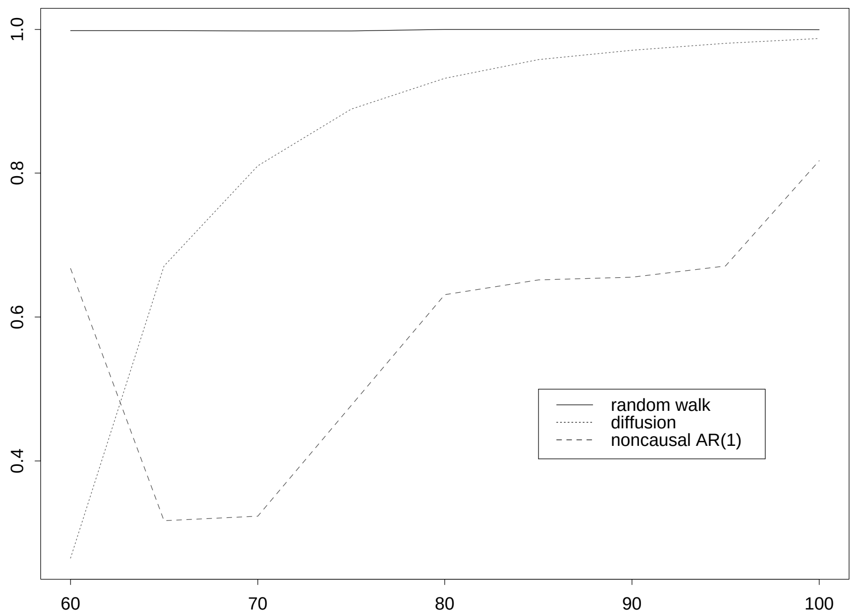

Figure 7: Shrinkaged OLS EStimate as a Function of a

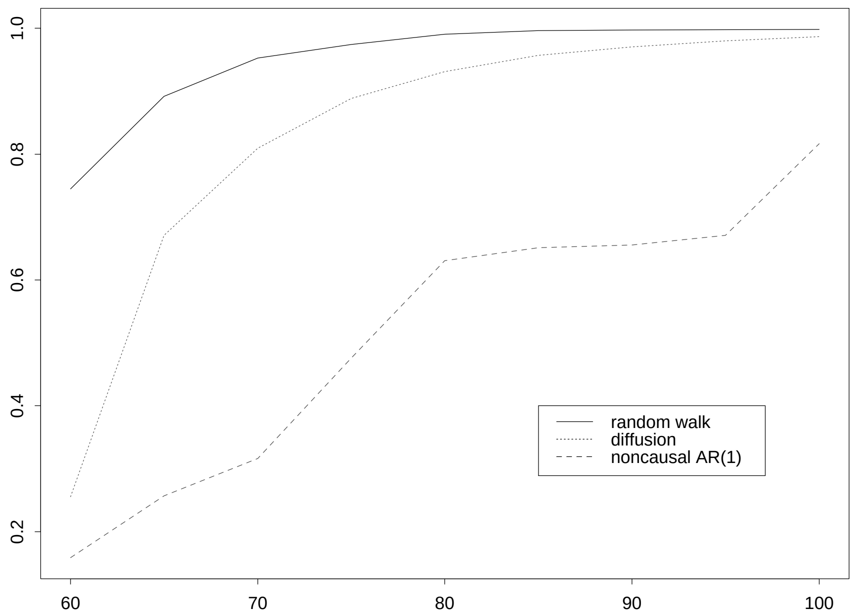

Figure 8: Shrinkaged OLS with constant as a Function of a

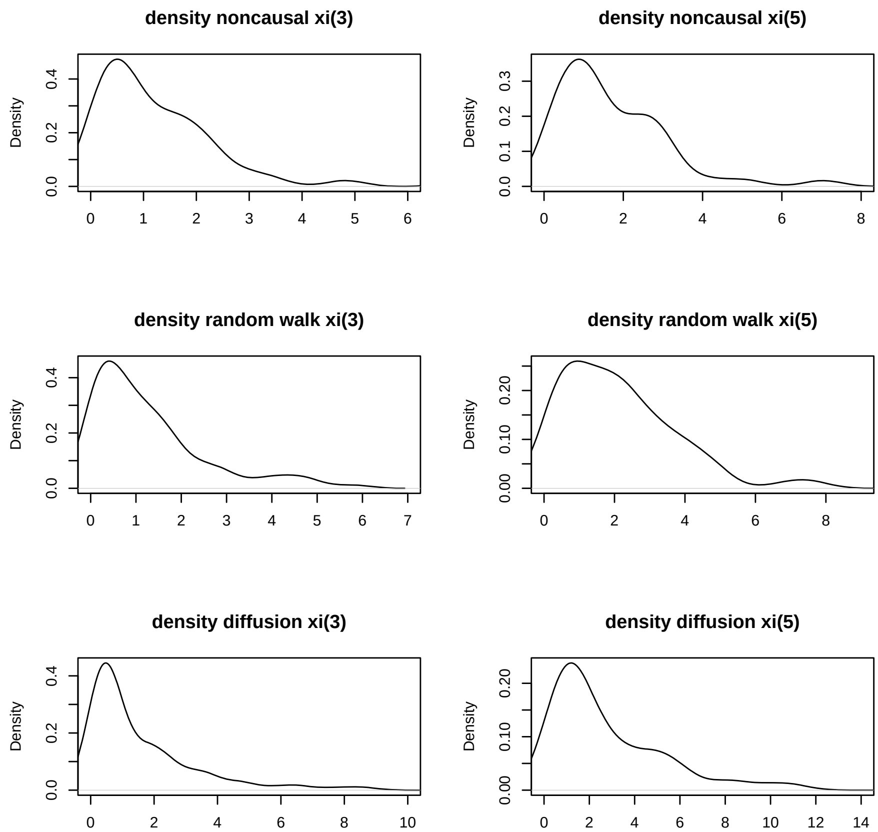

Figure 9: Finite sample density of test statistics, T=400