Weak convergence of non-stationary multivariate marked processes with applications to martingale testing

J. Carlos Escanciano∗, 1

Department of Economics, Indiana University, Bloomington, IN 47405-7104, USA

Received 13 July 2005
Available online 30 March 2007

Abstract

This paper establishes the weak convergence of a class of marked empirical processes of possibly non-stationary and/or non-ergodic multivariate time series sequences under martingale conditions. The assumptions involved are similar to those in Brown’s martingale central limit theorem. In particular, no mixing conditions are imposed. As an application, we propose a test statistic for the martingale hypothesis and we derive its asymptotic null distribution. Finally, a Monte Carlo study shows that the asymptotic results provide good approximations for small and moderate sample sizes. An application to the S&P 500 is also considered.

© 2007 Elsevier Inc. All rights reserved.

AMS 1991 subject classification: 2M07; 62G09; 62G10

Keywords: Marked empirical processes; Weak convergence; Martingale hypothesis; Non-stationary time series

1. Introduction and weak convergence theorem

Let for each \( n \geq 1 \), \( X_{n,0}', \ldots, X_{n,n-1}' \), be an array of random vectors in \( \mathbb{R}^d \), \( d \in \mathbb{N} \), and let \( \mathcal{F}_{n,t} = \sigma(X_{n,t}', X_{n,t-1}', \ldots, X_{n,0}') \), \( 0 \leq t \leq n \), be the \( \sigma \)-field generated by the observations obtained up to time \( t \). Furthermore, let, for each \( n \geq 1 \), \( Z_{n,1}, \ldots, Z_{n,n} \), be an array of square integrable real

∗ Fax: +1 812 855 7925.
E-mail address: jescanci@indiana.edu
URL: http://mypage.iu.edu/~jescanci.

1 Research funded by the Spanish Ministerio de Educación y Ciencia, reference numbers SEJ2004-04583/ECON and SEJ2005-07657/ECON.

0047-259X/07 - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2007.03.004
random variables (r.v.) such that for each \( t, 1 \leq t \leq n \), \( Z_{n,t} \) is \( F_{n,t} \)-measurable and such that almost surely (a.s.)

\[
E[Z_{n,t} | F_{n,t-1}] = 0, \quad 1 \leq t \leq n, \quad \forall n \geq 1.
\]

Denote, for each \( n \geq 1 \), by \((\Omega_n, A_n, P_n)\) the probability space in which all the r.v. \( Z_{n,t}, X'_{n,t-1}, 1 \leq t \leq n \), are defined. Then, \( F_{n,t} \) is a double array of sub \( \sigma \)-fields of \( A_n \) such that \( F_{n,t-1} \subset F_{n,t} \), \( 1 \leq t \leq n \). The main goal of this paper is to establish the weak convergence of the multivariate marked empirical process

\[
z_n(x) = s_n^{-1} \sum_{t=1}^{n} Z_{n,t} 1(X_{n,t-1} \leq x) \quad x \in \mathbb{R}^d,
\]

where \( 1(A) \) denotes the indicator function of the event \( A \) and \( s_n^2 = \sum_{t=1}^{n} E[Z_{n,t}^2] \). The process \( z_n \) is a marked empirical process with marks given by the r.v. \( Z_{n,1}, \ldots, Z_{n,n} \), and with jumps at the points \( X'_{n,0}, \ldots, X'_{n,n-1} \). Examples of such processes can be found extensively in the statistical and econometric literature and have been shown useful for inference problems such as model checks, e.g. Koul and Stute [14], or in testing and estimating the threshold in time series models, see, e.g. Chan [4]. In addition, consistent goodness of fit tests for simple hypothesis in multivariate non-linear regressions with integrated regressors can be based on the weak convergence of processes like \( z_n \). The limit distribution theory for this problem has been restricted in the literature to univariate regressors. The weak convergence of \( z_n \) in the multivariate case provides a starting point for developing inference procedures in multivariate regressions with non-stationary processes.

In an influential work, Stute [22] derived a weak convergence theorem for \( z_n \) to a continuous limit process when one observes a random sample \((Z'_{n,t}, X_{n,t-1})'\), \( 1 \leq t \leq n \), of independent and identically distributed (i.i.d) r.v., distributed as \((Z', X')'\). He derived the weak convergence under square integrability of \( Z \) and a continuity condition on the cumulative distribution function (cdf) of \( X \). For strictly stationary, ergodic, Markov processes with \( d = 1 \), Koul and Stute [14] proved the weak convergence of the process \( z_n \) under slightly more than fourth moment and bounded densities assumptions. Domínguez and Lobato [8] have extended Koul and Stute’s weak convergence to a general \( d \in \mathbb{N} \). To the best of our knowledge, these are the weakest assumptions in the literature for the stationary and ergodic case. Our results improve upon these existing results even in the stationary and ergodic case, see the discussion after Theorem 1. Recently, Park and Whang [20] have established the weak convergence of a particular case of a marked process \( z_n \) under non-stationary and/or non-ergodic univariate processes. Their weak convergence result depends crucially on the univariate character of the regressor and its extension to the multivariate case using their techniques seems to be difficult, see Park and Phillips [19, p. 143] for a discussion on this issue. Here, we extend these results in several directions. We obtain a weak convergence theorem for \( z_n \) under multivariate jumps and possibly non-stationary and/or non-ergodic sequences \( Z_{n,t}, X'_{n,t-1}, 1 \leq t \leq n \). Our assumptions are similar to those considered in Brown’s martingale central limit theorem (MCLT) and are comparable to those considered in the i.i.d setup. Instead of using spatial integrals and the local properties of the underlying processes as in Park and Whang [20], we consider an empirical processes approach as in van der Vaart and Wellner [23] suitable under multivariate regressors. Our result belongs to the class of weak convergence theorems for dependent data. A nice exposition of this rather new field can be found in Dehling et al. [6] and references therein.

To elaborate our weak convergence theorem we need some notation. Throughout, \( A' \) and \( |A| \) denote the matrix transpose and the Euclidean norm of \( A \), respectively. Let \( \wedge \) denote the minimum,
i.e., \( a \wedge b = (\min\{a_1, b_1\}, \ldots, \min\{a_d, b_d\})' \) for \( a = (a_1, \ldots, a_d)' \) and \( b = (b_1, \ldots, b_d)' \) in \( \mathbb{R}^d \). Let \( \mathbb{R}_d^d \) be the extended \( \mathbb{R}_d^d \) Euclidean space, i.e., \( \mathbb{R}_d^d = [-\infty, \infty)^d \), and let \( \ell^\infty(\mathbb{R}_d^d) \) be the Banach space of real-valued bounded functions on \( \mathbb{R}_d^d \), equipped with the supremum norm \( \|z\|_{\ell^\infty(\mathbb{R}_d^d)} = \sup_{x \in \mathbb{R}_d^d} |z(x)| \). Note that, by defining \( z_n(-\infty) = 0 \) and \( z_n(+\infty) = s_n^{-1} \sum_{i=1}^n Z_{n,t}, \) we can consider \( z_n \) as a mapping from \( \Omega_n \) to \( \ell^\infty(\mathbb{R}_d^d) \). Let \( \implies \) denote weak convergence in \( \ell^\infty(\mathbb{R}_d^d) \), see Definition 1.3.3 in van der Vaart and Wellner [23]. Of course, the sample paths of \( z_n \) are contained in the much smaller cadlag space \( D(\mathbb{R}_d^d) \), but as long as this space is equipped with the supremum metric, this is irrelevant for the weak convergence theorem. Unless indicated all convergences are taken as the sample size \( n \to \infty \). Let us define the quantities \( \sigma_{n,t}^2 = E[Z_{n,t}^2 | \mathcal{F}_{n,t-1}], \) \( V_n^2 = \sum_{t=1}^n \sigma_{n,t}^2 \) and \( \tau_n^2(x) = s_n^{-2} \sum_{t=1}^n \sigma_{n,t}^2 1(X_{n,t-1} \leq x), x \in \mathbb{R}_d^d \). Now, we are in a position to state our main result, the weak convergence theorem for \( z_n \) to a Gaussian process in \( \ell^\infty(\mathbb{R}_d^d) \).

Theorem 1. In addition to (1) suppose that the following conditions hold:

(A) \( V_n^2 s_n^{-2} P_n \to 1; \)

(B) \( s_n^{-2} \sum_{t=1}^n E[Z_{n,t}^2 1(|Z_{n,t}| > \varepsilon s_n)] P_n \to 0 \) for every \( \varepsilon > 0; \)

(C) There exists a continuous non-decreasing function \( \tau^2 \) on \( \mathbb{R}_d^d \) to \([0, \infty)\) such that uniformly in \( x \in \mathbb{R}_d^d \),

\[
\tau_n^2(x) = \tau^2(x) + o_P(1).
\]

(D) \( |E[Z_{n,t}^2 1(x \leq X_{n,t-1} \leq y) | \mathcal{F}_{n,t-2}]| \leq C_{n,t} \left| \tau^2(y) - \tau^2(x) \right| \) with \( C_{n,t} \) such that \( s_n^{-2} \sum_{t=1}^n E[C_{n,t}] = O(1). \)

Then, it follows that

\[ z_n(\cdot) \implies z_\infty \quad \text{in} \quad \ell^\infty(\mathbb{R}_d^d), \]

where \( z_\infty \) is a zero-mean Gaussian process with covariance function \( \tau^2(x \wedge y) \).

Theorem 1 generalizes Theorem 3.1 in Ossiander [18] and Theorem 2.11.9 in van der Vaart and Wellner [23] to empirical processes under possibly non-stationary and/or non-ergodic martingale difference sequences. Earlier results in this direction can be found in Levental [15] and Bae and Levental [1] for stationary and ergodic sequences and in Nishiyama [17] for some classes of nonstationary martingales. Some comments on the assumptions of Theorem 2 are necessary. Assumptions (A) and (B) are considered in Brown [3]. Under (A), (B) is also a necessary condition for the convergence of the finite dimensional distributions (fids) to hold, see Brown [3, Theorem 1]. Lemma 2 in Brown [3] gives, under (A), equivalent conditions to (B). Condition (C) is assumed in Koul and Stute [14]. Under the assumption that \( \sup_{n,t} E[Z_{n,t}^2] \leq p < \infty \), for some \( p > 1 \), the pointwise convergence of \( \tau_n^2(x) \) to \( \tau^2(x) \) implies the uniform convergence, see the proof of Lemma 4.3 in Park and Whang [20], and hence, also that the function \( \tau^2(x) \) must be the pointwise limit of \( s_n^{-2} \sum_{t=1}^n E[Z_{n,t}^2 1(X_{n,t-1} \leq x)] \).

In the i.i.d framework of Stute [22], our assumptions in Theorem 1 reduce to \( E[Z_{n,t}^2] = E[Z^2] < \infty \) and the continuity of the function \( E[Z_{n,t}^2 1(X_{n,t-1} \leq x)] = E[Z^2 1(X \leq x)] \). The latter condition in turn holds true if the cdf. of \( X \) is continuous. Thus, in the i.i.d case the assumptions of Theorem 1 here reduce to those of Theorem 1.1 in Stute [22]. For the strictly
stationary, ergodic and Markov case of Koul and Stute [14] and Domínguez and Lobato [8] by the Ergodic Theorem and a Glivenko–Cantelli’s argument, see Koul and Stute [14, (4.1)], our assumptions reduce to \( E[Z^2_{n,t}] < \infty \) and the continuity assumption in (D). Therefore, even in this case Theorem 1 here gives weaker conditions than the corresponding weak convergence results in Koul and Stute [14] and Domínguez and Lobato [8].

To illustrate the usefulness of the new weak convergence theorem, we consider in the next section a test for the martingale hypothesis of possibly non-stationary time series processes, thereby providing an application of Theorem 1 which cannot be covered by other existing weak convergence theorems proposed in the literature.

2. Testing the martingale hypothesis

In this section we consider a test for the so-called martingale hypothesis (MH) and we apply Theorem 1 to study the asymptotic null distribution of the proposed test statistic. The MH has important implications in statistical applications and other fields, since it implies that the best predictor (in a mean squared sense) of future values of a time series given the current information set is just the current value of the time series. In economics, many theories in a dynamic context, such as the market efficiency hypothesis, rational expectations, asset pricing or optimal consumption smoothing lead to such dependence restrictions on the underlying economic variables, see e.g. Lo [16] or Cochrane [5] and references therein.

The MH, or better the martingale difference hypothesis (MDH), see (4) below, has been typically tested in the stationary case using the autocorrelations or autocovariances or in the spectral domain using the periodogram, see for instance Deo [7] for a recent reference. However, correlation-based tests are not appropriate when non-linearity and/or non-Gaussianity are presented. In fact, these tests are not consistent against non-martingale difference sequences with zero autocorrelations, e.g. some bilinear or non-linear moving average models. Recently, some omnibus tests for the MDH based on empirical processes theory have been proposed in the literature, see e.g. Domínguez and Lobato [8] or Escanciano and Velasco [9]. Related tests to the MDH are the goodness-of-fit tests for the conditional mean, which can be adapted as MDH tests, see for instance, those tests proposed by Bierens [2] or Stute [22]. However, within this literature only Park and Whang [20] have considered a test for the MH with non-stationary time series. They consider the univariate case, i.e. \( d = 1 \). Their approach to show the asymptotic theory is specific for the MH, i.e., it depends crucially on the fact that the \( Z^i \)'s in (2) are the increments \( Y_t - Y_{t-1} \). Their weak convergence theorem relies on writing the tests statistics as functionals of a partial sum of the increments \( Y_t - Y_{t-1} \) and on the local properties of such partial sum. Such arguments are only valid for univariate regressors. On the contrary our approach is valid under multivariate processes under fairly weak conditions on the underlying data generating process. Here, we apply our new weak convergence theorem to propose an alternative proof of their result, extending it to the multivariate case \( d \geq 1 \). In addition it is worth to remark that for their asymptotic analysis Park and Whang [20] needed to assume a bounded conditional fourth moment assumption. This assumption may look restrictive, it implies that the conditional variance is also bounded and it rules out most empirically relevant conditional heteroskedastic processes whose fourth moments are often found to be infinite. We remove the bounded fourth conditional moment in our analysis.

In this section we define \( X_{n,t} = n^{-1/2}(Y_t, \ldots, Y_{t-d+1})' \) and the \( \sigma \)-fields \( \mathcal{F}_{n,t} = \sigma(X_{n,t}', X_{n,t-1}', \ldots, X_{n,0}') = \sigma(n^{-1/2}Y_t, \ldots, n^{-1/2}Y_{t-d+1}) \) and \( \mathcal{F}_{t,\infty} = \sigma(Y_t, Y_{t-1}, \ldots) \). The factor \( n^{-1/2} \) in \( X_{n,t} \)
simplifies the limit theory in the test procedure. Then, the MH is expressed as

\[ E[Y_t \mid \mathcal{F}_{t-1, \infty}] = Y_{t-1} \ a.s. \quad \text{for all } t \geq 1, \tag{3} \]

that can be trivially written as

\[ E[\Delta Y_t \mid \mathcal{F}_{t-1, \infty}] = 0 \ a.s. \quad \text{for all } t \geq 1, \tag{4} \]

with \( \Delta Y_t = Y_t - Y_{t-1} \). In view of a sample \( Y_1, \ldots, Y_n, d \geq 1 \), our test is based on the marked empirical process

\[ z_n(x) = n^{-1/2} \sum_{t=1}^{n} \Delta Y_t \mathbb{1}(X_{n,t-1} \leq x) \quad x \in \mathbb{R}^d, \tag{5} \]

see Park and Whang [20, p. 3] for motivation of the use of \( z_n \) for testing the MH. Note that under the MH,

\[ E[\Delta Y_t \mid \mathcal{F}_n,t-1] = 0 \ a.s. \quad \text{for all } t \geq 1, \quad n \geq 1 \]

and \( \Delta Y_t \) is \( \mathcal{F}_n,t \)-measurable. Tests for the MH can be based on the distance from \( z_n \) in (5) to zero. Here, we consider a Kolmogorov–Smirnov test statistic

\[ K_{S_{n,d}} = \sup_{x \in \mathbb{R}^d} |z_n(x)|. \]

Hence, we reject the MH given in (3) if we obtain a “large” value of \( K_{S_{n,d}} \). In the next theorem we establish the asymptotic null distribution of \( K_{S_{n,d}} \), which is a direct application of Theorem 1 and the continuous mapping theorem (cf. [23, Theorem 1.3.6]).

**Theorem 2.** Under the MH given in (3) and assumptions (A-D) of Theorem 1 but with \( Z_{n,t} \equiv Z_t = \Delta Y_t \), we obtain

\[ z_n(\cdot) \Rightarrow z_\infty(\cdot) \quad \text{in } \ell^\infty(\mathbb{R}^d), \]

where \( z_\infty \) is a Gaussian process with zero mean and covariance function \( \tau^2(x \wedge y) \). Furthermore,

\[ K_{S_{n,d}} \overset{d}{\rightarrow} \sup_{x \in \mathbb{R}^d} |z_\infty(\cdot)|. \]

Theorem 2 establishes the null limit distribution of \( K_{S_{n,d}} \) for \( d \geq 1 \). It is worth to remark that for \( d > 1 \) the limit is not distribution-free, so critical values have to be computed for each data generating process. A similar problem appears in many empirical processes-based tests, see e.g. Khmaladze and Koul [13]. A solution to this problem seems to be difficult under our non-stationary framework. A wild bootstrap approximation should be useful in this multivariate situation. That is, we can approximate the null distribution of \( z_n(\cdot) \) by the bootstrap process

\[ z_n^*(x) = n^{-1/2} \sum_{t=1}^{n} \Delta Y_t w_t \mathbb{1}(X_{n,t-1} \leq x), \]

where \( \{w_t\} \) is a sequence of independent r. v. with zero mean, unit variance, bounded support and also independent of the sequence \( \{Y_t\}_{t=-d+1}^{n} \). The validity of the bootstrap approximation can be
proved straightforwardly along the lines in Domínguez and Lobato [8] or Escanciano and Velasco [9]. Details are omitted.

We now focus on the univariate case, \( d = 1 \). In that case, the limit distribution is \( \mathcal{z}_\infty (\cdot) = B \circ \tau^2 (\cdot) \) (in law), where \( B \) is standard Brownian motion on \([0, \infty)\). Note that by the scaling properties of the Brownian motion we have that under the conditions of Theorem 2,

\[
K_{S,1} \xrightarrow{d} \sup_{x \in \mathbb{R}} \left| B \circ \tau^2 (x) \right| = \tau (\infty) \sup_{0 \leq t \leq 1} |B(t)| \quad \text{in law.}
\]

If the sequence \( \{V^2_n, n \geq 1\} \), given now by \( V^2_n = \sum_{i=1}^n E[(\Delta Y_t)^2 | \mathcal{F}_{t-1}] \), is uniformly integrable, we have by Hall and Heyde [12, Theorem 2.23] that \( \hat{\tau}^{-1}_n (\infty) = 1/n \sum_{i=1}^n (\Delta Y_t)^2 \) is a consistent estimator for \( \tau^2(\infty) \). Hence, in that case under the assumptions of Theorem 2,

\[
\hat{\tau}^{-1}_n (\infty) K_{S,1} \xrightarrow{d} \sup_{0 \leq t \leq 1} |B(t)|.
\]

Thus, the asymptotic null distribution of \( \hat{\tau}^{-1}_n (\infty) K_{S,1} \) is distribution-free, and the test procedure consists in rejecting the MH if \( \hat{\tau}^{-1}_n (\infty) K_{S,1} \) exceeds an appropriate critical value obtained from the boundary crossing probability of a Brownian motion, which are readily available on the unit interval, see Shorack and Wellner [21]. Note that the proposed test is very simple to use: the test statistic is easy to compute and it neither depends upon any smoothing parameter nor requires any resampling procedure to simulate the null distributions. Although the knowledge of the whole function \( \tau^2(x) \) is not necessary for the test procedure, we discuss here what is the natural candidate for \( \tau^2 \) in a particular case. Let denote \( \tau^2(\infty) = \sigma^2 \). We give the result in the following:

**Lemma 1.** Under the notation of Theorem 2, in addition to (3) and (B) assume that \( Y_{-1} = 0 \), \( d = 1 \), and that for all \( x \in \mathbb{R} \)

\[
(A') \quad \left| n^{-1} \sum_{t=1}^n \{ E[(\Delta Y_t)^2 | \mathcal{F}_{n,t-1}]1(X_{n,t-1} \leq x) - E[(\Delta Y_t)^2 1(X_{n,t-1} \leq x)] \} \right| \xrightarrow{P_n} 0.
\]

\[
(B') \quad \text{The sequence } \{V^2_n, n \geq 1\}, \text{ is uniformly integrable. Then, under the conclusions of Theorem 2 or under the condition } \sup_{n,t} E[(\Delta Y_t)^4 | \mathcal{F}_{n,t-1}] < \infty \quad \text{a.s., condition (C) in Theorem 1 holds with}
\]

\[
\tau^2(x) = \sigma^2 \int_1^0 P(W(r) \leq x) \, dr,
\]

where \( W(r) \) is a Brownian motion on \([0, 1]\) with

\[
\text{Cov}(W(r_1), W(r_2)) = \sigma^2 (r_1 \wedge r_2).
\]

Park and Whang [20] considered a similar statistic to \( T_{n,1} = \tau^{-1}_n (\infty) K_{S,1} \). They did not standardize by the sample variance \( \hat{\tau}^{-1}_n (\infty) \) in the test statistic but in the observations \( Y_0, \ldots, Y_n \). In spite of these different standardizations, the limit process is the same in both cases. But surprisingly, instead of the critical values from the boundary crossing probability of a Brownian motion, which are 1.959, 2.241 and 2.807 at the 10%, 5% and 1%, respectively, they used 2.119, 2.388 and 2.911 as the asymptotic critical values at 10%, 5% and 1% nominal level, respectively. As
we shall show their computation of critical values is erroneous. They stated that the asymptotic null limit process of $\sigma_n$ in (5) is that of (with $\sigma^2 = 1$)

$$M(x) = \int_0^1 1(W(r) \leq x) \, dW(r),$$

where $W(r)$ is a standard Brownian motion on $[0, 1]$, see (14) in Park and Whang [20]. But this limit process is nothing else than a Brownian motion in proper time, $B \circ \tau^2(x)$, where

$$\tau^2(x) = \int_0^1 P(W(r) \leq x) \, dr.$$

Hence,

$$\sup_{x \in \mathbb{R}} |M(x)| = \sup_{x \in \mathbb{R}} |B \circ \tau^2(x)| = \tau(\infty) \sup_{0 \leq t \leq 1} |B(t)| = \sup_{0 \leq t \leq 1} |B(t)| \text{ in law.}$$

So, Park and Whang’s critical values should be those of the supremum of the standard Brownian motion on $[0,1]$. This fact may explain the low empirical power reported in Park and Whang [20] for some alternatives relative to those obtained here.

The consistency properties of the proposed test can be analyzed using exactly the same techniques as in Park and Whang [20]. Under the alternative to (3) is necessary to assume some mixing condition in order to establish a uniform law of large numbers, we do not give the details for the sake of clarity. Instead, next section shows that the asymptotic results provide good approximations for small sample sizes.

3. Finite sample performance of the MH test and data analysis

The aim of this section is to examine the finite sample performance of the proposed asymptotic distribution-free test. The finite sample performance of the proposed bootstrap test will appear elsewhere. To that end, we carry out a simulation experiment with some data generating processes (DGP) under the null and under the alternative. Our test statistic is $T_{n,1} = \tau^{-1}(\infty) K S_{n,1}$.

In the sequel $\epsilon_t$ is a sequences of i.i.d $N(0,1)$ r.v.. The models used in the simulations are those of Park and Whang [20]:

1. NULL: $Y_t = Y_{t-1} + u_t$, where $u_t = \sigma_t \epsilon_t$, $\sigma_t^2 = 1 + \theta_1 u_t^2 + \theta_2 \sigma_{t-1}^2$.
2. ARMA: $Y_t = \theta_1 Y_{t-1} - \theta_2 \epsilon_{t-1} + \epsilon_t$.
3. EXAR: $Y_t = \theta_1 Y_{t-1} - \theta_2 Y_{t-1} \exp(-0.1 |Y_{t-1}|) + \epsilon_t$.
4. TAR: $Y_t = \theta_1 Y_{t-1} 1(|Y_{t-1}| < \theta_2) + 0.9 Y_{t-1} 1(|Y_{t-1}| \geq \theta_2) + \epsilon_t$.
5. BL: $Y_t = \theta_1 Y_{t-1} - \theta_2 Y_{t-1} \epsilon_{t-1} + \epsilon_t$.
6. NLMA: $Y_t = \theta_1 Y_{t-1} - \theta_2 \epsilon_{t-1} \epsilon_{t-2} + \epsilon_t$.

Model NULL generates random walk processes possibly with GARCH errors and is considered to evaluate the size performance of our test. The other models allow us to see the empirical power performance. For some description of the models and the parameter values taken by $\theta_1$ and $\theta_2$ see Park and Whang [20, p. 11].

We use for the experiments the sample sizes $n = 100$ and 300 and the number of Monte Carlo experiments is 1000. We consider a nominal level of 5%, the results with other nominal levels are similar, and here, they are not reported. In all the replications 200 pre-sample data values were generated and discarded. Random numbers were generated using IMSL ggnml subroutine.
Table 1
Rejection probabilities

<table>
<thead>
<tr>
<th>$(\theta_1, \theta_2)$</th>
<th>NULL</th>
<th>ARMA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$T_{n,1}$</td>
</tr>
<tr>
<td>(0.0, 0.0)</td>
<td>100</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>5.5</td>
</tr>
<tr>
<td>(0.9, 0.0)</td>
<td>100</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>4.0</td>
</tr>
<tr>
<td>(0.2, 0.3)</td>
<td>100</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>5.0</td>
</tr>
<tr>
<td>(0.7, 0.2)</td>
<td>100</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>4.4</td>
</tr>
</tbody>
</table>

Table 2
Rejection probabilities

<table>
<thead>
<tr>
<th>$(\theta_1, \theta_2)$</th>
<th>EXAR</th>
<th>TAR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$T_{n,1}$</td>
</tr>
<tr>
<td>(0.6, 0.2)</td>
<td>100</td>
<td>74.5</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>100.0</td>
</tr>
<tr>
<td>(0.6, 0.4)</td>
<td>100</td>
<td>96.7</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>100.0</td>
</tr>
<tr>
<td>(0.9, 0.2)</td>
<td>100</td>
<td>7.8</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>98.8</td>
</tr>
<tr>
<td>(0.9, 0.4)</td>
<td>100</td>
<td>43.7</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>100.0</td>
</tr>
</tbody>
</table>

In Table 1 we report the empirical rejections probabilities (RP) associated with the model NULL. The results for $T_{n,1}$ show that the empirical size properties of the test are excellent and that $T_{n,1}$ is robust to thick tails.

In Tables 1, 2 and 3 we report the empirical power against the models 2–6 for the sample sizes $n = 100$ and 300. The results show that for almost all non-martingale alternatives $T_{n,1}$ has reasonable empirical power. However, for the near-unit root cases is somewhat unsatisfactory for the sample size $n = 100$. As Park and Whang [20] showed with their simulations, these near-unit root cases need of large sample sizes (as $n = 1000$) to be detected. Also, we note that the power performance of $T_{n,1}$ against most alternatives is better than that of the tests proposed by Park and Whang [20]. As previously explained, the different results are due to a wrong approximation of the asymptotic critical values of the test statistic in their paper. Note also that the size performance of $T_{n,1}$ here is more accurate than those of Park and Whang [20].

Now, we present an application of our martingale test to the daily closed S&P 500 stock index for the period from 3 January 1994 until 31 December 1997. The number of observations is 1011. We consider as $Y_t = \log(S_t)$, with $S_t$ the value of the index at time $t$. An application of the
Table 3

<table>
<thead>
<tr>
<th>$(\theta_1, \theta_2)$</th>
<th>NULL</th>
<th>$(\theta_1, \theta_2)$</th>
<th>ARMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>$T_{n,1}$</td>
<td>n</td>
<td>$T_{n,1}$</td>
</tr>
<tr>
<td>(0.4, 0.1)</td>
<td>100</td>
<td>77.5</td>
<td>(0.4, 0.2)</td>
</tr>
<tr>
<td>300</td>
<td>100.0</td>
<td>300</td>
<td>100.0</td>
</tr>
<tr>
<td>(0.4, 0.3)</td>
<td>100</td>
<td>63.4</td>
<td>(0.4, 0.6)</td>
</tr>
<tr>
<td>300</td>
<td>100.0</td>
<td>300</td>
<td>100.0</td>
</tr>
<tr>
<td>(0.8, 0.1)</td>
<td>100</td>
<td>2.5</td>
<td>(0.8, 0.2)</td>
</tr>
<tr>
<td>300</td>
<td>79.2</td>
<td>300</td>
<td>84.8</td>
</tr>
<tr>
<td>(0.8, 0.3)</td>
<td>100</td>
<td>1.5</td>
<td>(0.8, 0.6)</td>
</tr>
<tr>
<td>300</td>
<td>26.2</td>
<td>300</td>
<td>88.8</td>
</tr>
</tbody>
</table>

spectral domain test of Deo [7] reveals that the increments of $Y_t$ are uncorrelated (the p-value of Deo’s test is 0.730). Our test $T_{n,1}$ strongly rejects the MH with a p-value of 0.00. This shows the ability of our test to detect non-linear dependence. Further investigation is necessary to find the functional form of the conditional mean for the S&P 500 in this period. We defer this and the more general problem of model checks of non-stationary sequences for future research. Our Theorem 1 provides the first step toward the solution of such a challenging problem.

4. Proofs

To prove Theorem 1 we first consider the following lemmas. First lemma corresponds to Theorems 1.5.4 and 1.5.6 in van der Vaart and Wellner [23].

**Lemma A1.** Let $x_n$ be a mapping from $\Omega_n$ to $\ell^\infty(\mathbb{R}^d)$. Consider the following statements:

(i) $x_n$ converges weakly to a tight, Borel law;

(ii) every finite-dimensional marginal of $x_n$ converges weakly to a (tight) Borel law;

(iii) for every $\epsilon, \eta > 0$ there exists a finite partition $B = \{T_k; 1 \leq k \leq N\}$ of $\mathbb{R}^d$ such that

$$\limsup_{n \to \infty} P^* \left[ \max_{1 \leq k \leq N} \sup_{t, s \in T_k} |x_n(t) - x_n(s)| > \epsilon \right] \leq \eta.$$ 

Then, there is the equivalence (i) $\iff$ (ii) + (iii). Furthermore, if the marginals of a stochastic process $x$ have the same laws as the limits in (ii), then there exists a version $\tilde{x}$ of $x$ such that $x_n \equiv \tilde{x}$ in $\ell^\infty(\mathbb{R}^d)$.

Next lemma is the so-called Bernstein–Freedman inequality for a martingale difference array. See Freedman [10] for the proof.

**Lemma A2.** Let $\{M_{n,t}; 1 \leq t \leq n\}$ be an $\mathbb{R}$-valued martingale difference array with respect to the filtration $\mathcal{F}_{n,t-1}$, such that $|M_{n,t}| < a, \forall t \geq 1, n \geq 1$. Let $\sigma$ be a bounded stopping time.
Then for any $b > 0$, \[ P \left( \max_{1 \leq s \leq \sigma} \left| \sum_{t=1}^{s} M_{n,t} \right| > \varepsilon, \sum_{t=1}^{s} E[M_{n,t}^{2} | \mathcal{F}_{n,t-1}] \leq b \right) \leq 2 \exp \left( -\frac{\varepsilon^{2}}{2(a \varepsilon + b)} \right) \quad \forall \varepsilon > 0. \]

**Proof of Theorem 1.** To prove Theorem 1 we need to show that (ii) and (iii) in Lemma A1 hold. The convergence of the finite-dimensional distributions follows by applying the Brown’s MCLT and the Cramér–Wold device. To prove (iii), let us define the semimetric $d(x, y) = |\tau^{2}(y) - \tau^{2}(x)|$. Then, condition (C) guarantees that for any $\varepsilon > 0$ we can form a finite partition $\mathcal{B}_{\varepsilon} = \{ B_{k}; 1 \leq k \leq N(\varepsilon, \mathbb{R}^{d}, d) \}$ of $\mathbb{R}^{d}$ in $\varepsilon$-brackets $B_{k} = [x_{k}, y_{k}]$, i.e., $\{ B_{k}\}_{k=1}^{N_{\varepsilon}}$ covers $\mathbb{R}^{d}$, $x_{k} \leq y_{k}$, and $d(x_{k}, y_{k}) \leq \varepsilon$. Fix $v > 2$ and define for every $q \in \mathbb{N}$, $q \geq 1$, $\varepsilon = 2^{-qv}$. We denote the previous partition associated to $\varepsilon = 2^{-qu}$ by $\mathcal{B}_{q} = \{ B_{qk}; 1 \leq k \leq N_{q} \equiv N(2^{-qu}, \mathbb{R}^{d}, d) \}$. Without loss of generality we can assume that the finite partitions in the sequence $\{ \mathcal{B}_{q} \}$ are nested. From standard results on VC-classes, see van der Vaart and Wellner [23], we have
\[ \sum_{q=1}^{\infty} 2^{-q} \sqrt{\log N_{q}} < \infty. \]

Furthermore, because of the monotonicity of $1(X_{n,t-1} \leq x)$,
\[ z_{n}(\mathcal{B}_{q}) = \max_{1 \leq k \leq N_{q}} \left| s_{n}^{-2} \sum_{t=1}^{n} E[Z_{n,t}^{2} | \mathcal{F}_{n,t-1}] \sup_{x, y \in B_{qk}} \left| 1(X_{n,t-1} \leq x) - 1(X_{n,t-1} \leq y) \right|^{2} \right| = \max_{1 \leq k \leq N_{q}} d_{n}^{2}(x_{k}, y_{k}). \quad (6) \]

Define the event
\[ V_{n} = \left\{ \sup_{q \in \mathbb{N}} \max_{1 \leq k \leq N_{q}} \frac{d_{n}^{2}(x_{k}, y_{k})}{2^{-2q}} \geq \gamma \right\}. \]

We shall show that for all $\eta > 0$, there exists some $\gamma > 0$ such that $\limsup_{n \to \infty} P_{n}(\mathcal{V}_{n}) \leq \eta$. Note that
\[ P_{n}(\mathcal{V}_{n}) \leq \sum_{q=1}^{\infty} P_{n} \left( \max_{1 \leq k \leq N_{q}} \frac{d_{n}^{2}(x_{k}, y_{k})}{2^{-2q}} \geq \gamma \right) = \sum_{q=1}^{\infty} V_{nq}. \quad (7) \]

Now, define the process
\[ \tilde{z}_{n,w}(x) = s_{n}^{-2} \sum_{t=1}^{n} E[Z_{n,t}^{2} | \mathcal{F}_{n,t-1}]1(X_{n,t-1} \leq x) \]
and the quantities for $1 \leq t \leq n, \tilde{\beta}_{nt}(x) = E[Z_{n,t}^{2} | \mathcal{F}_{n,t-1}]1(X_{n,t-1} \leq x) - F_{n,w}(x)$ and $F_{nt}(x) = E[Z_{n,t}^{2} | 1(X_{n,t-1} \leq x) | \mathcal{F}_{n,t-2}]$. Hence
\[ \tilde{z}_{n,w}(x) = s_{n}^{-2} \sum_{t=1}^{n} \tilde{\beta}_{nt}(x) + s_{n}^{-2} \sum_{t=1}^{n} F_{nt}(x). \]
Then, we have
\[ V_n \lesssim P_n \left( \max_{1 \leq k \leq N_q} s_n^{-2} \sum_{t=1}^{n} \left| \tilde{\beta}_n(x_k) - \tilde{\beta}_n(y_k) \right| \right) \geq 2^{-2q} \gamma ) + P_n \left( \max_{1 \leq k \leq N_q} s_n^{-2} \sum_{t=1}^{n} |F_{nt}(x_k) - F_{nt}(y_k)| \right) \geq 2^{-2q} \gamma ) \equiv A_{1nq} + A_{2nq}. \] (8)

Notice that \( \{ \tilde{\beta}_{n,w}(x), \mathcal{F}_{n,t-2} \} \) is a martingale difference sequence for each \( x \in \mathbb{R}^d \), by construction. By a truncation argument, it can be assumed without loss of generality that \( \max_{1 \leq k \leq N_q} \left| Z_{n,t} \right| \leq s_n a_q - 1, \) where henceforth \( a_q = 2^{-q \rho} / \sqrt{\log(N_q + 1)} \) with \( 1 < \rho < 2 \). See the arguments below. Define the set
\[ \Gamma_n = \left\{ \left( s_n^{-2} \sum_{t=1}^{n} C_{n,t} \right) \leq K \right\}. \]

Now, by Freedman’s [10] inequality in Lemma A2 and Lemma 2.2.10 in van der Vaart and Wellner [23],
\[ E \max_{1 \leq k \leq N_q} s_n^{-2} \sum_{t=1}^{n} \left[ \tilde{\beta}_n(x_k) - \tilde{\beta}_n(y_k) \right] \leq C \left( a_{q-1}^2 \log(1 + N_q) + a_{q-1} 2^{-q/2} \sqrt{\log(1 + N_q)} \right). \]

Hence, by Markov’s inequality and the definition of \( a_q \), on the set \( \Gamma_n \),
\[ A_{1nq} \leq C a_{q-1}^2 \log(1 + N_q) + a_{q-1} 2^{-q/2} \sqrt{\log(1 + N_q)} \]
\[ = C \gamma^{-1} 2^{-2q(\rho-1)} + C \gamma^{-1} 2^{-q(\rho+v/2-1)}. \]

On the other hand, by (D) and by Markov’s inequality
\[ A_{2nq} \leq \gamma^{-1} s_n^{-2} \sum_{t=1}^{n} E \max_{1 \leq k \leq N_q} 2^q \left| s_n^{-2} \sum_{t=1}^{n} |F_{nt}(x_k) - F_{nt}(y_k)| \right| \leq \gamma^{-1} 2^{-q(v-2)} \left( s_n^{-2} \sum_{t=1}^{n} C_{n,t} \right) \leq K \gamma^{-1} 2^{-q(v-2)}, \]
on the set \( \Gamma_n \). Therefore, by (7), (8) and the last three bounds,
\[ P_n(V_n) \leq C \gamma^{-1} \sum_{\substack{q=1 \cr q \neq K}} 2^{-2q(\rho-1)} + 2^{-q(\rho+v/2-1)} + 2^{-q(v-2)} + P_n(\Gamma_n^c), \]

which can be made arbitrarily small by choosing a sufficiently large \( \gamma \) and \( K \).

The last inequality yields that for any \( \eta > 0 \) there exists a constant \( K = K_\eta > 0 \), such that
\[ \lim_{n \to \infty} P_n(\Omega_n \setminus \Omega_n^\eta) \leq \eta, \]
where
\[
\Omega_K^n = \left\{ \sup_{q \in \mathbb{N}} \frac{z_n(B_q)}{2^{-2q}} \leq K \right\}.
\]

Now, choose an element \(x_{qk}\) for each \(B_{qk}\) and define for every \(x \in [-\infty, \infty]^d\) the events
\[
\pi_q x = x_{qk} \quad \text{and} \quad B_q x = B_{qk} \quad \text{if} \quad x \in B_{qk}.
\]

To simplify notation define \(M^n_t(x) = s_n^{-1} Z_{n,t,1}(X_{n,t-1} \leq x)\). Then, by Lemma A1, see also the proof of Theorem 2.5.6 of van der Vaart and Wellner [23], it is sufficient to prove that for every \(\varepsilon, \eta > 0\) there exists a \(q_0 \in \mathbb{N}\) such that
\[
\lim \sup_{n \to \infty} \mathbb{P}^n \left[ \left\| \sum_{t=1}^{n} M^n_t(x) - \sum_{t=1}^{n} M^n_t(\pi_{q_0} x) \right\|_{\mathbb{R}^d} > \varepsilon \right] \leq \eta.
\]

To this end, fix any \(q_0\) for a while and let us define the quantities
\[
\Delta^n_t(B) = \sup_{x_1, x_2 \in B} \left| M^n_t(x_1) - M^n_t(x_2) \right|,
\]
and the events for \(q > q_0\)
\[
C^n_{t,q-1} = 1(\Delta^n_t(B_{q_0} x) \leq a_{q_0}, \ldots, \Delta^n_t(B_{q-1} x) \leq a_{q-1}),
\]
\[
D^n_{t,q} = 1(\Delta^n_t(B_{q_0} x) \leq a_{q_0}, \ldots, \Delta^n_t(B_{q-1} x) \leq a_{q-1}, \Delta^n_t(B_q x) > a_q),
\]
\[
D^n_{t,q_0} = 1(\Delta^n_t(B_{q_0} x) > a_{q_0}).
\]

Now, similarly to van der Vaart and Wellner [23, p. 131], we decompose
\[
M^n_t(x) - M^n_t(\pi_{q_0} x) = (M^n_t(x) - M^n_t(\pi_{q_0} x)) D^n_{t,q_0}
\]
\[
+ \sum_{q=q_0+1}^{\infty} (M^n_t(x) - M^n_t(\pi_q x)) D^n_{t,q}
\]
\[
+ \sum_{q=q_0+1}^{\infty} (M^n_t(\pi_q x) - M^n_t(\pi_{q-1} x)) C^n_{t,q}.
\]

On the other hand, by (1)
\[
0 = E[(M^n_t(x) - M^n_t(\pi_{q_0} x)) D^n_{t,q_0} | \mathcal{F}_{n,t-1}]
\]
\[
+ \sum_{q=q_0+1}^{\infty} E[(M^n_t(x) - M^n_t(\pi_q x)) D^n_{t,q} | \mathcal{F}_{n,t-1}]
\]
\[
+ \sum_{q=q_0+1}^{\infty} E[(M^n_t(\pi_q x) - M^n_t(\pi_{q-1} x)) C^n_{t,q} | \mathcal{F}_{n,t-1}].
\]

Now, by (9) and (10)
\[
\left\| \sum_{t=1}^{n} M^n_t(x) - \sum_{t=1}^{n} M^n_t(\pi_{q_0} x) \right\|_{\mathbb{R}^d} \leq I_1 + I_2 + I_1 + I_2 + III,
\]
where

\[ I_1 = \left\| \sum_{t=1}^{\infty} \Delta^*_n(B_{q_0}x)D^*_n\right\|_{\mathbb{R}^d}, \]

\[ I_2 = \left\| \sum_{t=1}^{\infty} n E[\Delta^*_n(B_{q_0}x)D^*_n | \mathcal{F}_{n,t-1}] \right\|_{\mathbb{R}^d}, \]

\[ II_1 = \left\| \sum_{t=1}^{\infty} \sum_{q=q_0+1}^{\infty} \Delta^*_n(B_qx)D^*_n\right\|_{\mathbb{R}^d}, \]

\[ II_2 = \left\| \sum_{t=1}^{\infty} \sum_{q=q_0+1}^{\infty} E[\Delta^*_n(B_qx)D^*_n | \mathcal{F}_{n,t-1}] \right\|_{\mathbb{R}^d}, \]

and

\[ III = \left\| \sum_{t=1}^{\infty} \sum_{q=q_0+1}^{\infty} (M^*_n(\pi_qx) - M^*_n(\pi_{q-1}x))C^*_n - E[(M^*_n(\pi_qx) - M^*_n(\pi_{q-1}x))C^*_n | \mathcal{F}_{n,t-1}] \right\|_{\mathbb{R}^d}. \]

Further, it holds that \( II_1 \leq II_3 + II_2 \) where

\[ II_3 = \left\| \sum_{t=1}^{\infty} \sum_{q=q_0+1}^{\infty} \Delta^*_n(B_qx)D^*_n - E[\Delta^*_n(B_qx)D^*_n | \mathcal{F}_{n,t-1}] \right\|_{\mathbb{R}^d}. \]

From \( \Delta^*_n(B_qx) \leq 2 \| M^*_n(x) \|_{\mathbb{R}^d} \) we have that

\[ \Delta^*_n(B_{q_0}x)D^*_n \leq 2s_n^{-1} |Z_n,t| 1(|Z_n,t| > s_n^{-1}a_{q_0}). \]

Thus, using the last displayed, (1) and (B) it can be easily proved that \( I_1 \) and \( I_2 \) converge in probability to zero for any fixed \( q_0 \). On the other hand, by (6)

\[ II_2 \leq \sup_{q \geq q_0+1} \left\| \sum_{t=1}^{\infty} E[\Delta^*_n(B_qx)D^*_n | \mathcal{F}_{n,t-1}] \right\|_{\mathbb{R}^d} \leq \sum_{q=q_0+1}^{\infty} \frac{2^{-2q}}{a_q} \leq K \sum_{q=q_0+1}^{\infty} 2^{-q(2-\rho)} \sqrt{\log N_{q+1}} \text{ a.s. on the set } \Omega^n_K. \]

As for \( III \), since the partitions are nested

\[ \left| \Delta^*_n(B_qx)D^*_n - E[\Delta^*_n(B_qx)D^*_n | \mathcal{F}_{n,t-1}] \right| \leq 2a_{q-1} \text{ identically} \]
and
\[ \sum_{t=1}^{n} E[|\Delta_t^n(B_q x)|^2 D_{t,q}^n | \mathcal{F}_{n,t-1}] \leq K 2^{-2q} \quad \text{a.s. on the set } \Omega_K^n. \]

It follows from the Freedman’s [10] inequality in Lemma A2, which plays here the same role as the Bernstein’s inequality does in the i.i.d. setup, and Lemma 2.11.17 of van der Vaart and Wellner [23] than for any measurable set \( A \in \mathcal{A}_n \),
\[
E \left| \sum_{t=1}^{n} \Delta_t^n(B_q x) D_{t,q}^n - E[\Delta_t^n(B_q x) D_{t,q}^n | \mathcal{F}_{n,t-1}] \right| 1(A \cap \Omega_K^n) \\
\leq C \left( 2a_{q-1} \log(N_q) + \sqrt{K} 2^{-q} \sqrt{\log(N_q)} \right) \left( P(A) + \frac{1}{N_q} \right) \\
\leq C \left( 2 + \sqrt{K} \right) 2^{-q} \sqrt{\log(N_q)} \left( P(A) + \frac{1}{N_q} \right).
\]

Thus, using the last inequality and defining for every \( q \in \mathbb{N}, q \geq 1 \), a partition \( \{\Omega_{q,k}^n : 1 \leq k \leq N_q \} \) of \( \Omega_n \) such that the maximum
\[
\left\| \sum_{t=1}^{n} \sum_{q=q_0+1}^{\infty} \Delta_t^n(B_q x) D_{t,q}^n - E[\Delta_t^n(B_q x) D_{t,q}^n | \mathcal{F}_{n,t-1}] \right\|_{\mathbb{P}^d}
\]

is achieved at \( B_{q,k} \) on the set \( \Omega_{q,k}^n \). Then, we have
\[
E \left| \Pi_3 \right| 1(\Omega_K^n) \\
\leq E \left\| \sum_{t=1}^{n} \sum_{q=q_0+1}^{\infty} \Delta_t^n(B_q x) D_{t,q}^n - E[\Delta_t^n(B_q x) D_{t,q}^n | \mathcal{F}_{n,t-1}] \right\|_{\mathbb{P}^d} 1(\Omega_K^n) \\
\leq \sum_{q=q_0+1}^{\infty} E \left\| \sum_{t=1}^{n} \Delta_t^n(B_q x) D_{t,q}^n - E[\Delta_t^n(B_q x) D_{t,q}^n | \mathcal{F}_{n,t-1}] \right\|_{\mathbb{P}^d} 1(\Omega_K^n) \\
\leq \sum_{q=q_0+1}^{\infty} \sum_{k=1}^{N_q} E \left\| \sum_{t=1}^{n} \Delta_t^n(B_q x) D_{t,q}^n - E[\Delta_t^n(B_q x) D_{t,q}^n | \mathcal{F}_{n,t-1}] \right\|_{\mathbb{P}^d} 1(\Omega_{q,k}^n \cap \Omega_K^n) \\
\leq C(2 + \sqrt{K}) \sum_{q=q_0+1}^{\infty} \sum_{k=1}^{N_q} 2^{-q \rho} \sqrt{\log(N_q)} \left( P(\Omega_{q,k}^n) + \frac{1}{N_q} \right) \\
\leq C(2 + \sqrt{K}) \sum_{q=q_0+1}^{\infty} 2^{-q \rho} \sqrt{\log(N_q)}.
\]

Finally, for the estimation of \( \Pi_3 \), by the same arguments as for \( \Pi_3 \) we obtain
\[
E \left| \Pi_3 \right| 1(\Omega_K^n) \leq C(2 + \sqrt{K}) \sum_{q=q_0+1}^{\infty} 2^{-q \rho} \sqrt{\log(N_q)}.
\]

The Theorem follows from choosing a large \( K \), a large \( q_0 \) and then, letting \( n \to \infty \).
Proof of Lemma 1. From (A′) and Lemma 4.3 in Park and Whang [20] the pointwise convergence in (A′) is uniform in $x \in \mathbb{R}^d$. Now, we shall show the pointwise convergence of $n^{-1} \sum_{t=1}^{n} E[(\Delta Y_t)^2 1(X_{n,t-1} \leq x)]$. Let us define the partial sum process $W_n(r)$, $r \in [0, 1]$, as

$$W_n(r) = n^{-1/2} \sum_{t=1}^{[nr]} u_t,$$

where $u_t = Y_t - Y_{t-1}, 1 \leq t \leq n$, and $u_0 = Y_0$. Then, it is easy to show that under the conditions of the Lemma 1 $W_n(r)$ converges weakly to the specified limit process $W(r)$. Also, note that $z_n$ in (5) can be written as

$$z_n(x) = \int_0^1 1(W_n(r) \leq x) dW_n(r), \quad x \in \mathbb{R}^d.$$

Then, $\sup_{n,t} E[(\Delta Y_t)^4 \mid \mathcal{F}_{n,t-1}] < \infty$ and (B) imply the weak convergence of $z_n(x)$, see Lemma 3.3 in Park and Whang [20]. Thus, from the weak convergence and (B′)

$$E[z_n(x_1)z_n(x_2)] \xrightarrow{p} E \left[ \int_0^1 1(W(r) \leq x_1) dW(r) \int_0^1 1(W(r) \leq x_2) dW(r) \right]$$

$$= \sigma^2 \int_0^1 P(W(r) \leq x) \, dr.$$

Therefore, Lemma 1 follows. □

Acknowledgments

Research funded by the Spanish Ministerio de Educación y Ciencia, reference numbers SEJ2004-04583/ECON and SEJ2005-07657/ECON. The author thanks the hospitality and the helpful comments received from Professor Winfred Stute during his visit to the Giessen University. The author also would like to thank an anonymous referee for helpful comments and suggestions which have improved a previous version of the paper.

References