APPLICATIONES MATHEMATICAE
39,3 (2012), pp. 273–282

Artur Bryk (Warszawa)

USING RANDOMIZATION TO IMPROVE PERFORMANCE OF A VARIANCE ESTIMATOR OF STRONGLY DEPENDENT ERRORS

Abstract. We consider a fixed-design regression model with long-range dependent errors which form a moving average or Gaussian process. We introduce an artificial randomization of grid points at which observations are taken in order to diminish the impact of strong dependence. We estimate the variance of the errors using the Rice estimator. The estimator is shown to exhibit weak (i.e. in probability) consistency. Simulation results confirm this property for moderate and large sample sizes when randomization is employed.

1. Introduction. Consider a fixed-design regression model (FDR)

$$Y_{i,n} = g(i/n) + \varepsilon_{i,n}, \quad i = 1, \dots, n, \tag{1.1}$$

where $g : [0, 1] \to \mathbb{R}$ is some function with smoothness properties described later. For each n, we observe the random variables $Y_{1,n}, Y_{2,n}, \dots, Y_{n,n}$ and the aim is to estimate the variance of the errors based on this information. Here $(\varepsilon_{i,n})$ is a triangular array such that for each n, the finite sequence $\{\varepsilon_{i,n}\}_{i=1}^n$ is stationary, $\mathbb{E}\varepsilon_{i,n} = 0$, $\mathbb{E}\varepsilon_{i,n}^2 = \sigma_\varepsilon^2 > 0$, $\mathrm{Cov}(\varepsilon_{i,n}, \varepsilon_{j,n}) = r(|i - j|)$, where $r(\cdot)$ is a covariance function which does not depend on n. We assume that $r(k) = L(k)k^{-\alpha}$, $k = 1, \dots, n - 1$, where $0 < \alpha < 1$ is a fixed constant and $L(\cdot)$ is a function defined on $[0,\infty)$, slowly varying at infinity and positive in some neighbourhood of infinity. The array $(\varepsilon_{i,n})$ is long-range dependent (LRD) in the sense that $\sum_{k=1}^\infty |r(k)| = \infty$.

In a nonparametric setting the regression function g at a given point x is usually estimated by one of many methods involving local polynomials, smoothing splines or kernel estimators. Any of these methods weighs

---

2010 *Mathematics Subject Classification*: 62G08, 62G20, 62M10.
Key words and phrases: Rice estimator, long-range dependence, randomization, linear process, fixed-design regression.

DOI: 10.4064/am39-3-2

concomitants of grid points around x in such a way that those closer to x contribute more to the value of the estimator. As the concomitants corresponding to a small neighbourhood of x form a block of consecutive observations which are strongly dependent, the resulting estimator is more variable than in a weakly dependent case. In order to alleviate the effect of dependence on variability of the regression estimator we consider a randomly chosen permutation $\sigma = \sigma_n$ of $\{1, \ldots, n\}$ and assume that the observations are taken consecutively at the points $\sigma(1)/n, \sigma(2)/n, \ldots, \sigma(n)/n$ instead of $1/n, 2/n, \ldots, 1$. As dependence of the observations reflects solely the temporal order in which they are taken, the appropriate model of this observational scheme is

$$Y_{i,n} = g\left(\frac{\sigma_n(i)}{n}\right) + \varepsilon_{i,n}, \qquad i = 1, \ldots, n. \tag{1.2}$$

The random permutation σ_n is chosen independently of $(\varepsilon_{i,n})$. We will refer to (1.2) as the *Randomized Fixed-Design Regression model* (RFDR). The idea of considering (1.2) is based on the observation made in [6] that the regression estimators in a random design regression model with LRD errors are less variable than in the fixed-design case, and is in line with a general discussion in [8]. For a thorough discussion of the influence of design type on regression estimation with LRD errors see [5]. We stress that plausibility of model (1.2) is based on the assumption that the dependence between the observations is due to their temporal and not spatial proximity. Thus, for example, dependence of two consecutive observations ($t = i, i+1$) will be the same regardless of grid points at which the observations are taken. Another insight into advantages of randomization can be found in [2].

In the following we suppress the dependence of $Y_{i,n}$ and $\varepsilon_{i,n}$ on n in notation.

In particular, consider the case when (ε_i) is a one-sided moving average process given by

$$\varepsilon_i = \sum_{t=0}^{\infty} c_t \eta_{i-t}, \qquad i = 1, \ldots, n, \tag{1.3}$$

where $(\eta_t)_{t=-\infty}^{\infty}$ is a sequence of independent, identically distributed innovations such that $\mathbb{E}\eta = 0$, $\mathbb{E}(\eta^2) = \sigma_\eta^2 < \infty$ and the c_t satisfy $\sum_{t=0}^{\infty} c_t^2 < \infty$. Let

$$c_t = L_c(t) t^{-\beta} \quad (c_0 = 1), \tag{1.4}$$

where $1/2 < \beta < 1$ and $L_c(\cdot)$ is a function defined on $[0, +\infty)$, slowly varying at infinity and positive in some neighbourhood of infinity. A routine calculation based on the Karamata theorem (see [7, p. 281]) implies that $r(k) \sim \sigma_\eta^2 C(\beta) L_c^2(k) k^{-\alpha}$, where $C(\beta) = \int_0^\infty (x + x^2)^{-\beta}\, dx$ and $\alpha = 2\beta - 1$. Thus, in this case, the sum of the absolute values of the covariances diverges.

Put $c_t = 0$ for $t < 0$. Then it is easily seen that

$$\operatorname{Var}\Big(\sum_{i=1}^{n} \varepsilon_i\Big) = \sigma_\eta^2 \sum_{k=-\infty}^{n} \Big(\sum_{t=1}^{n} c_{t-k}\Big)^2 \sim \sigma_\eta^2 D(\beta) n^{2-\alpha} L_c^2(n),$$

where $D(\beta) = [(2-2\beta)(3/2-\beta)]^{-1} C(\beta)$.

Another model is a Gaussian LRD sequence with $r(i) = L(i) i^{-\alpha}$, $0 < \alpha < 1$, and hence $\operatorname{Var}(\varepsilon_1 + \cdots + \varepsilon_n) \sim D(\alpha) L(n) n^{2-\alpha}$, where $D(\alpha) = 2/[(1-\alpha)(2-\alpha)]$.

Let Y_i, $i = 1, \ldots, n$, be the observations in the RFDR model defined in (1.2). We will use the Rice estimator (see [10]) to estimate the variance of the errors σ_ε^2. In the RFDR model it is defined as follows:

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (\bar{Y}_{i+1} - \bar{Y}_i)^2, \tag{1.5}$$

where $\bar{Y}_i = Y_{\sigma_n^{-1}(i)}$. We investigate conditions under which $\hat{\sigma}_\varepsilon^2$ is weakly consistent.

The paper concludes with a simulation example showing the effect of randomization in practice. It indicates that randomization has a non-negligible impact on the values of the Rice estimator for moderate and large sample sizes.

2. Results. We consider first how the randomization introduced affects the properties of long-range dependent errors.

PROPOSITION 1. *Let* $\bar{\varepsilon}_{i,n} = \varepsilon_{\sigma_n^{-1}(i)}$, $i = 1, \ldots, n$, $n \in \mathbb{N}$. *Then for the RFDR model:*

(i) $(\bar{\varepsilon}_{i,n})$, $i \le n$, $n \in \mathbb{N}$, *is a rowwise exchangeable array of random variables;*

(ii) $\operatorname{Cov}(\bar{\varepsilon}_{i,n}, \bar{\varepsilon}_{j,n}) \sim \dfrac{2L(n)n^{-\alpha}}{(1-\alpha)(2-\alpha)}$ *for* $i \neq j$ *and* $\operatorname{Var}(\bar{\varepsilon}_{i,n}) = \operatorname{Var}(\varepsilon_{i,n})$.

Point (ii) shows a fundamental advantage of randomization: large covariances of the initial sequence, $r(k) = L(k)k^{-\alpha}$ for small k, decrease to the level $CL(n)n^{-\alpha}$ in the randomized sequence.

Property (ii) follows by noting that $\operatorname{Cov}(\bar{\varepsilon}_{i,n}, \bar{\varepsilon}_{j,n})$ equals, for $i \neq j$,

$$\frac{1}{n(n-1)} \sum_{1 \le k \neq l \le n} \operatorname{Cov}(\bar{\varepsilon}_{i,n}, \bar{\varepsilon}_{j,n} \mid \sigma(k) = i,\, \sigma(l) = j)$$
$$= \frac{1}{n(n-1)} \sum_{1 \le k \neq l \le n} \operatorname{Cov}(\varepsilon_{k,n}, \varepsilon_{l,n}),$$

where the last equality follows from the independence of σ_n and $(\varepsilon_{i,n})$. A routine application of the Karamata theorem yields (ii).

Note that the Rice estimator given by (1.5) has the representation

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{2(n-1)}\bigg\{\sum_{i=1}^{n-1}\bigg(g\Big(\frac{i+1}{n}\Big) - g\Big(\frac{i}{n}\Big)\bigg)^2 + \sum_{i=1}^{n-1}(\bar{\varepsilon}_{i+1}-\bar{\varepsilon}_i)^2 + 2\sum_{i=1}^{n-1}\bigg(g\Big(\frac{i+1}{n}\Big) - g\Big(\frac{i}{n}\Big)\bigg)(\bar{\varepsilon}_{i+1}-\bar{\varepsilon}_i)\bigg\} =: \Theta_1+\Theta_2+\Theta_3. \tag{2.1}$$

We will prove below that Θ_1 and Θ_3 tend to 0 in probability as $g(\cdot)$ is Lipschitz, whereas $\Theta_2 \to \sigma_\varepsilon^2 = \mathbb{E}\varepsilon^2$ in probability. It will follow from the proofs that Θ_1 and Θ_3 behave just as the analogously defined Rice estimator in the FDR model, whereas, in this case, Θ_2 is of the form $\frac{1}{2(n-1)}\sum_{i=1}^{n-1}(\varepsilon_{i+1}-\varepsilon_i)^2$ and, in view of the ergodic theorem, tends a.s. to $2^{-1}\mathbb{E}(\varepsilon_2-\varepsilon_1)^2 = r(0) - r(1) = \sigma_\varepsilon^2 - r(1) \neq \sigma_\varepsilon^2$ if $r(1)\neq 0$. Thus, using randomization in the FDR model, we can construct a weakly consistent estimator of $\hat{\sigma}_\varepsilon^2$.

We consider first the case of positively correlated, LRD Gaussian errors, and then the case of a one-sided moving average process $(\varepsilon_i)_{i\in\mathbb{N}}$.

Theorem 1. *Let $(\varepsilon_i)_{i\in\mathbb{N}}$ be a Gaussian stationary sequence such that $r(k)\geq 0$ for $k\in\mathbb{N}$ and $r(k) = L(k)k^{-\alpha}$ for $0<\alpha<1$. Assume that g is Lipschitz continuous supported on $[0,1]$. Then $\hat{\sigma}_\varepsilon^2 \to \sigma_\varepsilon^2$ in probability.*

Theorem 2. *Let $(\varepsilon_i)_{i\in\mathbb{N}}$ be a linear process defined in (1.3) with coefficients given in (1.4). Assume that g is Lipschitz continuous supported on $[0,1]$. Then $\hat{\sigma}_\varepsilon^2 \to \sigma_\varepsilon^2$ in probability.*

3. Proofs. In all proofs C denotes a generic constant whose value may change.

Recall that $\bar{\varepsilon}_j = \bar{\varepsilon}_{j,n} = \varepsilon_{\sigma_n^{-1}(j)}$. In order to prove Theorem 1 we will need some properties of the moments of the randomized errors.

Lemma 1. *Let (ε_i) satisfy the assumptions of Theorem 1. Moreover, $i_1,\dots,i_l\in\mathbb{N}$ are different indices and $k_1,\dots,k_l\in\mathbb{N}$. Then*

(i) $\mathbb{E}(\bar{\varepsilon}_{i_1}^{k_1}\dots\bar{\varepsilon}_{i_l}^{k_l}) = 0$ *when* $k_1+\dots+k_l = 2k+1$ *with* $k\in\mathbb{N}$*;*

(ii) $\mathbb{E}(\bar{\varepsilon}_{i_1}^{k_1}\dots\bar{\varepsilon}_{i_l}^{k_l}) = \mathcal{O}(a_n^{2\lceil s/2\rceil})$ *when* $k_1+\dots+k_l = 2k$ *and* $s = \#\{k_j,\ 1\leq j\leq l : k_j = 1\}$, *where* $a_n^2 = \mathrm{Var}(n^{-1}\sum_{i=1}^n \varepsilon_i) = \mathrm{Var}(n^{-1}\sum_{i=1}^n\bar{\varepsilon}_i) \sim 2((1-\alpha)(2-\alpha))^{-1}L(n)n^{-\alpha}$.

The above lemma was proved in [3].

Proof of Theorem 1. We use representation (2.1). It is easy to check that $|\Theta_1| = \mathcal{O}(n^{-2})$ as g is Lipschitz. Using the Schwarz inequality, we have

$|\Theta_3| \le 2\Theta_1^{1/2}\Theta_2^{1/2}$. Thus if $\Theta_2 = \mathcal{O}_P(1)$ then $\Theta_3 \xrightarrow{\mathcal{P}} 0$, where $\xrightarrow{\mathcal{P}}$ denotes convergence in probability.

Note that a Gaussian stationary process such that $r(k) \to 0$ is mixing (see [4, Theorem 14 § 2.2]), so (ε_i^2) is mixing, and thus ergodic. Hence $n^{-1}\sum_{i=1}^n \bar\varepsilon_i^2 = n^{-1}\sum_{i=1}^n \varepsilon_i^2 \xrightarrow{\mathcal{P}} \sigma_\varepsilon^2$, which implies that $\frac{1}{n-1}\sum_{i=1}^{n-1}\bar\varepsilon_i^2 \xrightarrow{\mathcal{P}} \sigma_\varepsilon^2$. Thus it is enough to prove that

$$\frac{1}{n-1}\sum_{i=1}^{n-1}\bar\varepsilon_i\bar\varepsilon_{i+1} \xrightarrow{\mathcal{P}} 0. \tag{3.1}$$

The second moment of the lhs of (3.1) equals

$$\frac{1}{(n-1)^2}\sum_{k,i=1}^{n-1}\mathbb{E}(\bar\varepsilon_i\bar\varepsilon_{i+1}\bar\varepsilon_k\bar\varepsilon_{k+1}).$$

Let $m = \#\{i, i+1, k, k+1\}$. By Lemma 1 the terms of the last sum are $\mathcal{O}(L^2(n)n^{-2\alpha})$ for $m = 4$ and $\mathcal{O}(L(n)n^{-\alpha})$ for $m = 3$. By the Schwarz inequality $\mathbb{E}(\bar\varepsilon_i^2\bar\varepsilon_{i+1}^2)$ is bounded by a constant which does not depend on i. Hence

$$\mathbb{E}\biggl(\frac{1}{n-1}\sum_{i=1}^{n-1}\bar\varepsilon_i\bar\varepsilon_{i+1}\biggr)^2 = \mathcal{O}(n^{-2}(n^2L^2(n)n^{-2\alpha} + nL(n)n^{-\alpha} + Cn)) = o(1).$$

Thus (3.1) follows via the Markov inequality. ∎

Proof of Theorem 2. We follow the same argument after noting that, in view of Theorem 1.3.3 in [11], (ε_i^2) is ergodic. Thus, in order to prove (3.1), it is enough to show that the second moment of the lhs of (3.1) tends to 0. We will exploit the bound

$$\begin{aligned}\sum_{1\le k\ne l\le n}\sum_{i=-\infty}^{\infty} c_{k-i}^2 c_{l-i}^2 &= 2\sum_{1\le k<l\le n}\sum_{i=0}^{\infty} c_i^2 c_{i+(l-k)}^2 \\ &\le 2\sum_{k=1}^{n}\sum_{i=0}^{\infty} c_i^2 \sum_{j=1}^{n} c_{i+j}^2 = \mathcal{O}(n).\end{aligned} \tag{3.2}$$

Note that

$$\mathbb{E}\biggl(\frac{1}{n-1}\sum_{t=1}^{n-1}\bar\varepsilon_t\bar\varepsilon_{t+1}\biggr)^2 = \frac{1}{(n-1)^2}\sum_{1\le t,s\le n-1} W_{t,s},$$

where

$$W_{t,s} = \sum_{i_1,i_2,i_3,i_4=-\infty}^{\infty}\mathbb{E}(c_{\sigma_n^{-1}(t)-i_1}c_{\sigma_n^{-1}(t+1)-i_2}c_{\sigma_n^{-1}(s)-i_3}c_{\sigma_n^{-1}(s+1)-i_4})\mathbb{E}(\eta_{i_1}\eta_{i_2}\eta_{i_3}\eta_{i_4}).$$

Let $W_{t,s} = T_l$, where $l = \#\{\{t, t+1\} \cap \{s, s+1\}\}$. We first consider T_2:

$$\begin{aligned}
T_2 = & \sum_{i=-\infty}^{\infty} \mathbb{E}(c^2_{\sigma_n^{-1}(t)-i} c^2_{\sigma_n^{-1}(t+1)-i}) \mathbb{E}\eta_1^4 + \sum_{i\neq j} \mathbb{E}(c^2_{\sigma_n^{-1}(t)-i} c^2_{\sigma_n^{-1}(t+1)-j}) \mathbb{E}(\eta_1^2)^2 \\
& + 2\sum_{i\neq j} \mathbb{E}(c_{\sigma_n^{-1}(t)-i} c_{\sigma_n^{-1}(t+1)-i} c_{\sigma_n^{-1}(t)-j} c_{\sigma_n^{-1}(t+1)-j}) \mathbb{E}(\eta_1^2)^2 \\
=: & \Psi_1 + \Psi_2 + \Psi_3.
\end{aligned}$$

Note that using (3.2) we get

$$\Psi_1 = \frac{1}{n(n-1)} \sum_i \sum_{1\le k\neq l\le n} c^2_{k-i} c^2_{l-i} \mathbb{E}\eta_1^4 = \mathcal{O}\Big(\frac{1}{n}\Big).$$

Next, we show that Ψ_2 and Ψ_3 are bounded by a constant which does not depend on t. Namely,

$$\Psi_2 = \frac{1}{n(n-1)} \sum_{i\neq j} \sum_{1\le k\neq l\le n} c^2_{k-i} c^2_{l-j} \mathbb{E}(\eta_1^2)^2 \le \Big(\sum_i c_i^2\Big)^2 \mathbb{E}(\eta_1^2)^2 = \mathcal{O}(1)$$

and, by the Schwarz inequality,

$$\begin{aligned}
|\Psi_3| &\le \frac{2\mathbb{E}(\eta_1^2)^2}{n(n-1)} \sum_{i\neq j} \sum_{1\le k\neq l\le n} |c_{k-i} c_{l-i} c_{k-j} c_{l-j}| \\
&\le \frac{2\mathbb{E}(\eta_1^2)^2}{n(n-1)} \sum_{k\neq l} \Big(\sum_i |c_{k-i} c_{l-i}|\Big)^2 \\
&\le \frac{2\mathbb{E}(\eta_1^2)^2}{n(n-1)} \sum_{k\neq l} \Big(\sum_i c^2_{k-i}\Big)\Big(\sum_i c^2_{l-i}\Big) = \mathcal{O}(1).
\end{aligned}$$

Thus we have $T_2 = \mathcal{O}(1)$. Next, we show that T_1 is also $\mathcal{O}(1)$. Note that

$$\begin{aligned}
T_1 = & \sum_{i=-\infty}^{\infty} \mathbb{E}(c_{\sigma_n^{-1}(t)-i} c^2_{\sigma_n^{-1}(t+1)-i} c_{\sigma_n^{-1}(t+2)-i}) \mathbb{E}\eta_1^4 \\
& + 2\sum_{i\neq j} \mathbb{E}(c_{\sigma_n^{-1}(t)-i} c_{\sigma_n^{-1}(t+1)-i} c_{\sigma_n^{-1}(t+1)-j} c_{\sigma_n^{-1}(t+2)-j}) \mathbb{E}(\eta_1^2)^2 \\
& + \sum_{i\neq j} \mathbb{E}(c_{\sigma_n^{-1}(t)-i} c^2_{\sigma_n^{-1}(t+1)-j} c_{\sigma_n^{-1}(t+2)-i}) \mathbb{E}(\eta_1^2)^2 =: \overline{\Psi}_1 + \overline{\Psi}_2 + \overline{\Psi}_3.
\end{aligned}$$

Using the fact that $ab \le \frac{1}{2}(a^2+b^2)$ and (3.2), we have

$$\begin{aligned}
|\overline{\Psi}_1| &\le \frac{1}{n(n-1)(n-2)} \sum_i \sum_{1\le k\neq l\neq p\le n} |c_{k-i} c^2_{l-i} c_{p-i}| \mathbb{E}\eta_1^4 \\
&\le \frac{1}{n(n-1)(n-2)} \sum_{k\neq l\neq p} \frac{1}{2}\Big(\sum_i c^2_{k-i} c^2_{l-i} + \sum_i c^2_{l-i} c^2_{p-i}\Big) \mathbb{E}\eta_1^4 = \mathcal{O}\Big(\frac{1}{n}\Big).
\end{aligned}$$

Moreover

$$|\overline{\Psi}_2| = \frac{2}{n(n-1)(n-2)} \sum_{i\neq j} \sum_{1\le k\neq l\neq p\le n} |c_{k-i}c_{l-i}c_{l-j}c_{p-j}|\mathbb{E}(\eta_1^2)^2$$
$$\le \frac{2}{n(n-1)(n-2)} \sum_{k\neq l\neq p} \frac{1}{2}\sum_{i,j}(c_{k-i}^2c_{l-j}^2 + c_{l-i}^2c_{p-j}^2)\mathbb{E}(\eta_1^2)^2 = \mathcal{O}(1),$$

since $\sum_i c_i^2 < \infty$. Using a similar reasoning we get $|\overline{\Psi}_3| = \mathcal{O}(1)$.

Consider now the case of different indices $t, t+1, s, s+1$. We have

$$T_0 = \sum_{i=-\infty}^{\infty} \mathbb{E}(c_{\sigma_n^{-1}(t)-i}c_{\sigma_n^{-1}(t+1)-i}c_{\sigma_n^{-1}(s)-i}c_{\sigma_n^{-1}(s+1)-i})\mathbb{E}\eta_1^4$$
$$+3\sum_{i\neq j} \mathbb{E}(c_{\sigma_n^{-1}(t)-i}c_{\sigma_n^{-1}(t+1)-i}c_{\sigma_n^{-1}(s)-j}c_{\sigma_n^{-1}(s+1)-j})\mathbb{E}(\eta_1^2)^2 =: \widetilde{\Psi}_1 + \widetilde{\Psi}_2.$$

A reasoning analogous to that for $\overline{\Psi}_1$ yields $|\widetilde{\Psi}_1| = \mathcal{O}(n^{-1})$. Moreover

$$\widetilde{\Psi}_2 = \frac{3\mathbb{E}(\eta_1^2)^2}{n(n-1)(n-2)(n-3)} \sum_{1\le k\neq l\neq p\neq q\le n} \sum_{i\neq j} c_{k-i}c_{l-i}c_{p-j}c_{q-j}. \tag{3.3}$$

The rhs of (3.3) can be written as

$$\sum_{k\neq l\neq p\neq q} = \sum_{k,l,p,q} - \sum_{k=l=p=q} -4\sum_{k\neq l=p=q} -3\sum_{k=l\neq p=q} -6\sum_{k\neq l\neq p=q} =: \Upsilon_1 - \Upsilon_2 - \Upsilon_3 - \Upsilon_4 - \Upsilon_5. \tag{3.4}$$

Note that

$$\Upsilon_1 = \frac{3\mathbb{E}(\eta_1^2)^2}{n(n-1)(n-2)(n-3)}\Big\{\Big(\sum_{k,l}\sum_i c_{k-i}c_{l-i}\Big)^2 - \sum_{k,l,p,q}\sum_i c_{k-i}c_{l-i}c_{p-i}c_{q-i}\Big\}$$
$$= \mathcal{O}\bigg(\frac{1}{n^4}\{(L_c(n)n^{2-\alpha})^2 + n^3\}\bigg) = \mathcal{O}(L_c^2(n)n^{-2\alpha} + n^{-1}).$$

Moreover

$$|\Upsilon_2| \le \frac{3\mathbb{E}(\eta_1^2)^2}{n(n-1)(n-2)(n-3)} \sum_{k=1}^{n}\Big(\sum_i c_{k-i}^2\Big)\Big(\sum_j c_{k-j}^2\Big) = \mathcal{O}\Big(\frac{1}{n^3}\Big).$$

A similar reasoning to the case of Ψ_2 and Ψ_3 implies $|\Upsilon_4| = \mathcal{O}(n^{-2})$. The other terms of (3.4) are estimated analogously to $\overline{\Psi}_2$. We find that $|\Upsilon_3| = \mathcal{O}(n^{-2})$ and $|\Upsilon_5| = \mathcal{O}(n^{-1})$. Thus $T_0 = \mathcal{O}(L_c^2(n)n^{-2\alpha} + n^{-1})$. Hence

$$\mathbb{E}\bigg(\frac{1}{n-1}\sum_{i=1}^{n-1}\bar{\varepsilon}_i\bar{\varepsilon}_{i+1}\bigg)^2 = \mathcal{O}\bigg(\frac{1}{n^2}(n^2(L_c^2(n)n^{-2\alpha} + n^{-1}) + Cn)\bigg) = o(1)$$

and (3.1) follows via the Markov inequality. ■

4. Simulation results. We conducted a simulation study to investigate the effect of randomization of a fixed design regression in practice. We generated series (Y_i) of length $n = 500, 1000$ and 10000 with trend functions

(i) $g_1(x) = 2\sin(4\pi x)$;
(ii) $g_2(x) = 2 - 5x + 5\exp\{-100(x-0.5)^2\}$.

These regression functions were also used in [9]. Moreover, we present an estimation for non-Lipschitz trend functions to check the importance of the Lipschitz assumption.

The errors considered follow a fractional autoregressive integrated moving average process FARIMA$(0,d,0)$ with $d = 0, 0.1, 0.2, 0.3, 0.4$. It is known that $L(n) \sim C$ and a one-sided moving average representation exists in this case. For a FARIMA$(0,d,0)$ process $\varepsilon_t = (1-B)^{-d}\eta_t$, where (η_t) is a Gaussian white noise with marginal variance σ_η^2 and $B\eta_t = \eta_{t-1}$, we have $C = \sigma_\eta^2 \Gamma(1-2d)/\Gamma(d)\Gamma(1-d)$. We refer to [1] for more information on this process.

The number of replications of each experiment was 1000.

Table 1 indicates that, in the RFDR model, the average values of the Rice estimator correspond to theoretical values for $d \le 0.3$ even if $n = 500$. However, in the FDR model, where the Rice estimator estimates $\sigma_\varepsilon^2 - r(1)$, we observe a decrease in the mean value of the estimator with increasing d which is caused by an increase of the covariance $r(1)$. Moreover, the accuracy of estimation of both g_1 and g_2 is the same in both models.

Table 1. Means and standard deviations of Rice estimator in FDR and RFDR models

		$n = 500$				$n = 1000$				$n = 10000$			
		g_1		g_2		g_1		g_2		g_1		g_2	
d		FDR	RFDR	FDR	RFDR	FDR	RFDR	FDR	RFDR	FDR	RFDR	FDR	RFDR
0	mean	0.9986	0.9976	1.0027	1.0008	1.0003	1.0014	0.9985	1.0000	1.0003	1.0003	0.9990	0.9994
	SE	0.0781	0.0767	0.0779	0.0752	0.0560	0.0535	0.0537	0.0548	0.0174	0.0178	0.0172	0.0171
0.1	mean	0.9085	1.0135	0.9078	1.0144	0.9041	1.0148	0.9057	1.0162	0.9074	1.0193	0.9053	1.0178
	SE	0.0702	0.0800	0.0672	0.0816	0.0486	0.0565	0.0485	0.0568	0.0157	0.0185	0.0151	0.0182
0.2	mean	0.8253	1.0804	0.8207	1.0743	0.8234	1.0808	0.8260	1.0855	0.8236	1.0940	0.8242	1.0947
	SE	0.0625	0.0954	0.0583	0.0908	0.0445	0.0646	0.0441	0.0653	0.0137	0.0205	0.0138	0.0221
0.3	mean	0.7523	1.2213	0.7518	1.2181	0.7533	1.2446	0.7503	1.2434	0.7522	1.2876	0.7525	1.2878
	SE	0.0557	0.1357	0.0546	0.1244	0.0392	0.0990	0.0379	0.0980	0.0123	0.0396	0.0122	0.0397
0.4	mean	0.6895	1.5090	0.6912	1.5268	0.6902	1.5840	0.6887	1.5720	0.6902	1.7678	0.6900	1.7686
	SE	0.0486	0.2235	0.0488	0.2387	0.0352	0.2057	0.0353	0.1921	0.0106	0.1208	0.0111	0.1240

Table 2 indicates that, in both models, the accuracy of estimation does not change for all non-Lipschitz regression functions. The estimation results are similar for g_3, g_4, g_5 and g_6. The accuracy gets worse when the trend function is not continuous, but is still quite good for g_7 and g_8. However, when we use g_9 and g_{10} then the average values of the Rice estimator are very high and do not correspond to the theoretical values at all.

Table 2. Mean values of Rice estimator in FDR and RFDR models for non-Lipschitz trend functions supported on $[0,1]$ ($n = 1000$)

	g_3		g_4		g_5		g_6	
d	FDR	RFDR	FDR	RFDR	FDR	RFDR	FDR	RFDR
0	0.9993	1.0008	0.9988	1.0025	0.9972	0.9975	1.0003	0.9988
0.1	0.9078	1.0172	0.9067	1.0164	0.9048	1.0179	0.9055	1.0183
0.2	0.8247	1.0851	0.8227	1.0829	0.8234	1.0849	0.8242	1.0846
0.3	0.7514	1.2427	0.7504	1.2451	0.7510	1.2424	0.7512	1.2436
0.4	0.6884	1.5936	0.6885	1.5902	0.6895	1.5970	0.6907	1.5743

	g_7		g_8		g_9		g_{10}	
d	FDR	RFDR	FDR	RFDR	FDR	RFDR	FDR	RFDR
0	1.0035	1.0045	1.0496	1.0520	5.7048	5.6985	936.7202	936.6520
0.1	0.9116	1.0242	0.9604	1.0710	5.6136	5.7233	936.6218	936.8044
0.2	0.8303	1.0903	0.8794	1.1380	5.5259	5.7816	936.5948	936.8704
0.3	0.7585	1.2536	0.8072	1.2950	5.4562	5.9519	936.5458	936.9214
0.4	0.6942	1.5933	0.7447	1.6439	5.3922	6.2967	936.4752	937.2900

Here

$$g_3(x) = -2\sqrt{1-x}, \qquad g_4(x) = \sqrt[4]{(x-0.5)^2},$$
$$g_5(x) = x\log x, \quad x \neq 0, \qquad g_5(0) = 0,$$
$$g_6(x) = x^{3/2}\sin 1/x, \quad x \neq 0, \qquad g_6(0) = 0,$$
$$g_7(x) = \begin{cases} 0, & x = 0, \\ 1/\log x, & x \in (0, 0.5], \\ 1/x, & x \in (0.5, 1], \end{cases}$$
$$g_8(x) = \begin{cases} 0, & x = 0, \\ 1/\log x, & x \in (0, 0.1], \\ 1/x, & x \in (0.1, 1], \end{cases} \qquad g_9(x) = \begin{cases} 0, & x = 0, \\ 1/\log x, & x \in (0, 0.01], \\ 1/x, & x \in (0.01, 1], \end{cases}$$
$$g_{10}(x) = 1/x, \quad x \neq 0, \qquad g_{10}(0) = 0.$$

The theoretical marginal variance for a FARIMA$(0, d, 0)$ process with $\sigma_\eta^2 = 1$ is

$$\sigma_f^2 := \frac{\Gamma(1-2d)}{\Gamma^2(1-d)} = \begin{cases} 1 & \text{if } d = 0, \\ 1.0195 & \text{if } d = 0.1, \\ 1.0987 & \text{if } d = 0.2, \\ 1.3165 & \text{if } d = 0.3, \\ 2.0701 & \text{if } d = 0.4. \end{cases}$$

The theoretical values of $\sigma_f^2 - r(1)$ for a FARIMA$(0,d,0)$ process are

$$\sigma_f^2 - r(1) = \frac{\Gamma(1-2d)\Gamma(1+d)}{\Gamma(1-d)\Gamma(2-d)\Gamma(d)} = \begin{cases} 1 & \text{if } d = 0, \\ 0.9062 & \text{if } d = 0.1, \\ 0.8240 & \text{if } d = 0.2, \\ 0.7523 & \text{if } d = 0.3, \\ 0.6900 & \text{if } d = 0.4. \end{cases}$$

References

[1] J. Beran, *Statistics for Long-Memory Processes*, Chapman & Hall, London, 1994.

[2] A. Bryk and J. Mielniczuk, *Using randomization to improve performance of regression estimators under dependence*, Acta Sci. Math. (Szeged) 73 (2007), 817–838.

[3] A. Bryk and J. Mielniczuk, *Randomized fixed design regression under long-range dependent errors*, Comm. Statist. 37 (2008), 520–531.

[4] I. Cornfeld, S. Fomin and Y. Sinai, *Ergodic Theory*, Springer, 1982.

[5] S. Csörgő, *Smoothing dichotomy in nonparametric regression under long-range dependent errors*, Statist. Neerlandica 56 (2002), 132–142.

[6] S. Csörgő and J. Mielniczuk, *Random design regression under long-range dependent errors*, Bernoulli 5 (1999), 209–224.

[7] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 2, Wiley, New York, 1971.

[8] H. Künsch, J. Beran and F. Hampel, *Contrasts under long-range dependent correlations*, Ann. Statist. 21 (1993), 943–964.

[9] B. K. Ray and R. S. Tsay, *Iterative bandwidth estimation for nonparametric regression with long-range dependent errors*, in: Athens Conference on Applied Probability and Time Series, Vol. 2, P. M. Robinson and M. Rosenblatt (eds.), Lecture Notes in Statist. 115, Springer, 1996, 339–351.

[10] J. Rice, *Bandwidth choice for nonparametric regression*, Ann. Statist. 12 (1984), 1215–1230.

[11] M. Taniguchi and Y. Kakizawa, *Asymptotic Theory of Statistical Inference for Time Series*, Springer, 2000.

Artur Bryk
Department of Mathematics and Mathematical Economics
Warsaw School of Economics
Al. Niepodległości 162
02-554 Warszawa, Poland
E-mail: abryk@sgh.waw.pl

Received on 30.5.2011;
revised version on 28.7.2011 (2092)