Discussion Paper #2005-04

Nonparametric Estimation of Conditional Medians for Linear and Related Processes

(The former title is Least Absolute Deviation Regression for Long-range Dependent Processes)

Toshio HONDA

September 2005
October 2007(revised)

NONPARAMETRIC ESTIMATION OF CONDITIONAL MEDIANS FOR LINEAR AND RELATED PROCESSES

Toshio Honda

Graduate School of Economics, Hitotsubashi University

2-1 Naka, Kunitachi, Tokyo 186-8601, JAPAN

NONPARAMETRIC MEDIAN ESTIMATION

Abstract. We consider nonparametric estimation of conditional medians for time series data. The time series data are generated from two mutually independent linear processes. The linear processes may show long-range dependence. The estimator of the conditional medians is based on minimizing the locally weighted sum of absolute deviations for local linear regression. We present the asymptotic distribution of the estimator. The rate of convergence is independent of regressors in our setting. The result of a simulation study is also given.

Key words and phrases: Local linear estimator, least absolute deviation regression, conditional quantiles, linear processes, short-range dependence, long-range dependence, random design, martingale CLT, simulation study.

1. Introduction

Let $\{(X_i, Y_i)\}_{i=1}^{\infty}$ be a stationary bivariate process generated by

$$Y_i = u(X_i) + V_i, \tag{1.1}$$

where $V_i = V(X_i, Z_i)$, $\{X_i\}_{i=1}^{\infty}$ and $\{Z_i\}_{i=1}^{\infty}$ are mutually independent stationary linear processes, and the conditional median of V_i on X_i is 0. We consider the estimation of the conditional median of Y_i on $X_i = x_0$, $u(x_0)$, by local linear LAD (least absolute deviation) regression without any parametric assumptions on $u(x)$.

We specify $\{X_i\}_{i=1}^{\infty}$ and $\{Z_i\}_{i=1}^{\infty}$ later in this section. Technical assumptions on $V(x, z)$ will be stated in Section 2. Note that we can incorporate some heteroscedasticity into error terms, for example, $V(x, z) = \sigma(x)G(z - m)$, where m is the median of Z_i and $G(z)$ is a symmetric function. $G(Z_i - m)$ may have infinite variance. We deal with only conditional medians in this paper for simplicity of presentation. However, the same arguments apply to other conditional quantiles.

Nonparametric regression is often used, for example, when no parametric assumption on regression functions is available or when we want to check the parametric assumptions. There are so much literature on nonparametric regression that we cannot name all of them and we mention only recent or relevant papers. See Fan and Gijbels (1996) and Härdle et al. (2004) for surveys. As for nonparametric regression for weakly dependent observations, see Nze et al. (2002) and Fan and Yao (2003).

There have been a lot of studies on quantile regression for linear models since Koenker and Basset (1978). See Koenker (2005) for recent developments of quantile regression. Note that Quantile regression is reduced to LAD regression when we estimate conditional medians. Pollard (1991) presented a simple proof of the asymptotic normality of regression coefficient estimators.

Chaudhuri (1991) considered nonparametric estimation of conditional quantiles and Fan et al. (1994) applied the method of Pollard (1991) to nonparametric robust estimation including nonparametric estimation of conditional quantiles. We examine the estimator of Chaudhuri (1991) in our setting by applying the method of Pollard (1991).

The results of Chaudhuri (1991) are for i.i.d. observations. Many authors considered cases of weakly dependent observations and studied the asymptotic properties of the estimators since Chaudhuri (1991). For example, Truong and Stone (1992) considered local medians for α-mixing processes. Honda (2000a) and Hall et al. (2002) examined the asymptotic properties of the estimator of Chaudhuri (1991). Hall et al. (2002) also employed the method of Pollard (1991). Zhao and Wu (2006) considered another setting from α-mixing processes. The asymptotic distributions in the above papers are the same as for i.i.d. observations under random design.

Following the recent developments of research on time series with long-range dependence, some authors investigated robust or nonparametric estimation of regression functions for time series with long-range dependence. See Beran (1994), Chapter 5 of Taniguchi and Kakizawa (2000), Robinson (2003), and Doukhan et al. (2003) for empirical and theoretical studies of time series with long-range dependence. We give a brief exposition on long-range dependence later in this section.

Koul and Mukherjee (1993) and Giraitis et al. (1996) considered robust estimation for linear models with long-range dependent errors. Koul et al. (2001) examined cases of errors with infinite variance and long-range dependence. As for nonparametric estimation of conditional mean functions for time series with long-range dependence, there are, for example, Robinson (1997), Hidalgo (1997), Csörgő and Mielniczuk (2000), Mielniczuk and Wu (2004) and Guo and Koul (2007). Wu and Mielniczuk (2002) fully examined the asymptotic properties of kernel density estimators. Honda (2000b) considered nonparametric estimation of conditional quantiles under long-range dependence. In Honda (2000b), $V_i = Z_i$ in (1.1) and some restrictive moment assumptions are imposed since the proof depends on the approximation theorem of joint density functions in Giraitis et al. (1996). Peng and Yao (2004) deals with nonparametric quantile estimators in the cases of heavy-tailed errors.

It is known that the asymptotic distributions of nonparametric regression estimators depend on how strong the long-range dependence is in random design models. When the long-range dependence is rather weak, the estimators asymptotically behave in the same way as for i.i.d. observations. On the other hand, the estimators asymptotically behave in the same way as the sample means when the long-range dependence exceeds some level. This is true of kernel density estimation. Robinson (1997), Peng and Yao (2004), and Guo and Koul (2007) considered nonparametric estimation of trend functions under fixed design and the asymptotics are different from those under random design.

We apply the methods of Wu and Mielniczuk (2002) and Mielniczuk and Wu (2004) to nonparametric quantile regression, or nonparametric estimation of conditional quantiles, in this paper. The stochastic structure in (1.1) is motivated by Mielniczuk and Wu (2004). Mielniczuk and Wu (2004) allows for some dependence between $\{X_i\}_{i=1}^{\infty}$ and $\{Z_i\}_{i=1}^{\infty}$. However, we concentrate on the cases where $\{X_i\}_{i=1}^{\infty}$ and $\{Z_i\}_{i=1}^{\infty}$ are independent in order to avoid complicated assumptions on A_1 and B_1 in (4.7) of this paper. The results of the two papers by Wu and Mielniczuk and this paper also rely on the martingale decompositions of long-range dependent linear processes initiated by Ho and Hsing (1996, 1997). The results of the two papers are improved by Koul and Surgailis (2002) and Wu (2003). We also mention Honda (2006), which studies kernel density estimation for heavy-tailed linear processes.

We define (X_i, Z_i) for $i = 1, 2, \dots$ by

$$X_i = \sum_{j=0}^{\infty} b_j \epsilon_{i-j} \quad \textit{and} \quad Z_i = \sum_{j=0}^{\infty} c_j \zeta_{i-j}, \tag{1.2}$$

where $\{\epsilon_i\}_{i=-\infty}^{\infty}$ and $\{\zeta_i\}_{i=-\infty}^{\infty}$ are mutually independent mean-zero i.i.d. processes. We denote the variances of ϵ_1 and ζ_1 by σ_ϵ^2 and σ_ζ^2, respectively. We assume that for some $r > 2$,

$$\text{(1.3)} \qquad \mathrm{E}\{|\epsilon_1|^r\} < \infty \quad \text{and} \quad \mathrm{E}\{|\zeta_1|^r\} < \infty.$$

We also assume that

$$\text{(1.4)} \qquad b_j = j^{-(1+\gamma_X)/2} L_X(j) \quad \text{and} \quad c_j = j^{-(1+\gamma_Z)/2} L_Z(j), \quad j = 1, 2, \ldots,$$

where $L_X(j)$ and $L_Z(j)$ are slowly varying functions, $\gamma_X > 0$, and $\gamma_Z > 0$. We put $b_0 = 1$ and $c_0 = 1$ for convenience. Under the regularity conditions (see Assumptions A1-5 of Section 2), we derive the same kind of asymptotic distribution of the estimator as in Mielniczuk and Wu (2004). The convergence rate of the estimator is independent of $\{b_j\}_{j=0}^{\infty}$ in our setting (1.1)-(1.4).

We describe the definition of short-range dependence and long-range dependence. Some results on the autocovariances and the variances of partial sums are also stated. We denote the autocovariance function of $\{X_i\}_{i=0}^{\infty}$ by $r_X(j)$ and write $\sigma_{n,X}^2$ for $\mathrm{E}\{(\sum_{i=1}^n X_i)^2\}$. In this paper, n stands for the sample size.

When $\sum_{j=0}^{\infty} |b_j| < \infty$, $\sum_{j=-\infty}^{\infty} |r_X(j)| < \infty$ and this property is called short-range dependence. Then we have $\sigma_{n,X}^2 = O(n)$. Note that γ_X must be larger than or equal to 1. When $\{b_j\}_{j=0}^{\infty}$ does not meet restrictive conditions, it is difficult to establish that short-range dependent $\{X_i\}_{i=1}^{\infty}$ is an α-mixing process with sufficiently rapidly decaying mixing coefficients. See Doukhan (1994) about mixing processes. Thus the results on α-mixing processes do not cover some results on short-range dependent processes of this paper.

When $\sum_{j=0}^{\infty} |b_j| = \infty$, $\sum_{j=-\infty}^{\infty} |r_X(j)| = \infty$ and this property is called long-range dependence. Then we have $\lim_{n\to\infty}(\sigma_{n,X}^2/n) = \infty$. Then γ_X must be smaller than or equal to 1. Long-range dependent processes exhibit some different properties from short-range dependent processes.

It is well known that when $\gamma_X < 1$,

$$r_X(t) \sim C_{\gamma_X} t^{-\gamma_X} L_X^2(t)\sigma_\epsilon^2 \quad \text{and} \quad \sigma_{n,X}^2 \sim D_{\gamma_X} n^{2-\gamma_X} L_X^2(n)\sigma_\epsilon^2, \tag{1.5}$$

where

$$C_\gamma = \int_0^\infty (u+u^2)^{-(1+\gamma)/2} du \quad \text{and} \quad D_\gamma = \frac{2C_\gamma}{(1-\gamma)(2-\gamma)}.$$

$a_n \sim a_n'$ means $\lim_{n\to\infty}(a_n/a_n') = 1$ throughout this paper. The same results hold for $\{Z_i\}_{i=0}^\infty$ and we define $r_Z(j)$ and $\sigma_{n,Z}^2$ in the same way.

We carried out a simulation study to examine small sample properties of the estimators of conditional medians. The results show that the estimator does not work well when $\{Z_i\}_{i=0}^\infty$ is long-range dependent and that the effect of the long-range dependence of $\{X_i\}_{i=0}^\infty$ may not be negligible in small sample cases. However, the estimator seems to work well when $\gamma_Z = 1.5$, 2.5 and $\gamma_X = 2.5$.

This paper is organized as follows. In Section 2, we state assumptions and the asymptotic distribution of the estimator in Theorem 2.1. The theorem is verified in Section 4. We treat short-range dependent and long-range dependent processes in a unified manner in the proof of Theorem 2.1. The results of the simulation study are presented in Section 3.

We denote the Euclidean norm of $w \in R^k$ and the transpose of a matrix A by $|w|$ and A^T, respectively. We denote $(\mathrm{E}\{|W|^2\})^{1/2}$ by $\|W\|$ for a random variable W. Let C stand for generic positive constants. The values of d, δ, and M with no subscript also change from place to place. The sign function is defined by $\mathrm{sign}(v) = 1,\ v \geq 0,\ -1,\ v < 0$ and $\stackrel{d}{\to}$ and $\stackrel{p}{\to}$ stand for convergence in law and in probability, respectively.

First we state Assumptions A1-5 and related notations. Assumptions A1 ,A3, and A5 are necessary even for i.i.d. observations. Assumption A2 may be more restrictive. Our assumptions are much simpler than those in Mielniczuk and Wu (2004).

Assumption A1: $u(x)$ in (1.1) is twice continuously differentiable in a neighborhood of x_0.

Assumption A2: There exists a unique number m_0 satisfying $V(x_0, m_0) = 0$. In addition $V(x, z)$ is continuously differentiable in a neighborhood Ω of $(x_0, m_0)^T$ and $\frac{\partial V}{\partial z}(x_0, m_0) \neq 0$. There are also three positive constants, δ_1, δ_2, δ_3, such that $[x_0 - \delta_1, x_0 + \delta_1] \times [m_0 - \delta_2, m_0 + \delta_2] \subset \Omega$ and $\inf\{|V(x,z)| \,|\, |x - x_0| < \delta_1$ and $|z - m_0| \geq \delta_2\} > \delta_3$.

Assumption A2 and the implicit function theorem implies that there exists a unique function $m(x)$ in a neighborhood of x_0 such that $V(x, m(x)) = 0$ and $m_0 = m(x_0)$. We also have that $|V(x,z)| > d|z - m(x)|$ for some positive d in a neighborhood of $(x_0, m_0)^T$. Besides by the uniqueness of m_0, the continuity of $V(x, z)$, and the last condition in Assumption A2, $|z - m_0|$ is small when both $|V(x,z)|$ and $|x - x_0|$ is sufficiently small.

Assumption A3: The kernel function $K(\xi)$ is a bounded and symmetric density function and the support is included in $[-C_K, C_K]$ for some positive C_K. Let $h = cn^{-1/5}$ as bandwidths for nonparametric regression.

It is easy to see that almost the same results hold with some necessary modifications when we choose other bandwidths. However, the rate of convergence of the estimator is not improved by choosing $h = cn^{-d}(d \neq 1/5)$. We define κ_j and ν_j by

$$\kappa_j = \int \xi^j K(\xi) d\xi \qquad \text{and} \qquad \nu_j = \int \xi^j K^2(\xi) d\xi,$$

respectively. We omit the domain of integration when it is R or R^k or when there is no possibility of confusion.

The next assumption is imposed to deal with dependence among observations.

Assumption A4: Let $\phi_\epsilon(t)$ and $\phi_\zeta(t)$ denote the characteristic function of ϵ_1 and ζ_1, respectively. Then for some positive δ_1 and δ_2,

$$|\phi_\epsilon(t)| \le \delta_1(1+|t|)^{-\delta_2} \quad \text{and} \quad |\phi_\zeta(t)| \le \delta_1(1+|t|)^{-\delta_2}.$$

We also assume that both of ϵ_1 and ζ_1 have continuously differentiable density functions $f_1(x)$ and $g_1(z)$, respectively. In addition, they and their derivatives satisfy the following conditions.

$$|v(t)| \le C\frac{1}{1+t^2} \quad \text{and} \quad |v(s)-v(t)| \le C\frac{1}{1+t^2} \quad \text{for} \quad |s-t|<1, \tag{2.1}$$

where $v = f_1,\ f_1',\ g_1$, or g_1'. Remember that we have already assumed that $\mathrm{E}\{|\epsilon_1|^r\} < \infty$ and $\mathrm{E}\{|\zeta_1|^r\} < \infty$ for some $r > 2$.

Remark 1. All the necessary technical conditions on density functions are assured by (1.2)-(1.4) and Assumption A4. When we define $X_{i,j}$ and $Z_{i,j}$ by

$$X_{i,j} = \sum_{k=0}^{j-1} b_k \epsilon_{i-k}, \quad \text{and} \quad Z_{i,j} = \sum_{k=0}^{j-1} c_k \zeta_{i-k},$$

the arguments in Giraitis et al. (1996) and Koul and Surgailis (2001, 2002) imply that $X_{i,j}$ and $Z_{i,j}$ have continuously differentiable density functions for any positive integer j. Besides the density functions and their derivatives satisfy (2.1) with some common C. We denote the density functions by $f_j(x)$ and $g_j(z)$, respectively. This notation is conformable with those of Assumption A4 since $b_0 = c_0 = 1$. Write $f(x)$ and $g(z)$ for $f_\infty(x)$ and $g_\infty(z)$, respectively for notational simplicity.

The proofs of Proposition 2.1 and Theorems 3.1-3 of Honda (2006) imply that the part of (2.1) of Assumption A4 is not necessary. The assumptions on the characteristic functions and the moments are sufficient.

We need an assumption on $f(x_0)$ and $g(m_0)$.

Assumption A5: $f(x_0) > 0$ and $g(m_0) > 0$.

We introduce some more notations. Set

$$\tilde{X}_{i,j} = X_i - X_{i,j} = \sum_{k=j}^{\infty} b_k \epsilon_{i-k} \quad \text{and} \quad \tilde{Z}_{i,j} = Z_i - Z_{i,j} = \sum_{k=j}^{\infty} c_k \zeta_{i-k}.$$

When $\sum_{j=0}^{\infty} |b_j| = \infty$, $\mathrm{E}\{(\sum_{i=1}^{n}(X_i - \tilde{X}_{i,j}))^2\} = o(\sigma_{n,X}^2)$ for any j. We have the same result for Z_i and $\tilde{Z}_{i,j}$.

Next define filtrations, $\{\mathcal{S}_{0,i}\}_{i=-\infty}^{\infty}$, $\{\mathcal{S}_{1,i}\}_{i=-\infty}^{\infty}$, and $\{\mathcal{S}_{2,i}\}_{i=-\infty}^{\infty}$, by

$$\mathcal{S}_{0,i} = \sigma(\epsilon_i, \zeta_i, \epsilon_{i-1}, \zeta_{i-1}, \ldots,), \quad \mathcal{S}_{1,i} = \sigma(\epsilon_i, \epsilon_{i-1}, \ldots,), \quad \text{and} \quad \mathcal{S}_{2,i} = \sigma(\zeta_i, \zeta_{i-1}, \ldots,),$$

where $\sigma(\cdots)$ stands for the σ-field generated by the random variables inside the parentheses.

We define the local linear estimator of $u(x_0)$ as in Chaudhuri (1991). By the Taylor series expansion of $u(x)$ at x_0, we have

$$Y_i = u(x_0) + \frac{X_i - x_0}{h} h u'(x_0) + \frac{1}{2}\Big(\frac{X_i - x_0}{h}\Big)^2 h^2 u''(\bar{X}_i) + V_i,$$

where $\bar{X}_i$ is between x_0 and X_i. We define $V_i^* = V^*(X_i, Z_i)$ by

$$V_i^* = Y_i - (u(x_0), h u'(x_0))^T \eta_i = V_i + \frac{1}{2}\Big(\frac{X_i - x_0}{h}\Big)^2 h^2 u''(\bar{X}_i), \tag{2.2}$$

where $\eta_i = (1, (X_i - x_0)/h)^T$. We estimate $(u(x_0), h u'(x_0))^T$ by $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2)^T$ defined in (2.3).

$$\hat{\beta} = \mathrm{argmin}_{\beta \in R^2} \sum_{i=1}^{n} K_i |Y_i - \eta_i^T \beta|, \tag{2.3}$$

where $K_i = K((X_i - x_0)/h)$. When $\hat{\beta}$ is not uniquely determined, we should choose one from the candidates by some rule.

The asymptotic distribution of the estimator depends mainly on γ_Z and $L_Z(j)$ and there are the following two cases. The normalization constant τ_n is defined according to the cases.

Case 1: $n^{-\gamma_Z} L_Z^2(n) = o(n^{-4/5}) = o(1/(nh))$. Then we set $\tau_n = \sqrt{nh}$.

Case 2: $n^{-4/5} = o(n^{-\gamma_Z} L_Z^2(n))$. Then we set $\tau_n = (\sigma_{Z,n}^2/n^2)^{-1/2}$.

In Case 1, $\tau_n = c^{1/2}n^{2/5}$ and the asymptotic distribution of $\hat{\beta}$ is the same as for i.i.d. observations. In Cases 2, $\{Z_i\}_{i=1}^{\infty}$ have long-range dependence. If $L_Z(j)$ is a constant function, $\tau_n \sim dn^{\gamma_Z/2}$ for some positive d. The asymptotic distribution of $\hat{\beta}$ depends only on $\frac{\partial V}{\partial z}(x_0, z_0)$ and $\{Z_i\}_{i=1}^{\infty}$. When $n^{-\gamma_Z} L_Z^2(n)/n^{-4/5} \to d$ for some positive d, we just know the convergence rate and have not obtained the asymptotic distribution yet.

Normalize $\hat{\beta}$ and define $\hat{\theta}$, the normalized $\hat{\beta}$, by

$$\hat{\theta} = \tau_n \begin{pmatrix} \hat{\beta}_1 - u(x_0) \\ \hat{\beta}_2 - hu'(x_0) \end{pmatrix}. \tag{2.4}$$

We can represent $\hat{\theta}$ as

$$\hat{\theta} = \mathrm{argmin}_{\theta \in R^2} \sum_{i=1}^{n} K_i |V_i^* - \tau_n^{-1} \eta_i^T \theta|, \tag{2.5}$$

Before we state Theorem 2.1, we comment on the conditional density function of V_i on X_i, for which we write $f_V(v|x)$. It is easy to see from Assumptions A2 and Remark 1 below Assumption A4 that $f_V(v|x)$ exists in a neighborhood of $(0, x_0)$ and

$$f_V(0|x) = g(m(x))\Big(\frac{\partial V}{\partial z}(x, m(x))\Big)^{-1} \tag{2.6}$$

in a neighborhood of x_0.

THEOREM 2.1. *Suppose that Assumptions A1-A5 hold in our setting (1.1)-(1.4). Then as $n \to \infty$,*

Case 1($\tau_n = \sqrt{nh}$):

$$\hat{\theta} \xrightarrow{d} N\left(\begin{pmatrix} \frac{c^{5/2}u''(x_0)\kappa_2}{2} \\ \\ 0 \end{pmatrix}, \frac{1}{4f_V^2(0|x_0)f(x_0)}\begin{pmatrix} \nu_0 & 0 \\ \\ 0 & \kappa_2^{-2}\nu_2 \end{pmatrix}\right),$$

Case 2($\tau_n = (\sigma_{Z,n}^2/n^2)^{-1/2}$):

$$\hat{\theta} = -\frac{1}{2f_V(0|x_0)}\int \mathrm{sign}(V(x,z))g'(z)dz \begin{pmatrix} 1 \\ \\ 0 \end{pmatrix} \frac{1}{\sigma_{n,Z}}\sum_{i=1}^{n} Z_i + o_p(1).$$

□

The proof of Theorem 2.1 is given in Section 4.

(2.6) and some calculation yield that

$$\frac{1}{2f_V(0|x_0)}\int \mathrm{sign}(V(x,z))g'(z)dz = -\frac{\partial V}{\partial z}(x_0, m_0).$$

The asymptotic distribution does not depend on $\{X_i\}_{i=1}^{\infty}$ since the RHS of the above expression does not depend on $\{X_i\}_{i=1}^{\infty}$. It is also well known that $\sigma_{n,Z}^{-1}\sum_{i=1}^{n} Z_i \xrightarrow{d} N(0,1)$ as $n \to \infty$. See Theorem 5.2.3 of Taniguchi and Kakizawa for the proof. Hereafter we omit $n \to \infty$.

One might say that Theorem 2.1 focuses on some cases of Theorems 1 and 3 of Mielniczuk and Wu (2004). However, as mentioned before, we have proved that the rate of convergence does not depend on γ_X or $L_X(j)$ under the independence of $\{X_i\}_{i=1}^{\infty}$ and $\{Z_i\}_{i=1}^{\infty}$ and no complicated assumptions are imposed.

Remark 2. Suppose that we estimate $u(x_0)$ and $u(x_1)$ for $x_0 \neq x_1$. Then the proof of Theorem 2.1 implies that the two estimators are asymptotically independent in Case 1 since $\mathrm{E}\{K((X_i - x_0)/h)K((X_i - x_1)/h)\} = 0$ when $|x_0 - x_1| > 2hC_K$. On the other hand, the asymptotic correlation coefficient of the estimators is 1 in Case 2.

Remark 3. In Case 1, the same kind of statistical inference will be possible as for i.i.d. observations. On the other hand, in Case 2, the effect of long-range dependence appears and we have to estimate $\sigma^2_{n,Z}$ from $\{(X_i, Y_i)\}_{i=1}^n$. Even if we could observe V_i directly, this would be extremely difficult in the presence of heteroscedasticity. See also the last paragraph of Mielniczuk and Wu (2004) of p.1117 about this sort of difficulty.

3. Simulation study

We carried out a simulation study to examine small sample properties of the estimator. The results are given in Tables 2-9. We used R 2.2.1 and the rq function of the quantreg library. See R Development Core Team (2005) and Koenker (2006) about R and the quantreg library.

In this study, we set

$$Y_i = u(X_i) + Z_i, \tag{3.1}$$

where $u(x) = x^2 + x^4$ in Tables 2,3,6,7 and $u(x) = \cos x$ in Tables 4,5,8,9. The sample size is 200 and the replication number is 2000. The same random seed is used for each table.

We describe $\{X_i\}_{i=1}^\infty$ and $\{Z_i\}_{i=1}^\infty$ of this simulation study. $\{b_j\}_{j=0}^\infty$ and $\{c_j\}_{j=0}^\infty$ are given by

$$b_j = \begin{cases} (j+1)^{-(1+\gamma_X)/2}/(\sum_{j=0}^{999}(j+1)^{-(1+\gamma_X)})^{1/2}, & 0 \le j \le 999, \\ 0, & j \ge 1000, \end{cases}$$

$$c_j = \begin{cases} (j+1)^{-(1+\gamma_Z)/2}/(\sum_{j=0}^{999}(j+1)^{-(1+\gamma_Z)})^{1/2}, & 0 \le j \le 999, \\ 0, & j \ge 1000. \end{cases}$$

In Tables 2-5, ϵ_i and ζ_i follow the standard normal distribution. In Tables 6-9, $\sqrt{3}\epsilon_i$ and $\sqrt{3}\zeta_i$ follow the t-distribution with d.f. 3. Then it is easy to see that $\mathrm{E}\{X_i^2\} = \mathrm{E}\{Z_i^2\} = 1$. Note that the error term Z_i is relatively large compared to $u(X_i)$. We used the Epanechnikov kernel to estimate $u(x_0)$.

In each table, we give the simulation results for $x_0 = 0.0,\ 0.5,\ 1.0$, $h = 0.2,\ 0.3,\ 0.4$, and $\gamma_Z = 0.5,\ 1.5,\ 2.5$. Note that $0.0,\ 0.5,\ 1.0$ on the left margin mean $x_0 = 0.0,\ 0.5,\ 1.0$, respectively. We estimate $u(x_0)$ and they are

$$u(0.0) = 0.000, \quad u(0.5) = 0.3125, \quad u(1.0) = 2.000 \quad \text{in Tables } 2, 3, 6, 7 \text{ and}$$

$$u(0.0) = 1.000, \quad u(0.5) = 0.8776, \quad u(1.0) = 0.5403 \quad \text{in Tables } 4, 5, 8, 9.$$

As for the design density function,

$$f(0.0) = 0.399,\ f(0.5) = 0.352,\ f(1.0) = 0.242 \text{ Tables 2-5,} \tag{3.2}$$

$$f(0.0) = 0.581,\ f(0.5) = 0.408,\ f(1.0) = 0.179 \text{ in Tables } 6, 8 \text{ of } \gamma_X = 2.5, \text{ and} \tag{3.3}$$

$$f(0.0) = 0.461,\ f(0.5) = 0.384,\ f(1.0) = 0.219 \text{ in Tables } 7, 9 \text{ of } \gamma_X = 0.5. \tag{3.4}$$

The above values of $f(x_0)$ for Tables 6-9 are estimated by kernel density estimation from simulated 100,000 i.i.d. samples.

To see the effects of the long-range dependence of $\{X_i\}_{i=1}^{\infty}$, we set $\gamma_X = 0.5$ in Tables 3,5,7,9. We can see the effects by comparing Table i and Table $(i+1)$ for $i = 2, 4, 6, 8$.

Those parameters for tables are summarized in Table 1. In Table 1, t_3 in the dist. row implies that $\sqrt{3}\epsilon_i$ and $\sqrt{3}\zeta_i$ follow the t-distribution with d.f. 3.

In Tables 2-9, mean, bias, var, and mse stand for the sample mean, the sample mean $-u(x_0)$, the sample variance, and the sample mean squared error of the replications. NA4 stand for the number of replications for which there are only less than 4 observations available to estimate $u(x_0)$ in $[x_0-h, x_0+h]$. Then we did not use the rq function. Instead we estimated $u(x_0)$ by local medians if at least one observation is available. NA stands for the number of replications for which there is no observation available to estimate $u(x_0)$ in $[x_0-h, x_0+h]$. We just removed those replications. The numbers of the NA4 rows include the replications in the NA rows.

{ Tables 1-9 are around here. }

We obtained the following implications from the simulation:

1. We took γ_Z less than or equal to 2.5. However, the estimator worked well when $\{X_i\}_{i=1}^{\infty}$ and $\{Z_i\}_{i=1}^{\infty}$ are short-range dependent. See columns of $\gamma_Z = 1.5$ and 2.5 of Tables 2,4,6,8.

2. The columns of $\gamma_Z = 0.5$ of Tables 2-9 shows that the long-range dependence of $\{Z_i\}_{i=1}^{\infty}$ badly affects the properties of the estimator.

3. We can see the effect of the long-range dependence of $\{X_i\}_{i=1}^{\infty}$, especially in the rows of $x_0 = 1.0$ of Tables 7,9. It seems that there were many cases in which not enough observations were available to estimate $u(1.0)$. The phenomenon will be characteristic of long-range dependent processes since $f(1.0)$ is not so small. See (3.2)-(3.4). One remedy will be to select larger bandwidths by using local quadratic or cubic regression. However, if there is no observation around x_0, it is not possible to estimate $u(x_0)$.

4. Proof of Theorem 2.1

The proof of Theorem 2.1 is presented in this section. We begin with three results, Lemmas 4.1-4.3, which will be verified after the proof of Theorem 2.1. The setting (1.1)-(1.4) is assumed throughout this section. We omit a.s.(almost surely) for notational convenience.

The first result deals with the asymptotic distribution of $\tau_n(nh)^{-1}\sum_{i=1}^n K_i\eta_i \operatorname{sign}(V_i)$, which is the stochastic part of $\hat{\theta}$.

LEMMA 4.1. *Suppose that Assumptions A1-A4 hold. Then we have*

Case 1:

$$\frac{\tau_n}{nh}\sum_{i=1}^n K_i\eta_i \operatorname{sign}(V_i) \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ \\ 0 \end{pmatrix}, f(x_0)\begin{pmatrix} \nu_0 & 0 \\ \\ 0 & \nu_2 \end{pmatrix}\right),$$

Case 2:

$$\frac{\tau_n}{nh}\sum_{i=1}^n K_i\eta_i \operatorname{sign}(V_i) + \begin{pmatrix} 1 \\ \\ 0 \end{pmatrix} f(x_0)\int \operatorname{sign}(V(x,z))g'(z)dz\frac{1}{\sigma_{n,Z}}\sum_{i=1}^n Z_i = o_p(1).$$

□

The second result is used to evaluate the loss function in (2.5). See (2.2) for the definitions of η_i and V_i^*.

LEMMA 4.2. *Suppose that Assumptions A1-A4 hold. Then we have for any fixed* θ,

$$\begin{aligned}&\frac{\tau_n^2}{nh}\sum_{i=1}^n K_i(|V_i^* - \tau_n^{-1}\eta_i\theta| - |V_i^*|) \\ &= \theta^T\begin{pmatrix} 1 & 0 \\ \\ 0 & \kappa_2 \end{pmatrix}\theta f_V(0|x_0)f(x_0) - \Big(\frac{\tau_n}{nh}\sum_{i=1}^n K_i\eta_i \operatorname{sign}(V_i^*)\Big)^T\theta + o_p(1).\end{aligned}$$

□

The third result is used to replace V_i^* with V_i in the sign function and the result is related to the bias term of $\hat{\theta}$ in Case 1.

LEMMA 4.3. *Suppose that Assumptions A1-A4 hold. Then we have*

$$
\begin{aligned}
&\frac{\tau_n}{nh}\sum_{i=1}^{n} K_i\eta_i \operatorname{sign}(V_i^*) \\
&= \frac{\tau_n}{nh}\sum_{i=1}^{n} K_i\eta_i \operatorname{sign}(V_i) + \frac{\tau_n}{\sqrt{nh}}(c^{5/2}\kappa_2 u''(x_0) f_V(0|x_0) f(x_0), 0)^T + o_p(1).
\end{aligned}
$$

□

Now we prove Theorem 2.1.

PROOF OF THEOREM 2.1. Remember that $\tau_n/\sqrt{nh} = 1$ in Case 1 and $\tau_n/\sqrt{nh} = o(1)$ in Case 2. (2.5) is equivalent to

$$
\hat\theta = \mathrm{argmin}_{\theta\in R^2}\frac{\tau_n^2}{nh}\sum_{i=1}^{n} K_i(|V_i^* - \tau_n^{-1}\eta_i^T\theta| - |V_i^*|). \tag{4.1}
$$

By Lemmas 4.2-3, we have for any fixed $\theta \in R^2$,

$$
\begin{aligned}
&\frac{\tau_n^2}{nh}\sum_{i=1}^{n} K_i(|V_i^* - \tau_n^{-1}\eta_i\theta| - |V_i^*|) \\
&= \theta^T \begin{pmatrix} 1 & 0 \\ 0 & \kappa_2 \end{pmatrix} \theta f_V(0|x_0) f(x_0) - \Big(\frac{\tau_n}{nh}\sum_{i=1}^{n} K_i\eta_i \operatorname{sign}(V_i)\Big)^T\theta \\
&\quad - \frac{\tau_n}{\sqrt{nh}}(c^{5/2}\kappa_2 u''(x_0) f_V(0|x_0) f(x_0), 0)\theta + o_p(1).
\end{aligned} \tag{4.2}
$$

As in Pollard (1991), Fan et al. (1994), and Hall et al. (2002), the convexity lemma implies that (4.2) holds uniformly on $\{|\theta| < M\}$ for any positive M.

We consider the RHS of (4.2). Lemma 4.1 implies that

$$
\frac{\tau_n}{nh}\sum_{i=1}^{n} K_i\eta_i \operatorname{sign}(V_i) = O_p(1). \tag{4.3}
$$

Combining (4.3), $\tau_n/\sqrt{nh} = O(1)$, the uniformity of (4.2), and the convexity of the objective function in (4.1), we conclude that $|\hat\theta| = O_p(1)$ by appealing to the standard argument.

By using $|\hat{\theta}| = O_p(1)$ and the uniformity of (4.2) again, we obtain

$$\hat{\theta} = \frac{1}{2f_V(0|x_0)f(x_0)} \begin{pmatrix} 1 & 0 \\ 0 & \kappa_2 \end{pmatrix}^{-1} \times \Big\{ \frac{\tau_n}{nh} \sum_{i=1}^n K_i \eta_i \operatorname{sign}(V_i) + \frac{\tau_n}{\sqrt{nh}} (c^{5/2}\kappa_2 u''(x_0) f_V(0|x_0) f(x_0), 0)^T \Big\} + o_p(1). \tag{4.4}$$

The results of the theorem follow from (4.4) and Lemma 4.1. Hence the proof of the theorem is complete. □

We next consider the evaluation of $n^{-1}\sum_{i=1}^n A(\tilde{X}_{i,1})B(\tilde{Z}_{i,1})$, where $|A(\tilde{X}_{i,1})| \le M_1$, $|B(\tilde{Z}_{i,1})| \le M_1$, and $A(\tilde{X}_{i,1})$ and $B(\tilde{Z}_{i,1})$ depend on $\xi \in [-C_K, C_K]$. The evaluation is important in the proofs of Lemmas 4.1-3 and given in Lemma 4.4 below.

Define $U_{A,i,j}$ and $U_{B,i,j}$ respectively by

$$U_{A,i,j} = \mathrm{E}\{A(\tilde{X}_{i,1})|\mathcal{S}_{1,i-j}\} - \mathrm{E}\{A(\tilde{X}_{i,1})|\mathcal{S}_{1,i-j-1}\} - A_1 b_j \epsilon_{i-j}, \tag{4.5}$$

$$U_{B,i,j} = \mathrm{E}\{B(\tilde{Z}_{i,1})|\mathcal{S}_{2,i-j}\} - \mathrm{E}\{B(\tilde{Z}_{i,1})|\mathcal{S}_{2,i-j-1}\} - B_1 c_j \zeta_{i-j}, \tag{4.6}$$

where

$$A_1 = \frac{\partial}{\partial v}\mathrm{E}\{A(\tilde{X}_{i,1}+v)\}|_{v=0} \quad \text{and} \quad B_1 = \frac{\partial}{\partial v}\mathrm{E}\{B(\tilde{Z}_{i,1}+v)\}|_{v=0}. \tag{4.7}$$

Set $U_{A,i,0} = U_{B,i,0} = 0$. Then we have

$$\sum_{j=1}^{\infty} U_{A,i,j} = A(\tilde{X}_{i,1}) - \mathrm{E}\{A(\tilde{X}_{i,1})\} - A_1 \tilde{X}_{i,1}, \tag{4.8}$$

$$\sum_{j=1}^{\infty} U_{B,i,j} = B(\tilde{Z}_{i,1}) - \mathrm{E}\{B(\tilde{Z}_{i,1})\} - B_1 \tilde{Z}_{i,1}. \tag{4.9}$$

Write $W(\xi) = O_m(a_n)$ when $\sup_{\xi\in[-C_K,C_K]}\|W(\xi)\| \le Ca_n$ for convenience. When $\sup_{\xi\in[-C_K,C_K]}\|W(\xi)\| = o(a_n)$, $W(\xi) = o_m(a_n)$.

LEMMA 4.4. *Suppose that Assumption A4 holds and that we have, uniformly in*

$\xi \in [-C_K, C_K]$,

$$\text{(4.10)} \quad \mathrm{E}\{U_{A,i,j}^2\} \le M_2(j+1)^{-(\delta+1+\gamma_X)} \quad \text{and} \quad \mathrm{E}\{U_{B,i,j}^2\} \le M_2(j+1)^{-(\delta+1+\gamma_Z)}$$

for positive numbers M_2 and δ such that $\delta+\gamma_X \neq 1$, $\delta+\gamma_Z \neq 1$, and $2\delta < 1 \wedge \gamma_X$. Then

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} A(\tilde{X}_{i,1})B(\tilde{Z}_{i,1}) \\
&= \mathrm{E}\{A(\tilde{X}_{i,1})\}\mathrm{E}\{B(\tilde{Z}_{i,1})\} + A_1\mathrm{E}\{B(\tilde{Z}_{i,1})\}\frac{1}{n}\sum_{i=1}^{n}\tilde{X}_{i,1} + \mathrm{E}\{A(\tilde{X}_{i,1})\}B_1\frac{1}{n}\sum_{i=1}^{n}\tilde{Z}_{i,1} \\
&\quad +O_m(M_1M_2(n^{-(\delta+\gamma_Z)/2} \vee n^{-1/2})) + \mathrm{E}\{B(\tilde{Z}_{i,1})\}O_m(M_2(n^{-(\delta+\gamma_X)/2} \vee n^{-1/2})) \\
&\quad +A_1B_1O_m(n^{-(\delta+\gamma_Z)/2} \vee n^{-1/2}) + B_1O_m(M_2(n^{-(\delta+\gamma_Z)/2} \vee n^{-1/2})).
\end{aligned}$$

□

We verify Lemma 4.4 after Lemmas 4.1-3 are proved. A similar result to Lemma 4.4 is given Lemma 4.1 of Koul et al. (2004). However, we cannot use the result directly to prove Lemmas 4.1-3. Note that (4.10) have to be verified when we use Lemma 4.4. The proofs of (4.10) for Lemmas 4.1-3 are almost the same as the arguments of Section 6 of Koul and Surgailis (2002). We prove only the latter of (4.10) for Lemma 4.2. See Lemma 4.6 at the end of this section for the proof.

PROOF OF LEMMA 4.1. The former half of the lemma will be established as in Wu and Mielniczuk (2002) and Mielniczuk and Wu (2004).

We deal with only the first element. We can treat the second element in the same way and the joint distribution follows from the Cramér and Wold device.

Define T_i by $T_i = K_i \,\mathrm{sign}(V_i) - \mathrm{E}\{K_i \,\mathrm{sign}(V_i)|\mathcal{S}_{0,i-1}\}$ and notice that we have by (1.2)

$$\begin{aligned}
\text{(4.11)} \quad &\frac{1}{h}\mathrm{E}\{K_i \,\mathrm{sign}(V_i)|\mathcal{S}_{0,i-1}\} \\
&= \frac{1}{h}\int\int K\Big(\frac{x+\tilde{X}_{i,1}-x_0}{h}\Big)\,\mathrm{sign}(V(x+\tilde{X}_{i,1}, z+\tilde{Z}_{i,1}))f_1(x)g_1(z)dxdz \\
&= \int K(\xi)\Big\{f_1(x_0+\xi h-\tilde{X}_{i,1})\int \mathrm{sign}(V(x_0+\xi h, z))g_1(z-\tilde{Z}_{i,1})dz\Big\}d\xi.
\end{aligned}$$

Set $A(\tilde{X}_{i,1}) = f_1(x_0 + \xi h - \tilde{X}_{i,1})$ and $B(\tilde{Z}_{i,1}) = \int \text{sign}(V(x_0 + \xi h, z))g_1(z - \tilde{Z}_{i,1})dz$ in Lemma 4.4. Then, by Assumptions A2 and Remark 1, we have uniformly in $\xi \in [-C_K, C_K]$,

$$\begin{aligned}
\mathrm{E}\{A(\tilde{X}_{i,1})\} &= f(x_0 + \xi h) = f(x_0) + O(h), \\
\mathrm{E}\{B(\tilde{Z}_{i,1})\} &= \int \text{sign}(V(x_0 + \xi h, z)g(z)dz = 0, \\
A_1 &= -f'(x_0 + \xi h) = -f'(x_0) + O(h), \\
B_1 &= -\int \text{sign}\, V(x_0 + \xi h, z)g'(z)dz = -\int \text{sign}\, V(x_0, z)g'(z)dz + o(1).
\end{aligned}$$

It follows from (1.2)-(1.4), Assumptions A2 and A4, and the arguments of Section 6 of Koul and Surgailis (2002) that (4.10) holds. The details are omitted. Now by Lemma 4.4, there is a positive constant δ such that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} A(\tilde{X}_{i,1})B(\tilde{Z}_{i,1}) \\
&= -f(x_0)\int \text{sign}(V(x_0, z))g'(z)dz\frac{1}{n}\sum_{i=1}^{n} \tilde{Z}_{i,1} + O_m(n^{-(\delta+\gamma_Z)/2} \vee n^{-1/2}).
\end{aligned} \tag{4.12}$$

By (4.12), Jensen's inequality regarding $\int \cdot K(\xi)d\xi$, and Fubini's theorem, we have

$$\begin{aligned}
&\frac{1}{nh}\sum_{i=1}^{n}\{K_i\, \text{sign}(V_i)|\mathcal{S}_{0,i-1}\} \\
&= -f(x_0)\int K(\xi)\Big\{\int \text{sign}(V(x_0 + \xi h, z))g'(z)dz\Big\}d\xi\frac{1}{n}\sum_{i=1}^{n} \tilde{Z}_{i,1} \\
&\quad +O_p(n^{-(\delta+\gamma_Z)/2} \vee n^{-1/2}).
\end{aligned} \tag{4.13}$$

Since $\tau_n n^{-1}\sigma_{n,Z} = o(1)$ in Case 1,

$$\begin{aligned}
&\frac{\tau_n}{nh}\sum_{i=1}^{n}\{K_i\, \text{sign}(V_i)|\mathcal{S}_{0,i-1}\} \\
&= \begin{cases} o_p(1) & \text{in Case1,} \\ -f(x_0)\int \text{sign}(V(x_0, z))g'(z)dz\frac{\tau_n}{n}\sum_{i=1}^{n} Z_i + o_p(1) & \text{in Case2.} \end{cases}
\end{aligned} \tag{4.14}$$

Next we apply the martingale central limit theorem (e.g. Theorem 9.5.2 of Chow and Teicher (1988)) to $(nh)^{-1/2}\sum_{i=1}^n T_i$. Since $|T_i| \le C$, we have only to verify the convergence in probability of the conditional variance. Noticing that $|\mathrm{E}\{K_i \operatorname{sign}(V_i)|\mathcal{S}_{0,i-1}\}| \le Ch$ and $|\operatorname{sign}(V_i)| = 1$, we can evaluate the conditional variance in the following way.

$$\begin{aligned}
&\frac{1}{nh}\sum_{i=1}^n \mathrm{E}\{T_i^2|\mathcal{S}_{0,i-1}\} \\
&= \frac{1}{nh}\sum_{i=1}^n \int\int K^2\Big(\frac{x+\tilde{X}_{i,1}-x_0}{h}\Big) f_1(x)g_1(z)dxdz + O_p(h) \\
&= \frac{1}{n}\sum_{i=1}^n \int K^2(\xi) f_1(x_0+\xi h-\tilde{X}_{i,1})d\xi + O_p(h) \\
&= \frac{\nu_0}{n}\sum_{i=1}^n f_1(x_0-\tilde{X}_{i,1}) + O_p(h) \xrightarrow{p} \nu_0 f(x_0).
\end{aligned}$$

The last line follows from the ergodic theorem.

Since the convergence of the conditional variance is established, the martingale CLT implies that

$$\frac{1}{\sqrt{nh}}\sum_{i=1}^n T_i \xrightarrow{d} N(0, f(x_0)\nu_0).$$

Therefore, recalling that $\tau_n/\sqrt{nh} = o(1)$ in Case 2, we have

$$\frac{\tau_n}{nh}\sum_{i=1}^n T_i \begin{cases} \xrightarrow{d} N(0, f(x_0)\nu_0) & \text{in Case1,} \\ = o_p(1) & \text{in Case2.} \end{cases} \tag{4.15}$$

The desired result follows from (4.14) and (4.15). Hence the proof is complete. □

Proof of Lemma 4.2. Define $S_\theta(X_i, Z_i)$ by

$$S_\theta(X_i,Z_i) = |V_i^* - \tau_n^{-1}\eta_i^T\theta| - |V_i^*| + \tau_n^{-1}\eta_i^T\theta \operatorname{sign}(V_i^*). \tag{4.16}$$

Since $|V_i^* - V_i| \le Ch^2$ and $\tau_n = O(h^{-2})$, it is easy to see that

$$S_\theta(X_i,Z_i)| \le 2|\tau_n^{-1}\eta_i^T\theta| I(|V_i| \le C\tau_n^{-1}|\theta|). \tag{4.17}$$

Set

$$T_i = K_i S_\theta(X_i,Z_i) - \mathrm{E}\{K_i S_\theta(X_i,Z_i)|\mathcal{S}_{0,i-1}\}.$$

Since

$$\mathrm{E}\Big\{\Big(\frac{\tau_n^2}{nh}\sum_{i=1}^n T_i\Big)^2\Big\} \le C\frac{\tau_n^2|\theta|^2}{(nh)^2}\sum_{i=1}^n \mathrm{E}\{K_i^2 I(|V_i| \le C\tau_n^{-1}|\theta|)\} \le C\frac{\tau_n|\theta|^3}{nh} \to 0,$$

we have

$$\frac{\tau_n^2}{nh}\sum_{i=1}^n T_i = o_p(1). \tag{4.18}$$

Assumption A2 implies that $d|Z_i - m(X_i)| \le |V_i|$ when both $|V_i|$ and $|X_i - x_0|$ is sufficiently small. We used this inequality to evaluate $\mathrm{E}\{K_i^2 I(|V_i| \le C\tau_n^{-1}|\theta|)\}$ and this inequality will be also used to verify (4.10) for Lemmas 4.2-3.

Next notice that we have by (1.2),

$$\begin{aligned}
&\frac{\tau_n^2}{h}\mathrm{E}\{K_i S_\theta(X_i, Z_i)|\mathcal{S}_{0,i-1}\} \\
&= \frac{\tau_n^2}{h}\int\int K\Big(\frac{x + \tilde{X}_{i,1} - x_0}{h}\Big)S_\theta(x + \tilde{X}_{i,1}, z + \tilde{Z}_{i,1})f_1(x)g_1(z)dxdz \\
&= \int K(\xi)\Big\{f_1(x_0 + \xi h - \tilde{X}_{i,1})\tau_n^2\int S_\theta(x_0 + \xi h, z)g_1(z - \tilde{Z}_{i,1})dz\Big\}d\xi.
\end{aligned}$$

Then set $A(\tilde{X}_{i,1}) = f_1(x_0 + \xi h - \tilde{X}_{i,1})$ as in the proof of Lemma 4.1 and $B(\tilde{Z}_{i,1}) = \tau_n^2\int S_\theta(x_0 + \xi h, z)g_1(z - \tilde{Z}_{i,1})dz$ in Lemma 4.4. By Assumptions A1-A2 and Remark 1, we have uniformly in $\xi \in [-C_K, C_K]$,

$$\mathrm{E}\{B(\tilde{Z}_{i,1})\} = ((1,\xi)\theta)^2 f_V(0|x_0) + o(1),$$

$$B_1 = -\tau_n^2\int S_\theta(x_0 + \xi h, z)g'(z)dz = O(1).$$

(4.10) follows from (1.2)-(1.4), Assumptions A1, A2, and A4, and the arguments of Section 6 of Koul and Surgailis (2002). We prove only the latter of (4.10) in Lemma 4.6 at the end of this section for reference. Then by Lemma 4.4, we have

$$\frac{1}{n}\sum_{i=1}^n A(\tilde{X}_{i,1})B(\tilde{Z}_{i,1}) = ((1,\xi)\theta)^2 f_V(0|x_0)f(x_0) + o_m(1).$$

Therefore in the same way as (4.13), we obtain

$$\begin{aligned}
(4.19)\qquad &\frac{\tau_n^2}{nh}\mathrm{E}\{K_i S_\theta(X_i,Z_i)|\mathcal{S}_{0,i-1}\}\\
&= \int K(\xi)((1,\xi)\theta)^2 d\xi f_V(0|x_0)f(x_0) + o_p(1)\\
&= \theta^T \begin{pmatrix} 1 & 0 \\ 0 & \kappa_2 \end{pmatrix} \theta f_V(0|x_0)f(x_0) + o_p(1).
\end{aligned}$$

The desired result follows from (4.18) and (4.19). Hence the proof is complete. □

PROOF OF LEMMA 4.3. The proof is similar to that of Lemma 4.2 and we consider only the first element.

Define T_i by

$$T_i = K_i(\mathrm{sign}(V_i^*) - \mathrm{sign}(V_i)) - \mathrm{E}\{K_i(\mathrm{sign}(V_i^*) - \mathrm{sign}(V_i))|\mathcal{S}_{0,i-1}\}.$$

Since $|V_i^* - V_i| \le Ch^2$,

$$\begin{aligned}
(4.20)\qquad &\mathrm{E}\Big\{\Big(\frac{\tau_n}{nh}\sum_{i=1}^n T_i\Big)^2\Big\}\\
&\le \frac{4\tau_n^2}{(nh)^2}\sum_{i=1}^n \mathrm{E}\{K_i^2 I(|V_i| \le Ch^2)\} \le C\frac{\tau_n^2 h^3}{nh^2} = C\frac{\tau_n^2 h}{n} \to 0.
\end{aligned}$$

Next notice that we have by (1.2),

$$\begin{aligned}
&\frac{\tau_n}{h}\mathrm{E}\{K_i(\mathrm{sign}(V_i^*) - \mathrm{sign}(V_i))|\mathcal{S}_{0,i-1}\}\\
&= \frac{\tau_n}{h}\int\int K\Big(\frac{x + \tilde{X}_{i,1} - x_0}{h}\Big)\\
&\qquad \times(\mathrm{sign}(V^*(x+\tilde{X}_{i,1}, z+\tilde{Z}_{i,1}) - \mathrm{sign}(V(x+\tilde{X}_{i,1}, z+\tilde{Z}_{i,1}))f_1(x)g_1(z)dxdz\\
&= \int K(\xi)\Big\{f_1(x_0+\xi h - \tilde{X}_{i,1})\\
&\qquad \times\tau_n\int(\mathrm{sign}(V^*(x_0+\xi h, z) - \mathrm{sign}(V(x_0+\xi h, z))g_1(z-\tilde{Z}_{i,1})dz\Big\}d\xi.
\end{aligned}$$

Set $A(\tilde{X}_{i,1}) = f_1(x_0 + \xi h - \tilde{X}_{i,1})$ as in the proof of Lemma 4.1 and $B(\tilde{Z}_{i,1}) = \tau_n\int(\mathrm{sign}(V^*(x_0+\xi h, z) - \mathrm{sign}(V(x_0+\xi h, z))g_1(z-\tilde{Z}_{i,1})dz$ in Lemma 4.4. Then by

(2.2), Assumptions A1 and A2, and Remark 1, we have uniformly in $\xi \in [-C_K, C_K]$,

$$\mathrm{E}\{B(\tilde{Z}_{i,1})\} = \frac{\tau_n}{\sqrt{nh}} c^{5/2}\xi^2 u''(x_0) f_V(0|x_0) + o(1),$$

$$B_1 = -\tau_n \int (\mathrm{sign}(V^*(x_0+\xi h, z) - \mathrm{sign}(V(x_0+\xi h, z))g'(z)dz = O(1).$$

(4.10) follows from (1.2)-(1.4), Assumptions A1, A2, and A4, and the arguments of Section 6 of Koul and Surgailis (2002). The details are omitted. Therefore by Lemma 4.4, we obtain

$$\frac{1}{n}\sum_{i=1}^{n} A(\tilde{X}_{i,1})B(\tilde{Z}_{i,1}) = \frac{\tau_n}{\sqrt{nh}} c^{5/2}\xi^2 u''(x_0) f_V(0|x_0) f(x_0) + o_m(1).$$

Finally in the same way as (4.13) and (4.19), we can see that

$$\begin{aligned} &\frac{\tau_n}{nh}\sum_{i=1}^{n} \mathrm{E}\{K_i(\mathrm{sign}(V_i^*) - \mathrm{sign}(V_i))|\mathcal{S}_{0,i-1}\} \\ &= \frac{\tau_n}{\sqrt{nh}} c^{5/2}\kappa_2 u''(x_0) f_V(0|x_0) f(x_0) + o_p(1). \end{aligned} \tag{4.21}$$

The desired result follows from (4.20) and (4.21). Hence the proof is complete. □

Lemma 4.5 below is a tool to prove Lemma 4.4. Lemma 4.5 is just a modification of a familiar lemma which has been used explicitly or implicitly in papers on long-range dependent linear processes.

Lemma 4.5. *Let $\{\mathcal{F}_i\}_{i=-\infty}^{\infty}$ be a filtration. Suppose that we have a set of random variables $\{U_{i,j} \,|\, 1 \le i \le n,\ 0 \le j < \infty\}$ such that $U_{i,j}$ are $\mathcal{F}_{i-j}$-measurable and $\mathrm{E}\{U_{i,j}|\mathcal{F}_{i-j-1}\} = 0$. In addition we have another set of random variables $\{D_i\}_{i=1}^{\infty}$ which is independent of $\{U_{i,j} \,|\, 1 \le i \le n,\ 0 \le j < \infty\}$. Suppose that every D_i has the second moment and that*

$$\mathrm{E}\{U_{i,j}^2\} \le M(j+1)^{-(1+\gamma+\delta)} \tag{4.22}$$

for some positive numbers M, δ, and γ $(\delta + \gamma \neq 1)$. Then we have

$$\mathrm{E}\{(\sum_{i=1}^{n} D_i \sum_{j=0}^{\infty} U_{i,j})^2\} \tag{4.23}$$

$$\leq \sum_{j=0}^{\infty}(\sum_{i=1}^{n}\|D_i\|\|U_{i,i+j}\|)^2 + \sum_{j=-n}^{-1}(\sum_{i=-j}^{n}\|D_i\|\|U_{i,i+j}\|)^2$$
$$\leq CM \max_{1\leq i\leq n}\{\|D_i\|\}(n^{2-\gamma-\delta} \vee n).$$

$\square$

The first inequality in (4.23) is essentially given in Lemma 2 of Mielniczuk and Wu (2004). The last inequality follows from (4.22) and some standard calculation. The details of the proof are omitted.

PROOF OF LEMMA 4.4. Applying Lemma 4.5 twice with $\{\mathcal{F}_i\}_{i=-\infty}^{\infty} = \{\mathcal{S}_{1,i}\}_{i=-\infty}^{\infty}$ and $\{\mathcal{S}_{2,i}\}_{i=-\infty}^{\infty}$ respectively, we obtain

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} A(\tilde{X}_{i,1})B(\tilde{Z}_{i,1}) \\
&= \frac{1}{n}\sum_{i=1}^{n} A(\tilde{X}_{i,1})\{\mathrm{E}\{B(\tilde{Z}_{i,1})\} + B_1\tilde{Z}_{i,1}\} \\
&\quad +O_m(M_1M_2(n^{-(\delta+\gamma_Z)/2} \vee n^{-1/2})) \\
&= \mathrm{E}\{A(\tilde{X}_{i,1})\}\mathrm{E}\{B(\tilde{Z}_{i,1})\} + A_1\mathrm{E}\{B(\tilde{Z}_{i,1})\}\frac{1}{n}\sum_{i=1}^{n}\tilde{X}_{i,1} \\
&\quad +B_1\frac{1}{n}\sum_{i=1}^{n} A(\tilde{X}_{i,1})\tilde{Z}_{i,1} + O_m(M_1M_2(n^{-(\delta+\gamma_Z)/2} \vee n^{-1/2})) \\
&\quad +\mathrm{E}\{B(\tilde{Z}_{i,1})\}O_m(M_2(n^{-(\delta+\gamma_X)/2} \vee n^{-1/2})).
\end{aligned} \tag{4.24}$$

We evaluate $n^{-1}\sum_{i=1}^{n} A(\tilde{X}_{i,1})\tilde{Z}_{i,1}$ in the last expression of (4.24). It is rewritten as

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} A(\tilde{X}_{i,1})\tilde{Z}_{i,1} \\
&= \mathrm{E}\{A(\tilde{X}_{i,1})\}\frac{1}{n}\sum_{i=1}^{n}\tilde{Z}_{i,1} + A_1\frac{1}{n}\sum_{i=1}^{n}\tilde{X}_{i,1}\tilde{Z}_{i,1} \\
&\quad +\frac{1}{n}\sum_{i=1}^{n}\{A(\tilde{X}_{i,1}) - \mathrm{E}\{A(\tilde{X}_{i,1})\} - A_1\tilde{X}_{i,1}\}\tilde{Z}_{i,1}.
\end{aligned} \tag{4.25}$$

For the second term of the RHS of (4.25), we have by (1.2)-(1.4),

$$\mathrm{E}\Big\{\Big(\frac{1}{n}\sum_{i=1}^{n}\tilde{X}_{i,1}\tilde{Z}_{i,1}\Big)^2\Big\} \tag{4.26}$$

$$= \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \mathrm{E}\{\tilde{X}_{i,1}\tilde{X}_{j,1}\}\mathrm{E}\{\tilde{Z}_{i,1}\tilde{Z}_{j,1}\}$$
$$\le \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n (1+|i-j|)^{-2\delta}|\mathrm{E}\{\tilde{Z}_{i,1}\tilde{Z}_{j,1}\}| \le C(n^{-(\delta+\gamma_Z)}\vee n^{-1}).$$

Remember that we took δ which is smaller than $(\gamma_X\wedge 1)/2$.

To deal with the third term of the RHS of (4.25), we use the following fact owing to (1.2)-(1.4) and Assumption A4.

$$|\mathrm{E}[\{A(\tilde{X}_{i,1}) - \mathrm{E}\{A(\tilde{X}_{i,1})\} - A_1\tilde{X}_{i,1}\}\{A(\tilde{X}_{j,1}) - \mathrm{E}\{A(\tilde{X}_{j,1})\} - A_1\tilde{X}_{j,1}\}]|$$
$$\le \begin{cases} CM_2^2(|i-j|+1)^{-(\delta+\gamma_X)}, & \text{if } \delta+\gamma_X<1,\\ CM_2^2(|i-j|+1)^{-(1+\delta+\gamma_X)/2}, & \text{if } \delta+\gamma_X>1. \end{cases} \tag{4.27}$$

See Lemma 6.1 of Koul and Surgailis (2002). Then by using (1.5) with $\tilde{Z}_{i,1}$ and (4.27), we evaluate the varinance and obtain

$$\frac{1}{n}\sum_{i=1}^n\{A(\tilde{X}_{i,1}) - \mathrm{E}\{A(\tilde{X}_{i,1})\} - A_1\tilde{X}_{i,1}\}\tilde{Z}_{i,1} = O_m(M_2(n^{-(\delta+\gamma_Z)/2}\vee n^{-1/2})). \tag{4.28}$$

The desired result follows from (4.24)-(4.26), and (4.28). Hence the proof is complete.

□

The latter of (4.10) for Lemma 4.2 is demonstrated in Lemma 4.6 below.

Lemma 4.6. *Suppose that Assumptions A1, A2, and A4 hold. When*

$$B(\tilde{Z}_{i,1}) = \int \tau_n^2 S_\theta(x_0+\xi h, z)g_1(z-\tilde{Z}_{i,1})dz,$$

there are positive M and δ such that

$$\mathrm{E}\{U_{B,i,j}^2\} \le M(j+1)^{-(\delta+1+\gamma_Z)} \quad uniformly\ in\ \xi\in[-C_K, C_K].$$

Proof. We follow the arguments of Koul and Surgailis (2002) and they rely on (1.2)-(1.4) and Assumption A4. We suppress B of $U_{B,i,j}$ for notational convenience.

Define $U^{(1)}_{i,j}$, $U^{(2)}_{i,j}$, and $U^{(3)}_{i,j}$ by

$$\begin{aligned}
U^{(1)}_{i,j} &= \int \tau_n^2 S_\theta(x_0+\xi h, z)(g_{j+1}(z-\tilde{Z}_{i,j}) - g_{j+1}(z-\tilde{Z}_{i,j+1}))dz \\
&\qquad +c_j\zeta_{i-j}\int \tau_n^2 S_\theta(x_0+\xi h, z)g'_{j+1}(z-\tilde{Z}_{i,j+1})dz, \\
U^{(2)}_{i,j} &= \int \tau_n^2 S_\theta(x_0+\xi h, z)(g_j(z-\tilde{Z}_{i,j}) - g_{j+1}(z-\tilde{Z}_{i,j}))dz, \\
U^{(3)}_{i,j} &= -c_j\zeta_{i-j}\int \tau_n^2 S_\theta(x_0+\xi h, z)(g'_{j+1}(z-\tilde{Z}_{i,j+1}) - g'(z))dz.
\end{aligned}$$

Then we have

$$U_{i,j} = U^{(1)}_{i,j} + U^{(2)}_{i,j} + U^{(3)}_{i,j}. \tag{4.29}$$

Recall that the support of $S_\theta(x_0+\xi h, z)$ is contained in $\{z \mid |z-m(x_0+\xi h)| < C\tau_n^{-1}\}$ for some positive C because of Assumption A2. This is crucial to the evaluation of $U^{(l)}_{i,j}$, $l=1,2,3$. In addition when $\delta \ge 2$, we have by Rosenthal's inequality,

$$\mathrm{E}\{|\tilde{Z}_{i,j+1}|^\delta\} \le C\{(\sum_{l=j+1}^{\infty} c_l^2)^{\delta/2} + \sum_{l=j+1}^{\infty}|c_l|^\delta\}. \tag{4.30}$$

First write $U^{(1)}_{i,j}$ as

$$U^{(1)}_{i,j} = \int_0^{-c_j\zeta_{i-j}}\Big\{\int \tau_n^2 S_\theta(x_0+\xi h, z)(g'_{j+1}(z+u-\tilde{Z}_{i,j+1}) - g'_{j+1}(z-\tilde{Z}_{i,j+1}))dz\Big\}du.$$

(5.14) of Lemma 5.2 of Koul and Surgailis (2002) and Remark 1 of this paper imply that there is a positive γ_1 such that $1 < \gamma_1 \le r/2$ and

$$|U^{(1)}_{i,j}| \le C|c_j|^{\gamma_1}|\zeta_{i-j}|^{\gamma_1}(1 \vee |\tilde{Z}_{i,j+1}|^{\gamma_1}). \tag{4.31}$$

Hence we have

$$\mathrm{E}\{|U^{(1)}_{i,j}|^2\} \le C|c_j|^{2\gamma_1}. \tag{4.32}$$

Next we consider $U^{(2)}_{i,j}$. By Lemma 5.1 of Koul and Surgailis (2002), we have

$$|U^{(2)}_{i,j}| \le Cc_j^2\int|\tau_n^2 S_\theta(x_0+\xi h, z)|(1+|z-\tilde{Z}_{i,j}|)^{-2}dz \le Cc_j^2. \tag{4.33}$$

Finally we deal with $U_{i,j}^{(3)}$. We introduce a random variable $\bar{Z}_{i,j+1}$ which is an independent copy of $\tilde{Z}_{i,j+1}$. We denote the expectation with respect to $\bar{Z}_{i,j+1}$ by $\bar{\mathrm{E}}\{\cdot\}$. By using $\bar{Z}_{i,j+1}$ and $\bar{\mathrm{E}}\{\cdot\}$, we can represent $U_{i,j}^{(3)}$ as

$$\begin{aligned} U_{i,j}^{(3)} &= -c_j\zeta_{i-j}\Big[\int \tau_n^2 S_\theta(x_0+\xi h, z)(g'_{j+1}(z-\tilde{Z}_{i,j+1}) - g'_{j+1}(z))dz \\ &\qquad -\bar{\mathrm{E}}\Big\{\int \tau_n^2 S_\theta(x_0+\xi h, z)(g'_{j+1}(z-\bar{Z}_{i,j+1}) - g'_{j+1}(z))dz\Big\}\Big] \end{aligned}$$

By (5.13) of Lemma 5.2 of Koul and Surgailis (2002), there is a positive γ_2 such that $1 < \gamma_2 \leq r/2$,

$$\begin{aligned} &\Big|\bar{\mathrm{E}}\Big\{\int \tau_n^2 S_\theta(x_0+\xi h, z)(g'_{j+1}(z-\bar{Z}_{i,j+1}) - g'_{j+1}(z))dz\Big\}\Big| \\ &\leq C\bar{\mathrm{E}}\{|\bar{Z}_{i,j+1}| \vee |\bar{Z}_{i,j+1}|^{\gamma_2}\} \end{aligned} \tag{4.34}$$

and

$$\begin{aligned} &\Big|\int \tau_n^2 S_\theta(x_0+\xi h, z)(g'_{j+1}(z-\tilde{Z}_{i,j+1}) - g'_{j+1}(z))dz\Big| \\ &\leq C(|\tilde{Z}_{i,j+1}| \vee |\tilde{Z}_{i,j+1}|^{\gamma_2}). \end{aligned} \tag{4.35}$$

(4.34) and (4.35) yield

$$\mathrm{E}\{|U_{i,j}^{(3)}|^2\} \leq C|c_j|^2[\mathrm{E}\{|\tilde{Z}_{i,j+1}|^2\} + \mathrm{E}\{|\tilde{Z}_{i,j+1}|^{2\gamma_2}\}]. \tag{4.36}$$

The desired result follows from (4.29), (4.30), (4.32), (4.33), and (4.36). Hence the proof is complete. □

Acknowledgements

The author appreciates valuable comments of the associate editor and two referees very much.

Beran, J. (1994). *Statistics for long-memory processes.* New York:Chapman & Hall.

Chaudhuri, P. (1991). Nonparametric estimates of regression quantiles and their local Bahadur representation. *The Annals of Statistics*, **19**, 760–777.

Chow, Y. S., Teicher, H. (1988). *Probability theory* (2nd ed.). New York:Springer.

Csörgő, S. and Mielniczuk, J. (2000). The smoothing dichotomy in random-design regression with long-memory errors based on moving averages. *Statistica Sinica*, **10**, 771–787.

Doukhan, P. (1994). *Mixing: properties and examples.* Lecture Notes in Statistics, **85**. New York:Springer.

Doukhan, P., Oppenheim, G., Taqqu, M. S. (Ed.) (2003). *Theory and applications of long-range dependence.* Boston:Birkhäser.

Giraitis, L., Koul, H. L., Surgailis, D. (1996). Asymptotic normality of regression estimators with long memory errors. *Statistics & Probability Letters*, **29**, 317–335.

Fan, J., Gijbels, I. (1996). *Local polynomial modelling and its applications.* London:Chapman & Hall.

Fan, J., Hu, T.-C., Truong, Y. K. (1994). Robust nonparametric function estimation. *Scandinavian Journal of Statistics*, **21**, 867–885.

Fan, J., Yao, Q. (2003). *Nonlinear time series.* New York:Springer.

Guo, H., Koul, H. L. (2007). Nonparametric regression with heteroscedastic long memory errors. *Journal of Statistical Planning and Inference*, **137**, 379–404.

Hall, P., Peng, L., Yao, Q. (2002). Prediction and nonparametric estimation for time series with heavy tails. *Journal of Time Series Analysis*, **23**, 313–331.

Härdle, W., Müller, M., Sperlich, S., Werwatz, A. (2004). *Nonparametric and semipara-*

metric models. Berlin:Springer.

Hidalgo, F. J. (1997). Nonparametric estimation with strongly dependent multivariate time series. *Journal of Time Series Analysis*, **18**, 95–122.

Ho, H.-C., Hsing, T. (1996). On the asymptotic expansion of the empirical process of long-memory moving averages. *The Annals of Statistics*, **24**, 992–1024.

Ho, H.-C., Hsing, T. (1997). Limit theorems for functionals of moving averages. *The Annals of Probability*, **25**, 1636–1669.

Honda, T. (2000a). Nonparametric estimation of a conditional quantile for α-mixing processes. *The Annals of the Institute of Statistical Mathematics*, **52**, 459–470.

Honda, T. (2000b). Nonparametric estimation of the conditional median function for long-range dependent processes. *Journal of the Japan Statistical Society*, **30**, 129–142.

Honda, T. (2006). *Nonparametric density estimation for linear processes with infinite variance.* Discussion Paper #2005-13, Graduate School of Economics, Hitotsubashi University, Tokyo.

Koenker, R. (2005). *Quantile regression.* New York:Cambridge University Press.

Koenker, R. (2006). quantreg: Quantile Regression. R package version 3.90(http://www.r-project.org).

Koenker, R., Basset, G. (1978). Regression quantiles. *Econometrica*, **46**, 33–50.

Koul, H. L. (2002). *Weighted empirical processes in dynamic linear models*(2nd ed.). Lecture Notes in Statistics, **166**. New York:Springer.

Koul, H. L., Baillie, R. T., Surgailis, D. (2004). Regression model fitting with a long memory covariate process. *Econometric Theory*, **20**, 485–512.

Koul, H. L., Mukherjee, K. (1993). Asymptotics of R-, MD-, and LAD-estimators in linear regression models with long range dependent errors, *Probability Theory and Related Fields*, **95**, 535–553.

Koul, H. L., Surgailis, D. (2001). Asymptotics of empirical processes of long memory moving averages with infinite variance. *Stochastic Processes and their Applications*, **91**, 309–336.

Koul, H. L., Surgailis, D. (2002). Asymptotic expansion of the empirical process of long memory moving averages. In H. Dehling, T. Mikosch, and M. Sørensen (Ed.) *Empirical process techniques for dependent data* (pp.213–239). Boston:Birkhäser.

Mielniczuk, J., Wu, W. B. (2004). On random design model with dependent errors. *Statistica Sinica*, **14**, 1105–1126.

Nze, P. A., Bühlman, P., Doukhan, P. (2002). Weak dependence beyond mixing and asymptotics for nonparametric regression. *The Annals of Statistics*, **30**, 397–430.

Peng, L., Yao, Q. (2004). Nonparametric regression under dependent errors with infinite variance. *The Annals of the Institute of Statistical Mathematics*, **56**, 73–86.

Pollard, D. (1991). Asymptotics for least absolute deviation regression estimates. *Econometric Theory*, **7**, 186–98.

R Development Core Team (2005). *R: A language and environment for statistical computing.* Vienna:R Foundation for Statistical Computing (URL http://www.R-project.org).

Robinson, P. M. (1997). Large-sample inference for nonparametric regression with dependent errors. *The Annals of Statistics*, **25**, 2054–2083.

Robinson, P. M. (Ed.) (2003). *Time series with long memory.* New York:Oxford University Press.

Taniguchi, M., Kakizawa, Y. (2000). *Asymptotic theory of statistical inference for time* *series*. New York:Springer.

Truong, Y. K., Stone, C. J. (1992). Nonparametric function estimation involving time series. *The Annals of Statistics*, **20**, 77–97.

Wu, W. B. (2003). Empirical processes of long-memory sequences. *Bernoulli*, **9**, 809–831.

Wu, W. B., Mielniczuk, J. (2002). Kernel density estimation for linear processes. *The Annals of Statistics*, **30**, 1441–1459.

Zhao, Z., Wu, W. B. (2006). *Kernel quantile regression for nonlinear stochastic models*. Technical Report No.572, Department of Statistics, The University of Chicago, Chicago.

Table 1: Parameters of tables

	Table							
	2	3	4	5	6	7	8	9
γ_x	2.5	0.5	2.5	0.5	2.5	0.5	2.5	0.5
$u(x)$	x^2+x^4	x^2+x^4	$cos(x)$	$cos(x)$	x^2+x^4	x^2+x^4	$cos(x)$	$cos(x)$
dist.	N(0,1)	N(0,1)	N(0,1)	N(0,1)	t_3	t_3	t_3	t_3

Table 2: γ_x=2.5, N(0,1), $u(x)=x^2+x^4$

γ_z		0.5			1.5			2.5		
h		0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
0.0	mean	0.016	0.026	0.039	0.013	0.022	0.033	0.003	0.017	0.031
	bias	0.016	0.026	0.039	0.013	0.022	0.033	0.003	0.017	0.031
	var	0.271	0.258	0.248	0.088	0.070	0.062	0.074	0.052	0.043
	mse	0.271	0.258	0.249	0.088	0.071	0.063	0.074	0.053	0.044
	NA4	0	0	0	0	0	0	0	0	0
	NA	0	0	0	0	0	0	0	0	0
0.5	mean	0.347	0.370	0.403	0.332	0.361	0.394	0.325	0.347	0.383
	bias	0.035	0.058	0.091	0.020	0.049	0.082	0.013	0.035	0.071
	var	0.282	0.263	0.255	0.103	0.081	0.069	0.080	0.057	0.047
	mse	0.284	0.266	0.263	0.103	0.083	0.076	0.080	0.058	0.052
	NA4	0	0	0	0	0	0	0	0	0
	NA	0	0	0	0	0	0	0	0	0
1.0	mean	2.048	2.127	2.220	2.044	2.119	2.208	2.051	2.124	2.209
	bias	0.048	0.127	0.220	0.044	0.119	0.208	0.051	0.124	0.209
	var	0.321	0.283	0.272	0.148	0.109	0.095	0.121	0.085	0.068
	mse	0.323	0.299	0.320	0.150	0.123	0.138	0.124	0.100	0.112
	NA4	0	0	0	0	0	0	0	0	0
	NA	0	0	0	0	0	0	0	0	0

Table 3: γ_x=0.5, N(0,1), $u(x)=x^2+x^4$

γ_z		0.5			1.5			2.5		
h		0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
0.0	mean	0.022	0.028	0.041	0.014	0.020	0.031	0.000	0.009	0.026
	bias	0.022	0.028	0.041	0.014	0.020	0.031	0.000	0.009	0.026
	var	0.295	0.274	0.266	0.106	0.085	0.073	0.077	0.056	0.047
	mse	0.295	0.275	0.267	0.106	0.085	0.074	0.077	0.056	0.048
	NA4	2	1	0	0	0	0	1	0	0
	NA	1	0	0	0	0	0	0	0	0
0.5	mean	0.341	0.364	0.395	0.332	0.358	0.389	0.325	0.348	0.380
	bias	0.029	0.052	0.083	0.020	0.046	0.077	0.013	0.036	0.068
	var	0.298	0.284	0.279	0.131	0.103	0.088	0.099	0.072	0.055
	mse	0.299	0.286	0.285	0.132	0.105	0.094	0.099	0.073	0.060
	NA4	4	1	0	9	2	1	4	1	0
	NA	1	0	0	0	0	0	0	0	0
1.0	mean	2.058	2.114	2.199	2.041	2.111	2.196	2.036	2.103	2.198
	bias	0.058	0.114	0.199	0.041	0.111	0.196	0.036	0.103	0.198
	var	0.415	0.367	0.356	0.245	0.198	0.172	0.213	0.163	0.171
	mse	0.418	0.380	0.396	0.247	0.210	0.210	0.215	0.173	0.210
	NA4	94	41	17	87	37	18	89	43	23
	NA	8	3	3	10	5	2	10	3	0

Table 4: γ_x=2.5, N(0,1), $u(x)=cos(x)$

γ_z		0.5			1.5			2.5		
h		0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
0.0	mean	1.005	0.999	0.991	1.002	0.995	0.986	0.992	0.990	0.984
	bias	0.005	-0.001	-0.009	0.002	-0.005	-0.014	-0.008	-0.010	-0.016
	var	0.272	0.258	0.248	0.089	0.071	0.063	0.074	0.052	0.043
	mse	0.272	0.258	0.248	0.089	0.071	0.063	0.074	0.052	0.043
	NA4	0	0	0	0	0	0	0	0	0
	NA	0	0	0	0	0	0	0	0	0
0.5	mean	0.889	0.883	0.876	0.874	0.873	0.867	0.867	0.861	0.855
	bias	0.011	0.005	-0.002	-0.004	-0.005	-0.011	-0.011	-0.017	-0.023
	var	0.283	0.263	0.253	0.103	0.080	0.068	0.080	0.056	0.046
	mse	0.283	0.263	0.253	0.103	0.080	0.068	0.080	0.057	0.047
	NA4	0	0	0	0	0	0	0	0	0
	NA	0	0	0	0	0	0	0	0	0
1.0	mean	0.533	0.541	0.544	0.528	0.533	0.527	0.535	0.540	0.534
	bias	-0.007	0.001	0.004	-0.012	-0.007	-0.013	-0.005	0.000	-0.006
	var	0.320	0.281	0.271	0.146	0.109	0.091	0.122	0.084	0.064
	mse	0.320	0.281	0.271	0.146	0.109	0.091	0.122	0.084	0.065
	NA4	0	0	0	0	0	0	0	0	0
	NA	0	0	0	0	0	0	0	0	0

Table 5: γ_x=0.5, N(0,1), *u(x)=cos(x)*

γ_z		0.5			1.5			2.5		
h		0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
0.0	mean	1.011	1.001	0.991	1.002	0.993	0.982	0.988	0.982	0.979
	bias	0.011	0.001	-0.009	0.002	-0.007	-0.018	-0.012	-0.018	-0.021
	var	0.295	0.275	0.265	0.106	0.085	0.072	0.078	0.056	0.047
	mse	0.295	0.275	0.265	0.106	0.085	0.073	0.078	0.056	0.047
	NA4	2	1	0	0	0	0	1	0	0
	NA	1	0	0	0	0	0	0	0	0
0.5	mean	0.884	0.878	0.872	0.876	0.871	0.864	0.868	0.862	0.857
	bias	0.006	0.000	-0.006	-0.002	-0.007	-0.014	-0.010	-0.016	-0.021
	var	0.297	0.282	0.279	0.131	0.103	0.089	0.099	0.072	0.053
	mse	0.297	0.282	0.279	0.131	0.104	0.089	0.099	0.072	0.054
	NA4	4	1	0	9	2	1	4	1	0
	NA	1	0	0	0	0	0	0	0	0
1.0	mean	0.551	0.545	0.538	0.532	0.539	0.531	0.530	0.534	0.536
	bias	0.011	0.005	-0.002	-0.008	-0.001	-0.009	-0.010	-0.006	-0.004
	var	0.410	0.358	0.333	0.238	0.179	0.146	0.197	0.143	0.150
	mse	0.410	0.358	0.333	0.238	0.179	0.146	0.197	0.143	0.150
	NA4	94	41	17	87	37	18	89	43	23
	NA	8	3	3	10	5	2	10	3	0

Table 6: γ_x=2.5, t_3, $u(x)= x^2+x^4$

γ_z		0.5			1.5			2.5		
h		0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
0.0	mean	-0.002	0.005	0.019	0.000	0.012	0.026	0.004	0.016	0.029
	bias	-0.002	0.005	0.019	0.000	0.012	0.026	0.004	0.016	0.029
	var	0.224	0.213	0.208	0.049	0.041	0.037	0.027	0.020	0.017
	mse	0.224	0.213	0.209	0.049	0.041	0.038	0.027	0.020	0.018
	NA4	0	0	0	0	0	0	0	0	0
	NA	0	0	0	0	0	0	0	0	0
0.5	mean	0.321	0.345	0.377	0.334	0.357	0.387	0.334	0.358	0.388
	bias	0.009	0.033	0.065	0.022	0.045	0.075	0.022	0.046	0.076
	var	0.241	0.230	0.223	0.060	0.048	0.042	0.038	0.027	0.023
	mse	0.242	0.231	0.228	0.060	0.050	0.048	0.038	0.029	0.028
	NA4	0	0	0	0	0	0	0	0	0
	NA	0	0	0	0	0	0	0	0	0
1.0	mean	2.033	2.098	2.179	2.057	2.112	2.189	2.042	2.105	2.186
	bias	0.033	0.098	0.179	0.057	0.112	0.189	0.042	0.105	0.186
	var	0.313	0.270	0.261	0.139	0.091	0.081	0.117	0.070	0.057
	mse	0.314	0.279	0.293	0.143	0.104	0.116	0.118	0.081	0.092
	NA4	2	0	0	1	0	0	2	0	0
	NA	0	0	0	0	0	0	0	0	0

Table 7: γ_x=0.5, t_3, $u(x)= x^2+x^4$

γ_z		0.5			1.5			2.5		
h		0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
0.0	mean	-0.004	0.006	0.017	0.016	0.023	0.038	0.014	0.020	0.033
	bias	-0.004	0.006	0.017	0.016	0.023	0.038	0.014	0.020	0.033
	var	0.245	0.230	0.262	0.065	0.053	0.047	0.037	0.026	0.022
	mse	0.245	0.230	0.263	0.066	0.053	0.048	0.037	0.026	0.023
	NA4	4	1	0	2	1	0	1	0	0
	NA	1	1	0	0	0	0	0	0	0
0.5	mean	0.324	0.349	0.377	0.325	0.349	0.381	0.327	0.352	0.384
	bias	0.012	0.037	0.065	0.013	0.037	0.069	0.015	0.040	0.072
	var	0.293	0.269	0.254	0.088	0.066	0.067	0.075	0.046	0.038
	mse	0.293	0.270	0.258	0.088	0.068	0.072	0.075	0.048	0.043
	NA4	16	8	3	20	14	5	26	11	7
	NA	4	2	2	0	0	0	1	0	0
1.0	mean	2.029	2.090	2.162	2.040	2.092	2.174	2.051	2.100	2.174
	bias	0.029	0.090	0.162	0.040	0.092	0.174	0.051	0.100	0.174
	var	0.418	0.415	0.381	0.231	0.195	0.240	0.176	0.171	0.120
	mse	0.418	0.423	0.408	0.233	0.203	0.271	0.179	0.181	0.151
	NA4	193	89	53	185	84	52	184	87	58
	NA	27	16	12	30	13	5	28	16	6

Table 8: γ_x=2.5, t_3, $u(x)$=$cos(x)$

γ_z		0.5			1.5			2.5		
h		0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
0.0	mean	0.986	0.979	0.972	0.988	0.986	0.979	0.992	0.990	0.983
	bias	-0.014	-0.021	-0.028	-0.012	-0.014	-0.021	-0.008	-0.010	-0.017
	var	0.224	0.213	0.208	0.049	0.041	0.037	0.027	0.020	0.017
	mse	0.224	0.214	0.209	0.049	0.041	0.038	0.027	0.020	0.017
	NA4	0	0	0	0	0	0	0	0	0
	NA	0	0	0	0	0	0	0	0	0
0.5	mean	0.863	0.860	0.853	0.876	0.871	0.865	0.877	0.873	0.866
	bias	-0.015	-0.018	-0.025	-0.002	-0.007	-0.013	-0.001	-0.005	-0.012
	var	0.241	0.229	0.222	0.060	0.047	0.042	0.038	0.027	0.022
	mse	0.241	0.229	0.222	0.060	0.048	0.042	0.038	0.027	0.022
	NA4	0	0	0	0	0	0	0	0	0
	NA	0	0	0	0	0	0	0	0	0
1.0	mean	0.522	0.520	0.521	0.543	0.535	0.531	0.532	0.530	0.533
	bias	-0.018	-0.020	-0.019	0.003	-0.005	-0.009	-0.008	-0.010	-0.007
	var	0.313	0.269	0.256	0.139	0.086	0.072	0.115	0.066	0.050
	mse	0.313	0.270	0.257	0.139	0.086	0.072	0.115	0.066	0.050
	NA4	2	0	0	1	0	0	2	0	0
	NA	0	0	0	0	0	0	0	0	0

Table 9: γ_x=0.5, t_3, $u(x)=cos(x)$

γ_z		0.5			1.5			2.5		
h		0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
0.0	mean	0.985	0.980	0.969	1.004	0.997	0.991	1.002	0.993	0.986
	bias	-0.015	-0.020	-0.031	0.004	-0.003	-0.009	0.002	-0.007	-0.014
	var	0.245	0.231	0.260	0.065	0.053	0.047	0.037	0.026	0.022
	mse	0.245	0.231	0.261	0.065	0.053	0.047	0.037	0.026	0.022
	NA4	4	1	0	2	1	0	1	0	0
	NA	1	1	0	0	0	0	0	0	0
0.5	mean	0.867	0.865	0.856	0.868	0.867	0.862	0.870	0.869	0.863
	bias	-0.011	-0.013	-0.022	-0.010	-0.011	-0.016	-0.008	-0.009	-0.015
	var	0.295	0.269	0.254	0.087	0.064	0.065	0.074	0.045	0.037
	mse	0.295	0.269	0.254	0.087	0.064	0.065	0.074	0.045	0.037
	NA4	16	8	3	20	14	5	26	11	7
	NA	4	2	2	0	0	0	1	0	0
1.0	mean	0.536	0.531	0.525	0.547	0.536	0.544	0.562	0.549	0.552
	bias	-0.004	-0.009	-0.015	0.007	-0.004	0.004	0.022	0.009	0.012
	var	0.389	0.379	0.340	0.225	0.171	0.201	0.153	0.141	0.085
	mse	0.389	0.379	0.340	0.225	0.171	0.201	0.153	0.141	0.086
	NA4	193	89	53	185	84	52	184	87	58
	NA	27	16	12	30	13	5	28	16	6