<table>
<thead>
<tr>
<th>Title</th>
<th>Nonparametric estimation of conditional medians for linear and related processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>HONDA, Toshio</td>
</tr>
<tr>
<td>Citation</td>
<td>Annals of the Institute of Statistical Mathematics, 62(6): 995-1021</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2010-12</td>
</tr>
<tr>
<td>Type</td>
<td>Journal Article</td>
</tr>
<tr>
<td>Text Version</td>
<td>author</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10086/22991">http://hdl.handle.net/10086/22991</a></td>
</tr>
<tr>
<td>Copyright ©</td>
<td>The Institute of Statistical Mathematics, Tokyo 2008</td>
</tr>
</tbody>
</table>
Nonparametric estimation of conditional medians for linear and related processes

Toshio Honda

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.

Keywords 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

\[
Y_i = u(X_i) + V_i,
\]

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. Even when \( Z_i - m \) has finite variance, \( G(Z_i - m) \) may not have finite mean or finite variance and we may have to use some robust estimators as in this paper. 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.


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 \( \alpha \)-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 \( \alpha \)-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 prop-
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) 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) is motivated by Mielniczuk and Wu (2004). Mielniczuk and Wu (2004) allows for some dependence between \( \epsilon_1 \) and \( \zeta_1 \) in (23) 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 (2008), which studies kernel density estimation for heavy-tailed linear processes.

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

\[
X_i = \sum_{j=0}^{\infty} b_j \epsilon_{i-j} \quad \text{and} \quad Z_i = \sum_{j=0}^{\infty} c_j \zeta_{i-j},
\]

(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 \(\epsilon_1\) and \(\zeta_1\) by \(\sigma_\epsilon^2\) and \(\sigma_\zeta^2\), respectively. We assume that for some \(r > 2\),

\[
E(\epsilon_1^r) < \infty \quad \text{and} \quad E(\zeta_1^r) < \infty.
\]

We also assume that

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

(4)

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)-(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 \( \text{E}\{(\sum_{i=1}^{n} X_i)^2\} \). In this paper, \( n \) stands for the sample size.

When \( \sum_{j=0}^{\infty} |b_j| < \infty, \sum_{i=0}^{\infty} |r_X(j)| < \infty \) and this property is called short-range dependence. Then we have \( \sigma_{n,X}^2 = O(n) \). Note that \( \gamma_X \) 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 \( \{X_i\}_{i=0}^{\infty} \) is an \( \alpha \)-mixing process with sufficiently rapidly decaying mixing coefficients. See Doukhan (1994) about mixing processes. Thus the results on \( \alpha \)-mixing processes do not cover some results on short-range dependent processes of this paper.

When \( \sum_{j=0}^{\infty} |b_j| = \infty, \sum_{i=0}^{\infty} |r_X(j)| = \infty \) and this property is called long-range dependence. Then we have \( \lim_{n \to \infty}(\sigma_{n,X}^2/n) = \infty \). Then \( \gamma_X \) 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 \( \gamma_X < 1, \)

\[
r_X(t) \sim C_{\gamma_X} t^{-\gamma_X} L_X(t) \sigma_x^2 \quad \text{and} \quad \sigma_{n,X}^2 \sim D_{\gamma_X} n^{2-\gamma_X} L_X^2(n) \sigma_x^2,
\]

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 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 1. The results of the simulation study are presented in Section 3.

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

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) 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 \(\Omega\) of \((x_0, m_0)^T\) \(\frac{\partial V}{\partial z}(x_0, m_0) \neq 0\). There are also three positive constants, \(\delta_1, \delta_2, \delta_3\), such that \([x_0 - \delta_1, x_0 + \delta_1] \times [m_0 - \delta_2, m_0 + \delta_2] \subset \Omega\) and \(\inf\{|V(x, z)| : |x - x_0| < \delta_1\ \text{and} \ |z - m_0| \geq \delta_2\} > \delta_3\).

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 \(\kappa_j\) and \(\nu_j\) by

\[\kappa_j = \int \xi^j K(\xi) d\xi \quad \text{and} \quad \nu_j = \int \xi^j K^2(\xi) d\xi,\]

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 \(\epsilon_1\) and \(\zeta_1\), respectively. Then for some positive \(\delta_1\) and \(\delta_2\),

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

We also assume that both of \(\epsilon_1\) and \(\zeta_1\) have continuously differentiable density functions \(f_1(x)\) and \(g_1(z)\), respectively. In addition, they and their derivatives satisfy the following conditions.

\[|v(t)| \leq C_2 \frac{1}{1 + t^2} \quad \text{and} \quad |v(s) - v(t)| \leq C_1 \frac{1}{1 + t^2} \quad \text{for} \quad |s - t| < 1, \quad (6)
\]

where \(v = f_1, f_1', g_1, \) or \(g_1'\). Remember that we have already assumed that \(\mathbb{E}[\{\epsilon_1\}^r] < \infty\) and \(\mathbb{E}[\{\zeta_1\}^r] < \infty\) for some \(r > 2\).
Remark 1  All the necessary technical conditions on density functions are assured by (2)-(4) and Assumption A4. When we define $X_{i,j}$ and $Z_{i,j}$ by

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

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 (6) 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 Theorems 1-3 and Lemma 1 of Honda (2008) imply that the part of (6) 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

$$\bar{X}_{i,j} = X_i - X_{i,j} = \sum_{k=0}^{\infty} b_k \varepsilon_{i-k} \quad \text{and} \quad \bar{Z}_{i,j} = Z_i - Z_{i,j} = \sum_{k=0}^{\infty} c_k \zeta_{i-k}. \quad (7)$$

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

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

$$\mathcal{S}_{0,i} = \sigma(\varepsilon_i, \zeta_i, \varepsilon_{i-1}, \zeta_{i-1}, \ldots.), \quad \mathcal{S}_{1,i} = \sigma(\varepsilon_i, \varepsilon_{i-1}, \ldots.),$$

and

$$\mathcal{S}_{2,i} = \sigma(\zeta_i, \zeta_{i-1}, \ldots.),$$

where $\sigma(\cdots)$ stands for the $\sigma$-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

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

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

$$V_i^* = Y_i - (u(x_0), h u'(x_0))^T \eta_i = V_i + \frac{1}{2} \left( \frac{X_i - x_0}{h} \right)^2 h^2 u''(\bar{X}_i), \quad (8)$$

where $\eta_i = (1, (X_i - x_0)/h)^T$. We estimate $(u(x_0), h u'(x_0))$ by $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2)^T$ defined in (9).

$$\hat{\beta} = \arg\min_{\beta \in \mathbb{R}^2} \sum_{i=1}^{n} K_i |Y_i - \eta_i^T \beta|, \quad (9)$$
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 \( \gamma_Z \) and \( L_Z(j) \) and there are the following two cases. The normalization constant \( \tau_n \) 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 \( \tau_n = \sqrt{n h} \).

**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, \( \tau_n = c_{1/2} n^{3/5} \) and the asymptotic distribution of \( \hat{\beta} \) is the same as for i.i.d. observations. In Cases 2, \( Z_i \) have long-range dependence. If \( L_Z(j) \) is a constant function, \( \tau_n \log n \) for some positive \( d \). The asymptotic distribution of \( \hat{\beta} \) depends only on \( \partial V / \partial z(x_0, m(x_0)) \) and \( f_Z(i) \).

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

\[
\hat{\theta} = \tau_n \left( \frac{\hat{\beta}_1 - u(x_0)}{\hat{\beta}_2 - hu(x_0)} \right) .
\]

We can represent \( \hat{\theta} \) as

\[
\hat{\theta} = \arg\min_{\theta \in \mathbb{R}^2} \sum_{i=1}^{n} K_i |V_i - \tau_n^{-1} \eta_i^\top \theta|, \tag{11}
\]

Before we state Theorem 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

\[
f_V(0|x) = g(m(x))(\frac{\partial V}{\partial z} (x, m(x)))^{-1} \tag{12}
\]
in a neighborhood of \( x_0 \).

**Theorem 1** Suppose that Assumptions A1-A5 hold in our setting (1)-(4). Then as \( n \to \infty \),

**Case 1** \( \tau_n = \sqrt{n h} \):

\[
\hat{\theta} \overset{d}{\rightarrow} N \left( \begin{pmatrix} \frac{c_{3/2} u''(x_0)x_2}{2} \\ 0 \end{pmatrix}, \frac{1}{4f_V^2(0|x_0)f(x_0)} \begin{pmatrix} 0 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \right) ;
\]

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

\[
\hat{\theta} = -\frac{1}{2f_V(0|x_0)} \int \text{sign}(V(x, z)) \frac{g'(z)dz}{\sigma_{n,Z}} \sum_{i=1}^{n} Z_i + o_p(1).
\]
The proof of Theorem 1 is given in Section 4. It follows from (12) that
\[
\frac{1}{2f_V(0|x_0)} \int \text{sign}(V(x, z))g'(z)dz = -\frac{\partial V}{\partial z}(x_0, m_0).
\]

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 \(\sigma^{-1}_n \sum_{i=1}^{n} Z_i \xrightarrow{d} N(0, 1)\) 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 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 \(\gamma_X\) 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 1 implies that the two estimators are asymptotically independent in Case 1 since \(E\{K((X_i - x_0)/h)K((X_i - x_1)/h)\} = 0\) 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 \(\{X_i, Y_i\}_{i=1}^{\infty}\). 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 2.7.1 and the rq function of the quantreg library. See R Development Core Team (2008) and Koenker (2008) about R and the quantreg library.

In this study, we set
\[
Y_i = u(X_i) + Z_i, \quad (13)
\]
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
\[
b_j = \begin{cases} (j + 1)^{-1+\gamma_x}/(\sum_{j=0}^{999}(j + 1)^{-1+\gamma_x})^{1/2}, & 0 \leq j \leq 999, \\ 0, & j \geq 1000, \end{cases}
\]

\(c_j = \{b_j\}_{j=0}^{\infty}\) and \(\{c_j\}_{j=0}^{\infty}\) are given by
\[
c_j = \begin{cases} (j + 1)^{-\gamma_x}/(\sum_{j=0}^{999}(j + 1)^{-\gamma_x})^{1/2}, & 0 \leq j \leq 999, \\ 0, & j \geq 1000, \end{cases}
\]
In Tables 2-5, \( e_i \) and \( \zeta_i \) follow the standard normal distribution. In Tables 6-9, \( \sqrt{3}e_i \) and \( \sqrt{3}\zeta_i \) follow the \( t \)-distribution with d.f. 3. Then it is easy to see that 
\[
E(X_i^2) = E(Z_i^2) = 1.
\]
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

\[
\begin{align*}
    u(0.0) &= 0.000, \\
    u(0.5) &= 0.3125, \\
    u(1.0) &= 2.000 \quad \text{in Tables 2, 3, 6, 7, and} \\
    u(0.0) &= 1.000, \\
    u(0.5) &= 0.8776, \\
    u(1.0) &= 0.5403 \quad \text{in Tables 4, 5, 8, 9.}
\end{align*}
\]

As for the design density function,

\[
\begin{align*}
    f(0.0) &= 0.399, \\
    f(0.5) &= 0.352, \\
    f(1.0) &= 0.242 \quad \text{in Tables 2-5,} \quad (14) \\
    f(0.0) &= 0.581, \\
    f(0.5) &= 0.408, \\
    f(1.0) &= 0.179 \quad \text{in Tables 6, 8 of } \gamma X = 2.5, \quad (15) \\
    f(0.0) &= 0.461, \\
    f(0.5) &= 0.384, \\
    f(1.0) &= 0.219 \quad \text{in Tables 7, 9 of } \gamma X = 0.5. \quad (16)
\end{align*}
\]

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 1 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 \( \sqrt{3}e_i \) 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 \( \gamma Z \) 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 (14)-(16). 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 1

The proof of Theorem 1 is presented in this section. We begin with three results, Lemmas 1-3, which will be verified after the proof of Theorem 1. The setting (1)-(4) is assumed throughout this section. We omit a.s.(almost surely) for notational convenience.

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

**Lemma 1** Suppose that Assumptions A1-A4 hold. Then we have

**Case 1:**
\[
\frac{\tau_n}{nh} \sum_{i=1}^{n} K_i \eta_i \text{sign}(V_i) \overset{d}{\to} N \left( \begin{pmatrix} 0 \\ \nu_0 \end{pmatrix}, \nu_0^2 \right).
\]

**Case 2:**
\[
\frac{\tau_n}{nh} \sum_{i=1}^{n} K_i \eta_i \text{sign}(V_i) + \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) f(x_0) \int \text{sign}(V(x,z)) g'(z) dz \frac{1}{\sigma_n, Z} \sum_{i=1}^{n} Z_i = o_p(1).
\]

The second result is used to evaluate the loss function in (11). See (8) for the definitions of \( \eta_i \) and \( V_i^* \).

**Lemma 2** Suppose that Assumptions A1-A4 hold. Then we have for any fixed \( \theta \),
\[
\frac{\tau_n^2}{nh} \sum_{i=1}^{n} K_i (|V_i^* - \tau_n^{-1} \eta \theta| - |V_i^*|)) = \theta^T \begin{pmatrix} 1 \\ 0 \end{pmatrix} \kappa_2 \theta f_V(0|x_0)f(x_0) - \left( \frac{\tau_n}{nh} \sum_{i=1}^{n} K_i \eta_i \text{sign}(V_i^*) \right)^T \theta + o_p(1).
\]

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 3** Suppose that Assumptions A1-A4 hold. Then we have
\[
\frac{\tau_n}{nh} \sum_{i=1}^{n} K_i \eta_i \text{sign}(V_i^*)
= \frac{\tau_n}{nh} \sum_{i=1}^{n} K_i \eta_i \text{sign}(V_i) + \frac{\tau_n}{\sqrt{nh}} (c^{3/2} \kappa_2 u''(x_0) f_V(0|x_0) f(x_0), 0)^T + o_p(1).
\]
Now we prove Theorem 1.

**Proof of Theorem 1** Recall that \( \tau_n / \sqrt{n}h = 1 \) in Case 1 and \( \tau_n / \sqrt{n}h = o(1) \) in Case 2. \((11) \) is equivalent to

\[
\hat{\theta} = \arg\min_{\theta \in \mathbb{R}^{2}} \frac{\tau_n^2}{nh} \sum_{i=1}^{n} K_i(|V^*_i - \tau_n^{-1} \eta \theta| - |V^*_i|). \tag{17}
\]

By Lemmas 2-3, we have for any fixed \( \theta \in \mathbb{R}^{2} \),

\[
\frac{\tau_n^2}{nh} \sum_{i=1}^{n} K_i(|V^*_i - \tau_n^{-1} \eta \theta| - |V^*_i|) = \theta^T \begin{pmatrix} 1 & 0 \\ 0 & \kappa_2 \end{pmatrix} f^T(x_0) f(x_0) - \left( \frac{\tau_n}{nh} \sum_{i=1}^{n} K_i \eta \text{sign}(V_i) \right)^T \theta - \frac{\tau_n}{\sqrt{nh}} (\epsilon^{5/2} \kappa_2 u''(x_0) f^T(0|x_0)) f(x_0), 0) \theta + o_p(1). \tag{18}
\]

As in Pollard (1991), Fan et al. (1994), and Hall et al. (2002), the convexity lemma implies that \((18) \) holds uniformly on \( \{ |\theta| < M \} \) for any positive \( M \).

We consider the RHS of \((18) \). Lemma 1 implies that

\[
\frac{\tau_n}{nh} \sum_{i=1}^{n} K_i \eta \text{sign}(V_i) = O_p(1). \tag{19}
\]

Combining \((19) \), \( \tau_n / \sqrt{n}h = O(1) \), the uniformity of \((18) \), and the convexity of the objective function in \((17) \), 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 \((18) \) again, we obtain

\[
\hat{\theta} = \frac{1}{2f^T(0|x_0) f(x_0)} \begin{pmatrix} 1 & 0 \\ 0 & \kappa_2 \end{pmatrix}^{-1} \left\{ \frac{\tau_n}{nh} \sum_{i=1}^{n} K_i \eta \text{sign}(V_i) + \frac{\tau_n}{\sqrt{nh}} (\epsilon^{5/2} \kappa_2 u''(x_0) f^T(0|x_0)) f(x_0), 0)^T \right\} + o_p(1). \tag{20}
\]

The results of the theorem follow from \((20) \) and Lemma 1. Hence the proof of the theorem is complete.

We next consider the evaluation of \( n^{-1} \sum_{i=1}^{n} A(\hat{X}_{i,1}) B(\hat{Z}_{i,1}) \), where \( \hat{X}_{i,1} \) and \( \hat{Z}_{i,1} \) are defined in \((7) \), \( |A(\hat{X}_{i,1})| \leq M_1, |B(\hat{Z}_{i,1})| \leq M_1, \) and \( A(\hat{X}_{i,1}) \) and \( B(\hat{Z}_{i,1}) \) depend on \( \xi \in [-C_K, C_K] \). The evaluation is important in the proofs of Lemmas 1-3 and given in Lemma 4 below.

Define \( U_{A,i,j} \) and \( U_{B,i,j} \) respectively by

\[
U_{A,i,j} = E\{A(\hat{X}_{i,1})|S_{1,i,j} - 1\} - E\{A(\hat{X}_{i,1})|S_{1,i,j} - 1\} - A_i b_j \epsilon_{i,j}, \tag{21}
\]

\[
U_{B,i,j} = E\{B(\hat{Z}_{i,1})|S_{2,i,j} - 1\} - E\{B(\hat{Z}_{i,1})|S_{2,i,j} - 1\} - B_i c_j \zeta_{i,j}. \tag{22}
\]
where
\[ A_1 = \frac{\partial}{\partial v} E\{A(\tilde{X}_{1,1} + v)\}\big|_{v=0} \quad \text{and} \quad B_1 = \frac{\partial}{\partial v} E\{B(\tilde{Z}_{1,1} + v)\}\big|_{v=0}. \] (23)

Set \( U_{A,i,0} = U_{B,i,0} = 0. \) Then we have
\[ \sum_{j=1}^{\infty} U_{A,i,j} = A(\tilde{X}_{1,1}) - E\{A(\tilde{X}_{1,1})\} - A_1 \tilde{X}_{1,1}, \] (24)
\[ \sum_{j=1}^{\infty} U_{B,i,j} = B(\tilde{Z}_{1,1}) - E\{B(\tilde{Z}_{1,1})\} - B_1 \tilde{Z}_{1,1}. \] (25)

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

Lemma 4: Suppose that Assumption A4 holds and that we have, uniformly in \( \xi \in [-C_K, C_K] \),
\[ E\{U_{A,i,j}^2\} \leq M_2(j+1)^{-(\delta+1+\gamma_X)} \quad \text{and} \quad E\{U_{B,i,j}^2\} \leq M_2(j+1)^{-(\delta+1+\gamma_Z)} \] (26)
for positive numbers \( M_2 \) and \( \delta \) such that \( \delta + \gamma_X \neq 1, \delta + \gamma_Z \neq 1, \) and \( 2\delta < 1 \land \gamma_X \). Then
\[ \frac{1}{n} \sum_{i=1}^{n} A(\tilde{X}_{1,1})B(\tilde{Z}_{1,1}) = E\{A(\tilde{X}_{1,1})\}E\{B(\tilde{Z}_{1,1})\} + A_1E\{B(\tilde{Z}_{1,1})\} \frac{1}{n} \sum_{i=1}^{n} \tilde{X}_{1,1} \]
\[ + E\{A(\tilde{X}_{1,1})\}B_1 \frac{1}{n} \sum_{i=1}^{n} \tilde{Z}_{1,1} + O_m(M_1 M_2(n^{-(\delta+\gamma_X)/2} \lor n^{-1/2})) \]
\[ + A_1 B_1 O_m(n^{-(\delta+\gamma_Z)/2} \lor n^{-1/2}) + B_1 O_m(M_2(n^{-(\delta+\gamma_X)/2} \lor n^{-1/2})) \]
\[ + E\{B(\tilde{Z}_{1,1})\} O_m(M_2(n^{-(\delta+\gamma_X)/2} \lor n^{-1/2})). \]

We verify Lemma 4 after Lemmas 1-3 are proved. A similar result to Lemma 4 is given Lemma 4.1 of Koul et al. (2004). However, we cannot use the result directly to prove Lemmas 1-3. Note that (26) have to be verified when we use Lemma 4. The proofs of (26) for Lemmas 1-3 are almost the same as the arguments of Section 6 of Koul and Surgailis (2002). We prove only the latter of (26) for Lemma 2. See Lemma 6 at the end of this section for the proof.

Proof of Lemma 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 $T_i = K_1 \text{sign}(V_i) - E\{K_1 \text{sign}(V_i) | \mathcal{S}_{0,i-1}\}$ and notice that we have by (2)

$$\frac{1}{h} E\{K_1 \text{sign}(V_i) | \mathcal{S}_{0,i-1}\}$$

$$= \frac{1}{h} \int \int K\left(\frac{x + \hat{X}_{i,1} - X_0}{h}\right) \text{sign}(V(x + \hat{X}_{i,1}, z + \hat{Z}_{i,1})) f_1(x) g_1(z) dx dz$$

$$= \int K(\xi) \left\{ f_1(x_0 + \xi h - \hat{X}_{i,1}) \int \text{sign}(V(x_0 + \xi h, z)) g_1(z - \hat{Z}_{i,1}) dz \right\} d\xi.$$

Set $A(\hat{X}_{i,1}) = f_1(x_0 + \xi h - \hat{X}_{i,1})$ and $B(\hat{Z}_{i,1}) = \int \text{sign}(V(x_0 + \xi h, z)) g_1(z - \hat{Z}_{i,1}) dz$ in Lemma 4. Then, by Assumptions A2 and A4, and the arguments of Section 6 of Koul and Surgailis (2002) that (26) holds. The details are omitted. Now by Lemma 4, there is a positive constant $\delta$ such that

$$\frac{1}{n} \sum_{i=1}^{n} A(\hat{X}_{i,1}) B(\hat{Z}_{i,1})$$

$$= -f(x_0) \int \text{sign}(V(x_0, z)) g'(z) dz \frac{1}{n} \sum_{i=1}^{n} \hat{Z}_{i,1} + O_p(n^{-1/2}) + O(n^{-1/2}).$$

By (28), Jensen’s inequality regarding $\int K(\xi) d\xi$, and Fubini’s theorem, we have

$$\frac{1}{n h} \sum_{i=1}^{n} \{K_1 \text{sign}(V_i) | \mathcal{S}_{0,i-1}\}$$

$$= -f(x_0) \int K(\xi) \left\{ \int \text{sign}(V(x_0 + \xi h, z)) g'(z) dz \right\} d\xi \frac{1}{n} \sum_{i=1}^{n} \hat{Z}_{i,1}$$

$$+ O_p(n^{-1/2}) + O(n^{-1/2}).$$

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

$$\frac{\tau_n}{n h} \sum_{i=1}^{n} \{K_1 \text{sign}(V_i) | \mathcal{S}_{0,i-1}\}$$

$$= \begin{cases} o_p(1) & \text{in Case 1,} \\ -f(x_0) \int \text{sign}(V(x_0, z)) g'(z) dz \frac{\tau_n}{n} \sum_{i=1}^{n} Z_i + o_p(1) & \text{in Case 2.} \end{cases}$$
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 \(|T_i| \leq C\), we have only to verify the convergence in probability of the conditional variance. Noticing that \(E\{K_i \cdot \text{sign}(V_i)|S_{0,i-1}\} \leq C h\) and \(|\text{sign}(V_i)| = 1\), we can evaluate the conditional variance in the following way.

\[
\frac{1}{nh} \sum_{i=1}^{n} E\{T_i^2 | S_{0,i-1}\} = \frac{1}{nh} \sum_{i=1}^{n} \int \int K^2 \left( \frac{x + \bar{X}_{i-1} - x_0}{h} \right) f_1(x) g_1(z) dx dz + O_p(h) \\
= \frac{1}{n} \sum_{i=1}^{n} \int K^2(\xi) f_1(x_0 + \xi - \bar{X}_{i-1}) d\xi + O_p(h) \\
= \frac{\nu_0}{n} \sum_{i=1}^{n} f_1(x_0 - \bar{X}_{i-1}) + O_p(h) \rightarrow \nu_0 f(x_0).
\]

The last line follows from the ergodic theorem.

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

\[
\frac{1}{\sqrt{nh}} \sum_{i=1}^{n} T_i \overset{d}{\rightarrow} N(0, f(x_0)\nu_0).
\]

Therefore, recalling that \(\tau_n = o(1)\) in Case 2, we have

\[
\frac{\tau_n}{nh} \sum_{i=1}^{n} T_i \left\{ \begin{array}{ll}
\overset{d}{\rightarrow} N(0, f(x_0)\nu_0) & \text{in Case 1}, \\
= o_p(1) & \text{in Case 2}.
\end{array} \right.
\]

The desired result follows from (30) and (31). Hence the proof is complete.

**Proof of Lemma 2** Define \(S_\theta(X_i, Z_i)\) by

\[
S_\theta(X_i, Z_i) = |V_i^* - \tau_n^{-1} \eta^T \theta| - |V_i| + \tau_n^{-1} \eta^T \theta \text{sign}(V_i^*).
\]

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

\[
S_\theta(X_i, Z_i) \leq 2|\tau_n^{-1} \eta^T \theta| I(|V_i| \leq C\tau_n^{-1} |\theta|).
\]

Set

\[
T_i = K_i S_\theta(X_i, Z_i) - E\{K_i S_\theta(X_i, Z_i)|S_{0,i-1}\}.
\]

Since

\[
E\left\{ \left( \frac{\tau_n}{nh} \sum_{i=1}^{n} T_i \right)^2 \right\} \leq C \frac{\tau_n^2 |\theta|^2}{(nh)^2} \sum_{i=1}^{n} E\{K_i^2 I(|V_i| \leq C\tau_n^{-1} |\theta|)\} \leq C \frac{\tau_n |\theta|^3}{nh} \rightarrow 0,
\]

\[
\frac{\tau_n}{nh} \sum_{i=1}^{n} T_i \overset{d}{\rightarrow} N(0, f(x_0)\nu_0).
\]
we have
\[ \frac{\tau^2}{n h} \sum_{i=1}^{n} T_i = o_p(1). \] (34)

Assumption A2 implies that \( d|Z_i - m(X_i)| \leq |V_i| \) when both \(|V_i|\) and \(|X_i - x_0|\) is sufficiently small. We used this inequality to evaluate \( \mathbb{E}\{K_n^2 I(|V_i| \leq C\tau_n^{-1}|\theta|)\} \) and this inequality will be also used to verify (26) for Lemmas 2-3.

Next notice that we have by (2),
\[ \frac{\tau^2}{n h} \mathbb{E}\{K_i S_\theta(X_i, Z_i)|S_{0,i-1}\} \]
\[ = \frac{\tau^2}{n h} \int \int K\left(\frac{x + \tilde{X}_{i,1} - x_0}{h}\right) S_\theta(x + \tilde{X}_{i,1}, z + \tilde{Z}_{i,1}) f_1(x) g_1(z) dx dz \]
\[ = \int K(\xi) \left\{ f_1(x_0 + \xi h - \tilde{X}_{i,1}) \tau_n^2 \int S_\theta(x_0 + \xi h, z) g_1(z - \tilde{Z}_{i,1}) dz \right\} d\xi. \]

Then set \( A(\tilde{X}_{i,1}) = f_1(x_0 + \xi h - \tilde{X}_{i,1}) \) as in the proof of Lemma 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. By Assumptions A1-A2 and Remark 1, we have uniformly in \( \xi \in [-C_K, C_K] \),
\[ \mathbb{E}\{B(\tilde{Z}_{i,1})\} = ((1, \xi)\theta^2 f_V(0|x_0) + o(1), \]
\[ B_1 = -\tau_n^2 \int S_\theta(x_0 + \xi h, z) g'(z) dz = O(1). \]

(26) follows from (2)-(4), Assumptions A1, A2, and A4, and the arguments of Section 6 of Koul and Surgailis (2002). We prove only the latter of (26) in Lemma 6 at the end of this section for reference. Then by Lemma 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 (29), we obtain
\[ \frac{\tau^2}{n h} \mathbb{E}\{K_i S_\theta(X_i, Z_i)|S_{0,i-1}\} \]
\[ = \int K(\xi) ((1, \xi)\theta^2 d\xi f_V(0|x_0) f(x_0) + o_p(1) \]
\[ = \theta^T \begin{pmatrix} 1 & 0 \\ 0 & \kappa_2 \end{pmatrix} (f_V(0|x_0) f(x_0) + o_p(1). \]

The desired result follows from (34) and (35). Hence the proof is complete.

**Proof of Lemma 3** The proof is similar to that of Lemma 2 and we consider only the first element.

Define \( T_i \) by
\[ T_i = K_i (\text{sign}(V_i^*) - \text{sign}(V_i)) - \mathbb{E}\{K_i (\text{sign}(V_i^*) - \text{sign}(V_i))|S_{0,i-1}\}. \]
Since $|V_t^* - V_t| \leq Ch^2$, 

$$
E \left\{ \left( \frac{\tau_n}{nh} \sum_{i=1}^n T_i \right)^2 \right\} \leq \frac{4\tau_n^2}{(nh)^2} \sum_{i=1}^n E \{ K_i^2 I(|V_t| \leq Ch^2) \} \leq C \frac{\tau_n^2 h^3}{nh^2} = C \frac{\tau_n^2 h}{n} \to 0.
$$

(36)

Next notice that we have by (2),

$$
\frac{\tau_n}{h} E \{ K_i (\text{sign}(V_t^*) - \text{sign}(V_t)) | S_{0,i-1} \} = \frac{\tau_n}{h} \int \int K \left( \frac{x + \tilde{X}_{i,1} - x_0}{h} \right) \times (\text{sign}(V^*(x + \tilde{X}_{i,1}, z + \tilde{Z}_{i,1}) - \text{sign}(V(x + \tilde{X}_{i,1}, z + \tilde{Z}_{i,1})) f_1(x) g_1(z) dx dz \\
= \int K(\xi) \left\{ f_1(x_0 + \xi h - \tilde{X}_{i,1}) \right. \\
\times \left. \tau_n \int (\text{sign}(V^*(x_0 + \xi h, z) - \text{sign}(V(x_0 + \xi h, z))) g_1(z - \tilde{Z}_{i,1}) dz \right\} d\xi.
$$

Set $A(\tilde{X}_{i,1}) = f_1(x_0 + \xi h - \tilde{X}_{i,1})$ as in the proof of Lemma 1 and $B(\tilde{Z}_{i,1}) = \tau_n \int (\text{sign}(V^*(x_0 + \xi h, z) - \text{sign}(V(x_0 + \xi h, z))) g_1(z - \tilde{Z}_{i,1}) dz$ in Lemma 4. Then by (8), Assumptions A1 and A2, and Remark 1, we have uniformly in $\xi \in [-C_K, C_K]$, 

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

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

(26) follows from (2)-(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, we obtain 

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

Finally in the same way as (29) and (35), we can see that 

$$
\frac{\tau_n}{nh} \sum_{i=1}^n E \{ K_i (\text{sign}(V_t^*) - \text{sign}(V_t)) | S_{0,i-1} \} = \frac{\tau_n}{\sqrt{nh}} e^{5/2 \kappa_2 u''(x_0)} f_V(0|x_0) f(x_0) + o_p(1).
$$

(37)

The desired result follows from (36) and (37). Hence the proof is complete.
Lemma 5 below is a tool to prove Lemma 4. Lemma 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 5** Let \( \{F_i\}_{i=-\infty}^\infty \) be a filtration. Suppose that we have a set of random variables \( \{U_{i,j} \mid 1 \leq i \leq n, \ 0 \leq j < \infty \} \) such that \( U_{i,j} \) are \( F_{i-j} \)-measurable and \( E\{U_{i,j} \mid F_{i-j-1}\} = 0 \). In addition we have another set of random variables \( \{D_i\}_{i=-\infty}^\infty \) which is independent of \( \{U_{i,j} \mid 1 \leq i \leq n, \ 0 \leq j < \infty \} \). Suppose that every \( D_i \) has the second moment and that

\[
E\{U_{i,j}^2\} \leq M(j+1)^{-(1+\gamma+\delta)}
\]

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

\[
E\{\sum_{i=1}^n D_i \sum_{j=0}^\infty U_{i,j}\}^2 \leq \sum_{j=0}^\infty (\sum_{i=1}^n \|D_i\|\|U_{i,j+i}\|)^2 + \sum_{j=-n}^{0-n} (\sum_{i=-j}^n \|D_i\|\|U_{i,j+i}\|)^2
\]

\[
\leq CM \max_{1 \leq i \leq n} (\|D_i\|)(n^{-2-\gamma-\delta} \vee n).
\]

The first inequality in (39) is essentially given in Lemma 2 of Mielniczuk and Wu (2004). The last inequality follows from (38) and some standard calculation. The details of the proof are omitted.

**Proof of Lemma 4** By applying Lemma 5 twice with \( \{F_i\}_{i=-\infty}^\infty = \{S_{1,i}\}_{i=-\infty}^\infty \) and \( \{S_{2,i}\}_{i=-\infty}^\infty \), respectively, we obtain

\[
\frac{1}{n} \sum_{i=1}^n A(\tilde{X}_{i,1}) B(\tilde{Z}_{i,1}) = \frac{1}{n} \sum_{i=1}^n A(\tilde{X}_{i,1}) [E\{B(\tilde{Z}_{i,1})\} + B_1 \tilde{Z}_{i,1}] + O_m(M_1M_2(n^{-(\delta+\gamma_x)/2} \vee n^{-1/2}))
\]

\[
= E\{A(\tilde{X}_{i,1})\} [E\{B(\tilde{Z}_{i,1})\} + A_1 E\{B(\tilde{Z}_{i,1})\} \frac{1}{n} \sum_{i=1}^n \tilde{X}_{i,1} + B_1 \frac{1}{n} \sum_{i=1}^n A(\tilde{X}_{i,1}) \tilde{Z}_{i,1} + O_m(M_1M_2(n^{-(\delta+\gamma_x)/2} \vee n^{-1/2}))
\]

\[
+ E\{B(\tilde{Z}_{i,1})\} O_m(M_2(n^{-(\delta+\gamma_x)/2} \vee n^{-1/2}))]
\]

We evaluate \( n^{-1} \sum_{i=1}^n A(\tilde{X}_{i,1}) \tilde{Z}_{i,1} \) in the last expression of (40). It is rewritten as

\[
\frac{1}{n} \sum_{i=1}^n A(\tilde{X}_{i,1}) \tilde{Z}_{i,1}
\]
\[ = \mathbb{E}(A(\tilde{X}_{i,1})) + \frac{1}{n} \sum_{i=1}^{n} \tilde{Z}_{i,1} + A_1 \frac{1}{n} \sum_{i=1}^{n} \tilde{X}_{i,1} \tilde{Z}_{i,1} \]

\[ + \frac{1}{n} \sum_{i=1}^{n} \{A(\tilde{X}_{i,1}) - \mathbb{E}(A(\tilde{X}_{i,1})) - A_1 \tilde{X}_{i,1}\} \tilde{Z}_{i,1}. \]

For the second term of the RHS of (41), we have by (2)-(4),

\[ \mathbb{E}\left\{ \left( \frac{1}{n} \sum_{i=1}^{n} \tilde{X}_{i,1} \tilde{Z}_{i,1} \right)^2 \right\} \]

\[ = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbb{E}(\tilde{X}_{i,1} \tilde{X}_{j,1}) \mathbb{E}(\tilde{Z}_{i,1} \tilde{Z}_{j,1}) \]

\[ \leq \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} (1 + |i - j|)^{-2\delta} |\mathbb{E}(\tilde{Z}_{i,1} \tilde{Z}_{j,1})| \leq C(n^{-\delta + \gamma X} \vee n^{-1}). \]

Remember that we took \( \delta \) which is smaller than \( (\gamma X \wedge 1)/2 \).

To deal with the third term of the RHS of (41), we use the following fact owing to (2)-(4) and Assumption A4.

\[ |\mathbb{E}[A(\tilde{X}_{i,1}) - \mathbb{E}(A(\tilde{X}_{i,1})) - A_1 \tilde{X}_{i,1}] [A(\tilde{X}_{j,1}) - \mathbb{E}(A(\tilde{X}_{j,1})) - A_1 \tilde{X}_{j,1}]| \]

\[ \leq \begin{cases} CM_2^2 (|i - j| + 1)^{-(\delta + \gamma X)}, & \text{if } \delta + \gamma X < 1, \\ CM_2^2 (|i - j| + 1)^{(1+\delta + \gamma X)/2}, & \text{if } \delta + \gamma X > 1. \end{cases} \quad (43) \]

See Lemma 6.1 of Koul and Surgailis (2002). Then by using (5) with \( \tilde{Z}_{i,1} \) and (43), we evaluate the variance and obtain

\[ \frac{1}{n} \sum_{i=1}^{n} \{A(\tilde{X}_{i,1}) - \mathbb{E}(A(\tilde{X}_{i,1})) - A_1 \tilde{X}_{i,1}\} \tilde{Z}_{i,1} = O_m(M_2(n^{-\delta + \gamma X} \vee n^{-1/2})). \]

(44)

The desired result follows from (40)-(42), and (44). Hence the proof is complete.

The latter of (26) for Lemma 2 is demonstrated in Lemma 6 below.

**Lemma 6** Suppose that Assumptions A1, A2, and A4 hold. When

\[ B(\tilde{Z}_{i,1}) = \int \tau^2 \phi(x_0 + \xi h, z) g_1(z - \tilde{Z}_{i,1}) dz, \]

there are positive \( M \) and \( \delta \) such that

\[ \mathbb{E}\{U^2_{B,i,j}\} \leq M(j + 1)^{-(\delta + 1 + \gamma X)} \text{ uniformly in } \xi \in [-C_K, C_K]. \]
Proof We follow the arguments of Koul and Surgailis (2002) and they rely on (2)-(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

\[
U^{(1)}_{i,j} = \int \tau^2_n S_\theta(x_0 + \xi h, z)(g_{j+1}(z - \tilde{Z}_{i,j}) - g_j(z - \tilde{Z}_{i,j+1}))dz + c_j \zeta_{i-j} \sum_{l=j+1}^{\infty} \left| q_l \right|^\gamma_1 \delta_{l+1}.
\]

\[
U^{(2)}_{i,j} = \int \tau^2_n S_\theta(x_0 + \xi h, z)(g_j(z - \tilde{Z}_{i,j}) - g_{j+1}(z - \tilde{Z}_{i,j+1}))dz,
\]

\[
U^{(3)}_{i,j} = -c_j \zeta_{i-j} \int \tau^2_n S_\theta(x_0 + \xi h, z)(g_{j+1}(z - \tilde{Z}_{i,j+1}) - g_j(z))dz.
\]

Then we have

\[
U_{i,j} = U^{(1)}_{i,j} + U^{(2)}_{i,j} + U^{(3)}_{i,j}.
\]

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 $\delta \geq 2$, we have by Rosenthal’s inequality,

\[
E[|\tilde{Z}_{i,j+1}|^\delta] \leq C \left\{ \sum_{l=j+1}^{\infty} c_l^{\delta/2} + \sum_{l=j+1}^{\infty} |q_l|^\gamma_1 \right\}.
\]

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

\[
|U^{(1)}_{i,j}| \leq C |c_j|^{\gamma_1} |\zeta_{i-j}|^{\gamma_1} (1 \lor |\tilde{Z}_{i,j+1}|^\gamma_1).
\]

Hence we have

\[
E[|U^{(1)}_{i,j}|^2] \leq C |c_j|^{2\gamma_1}.
\]

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

\[
|U^{(2)}_{i,j}| \leq C \gamma_2^2 \int \tau_n^2 S_\theta(x_0 + \xi h, z)(1 + |z - \tilde{Z}_{i,j}|)^{-2}dz \leq C \gamma_2^2.
\]
Finally we deal with $U^{(3)}_{i,j}$. We introduce a random variable $\tilde{Z}_{i,j+1}$ which is an independent copy of $\tilde{Z}_{i,j+1}$. We denote the expectation with respect to $\tilde{Z}_{i,j+1}$ by $\bar{E}\{\cdot\}$. By using $\tilde{Z}_{i,j+1}$ and $\bar{E}\{\cdot\}$, we can represent $U^{(3)}_{i,j}$ as

$$U^{(3)}_{i,j} = -c_j \zeta_{i-j} \left[ \int \tau_n^2 S_\theta(x_0 + \xi h, z)(g'_{j+1}(z - \tilde{Z}_{i,j+1}) - g'_{j+1}(z))dz - \bar{E}\left\{ \int \tau_n^2 S_\theta(x_0 + \xi h, z)(g'_{j+1}(z - \tilde{Z}_{i,j+1}) - g'_{j+1}(z))dz \right\} \right]$$

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

$$\bar{E}\left\{ \int \tau_n^2 S_\theta(x_0 + \xi h, z)(g'_{j+1}(z - \tilde{Z}_{i,j+1}) - g'_{j+1}(z))dz \right\} \leq C \bar{E}\{|\tilde{Z}_{i,j+1}|^2\} \quad (50)$$

and

$$\int \tau_n^2 S_\theta(x_0 + \xi h, z)(g'_{j+1}(z - \tilde{Z}_{i,j+1}) - g'_{j+1}(z))dz \leq C(|\tilde{Z}_{i,j+1}|^2). \quad (51)$$

(50) and (51) yield

$$E\{|U^{(3)}_{i,j}|^2\} \leq C|c_j|^2[E\{|\tilde{Z}_{i,j+1}|^2\} + E\{|\tilde{Z}_{i,j+1}|^{2\gamma_2}\}]. \quad (52)$$

The desired result follows from (45), (46), (48), (49), and (52). Hence the proof is complete.

Acknowledgements The author appreciates valuable comments of the associate editor and two referees very much. The research is partly supported by the Seimeikai Foundation.

References


Table 1: Parameters of tables

<table>
<thead>
<tr>
<th></th>
<th>Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma_x )</td>
<td>2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9</td>
</tr>
<tr>
<td>( u(x) )</td>
<td>( x^2+x^4 ) ( x^2+x^4 ) ( \cos(x) ) ( \cos(x) ) ( x^2+x^4 ) ( x^2+x^4 ) ( \cos(x) ) ( \cos(x) )</td>
</tr>
<tr>
<td>dist.</td>
<td>( N(0,1) ) ( N(0,1) ) ( N(0,1) ) ( N(0,1) ) ( t_3 ) ( t_3 ) ( t_3 ) ( t_3 )</td>
</tr>
<tr>
<td>$\gamma x$</td>
<td>0.5</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------</td>
</tr>
<tr>
<td>$h$</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>mean</td>
</tr>
<tr>
<td></td>
<td>bias</td>
</tr>
<tr>
<td></td>
<td>var</td>
</tr>
<tr>
<td></td>
<td>mse</td>
</tr>
<tr>
<td></td>
<td>NA4</td>
</tr>
<tr>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>0.5</td>
<td>mean</td>
</tr>
<tr>
<td></td>
<td>bias</td>
</tr>
<tr>
<td></td>
<td>var</td>
</tr>
<tr>
<td></td>
<td>mse</td>
</tr>
<tr>
<td></td>
<td>NA4</td>
</tr>
<tr>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>1.0</td>
<td>mean</td>
</tr>
<tr>
<td></td>
<td>bias</td>
</tr>
<tr>
<td></td>
<td>var</td>
</tr>
<tr>
<td></td>
<td>mse</td>
</tr>
<tr>
<td></td>
<td>NA4</td>
</tr>
<tr>
<td></td>
<td>NA</td>
</tr>
</tbody>
</table>
Table 3: $\gamma_x = 0.5$, $N(0,1)$, $u(x) = x^2 + x^4$

<table>
<thead>
<tr>
<th>$\gamma_x$</th>
<th>0.0</th>
<th>0.5</th>
<th>1.0</th>
<th>2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.2</td>
</tr>
<tr>
<td>mean</td>
<td>0.022</td>
<td>0.028</td>
<td>0.041</td>
<td>0.014</td>
</tr>
<tr>
<td>bias</td>
<td>0.022</td>
<td>0.028</td>
<td>0.041</td>
<td>0.014</td>
</tr>
<tr>
<td>var</td>
<td>0.295</td>
<td>0.274</td>
<td>0.266</td>
<td>0.106</td>
</tr>
<tr>
<td>mse</td>
<td>0.295</td>
<td>0.275</td>
<td>0.267</td>
<td>0.106</td>
</tr>
<tr>
<td>NA4</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NA</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>mean</td>
<td>0.341</td>
<td>0.364</td>
<td>0.395</td>
<td>0.332</td>
</tr>
<tr>
<td>bias</td>
<td>0.029</td>
<td>0.052</td>
<td>0.083</td>
<td>0.020</td>
</tr>
<tr>
<td>var</td>
<td>0.298</td>
<td>0.284</td>
<td>0.279</td>
<td>0.131</td>
</tr>
<tr>
<td>mse</td>
<td>0.299</td>
<td>0.286</td>
<td>0.285</td>
<td>0.132</td>
</tr>
<tr>
<td>NA4</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>NA</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>mean</td>
<td>2.058</td>
<td>2.114</td>
<td>2.199</td>
<td>2.041</td>
</tr>
<tr>
<td>bias</td>
<td>0.058</td>
<td>0.114</td>
<td>0.199</td>
<td>0.041</td>
</tr>
<tr>
<td>var</td>
<td>0.415</td>
<td>0.367</td>
<td>0.356</td>
<td>0.245</td>
</tr>
<tr>
<td>mse</td>
<td>0.418</td>
<td>0.380</td>
<td>0.396</td>
<td>0.247</td>
</tr>
<tr>
<td>NA4</td>
<td>94</td>
<td>41</td>
<td>17</td>
<td>87</td>
</tr>
<tr>
<td>NA</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>10</td>
</tr>
</tbody>
</table>
Table 4: $\gamma_x=2.5$, $N(0,1)$, $u(x)=\cos(x)$

<table>
<thead>
<tr>
<th>$\gamma_x$</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>1.005</td>
<td>0.999</td>
<td>0.991</td>
</tr>
<tr>
<td>bias</td>
<td>0.005</td>
<td>-0.001</td>
<td>-0.009</td>
</tr>
<tr>
<td>var</td>
<td>0.272</td>
<td>0.258</td>
<td>0.248</td>
</tr>
<tr>
<td>mse</td>
<td>0.272</td>
<td>0.258</td>
<td>0.248</td>
</tr>
<tr>
<td>NA4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.889</td>
<td>0.883</td>
<td>0.876</td>
</tr>
<tr>
<td>bias</td>
<td>0.011</td>
<td>0.005</td>
<td>-0.002</td>
</tr>
<tr>
<td>var</td>
<td>0.283</td>
<td>0.263</td>
<td>0.253</td>
</tr>
<tr>
<td>mse</td>
<td>0.283</td>
<td>0.263</td>
<td>0.253</td>
</tr>
<tr>
<td>NA4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.533</td>
<td>0.541</td>
<td>0.544</td>
</tr>
<tr>
<td>bias</td>
<td>-0.007</td>
<td>0.001</td>
<td>0.004</td>
</tr>
<tr>
<td>var</td>
<td>0.320</td>
<td>0.281</td>
<td>0.271</td>
</tr>
<tr>
<td>mse</td>
<td>0.320</td>
<td>0.281</td>
<td>0.271</td>
</tr>
<tr>
<td>NA4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\gamma_x$</td>
<td>0.5</td>
<td>1.5</td>
<td>2.5</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>$h$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>1.011</td>
<td>1.001</td>
<td>0.991</td>
</tr>
<tr>
<td>bias</td>
<td>0.011</td>
<td>0.001</td>
<td>-0.009</td>
</tr>
<tr>
<td>var</td>
<td>0.295</td>
<td>0.275</td>
<td>0.265</td>
</tr>
<tr>
<td>mse</td>
<td>0.295</td>
<td>0.275</td>
<td>0.265</td>
</tr>
<tr>
<td>NA4</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NA</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>mean</td>
<td>0.884</td>
<td>0.878</td>
<td>0.872</td>
</tr>
<tr>
<td>bias</td>
<td>0.006</td>
<td>0.000</td>
<td>-0.006</td>
</tr>
<tr>
<td>var</td>
<td>0.297</td>
<td>0.282</td>
<td>0.279</td>
</tr>
<tr>
<td>mse</td>
<td>0.297</td>
<td>0.282</td>
<td>0.279</td>
</tr>
<tr>
<td>NA4</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NA</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>mean</td>
<td>0.551</td>
<td>0.545</td>
<td>0.538</td>
</tr>
<tr>
<td>bias</td>
<td>0.011</td>
<td>0.005</td>
<td>-0.002</td>
</tr>
<tr>
<td>var</td>
<td>0.410</td>
<td>0.358</td>
<td>0.333</td>
</tr>
<tr>
<td>mse</td>
<td>0.410</td>
<td>0.358</td>
<td>0.333</td>
</tr>
<tr>
<td>NA4</td>
<td>94</td>
<td>41</td>
<td>17</td>
</tr>
<tr>
<td>NA</td>
<td>8</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 6: $\gamma_x=2.5$, $t_5$, $u(x)=x^2+x^4$

<table>
<thead>
<tr>
<th>$\gamma_x$</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>-0.002</td>
<td>0.005</td>
<td>0.019</td>
<td>0.000</td>
<td>0.012</td>
<td>0.026</td>
<td>0.004</td>
<td>0.016</td>
<td>0.029</td>
</tr>
<tr>
<td>bias</td>
<td>-0.002</td>
<td>0.005</td>
<td>0.019</td>
<td>0.000</td>
<td>0.012</td>
<td>0.026</td>
<td>0.004</td>
<td>0.016</td>
<td>0.029</td>
</tr>
<tr>
<td>var</td>
<td>0.224</td>
<td>0.213</td>
<td>0.208</td>
<td>0.049</td>
<td>0.041</td>
<td>0.037</td>
<td>0.027</td>
<td>0.020</td>
<td>0.017</td>
</tr>
<tr>
<td>mse</td>
<td>0.224</td>
<td>0.213</td>
<td>0.209</td>
<td>0.049</td>
<td>0.041</td>
<td>0.038</td>
<td>0.027</td>
<td>0.020</td>
<td>0.018</td>
</tr>
<tr>
<td>NA4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\gamma_x$</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.321</td>
<td>0.345</td>
<td>0.377</td>
<td>0.334</td>
<td>0.357</td>
<td>0.387</td>
<td>0.334</td>
<td>0.358</td>
<td>0.388</td>
</tr>
<tr>
<td>bias</td>
<td>0.009</td>
<td>0.033</td>
<td>0.065</td>
<td>0.022</td>
<td>0.045</td>
<td>0.075</td>
<td>0.022</td>
<td>0.046</td>
<td>0.076</td>
</tr>
<tr>
<td>var</td>
<td>0.241</td>
<td>0.230</td>
<td>0.223</td>
<td>0.060</td>
<td>0.048</td>
<td>0.042</td>
<td>0.038</td>
<td>0.027</td>
<td>0.023</td>
</tr>
<tr>
<td>mse</td>
<td>0.242</td>
<td>0.231</td>
<td>0.228</td>
<td>0.060</td>
<td>0.050</td>
<td>0.048</td>
<td>0.038</td>
<td>0.029</td>
<td>0.028</td>
</tr>
<tr>
<td>NA4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\gamma_x$</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>2.033</td>
<td>2.098</td>
<td>2.179</td>
<td>2.057</td>
<td>2.112</td>
<td>2.189</td>
<td>2.042</td>
<td>2.105</td>
<td>2.186</td>
</tr>
<tr>
<td>bias</td>
<td>0.033</td>
<td>0.098</td>
<td>0.179</td>
<td>0.057</td>
<td>0.112</td>
<td>0.189</td>
<td>0.042</td>
<td>0.105</td>
<td>0.186</td>
</tr>
<tr>
<td>var</td>
<td>0.313</td>
<td>0.270</td>
<td>0.261</td>
<td>0.139</td>
<td>0.091</td>
<td>0.081</td>
<td>0.117</td>
<td>0.070</td>
<td>0.057</td>
</tr>
<tr>
<td>mse</td>
<td>0.314</td>
<td>0.279</td>
<td>0.293</td>
<td>0.143</td>
<td>0.104</td>
<td>0.116</td>
<td>0.118</td>
<td>0.081</td>
<td>0.092</td>
</tr>
<tr>
<td>NA4</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 7: $\gamma = 0.5$, $u(x) = x^2 + x^4$

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$ 0.2</td>
<td>0.004</td>
<td>0.016</td>
<td>0.016</td>
</tr>
<tr>
<td>bias</td>
<td>0.006</td>
<td>0.023</td>
<td>0.023</td>
</tr>
<tr>
<td>var</td>
<td>0.017</td>
<td>0.038</td>
<td>0.038</td>
</tr>
<tr>
<td>mse</td>
<td>0.245</td>
<td>0.030</td>
<td>0.030</td>
</tr>
<tr>
<td>NA4</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>NA</td>
<td>0.324</td>
<td>0.325</td>
<td>0.326</td>
</tr>
<tr>
<td>bias</td>
<td>0.037</td>
<td>0.067</td>
<td>0.067</td>
</tr>
<tr>
<td>var</td>
<td>0.069</td>
<td>0.086</td>
<td>0.086</td>
</tr>
<tr>
<td>mse</td>
<td>0.263</td>
<td>0.289</td>
<td>0.289</td>
</tr>
<tr>
<td>NA4</td>
<td>16</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>NA</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

| $\gamma$ | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| $h$ 0.2  | 2.029 | 2.040 | 2.040 | 2.040 | 2.040 | 2.040 |
| bias | 0.090 | 0.092 | 0.092 | 0.092 | 0.092 | 0.092 |
| var | 0.212 | 0.240 | 0.240 | 0.240 | 0.240 | 0.240 |
| mse | 0.423 | 0.430 | 0.430 | 0.430 | 0.430 | 0.430 |
| NA4 | 193 | 185 | 185 | 185 | 185 | 185 |
| NA | 27 | 12 | 12 | 12 | 12 | 12 |
Table 8: $\gamma_x=2.5$, $t_3$, $u(x)=\cos(x)$

<table>
<thead>
<tr>
<th>$\gamma_x$</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>mean</td>
<td>0.986</td>
<td>0.979</td>
</tr>
<tr>
<td></td>
<td>bias</td>
<td>-0.014</td>
<td>-0.021</td>
</tr>
<tr>
<td></td>
<td>var</td>
<td>0.224</td>
<td>0.213</td>
</tr>
<tr>
<td></td>
<td>mse</td>
<td>0.224</td>
<td>0.214</td>
</tr>
<tr>
<td></td>
<td>NA4</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.5</td>
<td>mean</td>
<td>0.863</td>
<td>0.860</td>
</tr>
<tr>
<td></td>
<td>bias</td>
<td>-0.015</td>
<td>-0.018</td>
</tr>
<tr>
<td></td>
<td>var</td>
<td>0.241</td>
<td>0.229</td>
</tr>
<tr>
<td></td>
<td>mse</td>
<td>0.241</td>
<td>0.229</td>
</tr>
<tr>
<td></td>
<td>NA4</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1.0</td>
<td>mean</td>
<td>0.522</td>
<td>0.520</td>
</tr>
<tr>
<td></td>
<td>bias</td>
<td>-0.018</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td>var</td>
<td>0.313</td>
<td>0.269</td>
</tr>
<tr>
<td></td>
<td>mse</td>
<td>0.313</td>
<td>0.270</td>
</tr>
<tr>
<td></td>
<td>NA4</td>
<td>2.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Table 9: $\gamma_x=0.5$, $t_3$, $u(x) = \cos(x)$

<table>
<thead>
<tr>
<th>$\gamma_x$</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>mean</td>
<td>0.985</td>
<td>0.980</td>
</tr>
<tr>
<td></td>
<td>bias</td>
<td>-0.015</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td>var</td>
<td>0.245</td>
<td>0.231</td>
</tr>
<tr>
<td></td>
<td>mse</td>
<td>0.245</td>
<td>0.231</td>
</tr>
<tr>
<td>NA4</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NA</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0.5</td>
<td>mean</td>
<td>0.867</td>
<td>0.865</td>
</tr>
<tr>
<td></td>
<td>bias</td>
<td>-0.011</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>var</td>
<td>0.295</td>
<td>0.269</td>
</tr>
<tr>
<td></td>
<td>mse</td>
<td>0.295</td>
<td>0.269</td>
</tr>
<tr>
<td>NA4</td>
<td>16</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>NA</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1.0</td>
<td>mean</td>
<td>0.536</td>
<td>0.531</td>
</tr>
<tr>
<td></td>
<td>bias</td>
<td>-0.004</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td>var</td>
<td>0.389</td>
<td>0.379</td>
</tr>
<tr>
<td></td>
<td>mse</td>
<td>0.389</td>
<td>0.379</td>
</tr>
<tr>
<td>NA4</td>
<td>193</td>
<td>89</td>
<td>53</td>
</tr>
<tr>
<td>NA</td>
<td>27</td>
<td>16</td>
<td>12</td>
</tr>
</tbody>
</table>