



**HAL**  
open science

## Average Derivative Estimation from Biased Data

Christophe Chesneau, Maher Kachour, Fabien Navarro

► **To cite this version:**

Christophe Chesneau, Maher Kachour, Fabien Navarro. Average Derivative Estimation from Biased Data. ISRN Probability and Statistics, 2014, 2014, pp.864530. 10.1155/2014/864530 . hal-01532972

**HAL Id: hal-01532972**

**<https://hal.science/hal-01532972>**

Submitted on 4 Jun 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Research Article

# Average Derivative Estimation from Biased Data

Christophe Chesneau,<sup>1</sup> Maher Kachour,<sup>2</sup> and Fabien Navarro<sup>3</sup>

<sup>1</sup>Département de Mathématiques, UFR de Sciences, LMNO, Université de Caen Basse-Normandie, 14032 Caen Cedex, France

<sup>2</sup>École Supérieure de Commerce IDRAC, 47 rue Sergent Michel Berthet CP 607, 69258 Lyon Cedex 09, France

<sup>3</sup>Laboratoire de Mathématiques Jean Leray, UFR Sciences et Techniques, Université de Nantes, 2 rue de la Houssinière, BP 92208, F-44322 Nantes Cedex 3, France

Correspondence should be addressed to Christophe Chesneau; christophe.chesneau@gmail.com

Received 18 January 2014; Accepted 15 February 2014; Published 11 March 2014

Academic Editors: N. W. Hengartner, O. Pons, C. A. Tudor, and Y. Wu

Copyright © 2014 Christophe Chesneau et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We investigate the estimation of the density-weighted average derivative from biased data. An estimator integrating a plug-in approach and wavelet projections is constructed. We prove that it attains the parametric rate of convergence  $1/n$  under the mean squared error.

## 1. Introduction

The standard density-weighted average derivative estimation problem is the following. We observe  $n$  *i.i.d.* bivariate random variables  $(X_1, Y_1), \dots, (X_n, Y_n)$  defined on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . Let  $p$  be the unknown density function of  $X_1$ , and let  $\varphi$  be the unknown regression function given by

$$\varphi(x) = \mathbb{E}(Y_1 | X_1 = x), \quad x \in \mathbb{R}. \quad (1)$$

The density-weighted average derivative is defined by

$$\gamma = \mathbb{E}(p(X_1)\varphi'(X_1)) = \int p^2(x)\varphi'(x)dx. \quad (2)$$

The estimation of  $\gamma$  is of interest in some econometric problems, especially in the context of estimation of coefficients in index models (see, e.g., Stoker [1, 2], Powell et al. [3], and Härdle and Stoker [4]). Among the popular approaches, there are the nonparametric techniques based on kernel estimators (see, e.g., Härdle and Stoker [4], Powell et al. [3], Härdle et al. [5], and Stoker [6]) or orthogonal series methods introduced in Rao [7]. Recently, Chesneau et al. [8] have developed an estimator based on a new plug-in approach and a wavelet series method. We refer to Antoniadis [9], Härdle et al. [10], and Vidakovic [11] for further details about wavelets and their applications in nonparametric statistics.

In this paper, we extend this estimation problem to the biased data. It is based on the “biased regression model” which is described as follows. We observe  $n$  *i.i.d.* bivariate random variables  $(X_1, Y_1), \dots, (X_n, Y_n)$  defined on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  with the common density function:

$$f(x, y) = \frac{w(x, y)g(x, y)}{\mu}, \quad (x, y) \in \mathbb{R}^2, \quad (3)$$

where  $w$  is a known positive function,  $g$  is the density function of an unobserved bivariate random variable  $(U, V)$ , and  $\mu = \mathbb{E}(w(U, V)) < \infty$  (which is an unknown real number). This model has potential applications in biology, economics, and many other fields. Important results on methods and applications can be found in, for example, Ahmad [12], Sköld [13], Cristóbal and Alcalá [14], Wu [15], Cristóbal and Alcalá [16], Cristóbal et al. [17], Ojeda et al. [18], Cabrera and Van Keilegom [19], and Chaubey et al. [20]. Wavelet methods related to this model can be found in Chesneau and Shirazi [21], Chaubey et al. [22], and Chaubey and Shirazi [23].

Let  $h$  be the density of  $U$ ; that is,  $h(x) = \int g(x, y)dy$ ,  $x \in \mathbb{R}$ , and let  $\varphi$  be the unknown regression function given by

$$\varphi(x) = \mathbb{E}(V | U = x), \quad x \in \mathbb{R}. \quad (4)$$

The density-weighted average derivative is defined by

$$\delta = \mathbb{E} \left( h(U) \varphi'(U) \right) = \int h^2(x) \varphi'(x) dx. \quad (5)$$

We aim to estimate  $\delta$  from  $(X_1, Y_1), \dots, (X_n, Y_n)$ . To reach this goal, we adapt the methodology of Chesneau et al. [8] to this problem and develop new technical arguments derived to those developed in Chesneau and Shirazi [21] for the estimation of (4). A new wavelet estimator is thus constructed. We prove that it attains the parametric rate of convergence  $1/n$  under the mean squared error, showing the consistency of our estimator.

The paper is organized as follows. In Section 2, we introduce our wavelet methodology including our estimator. Additional assumptions on the model and our main theoretical result are set in Section 3. Finally, the proofs are postponed to Section 4.

## 2. Wavelet Methodology

In this section, after a brief description of the considered wavelet basis, we present our wavelet estimator of  $\delta$  (5).

*2.1. Compactly Supported Wavelet Basis.* Let us consider the following set of functions:

$$\mathbb{L}^2([0, 1]) = \left\{ u : [0, 1] \rightarrow \mathbb{R}; \|u\|_2^2 = \int_0^1 (u(x))^2 dx \right\}. \quad (6)$$

For the purposes of this paper, we use the compactly supported wavelet bases on  $[0, 1]$  briefly described below.

Let  $N \geq 5$  be a fixed integer, and let  $\phi$  and  $\psi$  be the initial wavelet functions of the Daubechies wavelets  $db2N$  (see, e.g., Mallat [24]). These functions have the features to be compactly supported and derivable.

Set

$$\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k), \quad \psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k), \quad (7)$$

and  $\Lambda_j = \{0, \dots, 2^j - 1\}$ . Then, with a specific treatments at the boundaries, there exists an integer  $\tau$  such that the collection

$$\mathcal{B} = \left\{ \phi_{\tau,k}, k \in \Lambda_\tau; \psi_{j,k}; j \in \mathbb{N} - \{0, \dots, \tau - 1\}, k \in \Lambda_j \right\} \quad (8)$$

is an orthonormal basis of  $\mathbb{L}^2([0, 1])$ .

Hence, a function  $u \in \mathbb{L}^2([0, 1])$  can be expanded on  $\mathcal{B}$  as

$$u(x) = \sum_{k \in \Lambda_\tau} \alpha_{\tau,k} \phi_{\tau,k}(x) + \sum_{j=\tau}^{\infty} \sum_{k \in \Lambda_j} \beta_{j,k} \psi_{j,k}(x), \quad (9)$$

where

$$\alpha_{\tau,k} = \int_0^1 u(x) \phi_{\tau,k}(x) dx, \quad \beta_{j,k} = \int_0^1 u(x) \psi_{j,k}(x) dx. \quad (10)$$

Further details about wavelet basis can be found in, for example, Meyer [25], Cohen et al. [26], and Mallat [24].

*2.2. Wavelet Estimator.* Proposition 1 provides another expression of the density-weighted average derivative (5) in terms of wavelet coefficients.

**Proposition 1.** *Let  $\delta$  be given by (5). Suppose that  $\text{supp}(X_1) = [0, 1]$ ,  $\phi h \in \mathbb{L}^2([0, 1])$ ,  $h' \in \mathbb{L}^2([0, 1])$ , and  $h(0) = h(1) = 0$ . Then the density-weighted average derivative (5) can be expressed as*

$$\delta = -2 \left( \sum_{k \in \Lambda_\tau} \alpha_{\tau,k} c_{\tau,k} + \sum_{j=\tau}^{\infty} \sum_{k \in \Lambda_j} \beta_{j,k} d_{j,k} \right), \quad (11)$$

where

$$\alpha_{\tau,k} = \int_0^1 \varphi(x) h(x) \phi_{\tau,k}(x) dx, \quad (12)$$

$$c_{\tau,k} = \int_0^1 h'(x) \phi_{\tau,k}(x) dx,$$

$$\beta_{j,k} = \int_0^1 \varphi(x) h(x) \psi_{j,k}(x) dx, \quad (13)$$

$$d_{j,k} = \int_0^1 h'(x) \psi_{j,k}(x) dx.$$

In view of Proposition 1, adapting the ideas of Chesneau et al. [8] and Chesneau and Shirazi [21] to our framework, we consider the following plug-in estimator for  $\delta$ :

$$\hat{\delta} = -2 \left( \sum_{k \in \Lambda_\tau} \hat{\alpha}_{\tau,k} \hat{c}_{\tau,k} + \sum_{j=\tau}^{j_0} \sum_{k \in \Lambda_j} \hat{\beta}_{j,k} \hat{d}_{j,k} \right), \quad (14)$$

where

$$\hat{\alpha}_{\tau,k} = \frac{\hat{\mu}}{n} \sum_{i=1}^n \frac{Y_i}{w(X_i, Y_i)} \phi_{\tau,k}(X_i), \quad (15)$$

$$\hat{c}_{\tau,k} = -\frac{\hat{\mu}}{n} \sum_{i=1}^n \frac{1}{w(X_i, Y_i)} (\phi_{\tau,k})'(X_i),$$

$$\hat{\beta}_{j,k} = \frac{\hat{\mu}}{n} \sum_{i=1}^n \frac{Y_i}{w(X_i, Y_i)} \psi_{j,k}(X_i), \quad (16)$$

$$\hat{d}_{j,k} = -\frac{\hat{\mu}}{n} \sum_{i=1}^n \frac{1}{w(X_i, Y_i)} (\psi_{j,k})'(X_i),$$

$$\hat{\mu} = \left( \frac{1}{n} \sum_{i=1}^n \frac{1}{w(X_i, Y_i)} \right)^{-1}, \quad (17)$$

and  $j_0$  is an integer which will be chosen a posteriori.

Proposition 2 provides some theoretical results explaining the constructions of the above estimators.

**Proposition 2.** *Suppose that  $\text{supp}(X_1) = [0, 1]$ . Then*

(i)  $1/\hat{\mu}$  (17) is an unbiased estimator for  $1/\mu$ ;

(ii)  $(\mu/\hat{\mu})\hat{\alpha}_{\tau,k}$  (15) and  $(\mu/\hat{\mu})\hat{\beta}_{j,k}$  (16) are unbiased estimators for  $\alpha_{\tau,k}$  (12) and  $\beta_{j,k}$  (13), respectively;

- (iii) under  $h(0) = h(1) = 0$ ,  $(\mu/\widehat{\mu})\widehat{c}_{\tau,k}$  (15) and  $(\mu/\widehat{\mu})\widehat{d}_{j,k}$  (16) are unbiased estimators for  $c_{\tau,k}$  (12) and  $d_{j,k}$  (13), respectively.

### 3. Assumptions and Result

After the presentation of some additional assumptions on the model, we describe our main result on the asymptotic properties of  $\widehat{\delta}$  (14).

3.1. *Assumptions.* We formulate the following assumptions on  $\varphi$ ,  $h$ , and  $w$ .

- (H1) The support of  $X_1$ , denoted by  $\text{supp}(X_1)$ , is compact. In order to fix the notations, we suppose that  $\text{supp}(X_1) = [0, 1]$ .
- (H2) There exists a constant  $C > 0$  such that

$$\sup_{x \in [0,1]} |\varphi(x)| \leq C. \quad (18)$$

- (H3) The function  $h$  satisfies  $h(0) = h(1) = 0$  and there exists a constant  $C > 0$  such that

$$\sup_{x \in [0,1]} h(x) \leq C, \quad \sup_{x \in [0,1]} |h'(x)| \leq C. \quad (19)$$

- (H4) There exists a constant  $C > 0$  such that

$$\sup_{x \in [0,1]} \int_{-\infty}^{\infty} y^4 g(x, y) dy \leq C, \quad (20)$$

$$\sup_{x \in [0,1]} \int_{-\infty}^{\infty} g(x, y) dy \leq C.$$

- (H5) There exist two constants  $C > 0$  and  $c > 0$  such that

$$\inf_{(x,y) \in [0,1] \times \mathbb{R}} w(x, y) \geq c, \quad \sup_{(x,y) \in [0,1] \times \mathbb{R}} w(x, y) \leq C. \quad (21)$$

Let  $s_1 > 0$ ,  $s_2 > 0$ , and  $\beta_{j,k}$  and  $d_{j,k}$  be given by (13). We formulate the following assumptions on  $\beta_{j,k}$  and  $d_{j,k}$ .

- (H6 ( $s_1$ )) There exists a constant  $C > 0$  such that

$$|\beta_{j,k}| \leq C 2^{-j(s_1+1/2)}. \quad (22)$$

- (H7 ( $s_2$ )) There exists a constant  $C > 0$  such that

$$|d_{j,k}| \leq C 2^{-j(s_2+1/2)}. \quad (23)$$

Note that (H2)–(H5) are boundedness assumptions, whereas (H6 ( $s_1$ )) and (H7 ( $s_2$ )) are related to the smoothness of  $\varphi h$  and  $h'$  represented by  $s_1$  and  $s_2$ . There exist deep connections between (H6 ( $s_1$ )) and (H7 ( $s_2$ )) and balls of Hölder spaces (see [10, Chapter 8]).

3.2. *Main Result.* The following theorem establishes the upper bound of the MSE of our estimator.

**Theorem 3.** Assume that (H1)–(H5), (H6 ( $s_1$ )) with  $s_1 > 3/2$ , and (H7 ( $s_2$ )) with  $s_2 > 1/2$  hold. Let  $\delta$  be given by (5), and let  $\widehat{\delta}$  be given by (14) with  $j_0$  such that  $n^{1/4} < 2^{j_0+1} \leq 2n^{1/4}$ . Then there exists a constant  $C > 0$  such that

$$\mathbb{E} \left( (\widehat{\delta} - \delta)^2 \right) \leq C \frac{1}{n}. \quad (24)$$

Theorem 3 proves that  $\widehat{\delta}$  attains the parametric rate of convergence  $1/n$  under the MSE. This implies the consistency of our estimator. This result provides a first theoretical aspect to the estimation of  $\delta$  from biased data.

### 4. Proofs

#### 4.1. On the Construction of $\widehat{\delta}$

*Proof of Proposition 1.* We follow the approach of Chesneau et al. [8]. Using  $\text{supp}(X_1) = [0, 1]$ ,  $h(0) = h(1) = 0$ , and an integration by parts, we obtain

$$\begin{aligned} \delta &= \left[ h^2(x) \varphi(x) \right]_0^1 - 2 \int_0^1 \varphi(x) h(x) h'(x) dx \\ &= -2 \int_0^1 \varphi(x) h(x) h'(x) dx. \end{aligned} \quad (25)$$

Moreover,

- (i) since  $\varphi h \in \mathbb{L}^2([0, 1])$ , we can expand it on  $\mathcal{B}$  as (9):

$$\varphi(x) h(x) = \sum_{k \in \Lambda_\tau} \alpha_{\tau,k} \phi_{\tau,k}(x) + \sum_{j=\tau}^{\infty} \sum_{k \in \Lambda_j} \beta_{j,k} \psi_{j,k}(x), \quad (26)$$

where  $\alpha_{\tau,k}$  and  $\beta_{j,k}$  are (12),

- (ii) since  $h' \in \mathbb{L}^2([0, 1])$ , we can expand it on  $\mathcal{B}$  as (9):

$$h'(x) = \sum_{k \in \Lambda_\tau} c_{\tau,k} \phi_{\tau,k}(x) + \sum_{j=\tau}^{\infty} \sum_{k \in \Lambda_j} d_{j,k} \psi_{j,k}(x), \quad (27)$$

where  $c_{\tau,k}$  and  $d_{j,k}$  are (13).

Thanks to (25) and the orthonormality of  $\mathcal{B}$  on  $\mathbb{L}^2([0, 1])$ , we get

$$\delta = -2 \left( \sum_{k \in \Lambda_\tau} \alpha_{\tau,k} c_{\tau,k} + \sum_{j=\tau}^{\infty} \sum_{k \in \Lambda_j} \beta_{j,k} d_{j,k} \right). \quad (28)$$

Proposition 1 is proved.  $\square$

*Proof of Proposition 2.* (i) We have

$$\begin{aligned} \mathbb{E}\left(\frac{1}{\widehat{\mu}}\right) &= \mathbb{E}\left(\frac{1}{n\sum_{i=1}^n \frac{1}{w(X_i, Y_i)}}\right) = \mathbb{E}\left(\frac{1}{w(X_1, Y_1)}\right) \\ &= \int_{-\infty}^{\infty} \int_0^1 \frac{1}{w(x, y)} f(x, y) dx dy \\ &= \frac{1}{\mu} \int_{-\infty}^{\infty} \int_0^1 g(x, y) dx dy = \frac{1}{\mu}. \end{aligned} \quad (29)$$

(ii) Using the identical distribution of  $(X_1, Y_1), \dots, (X_n, Y_n)$  and the definition of  $\varphi$  (4), we obtain

$$\begin{aligned} \mathbb{E}\left(\frac{\mu}{\widehat{\mu}} \widehat{\beta}_{j,k}\right) &= \mathbb{E}\left(\frac{\mu}{n\sum_{i=1}^n \frac{Y_i}{w(X_i, Y_i)}} \psi_{j,k}(X_i)\right) \\ &= \mu \mathbb{E}\left(\frac{Y_1}{w(X_1, Y_1)} \psi_{j,k}(X_1)\right) \\ &= \mu \int_{-\infty}^{\infty} \int_0^1 \frac{y}{w(x, y)} \psi_{j,k}(x) f(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_0^1 yg(x, y) \psi_{j,k}(x) dx dy \\ &= \int_0^1 \left(\int_{-\infty}^{\infty} yg(x, y) dy\right) \psi_{j,k}(x) dx \\ &= \int_0^1 \varphi(x) h(x) \psi_{j,k}(x) dx = \beta_{j,k}. \end{aligned} \quad (30)$$

Similarly, we prove that  $\mathbb{E}((\mu/\widehat{\mu})\widehat{\alpha}_{\tau,k}) = \alpha_{\tau,k}$ .

(iii) Using the identical distribution of  $X_1, \dots, X_n$ , an integration by parts, and  $h(0) = h(1) = 0$ , we obtain

$$\begin{aligned} \mathbb{E}\left(\frac{\mu}{\widehat{\mu}} \widehat{d}_{j,k}\right) &= -\mathbb{E}\left(\frac{\mu}{n\sum_{i=1}^n \frac{1}{w(X_i, Y_i)}} (\psi_{j,k})'(X_i)\right) \\ &= -\mu \mathbb{E}\left(\frac{1}{w(X_1, Y_1)} (\psi_{j,k})'(X_1)\right) \\ &= -\mu \int_{-\infty}^{\infty} \int_0^1 \frac{1}{w(x, y)} (\psi_{j,k})'(x) f(x, y) dx dy \\ &= -\int_{-\infty}^{\infty} \int_0^1 g(x, y) (\psi_{j,k})'(x) dx dy \\ &= -\int_0^1 \left(\int_{-\infty}^{\infty} g(x, y) dy\right) (\psi_{j,k})'(x) dx \\ &= -\int_0^1 h(x) (\psi_{j,k})'(x) dx \\ &= -\left([h(x) \psi_{j,k}(x)]_0^1 - \int_0^1 h'(x) \psi_{j,k}(x) dx\right) \\ &= \int_0^1 h'(x) \psi_{j,k}(x) dx = d_{j,k}. \end{aligned} \quad (31)$$

Similarly, we prove that  $\mathbb{E}((\mu/\widehat{\mu})\widehat{c}_{\tau,k}) = c_{\tau,k}$ . This ends the proof of Proposition 2.  $\square$

#### 4.2. Proof of Some Intermediate Results

**Proposition 4.** Suppose that (H1)–(H5) hold. Let  $\beta_{j,k}$  and  $d_{j,k}$  be given by (13), and let  $\widehat{\beta}_{j,k}$  and  $\widehat{d}_{j,k}$  be given by (16) with  $j$  such that  $2^j \leq n$ . Then

(i) there exists a constant  $C > 0$  such that

$$\mathbb{E}\left(\left(\frac{1}{\widehat{\mu}} - \frac{1}{\mu}\right)^4\right) \leq C \frac{1}{n^2}, \quad (32)$$

(ii) there exists a constant  $C > 0$  such that

$$\mathbb{E}\left(\left(\widehat{\beta}_{j,k} - \beta_{j,k}\right)^4\right) \leq C \frac{1}{n^2}, \quad (33)$$

(iii) there exists a constant  $C > 0$  such that

$$\mathbb{E}\left(\left(\widehat{d}_{j,k} - d_{j,k}\right)^4\right) \leq C \frac{2^{4j}}{n^2}. \quad (34)$$

These inequalities hold with  $(\widehat{\alpha}_{\tau,k}, \widehat{c}_{\tau,k})$  in (15) instead of  $(\widehat{\beta}_{j,k}, \widehat{d}_{j,k})$  and  $(\alpha_{\tau,k}, c_{\tau,k})$  in (12) instead of  $(\beta_{j,k}, d_{j,k})$  for  $j = \tau$ .

*Proof of Proposition 4.* We use the following version of the Rosenthal inequality. The proof can be found in Rosenthal [27].

**Lemma 5.** Let  $n$  be a positive integer,  $p \geq 2$ , and let  $U_1, \dots, U_n$  be  $n$  zero mean independent random variables such that  $\sup_{i \in \{1, \dots, n\}} \mathbb{E}(|U_i|^p) < \infty$ . Then there exists a constant  $C > 0$  such that

$$\mathbb{E}\left(\left|\sum_{i=1}^n U_i\right|^p\right) \leq C \left(\sum_{i=1}^n \mathbb{E}(|U_i|^p) + \left(\sum_{i=1}^n \mathbb{E}(U_i^2)\right)^{p/2}\right). \quad (35)$$

(i) Note that

$$\mathbb{E}\left(\left(\frac{1}{\widehat{\mu}} - \frac{1}{\mu}\right)^4\right) = \frac{1}{n^4} \mathbb{E}\left(\left(\sum_{i=1}^n U_i\right)^4\right), \quad (36)$$

with

$$U_i = \frac{1}{w(X_i, Y_i)} - \frac{1}{\mu}, \quad i \in \{1, \dots, n\}. \quad (37)$$

Since  $(X_1, Y_1), \dots, (X_n, Y_n)$  are *i.i.d.*, we get that  $U_1, \dots, U_n$  are also *i.i.d.*. Moreover, from Proposition 2, we have  $\mathbb{E}(U_1) = 0$ . Using (H5), for any  $u \in \{2, 4\}$ , we have  $\mathbb{E}(U_1^u) \leq C$ . Thus, Lemma 5 with  $p = 4$  yields

$$\mathbb{E}\left(\left(\frac{1}{\widehat{\mu}} - \frac{1}{\mu}\right)^4\right) \leq C \frac{1}{n^4} (n\mathbb{E}(U_1^4) + n^2(\mathbb{E}(U_1^2))^2) \leq C \frac{1}{n^2}. \quad (38)$$

(ii) We have

$$\begin{aligned} \widehat{\beta}_{j,k} - \beta_{j,k} &= \frac{\widehat{\mu}}{\mu} \left( \frac{\mu}{n} \sum_{i=1}^n \frac{Y_i}{w(X_i, Y_i)} \psi_{j,k}(X_i) - \beta_{j,k} \right) \\ &\quad + \beta_{j,k} \widehat{\mu} \left( \frac{1}{\mu} - \frac{1}{\widehat{\mu}} \right). \end{aligned} \quad (39)$$

By (H5), we have  $|\widehat{\mu}/\mu| \leq C$ ,  $|\widehat{\mu}| \leq C$ , and by (H2) and (H3),  $|\beta_{j,k}| \leq C$ . Hence, by the triangular inequality,

$$\begin{aligned} |\widehat{\beta}_{j,k} - \beta_{j,k}| &\leq C \left( \left| \frac{\mu}{n} \sum_{i=1}^n \frac{Y_i}{w(X_i, Y_i)} \psi_{j,k}(X_i) - \beta_{j,k} \right| + \left| \frac{1}{\widehat{\mu}} - \frac{1}{\mu} \right| \right). \end{aligned} \quad (40)$$

The elementary inequality,  $(x+y)^4 \leq 8(x^4 + y^4)$ ,  $(x, y) \in \mathbb{R}^2$ , implies that

$$\mathbb{E} \left( \left( \widehat{\beta}_{j,k} - \beta_{j,k} \right)^4 \right) \leq C(I_1 + I_2), \quad (41)$$

where

$$\begin{aligned} I_1 &= \frac{1}{n^4} \mathbb{E} \left( \left( \sum_{i=1}^n U_i \right)^4 \right), \\ U_i &= \mu \frac{Y_i}{w(X_i, Y_i)} \psi_{j,k}(X_i) - \beta_{j,k}, \quad i \in \{1, \dots, n\}, \quad (42) \\ I_2 &= \mathbb{E} \left( \left( \frac{1}{\widehat{\mu}} - \frac{1}{\mu} \right)^4 \right). \end{aligned}$$

*Upper bound for  $I_1$ .* Since  $(X_1, Y_1), \dots, (X_n, Y_n)$  are *i.i.d.*, we get that  $U_1, \dots, U_n$  are also *i.i.d.* Moreover, from Proposition 2, we have  $\mathbb{E}(U_1) = 0$ . Now, using the Hölder inequality, (H4), (H5) and  $2^j \leq n$ , for any  $u \in \{2, 4\}$ , observe that

$$\begin{aligned} \mathbb{E}(U_1^u) &\leq C \mathbb{E} \left( \left( \mu \frac{Y_1}{w(X_1, Y_1)} \psi_{j,k}(X_1) \right)^u \right) \\ &\leq C \mu \mathbb{E} \left( \left( Y_1 \psi_{j,k}(X_1) \right)^u \frac{1}{w(X_1, Y_1)} \right) \\ &= C \mu \int_{-\infty}^{\infty} \int_0^1 (y \psi_{j,k}(x))^u \frac{1}{w(x, y)} f(x, y) dx dy \\ &= C \int_{-\infty}^{\infty} \int_0^1 (y \psi_{j,k}(x))^u g(x, y) dx dy \\ &= C \int_0^1 \left( \int_{-\infty}^{\infty} y^u g(x, y) dy \right) (\psi_{j,k}(x))^u dx \\ &\leq C \int_0^1 (\psi_{j,k}(x))^u dx = C 2^{j(u-2)/2} \int_0^1 (\psi(x))^u dx \\ &\leq C n^{(u-2)/2}. \end{aligned} \quad (43)$$

Combining Lemma 5 with  $p = 4$  with the previous inequality, we obtain

$$I_1 \leq C \frac{1}{n^4} \left( n \mathbb{E}(U_1^4) + n^2 (\mathbb{E}(U_1^2))^2 \right) \leq C \frac{1}{n^2}. \quad (44)$$

*Upper bound for  $I_2$ .* The point (i) yields

$$I_2 \leq C \frac{1}{n^2}. \quad (45)$$

It follows from (41), (44), and (45) that

$$\mathbb{E} \left( \left( \widehat{\beta}_{j,k} - \beta_{j,k} \right)^4 \right) \leq C \frac{1}{n^2}, \quad (46)$$

(iii) Similar arguments to the beginning of (ii) give

$$\mathbb{E} \left( \left( \widehat{d}_{j,k} - d_{j,k} \right)^4 \right) \leq C(I_1 + I_2), \quad (47)$$

where

$$\begin{aligned} I_1 &= \frac{1}{n^4} \mathbb{E} \left( \left( \sum_{i=1}^n U_i \right)^4 \right), \\ U_i &= -\mu \frac{1}{w(X_i, Y_i)} (\psi_{j,k})'(X_i) - d_{j,k}, \quad i \in \{1, \dots, n\}, \\ I_2 &= \mathbb{E} \left( \left( \frac{1}{\widehat{\mu}} - \frac{1}{\mu} \right)^4 \right). \end{aligned} \quad (48)$$

*Upper bound for  $I_1$ .* Since  $(X_1, Y_1), \dots, (X_n, Y_n)$  are *i.i.d.*, we get that  $U_1, \dots, U_n$  are also *i.i.d.* Moreover, from Proposition 2, we have  $\mathbb{E}(U_1) = 0$ . Now, using the Hölder inequality, (H4), (H5),  $(\psi_{j,k})'(x) = 2^j 2^{j/2} \psi'(2^j x - k)$ , and  $2^j \leq n$ , for any  $u \in \{2, 4\}$ , observe that

$$\begin{aligned} \mathbb{E}(U_1^u) &\leq C \mathbb{E} \left( \left( \mu \frac{1}{w(X_1, Y_1)} (\psi_{j,k})'(X_1) \right)^u \right) \\ &\leq C \mu \mathbb{E} \left( \left( (\psi_{j,k})'(X_1) \right)^u \frac{1}{w(X_1, Y_1)} \right) \\ &= C \mu \int_{-\infty}^{\infty} \int_0^1 \left( (\psi_{j,k})'(x) \right)^u \frac{1}{w(x, y)} f(x, y) dx dy \\ &= C \int_{-\infty}^{\infty} \int_0^1 \left( (\psi_{j,k})'(x) \right)^u g(x, y) dx dy \\ &= C \int_0^1 \left( \int_{-\infty}^{\infty} g(x, y) dy \right) \left( (\psi_{j,k})'(x) \right)^u dx \\ &\leq C \int_0^1 \left( (\psi_{j,k})'(x) \right)^u dx \\ &= C 2^{ju} 2^{j(u-2)/2} \int_0^1 (\psi'(x))^u dx \\ &\leq C 2^{ju} n^{(u-2)/2}. \end{aligned} \quad (49)$$

Owing to Lemma 5 with  $p = 4$  and the previous inequality, we obtain

$$\begin{aligned} I_1 &\leq C \frac{1}{n^4} \left( n \mathbb{E} (U_1^4) + n^2 (\mathbb{E} (U_1^2))^2 \right) \\ &\leq C \frac{1}{n^4} \left( n 2^{4j} n + n^2 (2^{2j})^2 \right) \leq C \frac{2^{4j}}{n^2}. \end{aligned} \quad (50)$$

*Upper bound for  $I_2$ .* The point (i) yields

$$I_2 \leq C \frac{1}{n^2}. \quad (51)$$

It follows from (47), (50), and (51) that

$$\mathbb{E} \left( (\widehat{d}_{j,k} - d_{j,k})^4 \right) \leq C \frac{2^{4j}}{n^2}. \quad (52)$$

Proposition 4 is proved.  $\square$

Proposition 6 below is a consequence of [8, Proposition 5.2] and the results of Proposition 4 above.

**Proposition 6.** (i) Suppose that (H1)–(H5), (H6 ( $s_1$ )), and (H7 ( $s_2$ )) hold. Let  $\beta_{j,k}$  and  $d_{j,k}$  be given by (13), and let  $\widehat{\beta}_{j,k}$  and  $\widehat{d}_{j,k}$  be given by (16) with  $j$  such that  $2^j \leq n$ . Then there exists a constant  $C > 0$  such that

$$\begin{aligned} &\mathbb{E} \left( (\widehat{\beta}_{j,k} \widehat{d}_{j,k} - \beta_{j,k} d_{j,k})^2 \right) \\ &\leq C \left( \frac{2^{-j(2s_1-1)}}{n} + \frac{2^{-j(2s_2+1)}}{n} + \frac{2^{2j}}{n^2} \right). \end{aligned} \quad (53)$$

(ii) Suppose that (H1)–(H5) hold. Let  $\alpha_{\tau,k}$  and  $c_{\tau,k}$  be given by (12), and let  $\widehat{\alpha}_{\tau,k}$  and  $\widehat{c}_{\tau,k}$  be given by (15). Then there exists a constant  $C > 0$  such that

$$\mathbb{E} \left( (\widehat{\alpha}_{\tau,k} \widehat{c}_{\tau,k} - \alpha_{\tau,k} c_{\tau,k})^2 \right) \leq C \frac{1}{n}. \quad (54)$$

#### 4.3. Proof of the Main Result

*Proof of Theorem 3.* Using the intermediary results above, the proof follows the lines of [8, Theorem 5.1]. It follows from Proposition 1 and the elementary inequality,  $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$ ,  $(a, b, c) \in \mathbb{R}^3$ , that

$$\mathbb{E} \left( (\widehat{\delta} - \delta)^2 \right) \leq 12 (W_1 + W_2 + W_3), \quad (55)$$

where

$$\begin{aligned} W_1 &= \mathbb{E} \left( \left( \sum_{k \in \Lambda_\tau} |\widehat{\alpha}_{\tau,k} \widehat{c}_{\tau,k} - \alpha_{\tau,k} c_{\tau,k}| \right)^2 \right), \\ W_2 &= \mathbb{E} \left( \left( \sum_{j=\tau}^{j_0} \sum_{k \in \Lambda_j} |\widehat{\beta}_{j,k} \widehat{d}_{j,k} - \beta_{j,k} d_{j,k}| \right)^2 \right), \\ W_3 &= \left( \sum_{j=j_0+1}^{\infty} \sum_{k \in \Lambda_j} |\beta_{j,k}| |d_{j,k}| \right)^2. \end{aligned} \quad (56)$$

*Upper bound for  $W_1$ .* Using the Cauchy-Schwarz inequality, the second point of Proposition 6, and  $\text{Card}(\Lambda_\tau) \leq C$ , we obtain

$$W_1 \leq \left( \sum_{k \in \Lambda_\tau} \sqrt{\mathbb{E} \left( (\widehat{\alpha}_{\tau,k} \widehat{c}_{\tau,k} - \alpha_{\tau,k} c_{\tau,k})^2 \right)} \right)^2 \leq C \frac{1}{n}. \quad (57)$$

*Upper bound for  $W_2$ .* It follows from the Cauchy-Schwarz inequality, the first point of Proposition 6,  $\text{Card}(\Lambda_j) \leq C 2^j$ , the elementary inequality,  $\sqrt{a + b + c} \leq \sqrt{a} + \sqrt{b} + \sqrt{c}$ ,  $s_1 > 3/2$ ,  $s_2 > 1/2$ , and  $2^{j_0} \leq n^{1/4}$ , that

$$\begin{aligned} W_2 &\leq \left( \sum_{j=\tau}^{j_0} \sum_{k \in \Lambda_j} \sqrt{\mathbb{E} \left( (\widehat{\beta}_{j,k} \widehat{d}_{j,k} - \beta_{j,k} d_{j,k})^2 \right)} \right)^2 \\ &\leq C \left( \sum_{j=\tau}^{j_0} 2^j \sqrt{\frac{2^{-j(2s_1-1)}}{n} + \frac{2^{-j(2s_2+1)}}{n} + \frac{2^{2j}}{n^2}} \right)^2 \\ &\leq C \left( \frac{1}{\sqrt{n}} \sum_{j=\tau}^{j_0} 2^{-j(s_1-3/2)} + \frac{1}{\sqrt{n}} \sum_{j=\tau}^{j_0} 2^{-j(s_2-1/2)} + \frac{1}{n} \sum_{j=\tau}^{j_0} 2^{2j} \right)^2 \\ &\leq C \left( \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{n}} + \frac{2^{2j_0}}{n} \right)^2 \leq C \frac{1}{n}. \end{aligned} \quad (58)$$

*Upper bound for  $W_3$ .* By (H6 ( $s_1$ )) with  $s_1 > 3/2$ , (H7 ( $s_2$ )) with  $s_2 > 1/2$ , and  $2^{j_0+1} > n^{1/4}$ , we have

$$\begin{aligned} W_3 &\leq C \left( \sum_{j=j_0+1}^{\infty} 2^j 2^{-j(s_1+1/2)} 2^{-j(s_2+1/2)} \right)^2 \\ &\leq C 2^{-2j_0(s_1+s_2)} \leq C 2^{-4j_0} \leq C \frac{1}{n}. \end{aligned} \quad (59)$$

Putting (55), (57), (58), and (59) together, we obtain

$$\mathbb{E} \left( (\widehat{\delta} - \delta)^2 \right) \leq C \frac{1}{n}. \quad (60)$$

This ends the proof of Theorem 3.  $\square$

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] T. M. Stoker, "Consistent estimation of scaled coefficients," *Econometrica*, vol. 54, no. 6, pp. 1461–1481, 1986.
- [2] T. M. Stoker, "Tests of additive derivative constraints," *Review of Economic Studies*, vol. 56, no. 4, pp. 535–552, 1989.
- [3] J. L. Powell, J. H. Stock, and T. M. Stoker, "Semiparametric estimation of index coefficients," *Econometrica*, vol. 57, no. 6, pp. 1403–1430, 1989.

- [4] W. Härdle and T. M. Stoker, "Investigating smooth multiple regression by the method of average derivatives," *Journal of the American Statistical Association*, vol. 84, no. 408, pp. 986–995, 1989.
- [5] W. Härdle, J. Hart, J. S. Marron, and A. B. Tsybakov, "Bandwidth choice for average derivative estimation," *Journal of the American Statistical Association*, vol. 87, no. 417, pp. 218–226, 1992.
- [6] T. M. Stoker, "Equivalence of direct, indirect, and slope estimators of average derivatives," in *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, W. A. Barnett, J. Powell, and G. Tauchen, Eds., pp. 99–118, Cambridge University Press, Cambridge, UK, 1991.
- [7] B. L. S. P. Rao, "Consistent estimation of density-weighted average derivative by orthogonal series method," *Statistics & Probability Letters*, vol. 22, no. 3, pp. 205–212, 1995.
- [8] C. Chesneau, M. Kachour, and F. Navarro, "On the estimation of density-weighted average derivative by wavelet methods under various dependence structures," *Sankhya A*, vol. 76, no. 1, pp. 48–76, 2014.
- [9] A. Antoniadis, "Wavelets in statistics: a review," *Journal of the Italian Statistical Society B*, vol. 6, no. 2, pp. 97–144, 1997.
- [10] W. Härdle, G. Kerkycharian, D. Picard, and A. Tsybakov, *Wavelets, Approximation, and Statistical Applications*, vol. 129 of *Lecture Notes in Statistics*, Springer, New York, NY, USA, 1998.
- [11] B. Vidakovic, *Statistical Modeling by Wavelets*, John Wiley & Sons, New York, NY, USA, 1999.
- [12] I. A. Ahmad, "On multivariate kernel estimation for samples from weighted distributions," *Statistics & Probability Letters*, vol. 22, no. 2, pp. 121–129, 1995.
- [13] M. Sköld, "Kernel regression in the presence of size-bias," *Journal of Nonparametric Statistics*, vol. 12, no. 1, pp. 41–51, 1999.
- [14] J. A. Cristóbal and J. T. Alcalá, "Nonparametric regression estimators for length biased data," *Journal of Statistical Planning and Inference*, vol. 89, no. 1-2, pp. 145–168, 2000.
- [15] C. O. Wu, "Local polynomial regression with selection biased data," *Statistica Sinica*, vol. 10, no. 3, pp. 789–817, 2000.
- [16] J. A. Cristóbal and J. T. Alcalá, "An overview of nonparametric contributions to the problem of functional estimation from biased data," *Test*, vol. 10, no. 2, pp. 309–332, 2001.
- [17] J. A. Cristóbal, J. L. Ojeda, and J. T. Alcalá, "Confidence bands in nonparametric regression with length biased data," *Annals of the Institute of Statistical Mathematics*, vol. 56, no. 3, pp. 475–496, 2004.
- [18] J. L. Ojeda, W. González-Manteiga, and J. A. Cristobal, "A bootstrap based model checking for selection-biased data," Tech. Rep. 07–05, Universidade de Santiago de Compostela, 2007.
- [19] J. L. O. Cabrera and I. van Keilegom, "Goodness-of-fit tests for parametric regression with selection biased data," *Journal of Statistical Planning and Inference*, vol. 139, no. 8, pp. 2836–2850, 2009.
- [20] Y. P. Chaubey, N. Laïb, and J. Li, "Generalized kernel regression estimator for dependent size-biased data," *Journal of Statistical Planning and Inference*, vol. 142, no. 3, pp. 708–727, 2012.
- [21] C. Chesneau and E. Shirazi, "Nonparametric wavelet regression based on biased data," *Communications in Statistics*. In press.
- [22] Y. P. Chaubey, C. Chesneau, and E. Shirazi, "Wavelet-based estimation of regression function for dependent biased data under a given random design," *Journal of Nonparametric Statistics*, vol. 25, no. 1, pp. 53–71, 2013.
- [23] Y. P. Chaubey and E. Shirazi, "On MISE of a nonlinear wavelet estimator of the regression function based on biased data under strong mixing," *Communications in Statistics*. In press.
- [24] S. Mallat, *A Wavelet Tour of Signal Processing*, Elsevier/Academic Press, Amsterdam, The Netherlands, 3rd edition, 2009.
- [25] Y. Meyer, *Wavelets and Operators*, vol. 37, Cambridge University Press, Cambridge, UK, 1992.
- [26] A. Cohen, I. Daubechies, and P. Vial, "Wavelets on the interval and fast wavelet transforms," *Applied and Computational Harmonic Analysis*, vol. 1, no. 1, pp. 54–81, 1993.
- [27] H. P. Rosenthal, "On the subspaces of  $L^p(p > 2)$  spanned by sequences of independent random variables," *Israel Journal of Mathematics*, vol. 8, no. 3, pp. 273–303, 1970.



