Regularization parameter selection in indirect regression by residual based bootstrap

Nicolai Bissantz, Justin Chown and Holger Dette

Ruhr-Universität Bochum

Supplementary Material

S1 Technical Details

Recall the estimator $\hat{R}(k) = (2n+1)^{-1}\sum_{j=-n}^{n} Y_j \exp(-i2\pi k x_j)$ $(k \in \mathbb{Z})$ from the article. This estimator is biased only in the design points, which asymptotically exhaust the interval $[-1/2, 1/2]$ at the rate n^{-1}. We arrive at the following result concerning the bias of $\hat{R}$.

Lemma 1. *Let $\theta \in \mathscr{R}_s$, with $s \geq 1$. Then*

$$\max_{k\in\mathbb{Z}} \Big| E\big[\hat{R}(k)\big] - R(k) \Big| = O\big(n^{-1}\big).$$

Proof. For any $s_1 \leq s_2$, we have the inclusion $\mathscr{R}_{s_2} \subset \mathscr{R}_{s_1}$, and, therefore, we only need to prove the result for $s = 1$. For $r = K\theta$, we can write

$$E\big[\hat{R}(k)\big] = \frac{1}{2n+1}\sum_{j=-n}^{n} r(x_j)e^{-i2\pi k x_j}. \tag{S1.1}$$

(S1.1) shows that $\hat{R}$ is, on the average, estimating the discrete Fourier transform of r calculated on the design points, which is expected.

We can relate the discrete Fourier transform of r to its Fourier coefficients $\{R(k)\}_{k\in\mathbb{Z}}$ as follows. Partitioning the interval $[-1/2,\, 1/2]$ into

$$\bigcup_{j=-n}^{n}\left[\frac{2j-1}{4n+2},\, \frac{2j+1}{4n+2}\right)\bigcup\left\{\frac{1}{2}\right\}$$

allows a decomposition of $R(k)$ into an average corresponding to the design points $x_j = j/(2n)$; that is,

$$\begin{aligned} R(k) &= \int_{-1/2}^{1/2} r(x)e^{-i2\pi kx}\, dx \\ &= \frac{1}{2n+1}\sum_{j=-n}^{n}\int_{-1/2}^{1/2} r\left(x_j + \frac{v-x_j}{2n+1}\right)\exp\left(-i2\pi k\left(x_j+\frac{v-x_j}{2n+1}\right)\right)dv. \end{aligned}$$

We can see that $|E[\hat{R}(k)] - R(k)|$ is bounded by the sum $R_1(k) + R_2(k)$, with

$$R_1(k) = \frac{1}{2n+1}\sum_{j=-n}^{n}\int_{-1/2}^{1/2}\left|r(x_j) - r\left(x_j+\frac{v-x_j}{2n+1}\right)\right| dv$$

and $R_2(k)$ is equal to

$$\frac{1}{2n+1}\sum_{j=-n}^{n}\left|\int_{-1/2}^{1/2} r\left(x_j+\frac{v-x_j}{2n+1}\right)\left\{\exp\left(-i2\pi kx_j\right) - \exp\left(-i2\pi k\left(x_j+\frac{v-x_j}{2n+1}\right)\right)\right\}dv\right|.$$

The result follows, if we can show that $\max_{k\in\mathbb{Z}} R_i(k) = O(n^{-1})$, for each $i = 1, 2$.

To continue, use Euler's formula to write

$$\exp\left(-i2\pi kx_j\right) = \cos\left(2\pi kx_j\right) - i\sin\left(2\pi kx_j\right).$$

Since sine and cosine are each Lipschitz functions with constant equal to 1, it follows that

$$\left|\exp\left(-i2\pi kx_j\right) - \exp\left(-i2\pi k\left(x_j+\frac{v-x_j}{2n+1}\right)\right)\right|^2 \le 2^3\pi^2k^2\left(\frac{v-x_j}{2n+1}\right)^2.$$

Therefore, we have the bound

$$\left|\exp\left(-i2\pi kx_j\right) - \exp\left(-i2\pi k\left(x_j+\frac{v-x_j}{2n+1}\right)\right)\right| \le 2^{3/2}\pi|k|\frac{|v-x_j|}{2n+1},$$

which will be used throughout the proof.

Beginning with $R_1(k)$, it follows from both $\theta \in \mathscr{R}_1$ and the equivalence $r = K\theta$ that $r \in \mathscr{R}_1$ as well. Using the Fourier inversion formula, write

$$\Big| r(x_j) - r\Big(x_j + \frac{v - x_j}{2n+1}\Big)\Big| \le \frac{2^{3/2}\pi}{2n+1}\big|v - x_j\big| \sum_{k=-\infty}^{\infty} |k|\big|R(k)\big|.$$

Hence, we can find an appropriate constant $C > 0$, such that $R_1(k)$ is bounded by

$$Cn^{-2} \sum_{j=-n}^{n} \int_{-1/2}^{1/2} \big|v - x_j\big|\, dv,$$

which does not depend on k, and this is easily seen to be of order $O(n^{-1})$. This implies that $\max_{k\in\mathbb{Z}} R_1(k) = O(n^{-1})$.

Turning our attention to $R_2(k)$, we can assume without loss of generality that $|k| > 0$ as this term is equal to zero whenever $k = 0$. The integral in $R_2(k)$ is equal to the sum of

$$\int_{-1/2}^{1/2} \Big\{ r\Big(x_j + \frac{v - x_j}{2n+1}\Big) - r(x_j)\Big\}\, dv \exp\big(-i2\pi k x_j\big)$$

and

$$\int_{-1/2}^{1/2} \Big\{ r(x_j)\exp\big(-i2\pi k x_j\big) - r\Big(x_j + \frac{v - x_j}{2n+1}\Big)\exp\Big(-i2\pi k\Big(x_j + \frac{v - x_j}{2n+1}\Big)\Big)\Big\}\, dv.$$

Therefore, $R_2(k)$ is bounded by the sum of $\max_{k\in\mathbb{Z}} R_1(k)$ and the quantity

$$\begin{aligned} \frac{1}{2n+1} \sum_{j=-n}^{n} \Big| \int_{-1/2}^{1/2} \Big\{ & r(x_j)\exp\big(-i2\pi k x_j\big) \\ & - r\Big(x_j + \frac{v - x_j}{2n+1}\Big)\exp\Big(-i2\pi k\Big(x_j + \frac{v - x_j}{2n+1}\Big)\Big)\Big\}\, dv\Big|. \end{aligned} \tag{S1.2}$$

We can use the Fourier inversion formula to write

$$\begin{aligned} & r(x_j)\exp\big(-i2\pi k x_j\big) - r\Big(x_j + \frac{v - x_j}{2n+1}\Big)\exp\Big(-i2\pi k\Big(x_j + \frac{v - x_j}{2n+1}\Big)\Big) \\ & = \sum_{\xi=-\infty}^{\infty} R(\xi)\Big\{\exp\big(i2\pi(\xi - k)x_j\big) - \exp\Big(i2\pi(\xi - k)\Big(x_j + \frac{v - x_j}{2n+1}\Big)\Big)\Big\}. \end{aligned} \tag{S1.3}$$

From (S1.3), we can see that (S1.2) is further bounded by

$$\frac{1}{2n+1}\sum_{j=-n}^{n}\int_{-1/2}^{1/2}\sum_{\xi=-\infty}^{\infty}|R(\xi)|\left|\exp\big(i2\pi(\xi-k)x_j\big)-\exp\left(i2\pi(\xi-k)\left(x_j+\frac{v-x_j}{2n+1}\right)\right)\right|dv$$
$$\leq \frac{2^{3/2}\pi}{2n+1}\left\{\frac{1}{2n+1}\sum_{j=-n}^{n}\int_{-1/2}^{1/2}|v-x_j|\,dv\right\}\left\{\sum_{|\xi-k|>0}|\xi-k||R(\xi)|\right\}.$$

Since we have already shown that $r\in\mathscr{R}_1$, we have, for $\zeta=\xi-k$, $\max_{|k|>0}\sum_{|\zeta|>0}|\zeta||R(k+\zeta)|<\infty$. Thus, we can find an appropriate constant $C>0$ to see that (S1.2) is further bounded by

$$Cn^{-2}\sum_{j=-n}^{n}\int_{-1/2}^{1/2}|v-x_j|\,dv,$$

which does not depend on k, and this is easily seen to be of order $O(n^{-1})$. Combining this fact with the result that $\max_{k\in\mathbb{Z}}R_1(k)=O(n^{-1})$ implies that $\max_{k\in\mathbb{Z}}R_2(k)=O(n^{-1})$. □

With the result of Lemma 1, we can give the proof of Lemma 1 from the article:

Proof of Lemma 1 from the article. We begin with the decomposition

$$E\big[\hat{\theta}(x)\big]=\sum_{k=-\infty}^{\infty}\Lambda(h_nk)\Theta(k)\exp(i2\pi kx)+\sum_{k=-\infty}^{\infty}\frac{\Lambda(h_nk)}{\Psi(k)}\Big\{E\big[\hat{R}(k)\big]-R(k)\Big\}\exp(i2\pi kx)$$

so that $E[\hat{\theta}(x)]-\theta(x)$ is equal to

$$\sum_{k=-\infty}^{\infty}\big\{\Lambda(h_nk)-1\big\}\Theta(k)\exp(i2\pi kx)+\sum_{k=-\infty}^{\infty}\frac{\Lambda(h_nk)}{\Psi(k)}\Big\{E\big[\hat{R}(k)\big]-R(k)\Big\}\exp(i2\pi kx).$$

We can see that $\sup_{x\in[-1/2,\,1/2]}|E[\hat{\theta}(x)]-\theta(x)|$ is bounded by

$$\sum_{k=-\infty}^{\infty}|\Lambda(h_nk)-1||\Theta(k)|+\max_{k\in\mathbb{Z}}\Big|E\big[\hat{R}(k)\big]-R(k)\Big|\sum_{k=-\infty}^{\infty}\frac{|\Lambda(h_nk)|}{|\Psi(k)|}. \tag{S1.4}$$

Partition $\mathbb{Z}$ into $I(h_n)\cup I^c(h_n)$, where $I(h_n)=\{z\in\mathbb{Z}:h_n|z|\leq M\}=\{z\in\mathbb{Z}:|z|\leq Mh_n^{-1}\}$. Hence, for every $k\in I^c(h_n)$, it follows that $|\Lambda(h_nk)|\leq 1$, which implies both statements $|\Lambda(h_n k)-1| \leq 2$ and $|k| > M h_n^{-1}$ hold. The first term in the right-hand side of (S1.4) is therefore bounded by

$$2 \sum_{k \in I^c(h_n)} |\Theta(k)| \leq 2 h_n^s M^{-s} \sum_{k=-\infty}^{\infty} |k|^s |\Theta(k)|. \tag{S1.5}$$

This implies the first term in (S1.4) is of order $O(h_n^s)$.

We now turn to the second term in (S1.4). It follows from Assumptions 1 and 2 for the series in this term to be bounded by

$$h_n^{-1}\left[\min_{k \in \{z \in \mathbb{Z}:\, |z| \leq \Gamma\}} |\Psi(k)|\right]^{-1}\left\{h_n \sum_{\omega \in h_n \mathbb{Z}} |\Lambda(\omega)|\right\} + h_n^{-b-1} C_\Psi^{-1}\left\{h_n \sum_{\omega \in h_n \mathbb{Z}} |\omega|^b |\Lambda(\omega)|\right\},$$

which is easily seen to be of order $O(h_n^{-b-1})$. The additional factor of h_n^{-1} appears in the bound above because we have a shrinkage of k by h_n. This implies that $\sum_{k=-\infty}^{\infty}\{|\Lambda(h_n k)|/|\Psi(k)|\}$ is of order $O(h_n^{-b-1})$. Now we only need to consider the term $\max_{k \in \mathbb{Z}} |\hat{R}(k) - R(k)|$. The assumptions of Lemma 1 are satisfied. It then follows for $\max_{k \in \mathbb{Z}} |\hat{R}(k) - R(k)| = O(n^{-1})$. Hence, the second term in (S1.4) is of order $O((n h_n^{b+1})^{-1})$. □

We are now prepared to state the proof of Lemma 2 from the article.

Proof of Lemma 2 from the article. Without loss of generality we can assume that $n \geq 3$. Our argument is similar to the arguments found in Masry (1993), who gives related results for an errors-in-variables model. We will employ truncation as follows. Let the stabilizing sequence $\{\eta_n\}_{n \geq 3}$ satisfy $\eta_n = O((n h_n^{2b+1})^{-1/2} \log^{1/2}(n))$ and the truncation sequence $\{t_n\}_{n \geq 3}$ satisfy $t_n = O((n \log(n)(\log\log(n))^{1+\delta})^{1/\kappa})$, with $\delta > 0$. Write $K_j = E^{1/\kappa}[|Y_j|^\kappa]$. We can decompose $\hat{\theta}(x) - E[\hat{\theta}(x)]$ into the sum of $D_1(x) = \hat{\theta}(x) - \hat{\theta}^t(x)$, $D_2(x) = E[\hat{\theta}^t(x)] - E[\hat{\theta}(x)]$ and $D_3(x) = \hat{\theta}^t(x) - E[\hat{\theta}^t(x)]$, where

$$\hat{\theta}^t(x) = \frac{1}{2n+1} \sum_{j=-n}^{n} Y_j \mathbf{1}\big[|Y_j| \leq K_j t_n\big] W_{j,h_n}(x), \qquad x \in [-1/2,\, 1/2].$$

Beginning with $D_1(x)$, it follows along the same lines as the arguments in the proof of Lemma 2.1 of Masry (1993) for $\sup_{x\in[-1/2,\,1/2]}|D_1(x)| = o(\eta_n)$, almost surely. It is easy to show that $\sup_{x\in[-1/2,\,1/2]}|W_{j,h_n}(x)|$ is bounded by the series $\sum_{k=-\infty}^{\infty}\{|\Lambda(h_n k)|/|\Psi(k)|\}$, and we have already shown that this series is of order $O(h_n^{-b-1})$ in the proof of Lemma 1 from the article. Hence, we have that $\sup_{x\in[-1/2,\,1/2]}|W_{j,h_n}(x)| = O(h_n^{-b-1})$. It follows that we can find an appropriate constant $C > 0$, such that we can bound $\sup_{x\in[-1/2,\,1/2]}|D_2(x)|$ by

$$Ch_n^{-b-1}\frac{1}{2n+1}\sum_{j=-n}^{n} E\Big[|Y_j|\mathbf{1}\big[|Y_j| > K_j t_n\big]\Big]. \tag{S1.6}$$

Since $\kappa > 1$, we can apply Markov's inequality to obtain

$$\max_{j=-n,\ldots,n} E[|Y_j|\mathbf{1}[|Y_j| > K_j t_n]] = \max_{j=-n,\ldots,n}\int_0^\infty P\big(|Y_j| > \max\{s,\, K_j t_n\}\big)\,ds \le \frac{\kappa}{\kappa-1}M_K t_n^{1-\kappa},$$

with $M_K = \max_{j=-n,\ldots,n} K_j$. Therefore, enlarging the constant C in (S1.6) implies that $\sup_{x\in[-1/2,\,1/2]}|D_2(x)| \le Ch_n^{-b-1}t_n^{1-\kappa} = o(\eta_n)$.

To continue, we will require an additional result. For any $u, v \in [-1/2,\, 1/2]$, we can repeat the arguments in the proof of Lemma 1 to see that

$$\Big|W_{j,h_n}(u) - W_{j,h_n}(v)\Big| \le \big|u-v\big|2^{3/2}\pi\sum_{k=-\infty}^{\infty}|k|\frac{|\Lambda(h_n k)|}{|\Psi(k)|}.$$

Hence, we can find an appropriate constant $C > 0$, such that

$$\big|W_{j,h_n}(u) - W_{j,h_n}(v)\big| \le Ch_n^{-b-2}|u-v|, \qquad u, v \in [-1/2,\, 1/2]. \tag{S1.7}$$

Now we consider $D_3(x)$. Let $\{s_n\}_{n\ge3}$ be a sequence satisfying $s_n = O(h_n^{b+2}\eta_n t_n^{-1}) = o(1)$, such that, when we partition the interval $[-1/2,\, 1/2]$ into s_n^{-1} many intervals of the form $(x_i,\, x_{i+1}]$, with the first interval defined to be $[-1/2,\, x_2] = \{-1/2\} \cup (-1/2,\, x_2]$, the end points of our intervals satisfy $\max_{i=1,\ldots,s_n^{-1}}|x_{i+1} - x_i| \le s_n$. For any $x \in [-1/2,\, 1/2]$, there is exactly one interval $(x_{i'},\, x_{i'+1}]$ that contains x. On this interval, we can write

$$D_3(x) = D_{4,i}(x) - D_{5,i}(x) + D_{6,i},$$

where $D_{4,i}(x) = \hat{\theta}^t(x) - \hat{\theta}^t(x_{i'})$, $D_{5,i}(x) = E[\hat{\theta}^t(x)] - E[\hat{\theta}^t(x_{i'})]$, and $D_{6,i} = \hat{\theta}^t(x_{i'}) - E[\hat{\theta}^t(x_{i'})]$. It follows that $\sup_{x \in [-1/2,\, 1/2]} |D_3(x)|$ is bounded by

$$\max_{i=1,\ldots,s_n^{-1}} \sup_{x \in (x_i,\, x_{i+1}]} \big|D_{4,i}(x)\big| + \max_{i=1,\ldots,s_n^{-1}} \sup_{x \in (x_i,\, x_{i+1}]} \big|D_{5,i}(x)\big| + \max_{i=1,\ldots,s_n^{-1}} \big|D_{6,i}\big|.$$

Hence, the proof is complete once we have shown

$$\max_{i=1,\ldots,s_n^{-1}} \sup_{x \in (x_i,\, x_{i+1}]} |D_{4,i}(x)| = O(\eta_n), \quad \text{a.s.}, \tag{S1.8}$$

$$\max_{i=1,\ldots,s_n^{-1}} \sup_{x \in (x_i,\, x_{i+1}]} |D_{5,i}(x)| = O(\eta_n), \tag{S1.9}$$

and

$$\max_{i=1,\ldots,s_n^{-1}} |D_{6,i}| = O(\eta_n), \quad \text{a.s.}. \tag{S1.10}$$

Beginning with (S1.8), fix an arbitrary interval $(x_i,\, x_{i+1}]$. $D_{4,i}(x)$ is equal to

$$\frac{1}{2n+1} \sum_{j=-n}^{n} Y_j \mathbf{1}\big[|Y_j| \le K_j t_n\big] \big\{W_{j,h_n}(x) - W_{j,h_n}(x_i)\big\}, \qquad x \in (x_i,\, x_{i+1}].$$

It follows from (S1.7) that the inequality $\sup_{x \in (x_i,\, x_{i+1}]} |D_{4,i}(x)| \le C t_n h_n^{-b-2} s_n$ holds, almost surely, independent of i, for some $C > 0$. Therefore, by construction of $\{s_n\}_{n \ge 3}$, (S1.8) holds. Observing that $D_{5,i}(x) = E[D_{4,i}(x)]$, (S1.9) holds as well.

To see the final statement (S1.10) holds, define the random variables $U_{j,i} = \{Y_j \mathbf{1}[|Y_j| \le K_j t_n] - E[Y_j \mathbf{1}[|Y_j| \le K_j t_n]]\} W_{j,h_n}(x_i)$, $j = -n, \ldots, n$. Standard arguments can then be used to show that $U_{-n,i}, \ldots, U_{n,i}$ are independent, have mean zero, variance bounded by $C_1 h_n^{-2b-1}$, and are bounded in absolute value by $C_2 t_n h_n^{-b-1}$, for some $C_1 > 0$ and $C_2 > 0$. Note that both bounds are independent of j and i. Applying Bernstein's Inequality [see, e.g., Lemma

2.2.11 in van der Vaart and Wellner (1996)], there is an appropriate constant $C > 0$, such that

$$P\Big(\max_{i=1,\dots,s_n^{-1}} |D_{6,i}| > \eta_n\Big) \le 2s_n^{-1} \exp\Big(-C\frac{n\eta_n^2}{h_n^{-2b-1} + t_n h_n^{-b-1}\eta_n}\Big). \tag{S1.11}$$

In light of the fact that $t_n h_n^{-b-1}\eta_n = o(h_n^{-2b-1})$, since $\kappa > 2 + 1/b$, we can enlarge C for the right-hand side of (S1.11) to be further bounded by a positive constant multiplied by

$$h_n^{-3/2} n^{(1/2)+(1/\kappa)-C} \log^{-(1/2-1/\kappa)}(n)\big(\log\log(n)\big)^{(1+\delta)/\kappa}.$$

This bound is summable in n, provided we take $C > (3/2)(1 + 1/(2b + 1)) + 1/\kappa$, where $1/(2b + 1)$ accounts for the expansion of h_n^{-1}; that is, $(n^{1/(2b+1)}h_n)^{-3/2} \to 0$, as $n \to \infty$. It then follows by the Borel-Cantelli lemma that (S1.10) holds. □

We can now state the proof of Theorem 1 from the article:

Proof of Theorem 1 from the article. The first two assertions follow immediately from the results of Lemma 1 and Lemma 2 from the article, with the choice of regularizing sequence as discussed in Section 2.1 of the article. This means that we only need to show the last assertion. We begin by calculating the Fourier coefficients $\{\hat\Theta(\xi)\}_{\xi\in\mathbb{Z}}$ of $\hat\theta$; that is,

$$\begin{aligned}\hat\Theta(\xi) &= \int_{-1/2}^{1/2} \hat\theta(x)e^{-i2\pi\xi x}\,dx = \sum_{k=-\infty}^{\infty} \frac{\Lambda(h_n k)}{\Psi(k)}\hat R(k)\int_{-1/2}^{1/2} e^{i2\pi(k-\xi)x}\,dx \\ &= \Lambda(h_n\xi)\Theta(\xi) + \Big\{E\big[\hat R(\xi)\big] - R(\xi) + \frac{1}{2n+1}\sum_{j=-n}^{n}\varepsilon_j e^{-i2\pi\xi x_j}\Big\}\frac{\Lambda(h_n\xi)}{\Psi(\xi)},\end{aligned}$$

where we have used the orthonormality of the basis $\{\exp(i2\pi kx) : x \in [-1/2,\, 1/2]\}_{k\in\mathbb{Z}}$ in the final equality. From the definition of $\mathscr{R}_{s-1/2}$, we show that

$$\sum_{\xi=-\infty}^{\infty} |\xi|^{s-1/2}\big|\hat\Theta(\xi)\big| < \infty \tag{S1.12}$$

We can see that $|\hat\Theta(\xi)|$ is bounded by

$$|\Theta(\xi)| + \Big\{\max_{k\in\mathbb{Z}}\Big|E\big[\hat R(k)\big] - R(k)\Big| + \Big|\frac{1}{2n+1}\sum_{j=-n}^{n}\varepsilon_j e^{i2\pi\xi x_j}\Big|\Big\}\frac{|\Lambda(h_n\xi)|}{|\Psi(\xi)|}. \tag{S1.13}$$

$\theta \in \mathscr{R}_s$ implies that $\theta \in \mathscr{R}_{s-1/2}$.

The assumptions of Lemma 1 are satisfied, and $\max_{k\in\mathbb{Z}} |E[\hat{R}(k)] - R(k)| = O(n^{-1})$. Additionally, the map $x \mapsto \exp(-i2\pi kx)$ is confined to the unit circle in the complex plane. A standard truncation argument shows that

$$\max_{k\in\mathbb{Z}} \left| \frac{1}{2n+1} \sum_{j=-n}^{n} \varepsilon_j e^{-i2\pi k x_j} \right| = O\big(n^{-1/2} \log^{1/2}(n)\big), \qquad \text{a.s.}$$

In the proof of Lemma 1 from the article, we have shown that $\sum_{k=-\infty}^{\infty} \{|\Lambda(h_n k)|/|\Psi(k)|\} = O(h_n^{-b-1})$, and a similar argument yields $\sum_{k=-\infty}^{\infty} \{|k|^{s-1/2} |\Lambda(h_n k)|/|\Psi(k)|\} = O(h_n^{-s-b-1/2})$, with the assumption $\int_{-\infty}^{\infty} |u|^{s+b-1/2} |\Lambda(u)|\, du < \infty$. Together, these results imply that

$$\left\{ \max_{k\in\mathbb{Z}} \left| E\big[\hat{R}(k)\big] - R(k) \right| + \max_{k\in\mathbb{Z}} \left| \frac{1}{2n+1} \sum_{j=-n}^{n} \varepsilon_j e^{-i2\pi k x_j} \right| \right\} \sum_{\xi=-\infty}^{\infty} |\xi|^{s-1/2} \frac{|\Lambda(h_n \xi)|}{|\Psi(\xi)|}$$

is of order $O(1 + (n\log(n))^{-1/2}) = O(1)$, almost surely, and the series condition (S1.12) stated for the last term in (S1.13) holds. It follows that $\hat{\theta} - \theta \in \mathscr{R}_{s-1/2}$, almost surely, for large enough n. This statement with the first assertion yields the third assertion. □

Nickl and Pötscher (2007) study classes of functions of Besov- and Sobolev-type. These authors derive results concerning the bracketing metric entropy of these spaces, and the related central limit theorems. Their Corollary 4 on bracketing numbers for weighted Sobolev spaces can be extended to our case (see page 189). We summarize this result in the following proposition:

Proposition 1. *For the function space $\mathscr{R}_{s,1}$, with $s > 1/2$, a finite constant $C > 0$ exists, such that*

$$\log N_{[\,]}\big(\epsilon, \mathscr{R}_{s,1}, \|\cdot\|_\infty\big) \le C\epsilon^{-1/s}, \qquad \epsilon > 0,$$

where $N_{[\,]}(\epsilon, \mathscr{R}_{s,1}, \|\cdot\|_\infty)$ is the number of brackets of length ϵ required to cover the metric space $(\mathscr{R}_{s,1}, \|\cdot\|_\infty)$.

In light of the results on the estimator $\hat{\theta}$, we can now state a result on the modulus of continuity relating $\hat{\mathbb{F}}(t)$ to $(2n+1)^{-1}\sum_{j=-n}^{n}\mathbf{1}[\varepsilon_j \leq t]$. Using results on Donsker classes of functions, we can show this modulus of continuity holds up to a negligible term of order $o_P(n^{-1/2})$. The proof of this result follows along the same lines as the proof of Lemma A.1 in Van Keilegom and Akritas (1999), and, therefore, it is omitted [see also Neumeyer (2009)].

Lemma 2. *Let the assumptions of Theorem 1 from the article be satisfied, with $s > 3/2$. In addition, assume that F admits a bounded Lebesgue density function f. Then $\sup_{t\in\mathbb{R}}|M_n(t)| = o_P(n^{-1/2})$, where*

$$M_n(t) = \frac{1}{2n+1}\sum_{j=-n}^{n}\mathbf{1}\big[\varepsilon_j \leq t + \big[K\big(\hat{\theta}-\theta\big)\big](x_j)\big] - \int_{-1/2}^{1/2} F\big(t + \big[K\big(\hat{\theta}-\theta\big)\big](x)\big)\,dx - \frac{1}{2n+1}\sum_{j=-n}^{n}\mathbf{1}[\varepsilon_j \leq t] + F(t).$$

In the following result, we provide an expansion for the indirect regression estimator $\hat{\theta}$. This property, combined with the modulus of continuity result above, shows that our residual-based empirical distribution function behaves similarly to that in the a direct estimation setting [see, e.g., Müller et al. (2007), who construct expansions for many residual-based empirical distribution functions based on direct regression function estimators].

Proposition 2. *Let the assumptions of Lemma 1 from the article be satisfied, and assume that $E[\varepsilon_j^2] < \infty$, $j = -n, \ldots, n$. Let the regularizing sequence $\{h_n\}_{n\geq 1}$ satisfy $h_n^{s+b+1} = o(n^{-1/2})$ and $(nh_n)^{-1} = o(n^{-1/2})$. Then,*

$$\left|\int_{-1/2}^{1/2}\big[K\big(\hat{\theta}-\theta\big)\big](x)\,dx - \frac{1}{2n+1}\sum_{j=-n}^{n}\varepsilon_j\right| = o_P(n^{-1/2}).$$

Proof. Note that $\hat{R}(k) - E[\hat{R}(k)] = (2n+1)^{-1}\sum_{j=-n}^{n}\varepsilon_j\exp(-i2\pi k x_j)$. The left-hand side of the assertion is bounded by $S_1 + S_2 + S_3$, where

$$S_1 = \left| \frac{1}{2n+1} \sum_{j=-n}^{n} \varepsilon_j \int_{-1/2}^{1/2} \left\{ \sum_{k=-\infty}^{\infty} \{\Lambda(h_n k) - 1\} e^{i2\pi k(x - x_j)} \right\} dx \right|,$$

$$S_2 = \sum_{k=-\infty}^{\infty} \big|\Lambda(h_n k) - 1\big| \big|R(k)\big| \left| \int_{-1/2}^{1/2} e^{i2\pi k x}\, dx \right|,$$

and

$$S_3 = \left[\max_{k \in \mathbb{Z}} \left| E\big[\hat{R}(k)\big] - R(k) \right| \right] \sum_{k=-\infty}^{\infty} |\Lambda(h_n k)|.$$

It follows that $S_1 = o_P(n^{-1/2})$ from Assumption 2 from the article and that

$$\frac{1}{2n+1} \sum_{j=-n}^{n} \left\{ \int_{-1/2}^{1/2} \left\{ \sum_{k=-\infty}^{\infty} \{\Lambda(h_n k) - 1\} e^{i2\pi k(x - x_j)} \right\} dx \right\}^2 = o(1).$$

The proof is complete, once we have shown that $S_1 = o_P(n^{-1/2})$, $S_2 = o(n^{-1/2})$, and $S_3 = o(n^{-1/2})$.

The convolution theorem for Fourier transformation implies that $|R(k)| = |\Theta(k)||\Psi(k)|$. Thus, the integral term in S_2 is bounded by a positive constant C multiplied by $|k|^{-1}$. This fact, the constant C_Ψ^* from Assumption 1 from the article, and that $|\Lambda(h_n k) - 1| \leq 2$ shows that we can enlarge C, such that S_2 is bounded by

$$C h_n^{s+b+1} \sum_{k=-\infty}^{\infty} |k|^s |\Theta(k)| = O\big(h_n^{s+b+1}\big) = o\big(n^{-1/2}\big).$$

The assumptions of Lemma 1 are satisfied, and $\max_{k\in\mathbb{Z}} |E[\hat{R}(k)] - R(k)| = O(n^{-1})$. One shows that the series term in S_3 is of order $O(h_n^{-1})$ using similar lines of argument to those in the proof of Lemma 1 from the article. It follows that S_3 is of order $O((nh_n)^{-1}) = o(n^{-1/2})$. □

We can now give the proof of Theorem 2 from the article.

Proof of Theorem 2 from the article. Recall $M_n(t)$ from Lemma 2. A straightforward calculation shows that

$$\frac{1}{2n+1}\sum_{j=-n}^{n}\Big\{\mathbf{1}\big[\hat{\varepsilon}_j \le t\big] - \mathbf{1}\big[\varepsilon_j \le t\big] - \varepsilon_j f(t)\Big\} = M_n(t) + H_n(t) + L_n(t),$$

with

$$H_n(t) = \int_{-1/2}^{1/2} F\big(t + \big[K\big(\hat{\theta}-\theta\big)\big](x)\big)\,dx - F(t) - f(t)\int_{-1/2}^{1/2}\big[K\big(\hat{\theta}-\theta\big)\big](x)\,dx$$

and

$$L_n(t) = f(t)\bigg\{\int_{-1/2}^{1/2}\big[K\big(\hat{\theta}-\theta\big)\big](x)\,dx - \frac{1}{2n+1}\sum_{j=-n}^{n}\varepsilon_j\bigg\}.$$

The assumptions of Lemma 2 are satisfied, and $\sup_{t\in\mathbb{R}}|M_n(t)| = o_P(n^{-1/2})$. The assertion follows, once we show that $\sup_{t\in\mathbb{R}}|H_n(t)| = o_P(n^{-1/2})$ and $\sup_{t\in\mathbb{R}}|L_n(t)| = o_P(n^{-1/2})$.

Using the Hölder continuity of f, it follows that

$$H_n(t) = \int_{-1/2}^{1/2}\big[K\big(\hat{\theta}-\theta\big)\big](x)\int_0^1\Big\{f\big(t + s\big[K\big(\hat{\theta}-\theta\big)\big](x)\big) - f(t)\Big\}\,ds\,dx,$$

writing $C_{f,\gamma}$ for the Hölder constant of f with exponent γ. Therefore, $\sup_{t\in\mathbb{R}}|H_n(t)|$ is bounded by

$$\frac{C_{f,\gamma}}{1+\gamma}\bigg[\sup_{x\in[-1/2,\,1/2]}\Big|\hat{\theta}(x)-\theta(x)\Big|\bigg]^{1+\gamma}.$$

The assumptions of Theorem 1 from the article are satisfied, and the second term in the bound above is $o(n^{-1/2})$, almost surely. This fact implies that $\sup_{t\in\mathbb{R}}|H_n(t)| = o_P(n^{-1/2})$.

Since f is bounded, $\sup_{t\in\mathbb{R}}|L_n(t)|$ is bounded by

$$\sup_{t\in\mathbb{R}}|f(t)|\bigg|\int_{-1/2}^{1/2}\big[K\big(\hat{\theta}-\theta\big)\big](x)\,dx - \frac{1}{2n+1}\sum_{j=-n}^{n}\varepsilon_j\bigg|.$$

The parameter sequence $\{h_n\}_{n\ge1}$ satisfies

$$h_n^{s+b+1} = O\big(n^{-1/2-1/(4s+4b+2)}\log^{(s+b+1)/(2s+2b+1)}(n)\big) = o\big(n^{-1/2}\big)$$

and

$$\left(nh_n\right)^{-1} = O\left(n^{-(2s+2b)/(2s+2b+1)}\log^{-1/(2s+2b+1)}(n)\right) = o\left(n^{-1/2}\right),$$

since $s+b > 1/2$. The assumptions of Proposition 2 are satisfied, which implies that the second term in the bound above is $o_P(n^{-1/2})$. This shows that $\sup_{t\in\mathbb{R}} |L_n(t)| = o_P(n^{-1/2})$. □

Here, we provide a short proof of Proposition 1 from the article.

Proof of Proposition 1 from the article. Write the integrated variance of $\hat{\theta}$ as

$$\int_{-1/2}^{1/2} E\left[\left\{\hat{\theta}(x) - E\left[\hat{\theta}(x)\right]\right\}^2\right] dx = \frac{\sigma^2}{2n+1} \sum_{k=-\infty}^{\infty} \frac{\Lambda^2(h_n k)}{\Psi^2(k)}.$$

Repeating the arguments in the proof of Lemma 1 from the article shows that

$$\sum_{k=-\infty}^{\infty} \frac{\Lambda^2(h_n k)}{\Psi^2(k)} = O\left(h_n^{-2b-1}\right).$$

Therefore, we can specify $C_\Lambda > 0$ for the first assertion to hold. The second assertion follows directly by an application of Lemma 1 from the article. □

The choice of scaling sequence $\{c_n\}_{n\geq 1}$, used for the contaminates $c_n U_j$, $j = -n, \ldots, n$, in the smooth bootstrap always satisfies $(nc_n)^{-1}\log(n) = o(1)$. Theorem A of Silverman (1978), the Hölder continuity of w, and the results of Theorem 1 from the article imply that f_n^* is strongly consistent for f, uniformly over the entire real line. The result only holds when the density function f is Hölder with exponent $2/3 < \gamma \leq 1$, the density function w is of similar smoothness, and the smoothness index s of the function space $\mathscr{R}_s$ satisfies $s > (1+\gamma)(2b+1)/(3\gamma - 2)$. This lower bound is larger than the lower bound on s required by the second statement of Theorem 1 from the article. The additional smoothness in θ is required due to the fact that residuals are used in the estimator f_n^* rather than the model errors. Arguments similar to those used to prove related results in Neumeyer (2009) can be used to prove the following result.

Proposition 3. *Let the assumptions of Theorem 3 from the article be satisfied. Assume that the densities f and w are Hölder continuous with exponent $2/3 < \gamma \le 1$. Let the smoothness index s of the function space $\mathscr{R}_s$ satisfy $s > (1+\gamma)(2b+1)/(3\gamma-2)$. Then,*

$$\sup_{t\in\mathbb{R}} \Big| f_n^*(t) - f(t) \Big| = o_P(1),$$

$$\sup_{t\in\mathbb{R}} \Big| F_n^*(t) - F(t) \Big| = o_P(1),$$

and

$$\sup_{t\in\mathbb{R}} \Big| E^*\big[\varepsilon^* \mathbf{1}\big[\varepsilon^* \le t\big]\big] - E\big[\varepsilon \mathbf{1}[\varepsilon \le t]\big] \Big| = o_P(1).$$

We omit proof of the following result because it is proven in exactly the same manner as Theorem 1 from the article.

Proposition 4. *Let the assumptions of Theorem 1 from the article be satisfied. Choose the regularizing sequence $\{g_n\}_{n\ge1}$ according to (2.3) from the article, and let the scaling sequence $\{c_n\}_{n\ge1}$ satisfy $c_n = O(n^{-\alpha})$, with $0 < \alpha < 1/2 + 1/\kappa$. Then, P^*-almost surely,*

$$\sup_{x\in[-1/2,\,1/2]} \Big| \hat{\theta}^*(x) - \hat{\theta}(x) \Big| = O\big(n^{-s/(2s+2b+1)} \log^{s/(2s+2b+1)}(n)\big).$$

If $s > (2b+1)/(2\gamma)$, for some $0 < \gamma \le 1$,

$$\Bigg[\sup_{x\in[-1/2,\,1/2]} \Big| \hat{\theta}^*(x) - \hat{\theta}(x) \Big| \Bigg]^{1+\gamma} = o(n^{-1/2}).$$

For large enough n,

$$\hat{\theta}^* - \hat{\theta} \in \mathscr{R}_{s-1/2,1}.$$

References

E. Masry. Multivariate regression estimation with errors-in-variables for stationary processes. *J. Nonparamtr. Stat.*, 3(1):13–36, 1993. doi: 10.1080/10485259308832569.

U.U. Müller, A. Shick, and W. Wefelmeyer. Estimating the error distribution function in semiparametric regression. *Statist. Decisions*, 25(1):1–18, 2007. doi: 10.1524/stnd.2007.25.1.1.

N. Neumeyer. Smooth residual bootstrap for empirical processes of non-parametric regression residuals. *Scand. J. Stat.*, 36(2):204–228, 2009. doi: 10.1111/j.1467-9469.2008.00628.x.

R. Nickl and B.M. Pötscher. Bracketing metric entropy rates and empirical central limit theorems for function classes of besov- and sobolev-type. *J. Theoret. Probab.*, 20(2):177–199, 2007. doi: 10.1007/s10959-007-0058-1.

B.W. Silverman. Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *Ann. Statist.*, 6(1):177–184, 1978. doi: 10.1214/aos/1176344076.

A.W. van der Vaart and J.A. Wellner. *Weak convergence and empirical processes. With applications to statistics.* Springer series in statistics. Springer-Verlag, New York, 1996.

I. Van Keilegom and M.G. Akritas. Transfer of tail information in censored regression models. *Ann. Statist.*, 27(5):1745–1784, 1999. doi: 10.1214/aos/1017939150.