*Econometrica Supplementary Material*

# SUPPLEMENT TO "ROBUST NONPARAMETRIC CONFIDENCE INTERVALS FOR REGRESSION-DISCONTINUITY DESIGNS"

(*Econometrica*, Vol. 82, No. 6, November 2014, 2295–2326)

BY SEBASTIAN CALONICO, MATIAS D. CATTANEO, AND ROCIO TITIUNIK[1]

This supplement to Calonico, Cattaneo, and Titiunik (2014c) contains mathematical proofs of our main theorems, other methodological and technical results, additional simulation evidence, and an empirical illustration employing household data from Progresa/Oportunidades. Companion software packages in R and STATA are described in Calonico, Cattaneo, and Titiunik (2014b, 2014d).

CONTENTS



[1]The authors gratefully acknowledge financial support from the National Science Foundation (SES 1357561).

 DOI: 10.3982/ECTA11757

## S.1. INTRODUCTION

THIS SUPPLEMENT TO CALONICO, CATTANEO, AND TITIUNIK (2014c, CCT hereafter) contains mathematical proofs of our main theorems, other methodological and technical results, additional simulation evidence, and an empirical illustration employing household data from Progresa/Oportunidades.

Section S.2 presents several results for local polynomial estimators, some of which may be of independent interest while others are well known in the literature. For a review on local polynomials, see Fan and Gijbels (1996), and for related theoretical results regarding nonparametric bias correction, see also Calonico, Cattaneo, and Farrell (2014) and references therein. This section includes proofs of Lemmas A.1–A.2 and Theorems A.1–A.2 in CCT, some generalizations, and consistency of the nearest-neighbor-based standard error estimators introduced in Section 5 of CCT. Section S.2 also includes a discussion of consistent bandwidth selection for sharp RD designs, and further details on Remarks 3 and 7 in CCT and generalizations thereof.

Section S.3 describes the details of our simulation study, including a complete discussion of the data generating processes employed and an outline of how the estimators were implemented. The results reported in this section also include other estimators and bandwidth selector procedures, including ad hoc undersmoothing, which were omitted in CCT to conserve space.

Section S.4 complements the numerical evidence on the performance of the results presented in CCT with an empirical application that studies the effects of Progresa/Oportunidades, a large-scale anti-poverty conditional cash transfer program in Mexico, on households' consumption outcomes. We explore the performance of our proposed confidence intervals, as well as several of the conventional alternatives. The empirical results show that in some, but not all, cases, the conclusions drawn from conventional methods are not supported when our robust inference procedures are employed.

### S.1.1. *Setup and Notation*

Let $|\cdot|$ denote the Euclidean matrix norm, that is, $|A|^2 = \operatorname{trace}(A'A)$ for scalar, vector, or matrix $A$. Let $a_n \lesssim b_n$ denote $a_n \leq Cb_n$ for positive constant $C$ not depending on $n$, and $a_n \asymp b_n$ denote $C_1 b_n \leq a_n \leq C_2 b_n$ for positive constants $C_1$ and $C_2$ not depending on $n$. Recall that $\nu, p, q \in \mathbb{Z}_+$ with $\nu \leq p < q$ unless explicitly noted otherwise.

#### S.1.1.1. *Local Polynomial Estimators*

Let $W$ and $Z$ denote two random variables. All objects are defined using the reference outcome variable; this extra generality is used in the fuzzy RD designs. Whenever there is no confusion, we will drop this subindex. We have

$$\hat{\tau}_{Z,\nu,p}(h_n) = \hat{\mu}^{(\nu)}_{Z+,p}(h_n) - \hat{\mu}^{(\nu)}_{Z-,p}(h_n),$$
$$\hat{\mu}^{(\nu)}_{Z+,p}(h_n) = e'_\nu \hat{\beta}_{Z+,p}(h_n), \quad \hat{\mu}^{(\nu)}_{Z-,p}(h_n) = e'_\nu \hat{\beta}_{Z-,p}(h_n),$$
$$\hat{\beta}_{Z+,p}(h_n) = \arg\min_{\beta\in\mathbb{R}^p} \sum_{i=1}^n \mathbf{1}(X_i \geq 0)\big(Z_i - r_p(X_i)'\beta\big)^2 K_{h_n}(X_i),$$
$$\hat{\beta}_{Z-,p}(h_n) = \arg\min_{\beta\in\mathbb{R}^p} \sum_{i=1}^n \mathbf{1}(X_i < 0)\big(Z_i - r_p(X_i)'\beta\big)^2 K_{h_n}(X_i),$$

where $r_p(x) = (1, x, \dots, x^p)'$, $e_\nu$ is the conformable $(\nu+1)$th unit vector (e.g., $e_1 = (0,1,0)'$ if $p = 2$), $K_h(u) = K(u/h)/h$ with $K(\cdot)$ a kernel function, and $h_n$ is a positive bandwidth sequence.

We set $Y = [Y_1, \dots, Y_n]'$, $T = [T_1, \dots, T_n]'$, $\mathcal{X}_n = [X_1, \dots, X_n]'$, $\varepsilon_Z = [\varepsilon_{Z,1}, \dots, \varepsilon_{Z,n}]'$ with $\varepsilon_{Z,i} = Z_i - \mu_Z(X_i)$, $\mu_Z(X) = \mathbb{E}[Z|X]$, and $\Sigma_{WZ} = \mathbb{E}[\varepsilon_W \varepsilon'_Z|\mathcal{X}_n] = \operatorname{diag}(\sigma^2_{WZ}(X_1), \dots, \sigma^2_{WZ}(X_n))$ with $\sigma^2_{WZ}(X) = \mathbb{C}\mathrm{ov}[W, Z|X]$, where $\operatorname{diag}(a_1, \dots, a_n)$ denotes the $(n \times n)$ diagonal matrix with diagonal elements $a_1, \dots, a_n$. We also set

$$X_p(h) = \big[r_p(X_1/h), \dots, r_p(X_n/h)\big]',$$
$$S_p(h) = \big[(X_1/h)^p, \dots, (X_n/h)^p\big]',$$
$$W_+(h) = \operatorname{diag}\big(\mathbf{1}(X_1 \geq 0)K_h(X_1), \dots, \mathbf{1}(X_n \geq 0)K_h(X_n)\big),$$
$$W_-(h) = \operatorname{diag}\big(\mathbf{1}(X_1 < 0)K_h(X_1), \dots, \mathbf{1}(X_n < 0)K_h(X_n)\big).$$

In addition, we define the following (scaled) matrices:

$$\Gamma_{+,p}(h) = X_p(h)'W_+(h)X_p(h)/n,$$
$$\Gamma_{-,p}(h) = X_p(h)'W_-(h)X_p(h)/n,$$
$$\vartheta_{+,p,q}(h) = X_p(h)'W_+(h)S_q(h)/n,$$
$$\vartheta_{-,p,q}(h) = X_p(h)'W_-(h)S_q(h)/n,$$
$$\Psi_{WZ+,p,q}(h,b) = X_p(h)'W_+(h)\Sigma_{WZ}W_+(b)X_q(b)/n,$$
$$\Psi_{WZ-,p,q}(h,b) = X_p(h)'W_-(h)\Sigma_{WZ}W_-(b)X_q(b)/n,$$
$$\Psi_{WZ+,p}(h) = \Psi_{WZ+,p,p}(h,h) = X_p(h)'W_+(h)\Sigma_{WZ}W_+(h)X_p(h)/n,$$
$$\Psi_{WZ-,p}(h) = \Psi_{WZ-,p,p}(h,h) = X_p(h)'W_-(h)\Sigma_{WZ}W_-(h)X_p(h)/n.$$

Letting $H_p(h) = \operatorname{diag}(1, h^{-1}, \dots, h^{-p})$, it follows that

$$\hat{\beta}_{Y+,p}(h_n) = H_p(h_n)\Gamma^{-1}_{+,p}(h_n)X_p(h_n)'W_+(h_n)Y/n,$$
$$\hat{\beta}_{Y-,p}(h_n) = H_p(h_n)\Gamma^{-1}_{-,p}(h_n)X_p(h_n)'W_-(h_n)Y/n.$$

Finally, recall that

$$\beta_{Z+,p} = [\mu_{Z+}, \mu_{Z+}^{(1)}, \mu_{Z+}^{(2)}/2, \dots, \mu_{Z+}^{(p)}/p!]',$$
$$\beta_{Z-,p} = [\mu_{-}, \mu_{Z-}^{(1)}, \mu_{Z-}^{(2)}/2, \dots, \mu_{Z-}^{(p)}/p!]',$$
$$\Gamma_p = \int_0^\infty K(u) r_p(u) r_p(u)' \, \mathrm{d}u,$$
$$\vartheta_{p,q} = \int_0^\infty K(u) u^q r_p(u) \, \mathrm{d}u,$$
$$\Psi_p = \int_0^\infty K(u)^2 r_p(u) r_p(u)' \, \mathrm{d}u.$$

S.1.1.2. *Sharp RD Designs*

Recall the notation introduced in the paper. The estimand and estimators are

$$\tau_\nu = \mu_+^{(\nu)} - \mu_-^{(\nu)}, \quad \mu_+^{(\nu)} = \nu! e_\nu' \beta_{+,p}, \quad \mu_-^{(\nu)} = \nu! e_\nu' \beta_{-,p},$$
$$\hat{\tau}_{\nu,p}(h_n) = \hat{\mu}_{+,p}^{(\nu)}(h_n) - \hat{\mu}_{-,p}^{(\nu)}(h_n),$$
$$\hat{\mu}_{+,p}^{(\nu)}(h_n) = \nu! e_\nu' \hat{\beta}_{+,p}(h_n), \quad \hat{\mu}_{-,p}^{(\nu)}(h_n) = \nu! e_\nu' \hat{\beta}_{-,p}(h_n),$$

where, for any random variables $W$ and $Z$, and $s \in \mathbb{N}$,

$$\mu_{Z+}^{(s)} = \lim_{x \to 0^+} \frac{\partial^s}{\partial x^s} \mu_Z(x), \quad \mu_{Z-}^{(s)} = \lim_{x \to 0^-} \frac{\partial^s}{\partial x^s} \mu_Z(x),$$
$$\mu_Z(x) = \mathbb{E}[Z|X = x],$$
$$\sigma_{Z+}^2 = \lim_{x \to 0^+} \sigma_Z^2(x), \quad \sigma_{Z-}^2 = \lim_{x \to 0^-} \sigma_Z^2(x),$$
$$\sigma_Z^2(x) = \mathbb{V}[Z|X = x],$$
$$\sigma_{WZ+}^2 = \lim_{x \to 0^+} \sigma_{WZ}^2(x), \quad \sigma_{WZ-}^2 = \lim_{x \to 0^-} \sigma_{WZ}^2(x),$$
$$\sigma_{WZ}^2(x) = \mathbb{C}[W, Z|X = x].$$

In this setting, $Y_i = Y_i(0) \cdot \mathbf{1}(X_i < 0) + Y_i(1) \cdot \mathbf{1}(X_i \geq 0)$.

S.1.1.3. *Fuzzy RD Designs*

In this case, the treatment status $T_i$ is no longer a deterministic function of the forcing variable, but $\mathbb{P}[T_i = 1|X_i = x]$ changes discontinuously at the RD threshold level $\bar{x} = 0$. Here

$$Y_i = Y_i(0) \cdot (1 - T_i) + Y_i(1) \cdot T_i,$$
$$T_i = T_i(0) \cdot \mathbf{1}(X_i < 0) + T_i(1) \cdot \mathbf{1}(X_i \geq 0).$$

The estimand of interest is

$$\begin{aligned}
\varsigma_\nu &= \frac{\left.\dfrac{\mathrm{d}^\nu}{\mathrm{d}x^\nu}\mathbb{E}[Y(1)|X_i=x]\right|_{x=\bar{x}} - \left.\dfrac{\mathrm{d}^\nu}{\mathrm{d}x^\nu}\mathbb{E}[Y(0)|X_i=x]\right|_{x=\bar{x}}}{\left.\dfrac{\mathrm{d}^\nu}{\mathrm{d}x^\nu}\mathbb{E}[T(1)|X_i=x]\right|_{x=\bar{x}} - \left.\dfrac{\mathrm{d}^\nu}{\mathrm{d}x^\nu}\mathbb{E}[T(0)|X_i=x]\right|_{x=\bar{x}}} \\
&= \frac{\lim_{x\to 0^+}\mu_Y^{(\nu)}(x) - \lim_{x\to 0^-}\mu_Y^{(\nu)}(x)}{\lim_{x\to 0^+}\mu_T^{(\nu)}(x) - \lim_{x\to 0^-}\mu_T^{(\nu)}(x)}.
\end{aligned}$$

Thus,

$$\varsigma_\nu = \frac{\tau_{Y,\nu}}{\tau_{T,\nu}}, \quad \tau_{Y,\nu} = \mu_{Y+}^{(\nu)} - \mu_{Y-}^{(\nu)}, \quad \tau_{T,\nu} = \mu_{T+}^{(\nu)} - \mu_{T-}^{(\nu)}.$$

The plug-in local polynomial estimator is

$$\begin{aligned}
&\hat{\varsigma}_{\nu,p}(h_n) = \frac{\hat{\tau}_{Y,\nu,p}(h_n)}{\hat{\tau}_{T,\nu,p}(h_n)}, \\
&\hat{\tau}_{Y,\nu,p}(h_n) = \hat{\mu}_{Y+,p}^{(\nu)}(h_n) - \hat{\mu}_{Y-,p}^{(\nu)}(h_n), \\
&\hat{\tau}_{T,\nu,p}(h_n) = \hat{\mu}_{T+,p}^{(\nu)}(h_n) - \hat{\mu}_{T-,p}^{(\nu)}(h_n), \\
&\hat{\mu}_{Y+,p}^{(\nu)}(h_n) = \nu! e_\nu' \hat{\beta}_{Y+,p}(h_n), \quad \hat{\mu}_{Y-,p}^{(\nu)}(h_n) = \nu! e_\nu' \hat{\beta}_{Y-,p}(h_n), \\
&\hat{\mu}_{T+,p}^{(\nu)}(h_n) = \nu! e_\nu' \hat{\beta}_{T+,p}(h_n), \quad \hat{\mu}_{T-,p}^{(\nu)}(h_n) = \nu! e_\nu' \hat{\beta}_{T-,p}(h_n).
\end{aligned}$$

## S.2. DERIVATIONS, PROOFS AND FURTHER RESULTS

In the following results, we drop the notation for the dependent variable $(Y)$ for simplicity. All the results also apply to the dependent variable $T$ (fuzzy design) under Assumption 3.

### S.2.1. *Preliminary Lemmas*

The following lemma establishes convergence in probability of the sample matrices $\Gamma_{-,p}(h_n)$, $\vartheta_{-,p,q}(h_n)$, $\Psi_{-,p}(h_n)$ and $\Gamma_{+,p}(h_n)$, $\vartheta_{+,p,q}(h_n)$, $\Psi_{+,p}(h_n)$ to their expectation counterparts, and characterizes those limits.

LEMMA S.A.1: *Suppose Assumptions* 1–2 *hold*, *and* $nh_n \to \infty$.

(a) *If* $\kappa h_n < \kappa_0$, *then*:

(a.1) $\Gamma_{+,p}(h_n) = \tilde{\Gamma}_p(h_n) + o_p(1)$ *with* $\tilde{\Gamma}_{+,p}(h_n) = \int_0^\infty K(u) r_p(u) r_p(u)' \times f(uh_n)\,\mathrm{d}u \asymp \Gamma_p$,

(a.2) $\Gamma_{-,p}(h_n) = H_p(-1)\tilde{\Gamma}_p(h_n)H_p(-1) + o_p(1)$ *with* $\tilde{\Gamma}_{-,p}(h_n) = \int_0^\infty K(u) \times r_p(u) r_p(u)' f(-uh_n)\,\mathrm{d}u \asymp \Gamma_p$, (a.3) $\vartheta_{+,p,q}(h_n) = \tilde{\vartheta}_{+,p,q}(h_n) + o_p(1)$ *with* $\tilde{\vartheta}_{+,p,q}(h_n) = \int_0^\infty K(u) r_p(u) u^q \times f(uh_n)\,\mathrm{d}u \asymp \vartheta_{p,q}$,

(a.4) $\vartheta_{-,p,q}(h_n) = (-1)^q H_p(-1)\tilde{\vartheta}_{-,p,q}(h_n) + o_p(1)$ *with* $\tilde{\vartheta}_{-,p,q}(h_n) = \int_0^\infty K(u) r_p(u) u^q f(-uh_n)\,\mathrm{d}u \asymp \vartheta_{p,q}$,

(a.5) $h_n \Psi_{+,p}(h_n) = \tilde{\Psi}_{+,p}(h_n) + o_p(1)$ *with* $\tilde{\Psi}_{+,p}(h_n) = \int_0^\infty K(u)^2 r_p(u) \times r_p(u)' \sigma_+^2(uh_n) f(uh_n)\,\mathrm{d}u \asymp \Psi_p$,

(a.6) $h_n \Psi_{-,p}(h_n) = H_p(-1)\tilde{\Psi}_{-,p}(h_n) H_p(-1) + o_p(1)$ *with* $\tilde{\Psi}_{-,p}(h_n) = \int_0^\infty K(u)^2 r_p(u) r_p(u)' \sigma_-^2(-uh_n) f(-uh_n)\,\mathrm{d}u \asymp \Psi_p$.

(b) *If* $h_n \to 0$, *then*

(b.1) $\tilde{\Gamma}_{+,p}(h_n) = f\Gamma_p + o(1)$ *and* $\tilde{\Gamma}_{-,p}(h_n) = f\Gamma_p + o(1)$,

(b.2) $\tilde{\vartheta}_{+,p,q}(h_n) = f\vartheta_{p,q} + o(1)$ *and* $\tilde{\vartheta}_{-,p,q}(h_n) = f\vartheta_{p,q} + o(1)$,

(b.3) $\tilde{\Psi}_{+,p}(h_n) = \sigma_+^2 f \Psi_p + o(1)$ *and* $\tilde{\Psi}_{-,p}(h_n) = \sigma_-^2 f \Psi_p + o(1)$.

PROOF: First, for $\Gamma_{+,p}(h_n)$, change of variables implies

$$\begin{aligned}
\mathbb{E}\big[\Gamma_{+,p}(h_n)\big] &= \mathbb{E}\left[\frac{1}{nh_n}\sum_{i=1}^{n} \mathbf{1}(X_i \geq 0) K\left(\frac{X_i}{h_n}\right) r_p\left(\frac{X_i}{h_n}\right) r_p\left(\frac{X_i}{h_n}\right)'\right] \\
&= \frac{1}{h_n}\int_0^\infty K\left(\frac{x}{h_n}\right) r_p\left(\frac{x}{h_n}\right) r_p\left(\frac{x}{h_n}\right)' f(x)\,\mathrm{d}x \\
&= \int_0^\infty K(u) r_p(u) r_p(u)' f(uh_n)\,\mathrm{d}u = \tilde{\Gamma}_p(h_n),
\end{aligned}$$

and, provided $\kappa h_n < \kappa_0$,

$$\begin{aligned}
&\mathbb{E}\big[\big|\Gamma_{+,p}(h_n) - \mathbb{E}\big[\Gamma_{+,p}(h_n)\big]\big|^2\big] \\
&\quad \lesssim \frac{1}{h_n^2}\mathbb{E}\left[\left|\mathbf{1}(X_i \geq 0) K\left(\frac{X_i}{h_n}\right) r_p\left(\frac{X_i}{h_n}\right) r_p\left(\frac{X_i}{h_n}\right)'\right|^2\right] \\
&\quad = \frac{1}{h_n}\int_0^\infty K(u)^2 \big|r_p(u)\big|^4 f(uh_n)\,\mathrm{d}u = O\big(n^{-1}h_n^{-1}\big) = o(1).
\end{aligned}$$

Thus, using Markov Inequality, $\Gamma_{+,p}(h_n) = \tilde{\Gamma}_{+,p}(h_n) + o_p(1)$. If $\kappa h_n < \kappa_0$, $\tilde{\Gamma}_{+,p}(h_n) \asymp \Gamma_p$ because the density is bounded and bounded away from zero, which verifies part (a.1). The proof of part (a.2) is similar, but note that

$$\begin{aligned}
\mathbb{E}\big[\Gamma_{-,p}(h_n)\big] &= \mathbb{E}\left[\frac{1}{nh_n}\sum_{i=1}^{n} \mathbf{1}(X_i < 0) K\left(\frac{X_i}{h_n}\right) r_p\left(\frac{X_i}{h_n}\right) r_p\left(\frac{X_i}{h_n}\right)'\right] \\
&= \frac{1}{h_n}\int_{-\infty}^0 K\left(\frac{x}{h_n}\right) r_p\left(\frac{x}{h_n}\right) r_p\left(\frac{x}{h_n}\right)' f(x)\,\mathrm{d}x
\end{aligned}$$

$$= \int_0^\infty K(-u) r_p(-u) r_p(-u)' f(-uh_n)\, \mathrm{d}u$$

$$= H_p(-1) \int_0^\infty K(u) r_p(u) r_p(u)' f(-uh_n)\, \mathrm{d}u H_p(-1)$$

$$= H_p(-1) \tilde{\Gamma}_{-,p}(h_n) H_p(-1),$$

because $K(-u) = K(u)$ and $r_p(-u) = H_p(-1) r_p(u)$. Also, note that $\tilde{\Gamma}_{+,p}(h_n) = f\Gamma_p + o(1)$ and $\tilde{\Gamma}_{-,p}(h_n) = f\Gamma_p + o(1)$ if $h_n \to 0$, by continuity of $f(x)$, which proves part (b.1).

For $\vartheta_{+,p,q}(h_n)$, we have

$$\mathbb{E}\big[\vartheta_{+,p,q}(h_n)\big] = \frac{1}{h_n} \int_0^\infty K\left(\frac{x}{h_n}\right) r_p\left(\frac{x}{h_n}\right) \left(\frac{x}{h_n}\right)^q f(x)\, \mathrm{d}x$$

$$= \int_0^\infty K(u) r_p(u) u^q f(uh_n)\, \mathrm{d}u = \tilde{\vartheta}_{+,p,q}(h_n),$$

and, provided $\kappa h_n < \kappa_0$,

$$\mathbb{E}\big[\big|\vartheta_{+,p,q}(h_n) - \mathbb{E}\big[\vartheta_{+,p,q}(h_n)\big]\big|^2\big]$$

$$= \frac{1}{nh_n} \int_0^\infty K(u)^2 \big|r_p(u)\big|^2 |u|^{2q} f(uh_n)\, \mathrm{d}u = O\big(n^{-1} h_n^{-1}\big) = o(1).$$

Thus, part (a.3) is verified. Similarly, as above,

$$\mathbb{E}\big[\vartheta_{-,p,q}(h_n)\big] = \frac{1}{h_n} \int_{-\infty}^0 K\left(\frac{x}{h_n}\right) r_p\left(\frac{x}{h_n}\right) \left(\frac{x}{h_n}\right)^q f(x)\, \mathrm{d}x$$

$$= \int_0^\infty K(-u) r_p(-u) (-u)^q f(-uh_n)\, \mathrm{d}u$$

$$= (-1)^q H_p(-1) \tilde{\vartheta}_{+,p,q}(h_n),$$

and $\tilde{\vartheta}_{+,p,q}(h_n) = f\vartheta_{p,q} + o(1)$ and $\tilde{\vartheta}_{-,p,q}(h_n) = f\vartheta_{p,q} + o(1)$, which gives parts (a.4) and (b.2).

Finally, for part (a.5), as above,

$$\mathbb{E}\big[h_n \Psi_{+,p}(h_n)\big] = \int_0^\infty K(u)^2 r_p(u) r_p(u)' \sigma_+^2(uh_n) f(uh_n)\, \mathrm{d}u$$

$$= \tilde{\Psi}_{+,p}(h_n),$$

and $h_n^2\mathbb{E}[|\Psi_{+,p}(h_n) - \mathbb{E}[\Psi_{+,p}(h_n)]|^2] = n^{-1}h_n^{-1}\int_0^\infty K(u)^4|r_p(u)|^4 f(uh_n)\,\mathrm{d}u = O(n^{-1}h_n^{-1})$, provided $\kappa h_n < \kappa_0$. For part (a.6),

$$\mathbb{E}\big[h_n\Psi_{-,p}(h_n)\big] = h_n^{-1}\int_{-\infty}^0 K(u/h_n)^2 r_p(u/h_n)r_p(u/h_n)'\sigma_-^2(u)f(u)\,\mathrm{d}u$$
$$= H_p(-1)\tilde{\Psi}_{-,p}(h_n)H_p(-1),$$

and the rest is proven as above. Part (b.6) is also verified by continuity of $\sigma_+^2(u)$, $\sigma_-^2(u)$ and $f(u)$. *Q.E.D.*

The next lemma establishes convergence in probability of the sample matrix $\Psi_{+,p,q}(h_n, b_n)$ to its population counterpart, and characterizes this limit.

LEMMA S.A.2: *Suppose Assumptions* 1–2 *hold. Let* $m_n = \min\{h_n, b_n\}$. *If* $m_n \to 0$ *and* $nm_n \to \infty$, *then*

$$\frac{h_n b_n}{m_n}\Psi_{+,p,q}(h_n, b_n) = \tilde{\Psi}_{+,p,q}(h_n, b_n) + o_p(1),$$
$$\tilde{\Psi}_{+,p,q}(h_n, b_n) = \int_0^\infty K\left(\frac{m_n u}{h_n}\right)K\left(\frac{m_n u}{b_n}\right) \times r_p\left(\frac{m_n u}{h_n}\right)r_q\left(\frac{m_n u}{b_n}\right)'\sigma^2(um_n)f(um_n)\,\mathrm{d}u,$$

*and*

$$\frac{h_n b_n}{m_n}\Psi_{-,p,q}(h_n, b_n) = \tilde{\Psi}_{-,p,q}(h_n, b_n) + o_p(1),$$
$$\tilde{\Psi}_{-,p,q}(h_n, b_n) = \int_{-\infty}^0 K\left(\frac{m_n u}{h_n}\right)K\left(\frac{m_n u}{b_n}\right) \times r_p\left(\frac{m_n u}{h_n}\right)r_q\left(\frac{m_n u}{b_n}\right)'\sigma^2(um_n)f(um_n)\,\mathrm{d}u.$$

PROOF: First, change of variables gives

$$\mathbb{E}\left[\frac{h_n b_n}{m_n}\Psi_{+,p,q}(h_n, b_n)\right]$$
$$= \frac{1}{m_n}\int_0^\infty K\left(\frac{x}{h_n}\right)K\left(\frac{x}{b_n}\right)r_p\left(\frac{x}{h_n}\right)r_q\left(\frac{x}{b_n}\right)'\sigma^2(x)f(x)\,\mathrm{d}x$$

$$
\begin{aligned}
&= \int_0^\infty K\big(h_n^{-1}m_n u\big)K\big(b_n^{-1}m_n u\big) \\
&\quad \times r_p\big(h_n^{-1}m_n u\big)r_q\big(b_n^{-1}m_n u\big)'\sigma^2(um_n)f(um_n)\,\mathrm{d}u \\
&= \tilde{\Psi}_{+,p,q}(h_n, b_n),
\end{aligned}
$$

which gives the second conclusion. Next, we also have

$$
\begin{aligned}
&\mathbb{E}\left[\left|\frac{h_n b_n}{m_n}\Psi_{+,p,q}(h_n, b_n) - \mathbb{E}\left[\frac{h_n b_n}{m_n}\Psi_{+,p,q}(h_n, b_n)\right]\right|^2\right] \\
&= \frac{1}{nm_n}\int_0^\infty K\big(h_n^{-1}m_n u\big)^2 K\big(b_n^{-1}m_n u\big)^2 \\
&\quad \times \big|r_p\big(h_n^{-1}m_n u\big)\big|^2\big|r_q\big(b_n^{-1}m_n u\big)\big|^2\sigma^2(um_n)f(um_n)\,\mathrm{d}x \\
&= O\left(\frac{1}{nm_n}\right) = o(1),
\end{aligned}
$$

and the first result follows by Markov Inequality.

The proof of $\Psi_{-,p,q}(h_n, b_n)$ is analogous. *Q.E.D.*

Let $s, \ell \in \mathbb{N}$ with $s \le \ell$. The following lemma gives the asymptotic bias, variance, and distribution for the $\ell$th-order local polynomial estimator of $\mu_+^{(s)}$ and $\mu_-^{(s)}$:

$$
\begin{aligned}
&\hat{\mu}_{+,\ell}^{(s)}(h_n) = s!e_s'\hat{\beta}_{+,\ell}(h_n), \\
&\hat{\beta}_{+,\ell}(h_n) = H_\ell(h_n)\Gamma_{+,\ell}^{-1}(h_n)X_\ell(h_n)'W_+(h_n)Y/n, \\
&\hat{\mu}_{-,\ell}^{(s)}(h_n) = s!e_s'\hat{\beta}_{-,\ell}(h_n), \\
&\hat{\beta}_{-,\ell}(h_n) = H_\ell(h_n)\Gamma_{-,\ell}^{-1}(h_n)X_\ell(h_n)'W_-(h_n)Y/n.
\end{aligned}
$$

LEMMA S.A.3: *Suppose Assumptions* 1–2 *hold with* $S \ge \ell + 2$, *and* $nh_n \to \infty$.

(B) *If* $h_n \to 0$, *then*

$$
\begin{aligned}
\mathbb{E}\big[\hat{\mu}_{+,\ell}^{(s)}(h_n)|\mathcal{X}_n\big] &= s!e_s'\beta_{+,\ell} + h_n^{1+\ell-s}\frac{\mu_+^{(\ell+1)}}{(\ell+1)!}\mathcal{B}_{+,s,\ell,\ell+1}(h_n) \\
&\quad + h_n^{2+\ell-s}\frac{\mu_+^{(\ell+2)}}{(\ell+2)!}\mathcal{B}_{+,s,\ell,\ell+2}(h_n) + o_p\big(h_n^{2+\ell-s}\big), \\
\mathcal{B}_{+,s,\ell,r}(h_n) &= s!e_s'\Gamma_{+,\ell}^{-1}(h_n)\vartheta_{+,\ell,r}(h_n) = s!e_s'\Gamma_\ell^{-1}\vartheta_{\ell,r} + o_p(1),
\end{aligned}
$$

*and*

$$\mathbb{E}\big[\hat{\mu}^{(s)}_{-,\ell}(h_n)|\mathcal{X}_n\big] = s!e_s'\beta_{-,\ell} + h_n^{1+\ell-s}\frac{\mu_-^{(\ell+1)}}{(\ell+1)!}\mathcal{B}_{-,s,\ell,\ell+1}(h_n) + h_n^{2+\ell-s}\frac{\mu_-^{(\ell+2)}}{(\ell+2)!}\mathcal{B}_{-,s,\ell,\ell+2}(h_n) + o_p\big(h_n^{2+\ell-s}\big),$$

$$\mathcal{B}_{-,s,\ell,r}(h_n) = s!e_s'\Gamma^{-1}_{-,\ell}(h_n)\vartheta_{-,\ell,r}(h_n) = (-1)^{s+r}s!e_s'\Gamma_\ell^{-1}\vartheta_{\ell,r} + o_p(1).$$

(V) *If* $h_n \to 0$, *then* $\mathbb{V}[\hat{\mu}^{(s)}_{+,\ell}(h_n)|\mathcal{X}_n] = \mathcal{V}_{+,s,\ell}(h_n)$ *with*

$$\mathcal{V}_{+,s,\ell}(h_n) = \frac{1}{nh_n^{2s}}s!^2e_s'\Gamma^{-1}_{+,\ell}(h_n)\Psi_{+,\ell}(h_n)\Gamma^{-1}_{+,\ell}(h_n)e_s = \frac{1}{nh_n^{1+2s}}\frac{\sigma_+^2}{f}s!^2e_s'\Gamma_\ell^{-1}\Psi_\ell\Gamma_\ell^{-1}e_s\big[1+o_p(1)\big],$$

*and* $\mathbb{V}[\hat{\mu}^{(s)}_{-,\ell}(h_n)|\mathcal{X}_n] = \mathcal{V}_{-,s,\ell}(h_n)$ *with*

$$\mathcal{V}_{-,s,\ell}(h_n) = \frac{1}{nh_n^{2s}}s!^2e_s'\Gamma^{-1}_{-,\ell}(h_n)\Psi_{-,\ell}(h_n)\Gamma^{-1}_{-,\ell}(h_n)e_s = \frac{1}{nh_n^{1+2s}}\frac{\sigma_-^2}{f}s!^2e_s'\Gamma_\ell^{-1}\Psi_\ell\Gamma_\ell^{-1}e_s\big[1+o_p(1)\big].$$

(D) *If* $nh_n^{2\ell+5} \to 0$, *then*

$$\frac{\hat{\mu}^{(s)}_{+,\ell}(h_n) - \mu_+^{(s)} - h_n^{1+\ell-s}\frac{\mu_+^{(\ell+1)}}{(\ell+1)!}\mathcal{B}_{+,s,\ell,\ell+1}(h_n)}{\sqrt{\mathcal{V}_{+,s,\ell}(h_n)}} \to_d \mathcal{N}(0,1)$$

*and*

$$\frac{\hat{\mu}^{(s)}_{-,\ell}(h_n) - \mu_-^{(s)} - h_n^{1+\ell-s}\frac{\mu_-^{(\ell+1)}}{(\ell+1)!}\mathcal{B}_{-,s,\ell,\ell+1}(h_n)}{\sqrt{\mathcal{V}_{-,s,\ell}(h_n)}} \to_d \mathcal{N}(0,1).$$

PROOF: For part (B), a Taylor series expansion gives

$$\mathbb{E}\big[s!\hat{\beta}_{+,\ell}(h_n)|\mathcal{X}_n\big] = s!\beta_{+,\ell} + h_n^{\ell+1}H_\ell(h_n)\Gamma^{-1}_{+,\ell}(h_n)X_\ell(h_n)W_+(h_n)S_{\ell+1}(h_n)s!\frac{\mu_+^{(\ell+1)}}{(\ell+1)!} + h_n^{\ell+2}H_\ell(h_n)\Gamma^{-1}_{+,\ell}(h_n)X_\ell(h_n)W_+(h_n)S_{\ell+2}(h_n)s!\frac{\mu_+^{(\ell+2)}}{(\ell+2)!}$$

$$
\begin{aligned}
&+ H_\ell(h_n) o_p\big(h_n^{\ell+2}\big) \\
&= s!\beta_{+,\ell} + h_n^{\ell+1} H_\ell(h_n) s! \frac{\mu_+^{(\ell+1)}}{(\ell+1)!} \Gamma_{+,\ell}^{-1}(h_n) \vartheta_{+,\ell,\ell+1}(h_n) \\
&\quad + h_n^{\ell+2} H_\ell(h_n) s! \frac{\mu_+^{(\ell+2)}}{(\ell+2)!} \Gamma_{+,\ell}^{-1}(h_n) \vartheta_{+,\ell,\ell+2}(h_n) \\
&\quad + H_\ell(h_n) o_p\big(h_n^{\ell+2}\big),
\end{aligned}
$$

and the result for $\mathbb{E}[\hat{\mu}_{+,\ell}^{(s)}(h_n)|\mathcal{X}_n]$ follows by $e_s' H_\ell(h_n) = h_n^{-s}$ and Lemma S.A.1. Next, for $\mathbb{E}[\hat{\mu}_{-,\ell}^{(s)}(h_n)|\mathcal{X}_n]$ the same calculations apply, with only a modification for $\mathcal{B}_{-,s,\ell,r}(h_n)$ because, by Lemma S.A.1,

$$
\begin{aligned}
\mathcal{B}_{-,s,\ell,r}(h_n) &= s! e_s' \Gamma_{-,\ell}^{-1}(h_n) \vartheta_{-,\ell,r}(h_n) \\
&= s! e_s' \big[H_\ell(-1) \tilde{\Gamma}_{-,\ell}^{-1}(h_n) H_\ell(-1)\big]\big[(-1)^r H_\ell(-1) \vartheta_{-,\ell,r}(h_n)\big] \\
&\quad + o_p(1) \\
&= (-1)^{s+r} s! e_s' \tilde{\Gamma}_{-,\ell}^{-1}(h_n) \vartheta_{-,\ell,r}(h_n) + o_p(1),
\end{aligned}
$$

because $e_s' H_\ell(-1) = (-1)^s$ and $H_\ell(-1) H_\ell(-1) = I_{\ell+1}$.

For part (V), simply note that

$$
\begin{aligned}
\mathbb{V}\big[s! e_s' \hat{\beta}_{+,\ell}(h_n)|\mathcal{X}_n\big] &= s!^2 e_s' H_\ell(h_n) \Gamma_{+,\ell}^{-1}(h_n) X_\ell(h_n) W_+(h_n) \Sigma \\
&\quad \times W_+(h_n) X_\ell(h_n) \Gamma_{+,\ell}^{-1}(h_n) H_\ell(h_n) e_s / n \\
&= h_n^{-2s} s!^2 e_s' \Gamma_{+,\ell}^{-1}(h_n) X_\ell(h_n) W_+(h_n) \Sigma \\
&\quad \times W_+(h_n) X_\ell(h_n) \Gamma_{+,\ell}^{-1}(h_n) e_s / n \\
&= n^{-1} h_n^{-2s} s!^2 e_s' \Gamma_{+,\ell}^{-1}(h_n) \Psi_{+,\ell}(h_n) \Gamma_{+,\ell}^{-1}(h_n) e_s \\
&= \mathcal{V}_{+,\ell,s}(h_n),
\end{aligned}
$$

and the result follows by Lemma S.A.1. The proof of $\mathbb{V}[s! e_s' \hat{\beta}_{-,\ell}(h_n)|\mathcal{X}_n]$ is analogous.

For part (D), using the previous results, we have

$$
\begin{aligned}
&\frac{\hat{\mu}_{+,\ell}^{(s)}(h_n) - \mu_+^{(s)} - h_n^{1+\ell-s} \frac{\mu_+^{(\ell+1)}}{(\ell+1)!} \mathcal{B}_{+,s,\ell,\ell+1}(h_n)}{\sqrt{\mathcal{V}_{+,s,\ell}(h_n)}} \\
&\quad = \frac{s! e_s' \hat{\beta}_{+,\ell}(h_n) - s! e_s' \beta_{+,\ell} - h_n^{1+\ell-s} \frac{\mu_+^{(\ell+1)}}{(\ell+1)!} \mathcal{B}_{+,s,\ell,\ell+1}(h_n)}{\sqrt{\mathcal{V}_{+,s,\ell}(h_n)}}
\end{aligned}
$$

$$= \xi_{1,n} + \xi_{2,n}$$
$$= \xi_{1,n} + o_p(1),$$

where

$$\begin{aligned}
\xi_{1,n} &= \big(\mathcal{V}_{+,s,\ell}(h_n)\big)^{-1/2} s! e_s' \big(\hat{\beta}_{+,\ell}(h_n) - \mathbb{E}\big[\hat{\beta}_{+,\ell}(h_n)|\mathcal{X}_n\big]\big) \\
&= \big(\mathcal{V}_{+,s,\ell}(h_n)\big)^{-1/2} s! e_s' H_\ell(h_n) \Gamma_{+,\ell}^{-1}(h_n) X_\ell(h_n) W_+(h_n)\varepsilon/n, \\
\xi_{2,n} &= \big(\mathcal{V}_{+,s,\ell}(h_n)\big)^{-1/2} \\
&\quad \times \big(\mathbb{E}\big[s! e_s' \hat{\beta}_{+,\ell}(h_n)|\mathcal{X}_n\big] - s! e_s' \beta_{+,\ell} - h_n^{1+\ell-s} \mu_+^{(\ell)} \mathcal{B}_{+,s,\ell,\ell+1}(h_n)/\ell!\big) \\
&= O_p\Big(\sqrt{n h_n^{1+2s}}\Big) O_p\big(h_n^{2+\ell-s}\big) = O_p\Big(\sqrt{n h_n^{5+2\ell}}\Big) = o_p(1),
\end{aligned}$$

under the conditions imposed. Next, note that by Lemma S.A.1, $\xi_{1,n} = \tilde{\xi}_{1,n} + o_p(1)$ with

$$\tilde{\xi}_{1,n} = \sum_{i=1}^{n} \omega_{n,i}\varepsilon_i,$$

$$\omega_{n,i} = \left(\frac{1}{n h_n^{1+2s}} e_s' \Gamma_{+,\ell}^{-1} \Psi_{+,\ell} \Gamma_{+,\ell}^{-1} e_s\right)^{-1/2} h_n^{-s} e_s' \Gamma_{+,\ell}^{-1} K_{h_n}(X_i) r_\ell(X_i/h_n)/n,$$

where $\{\omega_{n,i}\varepsilon_i : 1 \leq i \leq n\}$ is a triangular array of independent random variables with $\mathbb{E}[\tilde{\xi}_{1,n}] = 0$ and $\mathbb{V}[\tilde{\xi}_{1,n}] \to 1$. Thus, $\tilde{\xi}_{1,n} \to_d \mathcal{N}(0,1)$ by the Linderberg–Feller central limit theorem for triangular arrays because

$$\begin{aligned}
\sum_{i=1}^{n} \mathbb{E}\big[|\omega_{n,i}\varepsilon_i|^4\big] &\lesssim n^2 h_n^{2+4s} h_n^{-4s} \sum_{i=1}^{n} \mathbb{E}\big[\big|e_s' \Gamma_{+,\ell}^{-1} K_{h_n}(X_i) r_\ell(X_i/h_n)\big|^4\big]/n^4 \\
&\lesssim n^{-1} h_n^{-2} \int_0^\infty \big|K(x/h_n) e_s' \Gamma_{+,\ell}^{-1} r_\ell(x/h_n)\big|^4 f(x)\,\mathrm{d}x \\
&= O\left(\frac{1}{n h_n}\right) = o(1).
\end{aligned}$$

The result for $\hat{\mu}_{-,\ell}^{(s)}(h_n)$ can be established the same way. This concludes the proof. *Q.E.D.*

Let $\nu, p, q \in \mathbb{N}$ with $\nu \leq p < q$. The final preliminary lemma gives the asymptotic bias, variance, and distribution for the $p$th-order local polynomial estimator of $\mu_+^{(\nu)}$ and $\mu_-^{(\nu)}$ with bias correction constructed using a $q$th-order local polynomial:

$$\hat{\mu}^{(\nu)\text{bc}}_{+,p,q}(h_n, b_n) = \nu! e'_\nu \hat{\beta}_{+,p}(h_n) - h_n^{p+1-\nu}\big(e'_{p+1}\hat{\beta}_{+,q}(b_n)\big)\mathcal{B}_{+,\nu,p,p+1}(h_n),$$
$$\hat{\mu}^{(\nu)\text{bc}}_{-,p,q}(h_n, b_n) = \nu! e'_\nu \hat{\beta}_{-,p}(h_n) - h_n^{p+1-\nu}\big(e'_{p+1}\hat{\beta}_{-,q}(b_n)\big)\mathcal{B}_{-,\nu,p,p+1}(h_n).$$

LEMMA S.A.4: *Suppose Assumptions* 1–2 *hold with* $S \geq q+1$, *and* $n\min\{h_n, b_n\} \to \infty$.

(B) *If* $\max\{h_n, b_n\} \to 0$, *then*

$$\begin{aligned}
&\mathbb{E}\big[\hat{\mu}^{(\nu)\text{bc}}_{+,p,q}(h_n, b_n)|\mathcal{X}_n\big] \\
&\quad = \nu! e'_\nu \beta_{+,p} + h_n^{2+p-\nu}\frac{\mu_+^{(p+2)}}{(p+2)!}\mathcal{B}_{+,\nu,p,p+2}(h_n)\big\{1+o_p(1)\big\} \\
&\qquad - h_n^{p+1-\nu} b_n^{q-p}\frac{\mu_+^{(q+1)}}{(q+1)!}\mathcal{B}_{+,p+1,q,q+1}(b_n)\frac{\mathcal{B}_{+,\nu,p,p+1}(h_n)}{(p+1)!}\big\{1+o_p(1)\big\}
\end{aligned}$$

*and*

$$\begin{aligned}
&\mathbb{E}\big[\hat{\mu}^{(\nu)\text{bc}}_{-,p,q}(h_n, b_n)|\mathcal{X}_n\big] \\
&\quad = \nu! e'_\nu \beta_{-,p} + h_n^{2+p-\nu}\frac{\mu_-^{(p+2)}}{(p+2)!}\mathcal{B}_{-,\nu,p,p+2}(h_n)\big\{1+o_p(1)\big\} \\
&\qquad - h_n^{p+1-\nu} b_n^{q-p}\frac{\mu_-^{(q+1)}}{(q+1)!}\mathcal{B}_{-,p+1,q,q+1}(b_n)\frac{\mathcal{B}_{-,\nu,p,p+1}(h_n)}{(p+1)!}\big\{1+o_p(1)\big\}.
\end{aligned}$$

(V) *If* $n\min\{h_n, b_n\} \to \infty$, *then* $\mathbb{V}[\hat{\mu}^{(\nu)\text{bc}}_{+,p,q}(h_n, b_n)|\mathcal{X}_n] = \mathcal{V}^{\text{bc}}_{+,\nu,p,q}(h_n, b_n)$, *where*

$$\begin{aligned}
\mathcal{V}^{\text{bc}}_{+,\nu,p,q}(h, b) &= \mathcal{V}_{+,\nu,p}(h) + h_n^{2(p+1-\nu)}\mathcal{V}_{+,p+1,q}(b)\frac{\mathcal{B}_{+,\nu,p,p+1}(h)^2}{(p+1)!^2} \\
&\quad - 2h^{p+1-\nu}\mathcal{C}_{+,\nu,p,q}(h, b)\frac{\mathcal{B}_{+,\nu,p,p+1}(h)}{(p+1)!}, \\
\mathcal{C}_{+,\nu,p,q}(h, b) &= \frac{1}{nh^\nu b^{p+1}}\nu!(p+1)! e'_\nu \Gamma^{-1}_{+,p}(h)\Psi_{+,p,q}(h, b)\Gamma^{-1}_{+,q}(b) e_p,
\end{aligned}$$

*and* $\mathbb{V}[\hat{\mu}^{(\nu)\text{bc}}_{-,p,q}(h_n, b_n)|\mathcal{X}_n] = \mathcal{V}^{\text{bc}}_{-,\nu,p,q}(h_n, b_n)$, *where*

$$\begin{aligned}
\mathcal{V}^{\text{bc}}_{-,\nu,p,q}(h, b) &= \mathcal{V}_{-,\nu,p}(h) + h_n^{2(p+1-\nu)}\mathcal{V}_{-,p+1,q}(b)\frac{\mathcal{B}_{-,\nu,p,p+1}(h)^2}{(p+1)!^2} \\
&\quad - 2h^{p+1-\nu}\mathcal{C}_{-,\nu,p,q}(h, b)\frac{\mathcal{B}_{-,\nu,p,p+1}(h)}{(p+1)!},
\end{aligned}$$

$$\mathcal{C}_{-,\nu,p,q}(h,b) = \frac{1}{nh^{\nu}b^{p+1}}\nu!(p+1)!e_{\nu}'\Gamma_{-,p}^{-1}(h)\Psi_{-,p,q}(h,b)\Gamma_{-,q}^{-1}(b)e_p.$$

(D) *If* $n\min\{h_n^{2p+3}, b_n^{2p+3}\}\max\{h_n^2, b_n^{2(q-p)}\} \to 0$, *and* $\kappa\max\{h_n, b_n\} < \kappa_0$, *then*

$$\frac{\hat{\mu}_{+,p,q}^{(\nu)\mathrm{bc}}(h_n,b_n) - \nu!e_{\nu}'\beta_{+,p}}{\sqrt{\mathcal{V}_{+,\nu,p,q}^{\mathrm{bc}}(h_n,b_n)}} \to_d \mathcal{N}(0,1)$$

*and*

$$\frac{\hat{\mu}_{-,p,q}^{(\nu)\mathrm{bc}}(h_n,b_n) - \nu!e_{\nu}'\beta_{-,p}}{\sqrt{\mathcal{V}_{-,\nu,p,q}^{\mathrm{bc}}(h_n,b_n)}} \to_d \mathcal{N}(0,1).$$

PROOF: We only give the proof for the treatment group (subindex "+") because the proof for the control group (subindex "−") is analogous. For part (B), first note that $\mathbb{E}[\hat{\mu}_{+,p,q}^{(\nu)\mathrm{bc}}(h_n,b_n)|\mathcal{X}_n] = B_1 - B_2$ with $B_1 = \mathbb{E}[\nu!e_{\nu}'\hat{\beta}_{+,p}(h_n)|\mathcal{X}_n]$ and $B_2 = h_n^{p+1-\nu}\mathbb{E}[e_{p+1}'\hat{\beta}_{+,q}(b_n)|\mathcal{X}_n]\mathcal{B}_{+,\nu,p}(h_n)$. By Lemma S.A.3, with $s=\nu$ and $\ell = p$, we have

$$\begin{aligned} B_1 &= \nu!e_{\nu}'\beta_{+,p} + h_n^{1+p-\nu}\frac{\mu_+^{(p+1)}}{(p+1)!}\mathcal{B}_{+,\nu,p,p+1}(h_n) \\ &\quad + h_n^{2+p-\nu}\frac{\mu_+^{(p+2)}}{(p+2)!}\mathcal{B}_{+,\nu,p,p+2}(h_n) + o_p\big(h_n^{2+p-\nu}\big). \end{aligned}$$

Similarly, by Lemma S.A.3, with $s = p+1$ and $\ell = q$, we have

$$\begin{aligned} &\mathbb{E}\big[(p+1)!e_{p+1}'\hat{\beta}_{+,q}(b_n)|\mathcal{X}_n\big] \\ &\quad = (p+1)!e_{p+1}'\beta_{+,q} + b_n^{q-p}\frac{\mu_+^{(q+1)}}{(q+1)!}\mathcal{B}_{+,p+1,q,q+1}(b_n) + o_p\big(b_n^{q-p}\big), \end{aligned}$$

and hence

$$\begin{aligned} B_2 &= h_n^{p+1-\nu}\mathbb{E}\big[(p+1)!e_{p+1}'\hat{\beta}_{+,q}(b_n)|\mathcal{X}_n\big]\frac{\mathcal{B}_{+,\nu,p,p+1}(h_n)}{(p+1)!} \\ &= h_n^{p+1-\nu}\big(e_{p+1}'\beta_{+,q}\big)\mathcal{B}_{+,\nu,p,p+1}(h_n) \\ &\quad + h_n^{p+1-\nu}b_n^{q-p}\frac{\mu_+^{(q+1)}}{(q+1)!}\mathcal{B}_{+,p+1,q,q+1}(b_n)\frac{\mathcal{B}_{+,\nu,p,p+1}(h_n)}{(p+1)!} \\ &\quad + h_n^{p+1-\nu}o_p\big(b_n^{q-p}\big)\mathcal{B}_{+,\nu,p,p+1}(h_n). \end{aligned}$$

Collecting terms, the result in part (B) follows:

$$\begin{aligned}
&\mathbb{E}\big[\nu! e_\nu' \hat{\beta}^{\mathrm{bc}}_{+,p,q}(h_n, b_n)|\mathcal{X}_n\big] \\
&\quad = \nu! e_\nu' \beta_{+,p} + h_n^{2+p-\nu} \frac{\mu_+^{(p+2)}}{(p+2)!} \mathcal{B}_{+,\nu,p,p+2}(h_n)\big\{1 + o_p(1)\big\} \\
&\qquad - h_n^{p+1-\nu} b_n^{q-p} \frac{\mu_+^{(q+1)}}{(q+1)!} \mathcal{B}_{+,p+1,q,q+1}(b_n) \frac{\mathcal{B}_{+,\nu,p}(h_n)}{(p+1)!}\big\{1 + o_p(1)\big\}.
\end{aligned}$$

For part (V), first note that $\mathbb{V}[\hat{\mu}^{(\nu)\mathrm{bc}}_{+,p,q}(h_n, b_n)|\mathcal{X}_n] = V_1 + V_2 - 2C_{12}$ where, using Lemma S.A.3 with $s = \nu$ and $\ell = p$,

$$V_1 = \mathbb{V}\big[\nu! e_\nu' \hat{\beta}_{+,p}(h_n)|\mathcal{X}_n\big] = \mathbb{V}\big[\hat{\mu}^{(\nu)}_{+,p}(h_n)|\mathcal{X}_n\big] = \mathcal{V}_{+,\nu,p}(h_n),$$

and, using Lemma S.A.3 with $s = p + 1$ and $\ell = q$,

$$\begin{aligned}
V_2 &= \mathbb{V}\big[h_n^{p+1-\nu}\big(e_{p+1}' \hat{\beta}_{+,q}(b_n)\big)\mathcal{B}_{+,\nu,p,p+1}(h_n)|\mathcal{X}_n\big] \\
&= h_n^{2(p+1-\nu)} \mathbb{V}\big[(p+1)! e_{p+1}' \hat{\beta}_{+,q}(b_n)|\mathcal{X}_n\big] \frac{\mathcal{B}_{+,\nu,p,p+1}(h_n)^2}{(p+1)!^2} \\
&= h_n^{2(p+1-\nu)} \mathcal{V}_{+,p+1,q}(b_n) \frac{\mathcal{B}_{+,\nu,p,p+1}(h_n)^2}{(p+1)!^2},
\end{aligned}$$

and

$$\begin{aligned}
C_{12} &= \mathbb{C}\big[\nu! e_\nu' \hat{\beta}_{+,p}(h_n), h_n^{p+1-\nu}\big(e_{p+1}' \hat{\beta}_{+,q}(b_n)\big)\mathcal{B}_{+,\nu,p,p+1}(h_n)|\mathcal{X}_n\big] \\
&= h_n^{p+1-\nu} \mathbb{C}\big[\nu! e_\nu' \hat{\beta}_{+,p}(h_n), (p+1)! e_{p+1}' \hat{\beta}_{+,q}(b_n)|\mathcal{X}_n\big] \frac{\mathcal{B}_{+,\nu,p,p+1}(h_n)}{(p+1)!}
\end{aligned}$$

with

$$\begin{aligned}
&\mathbb{C}\big[e_\nu' \hat{\beta}_{+,p}(h_n), e_{p+1}' \hat{\beta}_{+,q}(b_n)|\mathcal{X}_n\big] \\
&\quad = h_n^{-\nu} e_\nu' \Gamma_{+,p}^{-1}(h_n) X_p(h_n) W_+(h_n) \mathbb{C}[Y, Y|\mathcal{X}_n] \\
&\qquad \times W_+(b_n) X_q(b_n) \Gamma_{+,q}^{-1}(b_n) e_{p+1} b_n^{-p-1}/n^2 \\
&\quad = \frac{1}{n h_n^\nu b_n^{p+1}} \nu! (p+1)! e_\nu' \Gamma_{+,\ell}^{-1}(h_n) \Psi_{+,p,q}(h_n, b_n) \Gamma_{+,q}^{-1}(b_n) e_{p+1}.
\end{aligned}$$

Thus, collecting terms, we obtain the result in part (V).

Finally, to establish (D), we proceed as in the proof of Lemma S.A.3. First, note that if $n\min\{h_n, b_n\} \to \infty$ and $\kappa \max\{h_n, b_n\} < \kappa_0$, then

$$\mathcal{V}^{\mathrm{bc}}_{+,\nu,p,q}(h_n, b_n) = O_p\big(n^{-1}h_n^{-1-2\nu} + n^{-1}b_n^{-3-2p}h_n^{2p+2-2\nu}\big),$$
$$\mathbb{E}\big[\hat{\mu}^{(\nu)\mathrm{bc}}_{+,p,q}(h_n, b_n)|\mathcal{X}_n\big] - \nu! e_\nu' \beta_{+,p} = O_p\big(h_n^{p+2-\nu} + h_n^{p+1-\nu}b_n^{q-p}\big).$$

Next, observe that

$$\frac{\hat{\mu}^{(\nu)\mathrm{bc}}_{+,p,q}(h_n, b_n) - \nu! e_\nu' \beta_{+,p}}{\sqrt{\mathcal{V}^{\mathrm{bc}}_{+,\nu,p,q}(h_n, b_n)}} = \xi_{1,n} + \xi_{2,n},$$

where

$$\xi_{1,n} = \big(\mathcal{V}^{\mathrm{bc}}_{+,\nu,p,q}(h_n, b_n)\big)^{-1/2}\big(\hat{\mu}^{(\nu)\mathrm{bc}}_{+,p,q}(h_n, b_n) - \mathbb{E}\big[\hat{\mu}^{(\nu)\mathrm{bc}}_{+,p,q}(h_n, b_n)|\mathcal{X}_n\big]\big)$$

and

$$\xi_{2,n} = \big(\mathcal{V}^{\mathrm{bc}}_{+,\nu,p,q}(h_n, b_n)\big)^{-1/2}\big(\mathbb{E}\big[\hat{\mu}^{(\nu)\mathrm{bc}}_{+,p,q}(h_n, b_n)|\mathcal{X}_n\big] - \nu! e_\nu' \beta_{+,p}\big).$$

For the bias, note that $\xi_{2,n} = o_p(1)$ because

$$\begin{aligned}
\xi^2_{2,n} &= O_p\left(\min\left\{nh_n^{1+2\nu}, \frac{nb_n^{3+2p}}{h_n^{2+2p-2\nu}}\right\}\right) \\
&\quad \times O_p\big(\max\big\{h_n^{2p+4-2\nu}, h_n^{2p+2-2\nu}b_n^{2(q-p)}\big\}\big) \\
&= O_p\big(n\min\big\{h_n^{3+2p}, b_n^{3+2p}\big\}\max\big\{h_n^2, b_n^{2(q-p)}\big\}\big) = o_p(1),
\end{aligned}$$

provided that $\kappa\max\{h_n, b_n\} < \kappa_0$. Thus, it remains to show that $\xi_{1,n} \to_d \mathcal{N}(0,1)$. By Lemma S.A.1, $\xi_{1,n} = \tilde{\xi}_{1,n} + o_p(1)$ with

$$\tilde{\xi}_{1,n} = \sum_{i=1}^{n} \omega_{n,i}\varepsilon_i, \qquad \omega_{n,i} = \frac{\omega_{1,n,i}}{\sqrt{\omega_{2,n}}},$$

where

$$\begin{aligned}
\omega_{1,n,i} &= h_n^{-\nu} e_\nu' \tilde{\Gamma}_{+,p}(h_n)^{-1} K_{h_n}(X_i) r_p(X_i/h_n)/n \\
&\quad - h_n^{p+1-\nu} b_n^{-p-1}\big(e_\nu' \tilde{\Gamma}_{+,p}(h_n)^{-1}\tilde{\vartheta}_{+,p,p+1}(h_n)\big) \\
&\quad \times \big(e_{p+1}' \tilde{\Gamma}_{+,q}(b_n)^{-1} K_{b_n}(X_i) r_q(X_i/b_n)\big)/n
\end{aligned}$$

and

$$\begin{aligned}
\omega_{2,n} &= \sum_{i=1}^{n} \mathbb{E}\big[\omega_{1,n,i}^2 \varepsilon_i^2\big] \\
&= \frac{1}{nh_n^{1+2\nu}} e_\nu' \tilde{\Gamma}_{+,p}(h_n)^{-1} \tilde{\Psi}_{+,p}(h_n) \tilde{\Gamma}_{+,p}(h_n)^{-1} e_\nu \\
&\quad + \frac{h_n^{2p+2-2\nu}}{nb_n^{3+2p}} \big(e_{p+1}' \tilde{\Gamma}_{+,q}(b_n)^{-1} \tilde{\Psi}_{+,q}(b_n) \tilde{\Gamma}_{+,q}(b_n)^{-1} e_{p+1}\big) \\
&\quad \times \big(e_\nu' \tilde{\Gamma}_{+,p}(h_n)^{-1} \tilde{\vartheta}_{+,p,p+1}(h_n)\big)^2 \\
&\quad - 2\frac{\rho_n^{p+1} \min\{1, \rho_n\}}{nh_n^{1+2\nu}} \big(e_\nu' \tilde{\Gamma}_{+,\ell}^{-1}(h_n) \tilde{\Psi}_{+,p,q}(h_n, b_n) \tilde{\Gamma}_{+,q}^{-1}(b_n) e_{p+1}\big) \\
&\quad \times \big(e_\nu' \tilde{\Gamma}_{+,p}(h_n)^{-1} \tilde{\vartheta}_{+,p,p+1}(h_n)\big),
\end{aligned}$$

provided that $\kappa \max\{h_n, b_n\} < \kappa_0$ and $n \min\{h_n, b_n\} \to \infty$. Note that $\{\omega_{n,i}\varepsilon_i : 1 \leq i \leq n\}$ is a triangular array of independent random variables with $\mathbb{E}[\tilde{\xi}_{1,n}] = 0$ and $\mathbb{V}[\tilde{\xi}_{1,n}] = 1$. Therefore, $\tilde{\xi}_{1,n} \to_d \mathcal{N}(0, 1)$ by the Linderberg–Feller central limit theorem for triangular arrays because

$$\begin{aligned}
\sum_{i=1}^{n} \mathbb{E}\big[|\omega_{n,i}\varepsilon_i|^4\big] &\lesssim \min\bigg\{n^2 h_n^{2+4\nu}, \frac{n^2 b_n^{6+4p}}{h_n^{4+4p-4\nu}}\bigg\} h_n^{-4\nu} \\
&\quad \times \sum_{i=1}^{n} \mathbb{E}\big[\big|e_\nu' \tilde{\Gamma}_{+,p}(h_n)^{-1} K_{h_n}(X_i) r_p(X_i/h_n)\big|^4\big]/n^4 \\
&\quad + \min\bigg\{n^2 h_n^{2+4\nu}, \frac{n^2 b_n^{6+4p}}{h_n^{4+4p-4\nu}}\bigg\} \frac{h_n^{4p+4-4\nu}}{b_n^{4p+4}} \\
&\quad \times \sum_{i=1}^{n} \mathbb{E}\big[\big|e_{p+1}' \tilde{\Gamma}_{+,q}(b_n)^{-1} K_{b_n}(X_i) r_q(X_i/b_n)\big|^4\big]/n^4 \\
&\lesssim n^{-1} \min\bigg\{h_n^2, \frac{b_n^{6+4p}}{h_n^{4+4p}}\bigg\} h_n^{-3} \\
&\quad + n^{-1} \min\bigg\{h_n^{2+4\nu}, \frac{b_n^{6+4p}}{h_n^{4+4p-4\nu}}\bigg\} \frac{h_n^{4p+4-4\nu}}{b_n^{4p+4}} b_n^{-3} \\
&= n^{-1} h_n^{-1} \min\big\{1, \rho_n^{-6-4p}\big\} + n^{-1} b_n^{-1} \min\big\{\rho_n^{6+4p}, 1\big\} \to 0,
\end{aligned}$$

provided that $n \min\{h_n, b_n\} \to \infty$. This concludes the proof. *Q.E.D.*

### S.2.2. *Proofs of Lemma A.1 and Theorem A.1*

We first provide a proof of Lemma A.1, which we restate here in a less compact way for completeness.

LEMMA A.1: *Suppose Assumptions* 1–2 *hold with* $S \geq p+2$, *and* $nh_n \to \infty$. *Let* $r \in \mathbb{N}$.

(B) *If* $h_n \to 0$, *then*

$$\mathbb{E}\big[\hat{\tau}_{\nu,p}(h_n)|\mathcal{X}_n\big] = \tau_\nu + h_n^{p+1-\nu}\mathsf{B}_{\nu,p,p+1}(h_n) + h_n^{p+2-\nu}\mathsf{B}_{\nu,p,p+2}(h_n) + o_p\big(h_n^{p+2-\nu}\big),$$

*where*

$$\mathsf{B}_{\nu,p,r}(h_n) = \frac{\mu_+^{(r)}}{r!}\mathcal{B}_{+,\nu,p,r}(h_n) - \frac{\mu_-^{(r)}}{r!}\mathcal{B}_{-,\nu,p,r}(h_n),$$
$$\mathcal{B}_{+,\nu,p,r}(h_n) = \nu! e_\nu' \Gamma_{+,p}^{-1}(h_n)\vartheta_{+,p,r}(h_n) = \nu! e_\nu' \Gamma_p^{-1}\vartheta_{p,r} + o_p(1),$$
$$\mathcal{B}_{-,\nu,p,r}(h_n) = \nu! e_\nu' \Gamma_{-,p}^{-1}(h_n)\vartheta_{-,p,r}(h_n) = (-1)^{\nu+r}\nu! e_\nu' \Gamma_p^{-1}\vartheta_{p,r} + o_p(1).$$

(V) *If* $h_n \to 0$, *then* $\mathsf{V}_{\nu,p}(h_n) = \mathbb{V}[\hat{\tau}_{\nu,p}(h_n)|\mathcal{X}_n] = \mathcal{V}_{+,\nu,p}(h_n) + \mathcal{V}_{-,\nu,p}(h_n)$, *where*

$$\begin{aligned}\mathcal{V}_{+,\nu,p}(h_n) &= \frac{1}{nh_n^{2\nu}}\nu!^2 e_\nu' \Gamma_{+,p}^{-1}(h_n)\Psi_{+,p}(h_n)\Gamma_{+,p}^{-1}(h_n)e_\nu \\ &= \frac{1}{nh_n^{1+2\nu}}\frac{\sigma_+^2}{f}\nu!^2 e_\nu'\Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\big\{1+o_p(1)\big\}, \\ \mathcal{V}_{-,\nu,p}(h_n) &= \frac{1}{nh_n^{2\nu}}\nu!^2 e_\nu' \Gamma_{-,p}^{-1}(h_n)\Psi_{-,p}(h_n)\Gamma_{-,p}^{-1}(h_n)e_\nu \\ &= \frac{1}{nh_n^{1+2\nu}}\frac{\sigma_-^2}{f}\nu!^2 e_\nu'\Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\big\{1+o_p(1)\big\}.\end{aligned}$$

(D) *If* $nh_n^{2p+5} \to 0$, *then*

$$\frac{\hat{\tau}_{\nu,p}(h_n) - \tau_\nu - h_n^{p+1-\nu}\mathsf{B}_{\nu,p,p+1}(h_n)}{\sqrt{\mathsf{V}_{\nu,p}(h_n)}} \to_d \mathcal{N}(0,1).$$

PROOF: Part (B) follows immediately from Lemma S.A.3(B), its analogue for the left-side estimator ($s!e_s'\hat{\beta}_{-,\ell}(h_n)$), and the linearity of conditional expectations. Part (V) also follows immediately from Lemma S.A.3(V), its analogue for the left-side estimator ($s!e_s'\hat{\beta}_{-,\ell}(h_n)$), and the conditional independence of observations at either side of the threshold ($x = 0$). Finally, part (D) follows by the same argument given in the proof of Lemma S.A.3(D), but now applied to the estimator $\hat{\tau}_{\nu,p}(h_n) = \hat{\mu}^{(\nu)}_{+,p}(h_n) - \hat{\mu}^{(\nu)}_{-,p}(h_n) = \nu! e'_\nu \hat{\beta}_{+,p}(h_n) - \nu! e'_\nu \hat{\beta}_{-,p}(h_n)$. This completes the proof. *Q.E.D.*

The proof of Theorem A.1, which we also restate here in a less compact way for completeness, is discussed next. Recall that

$$\hat{\tau}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n) = \hat{\mu}^{(\nu)\mathrm{bc}}_{+,p,q}(h_n,b_n) - \hat{\mu}^{(\nu)\mathrm{bc}}_{-,p,q}(h_n,b_n),$$

$$\hat{\mu}^{(\nu)\mathrm{bc}}_{+,p,q}(h_n,b_n) = \nu! e'_\nu \hat{\beta}_{+,p}(h_n) - h_n^{p+1-\nu}\big(e'_{p+1}\hat{\beta}_{+,q}(b_n)\big)\mathcal{B}_{+,\nu,p,p+1}(h_n),$$

$$\hat{\mu}^{(\nu)\mathrm{bc}}_{-,p,q}(h_n,b_n) = \nu! e'_\nu \hat{\beta}_{-,p}(h_n) - h_n^{p+1-\nu}\big(e'_{p+1}\hat{\beta}_{-,q}(b_n)\big)\mathcal{B}_{-,\nu,p,p+1}(h_n).$$

THEOREM A.1: *Suppose Assumptions* 1–2 *hold with* $S \geq q+1$, *and* $n\min\{h_n,b_n\} \to \infty$.

(B) *If* $\max\{h_n,b_n\} \to 0$, *then*

$$\mathbb{E}\big[\hat{\tau}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)|\mathcal{X}_n\big] = \tau_\nu + h_n^{p+2-\nu}\mathsf{B}_{\nu,p,p+2}(h_n)\big\{1+o_p(1)\big\} - h_n^{p+1-\nu} b_n^{q-p}\mathsf{B}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)\big\{1+o_p(1)\big\},$$

*where*

$$\mathsf{B}^{\mathrm{bc}}_{\nu,p,q}(h,b) = \frac{\mu_+^{(q+1)}}{(q+1)!}\mathcal{B}_{+,p+1,q,q+1}(b)\frac{\mathcal{B}_{+,\nu,p,p+1}(h)}{(p+1)!} - \frac{\mu_-^{(q+1)}}{(q+1)!}\mathcal{B}_{-,p+1,q,q+1}(b)\frac{\mathcal{B}_{-,\nu,p,p+1}(h)}{(p+1)!}.$$

(V) $\mathsf{V}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n) = \mathbb{V}[\hat{\tau}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)|\mathcal{X}_n] = \mathcal{V}^{\mathrm{bc}}_{+,\nu,p,q}(h_n,b_n) + \mathcal{V}^{\mathrm{bc}}_{-,\nu,p,q}(h_n,b_n)$, *where*

$$\mathcal{V}^{\mathrm{bc}}_{+,\nu,p,q}(h,b) = \mathcal{V}_{+,\nu,p}(h) - 2h^{p+1-\nu}\mathcal{C}_{+,\nu,p,q}(h,b)\frac{\mathcal{B}_{+,\nu,p,p+1}(h)}{(p+1)!} + h^{2(p+1-\nu)}\mathcal{V}_{+,p+1,q}(b)\frac{\mathcal{B}^2_{+,\nu,p,p+1}(h)}{(p+1)!^2},$$

$$\mathcal{V}^{\mathrm{bc}}_{-,\nu,p,q}(h,b) = \mathcal{V}_{-,\nu,p}(h) - 2h^{p+1-\nu}\mathcal{C}_{-,\nu,p,q}(h,b)\frac{\mathcal{B}_{-,\nu,p,p+1}(h)}{(p+1)!} + h^{2(p+1-\nu)}\mathcal{V}_{-,p+1,q}(b)\frac{\mathcal{B}^2_{-,\nu,p,p+1}(h)}{(p+1)!^2},$$

$$\mathcal{C}_{+,\nu,p,q}(h,b) = \frac{1}{nh^\nu b^{p+1}}\nu!(p+1)! e'_\nu \Gamma^{-1}_{+,p}(h)\Psi_{+,p,q}(h,b)\Gamma^{-1}_{+,q}(b)e_{p+1},$$

$$\mathcal{C}_{-,\nu,p,q}(h,b) = \frac{1}{nh^\nu b^{p+1}}\nu!(p+1)! e'_\nu \Gamma^{-1}_{-,p}(h)\Psi_{-,p,q}(h,b)\Gamma^{-1}_{-,q}(b)e_{p+1}.$$

(D) *If* $n\min\{h_n^{2p+3}, b_n^{2p+3}\}\max\{h_n^2, b_n^{2(q-p)}\}\to 0$, *and* $\kappa\max\{h_n, b_n\}<\kappa_0$, *then*

$$T^{\mathrm{rbc}}_{\nu,p,q}(h_n,b_n)=\frac{\hat{\tau}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)-\tau_\nu}{\sqrt{\mathsf{V}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)}}\to_d \mathcal{N}(0,1).$$

PROOF: As in the proof of Lemma A.1 (and its relationship to the proof of Lemma S.A.3), the proof of this theorem proceeds as in the proof of Lemma S.A.4 but now considering both estimators $\hat{\mu}^{(\nu)\mathrm{bc}}_{+,p,q}(h_n,b_n)$ and $\hat{\mu}^{(\nu)\mathrm{bc}}_{-,p,q}(h_n,b_n)$ together. *Q.E.D.*

### S.2.3. *Proofs of Lemma A.2 and Theorem A.2*

Recall that the $\nu$th fuzzy RD estimand ($\nu\le S$) is $\varsigma_\nu=\tau_{Y,\nu}/\tau_{T,\nu}$ with $\tau_{Y,\nu}=\mu^{(\nu)}_{Y+}-\mu^{(\nu)}_{Y-}$ and $\tau_{T,\nu}=\mu^{(\nu)}_{T+}-\mu^{(\nu)}_{T-}$. The corresponding estimator based on the two $p$th-order local polynomial estimators ($\nu\le p$) of the reduced-form equations is $\hat{\varsigma}_{\nu,p}(h_n)=\hat{\tau}_{Y,\nu,p}(h_n)/\hat{\tau}_{T,\nu,p}(h_n)$ with $\hat{\tau}_{Y,\nu,p}(h_n)=\hat{\mu}^{(\nu)}_{Y+,p}(h_n)-\hat{\mu}^{(\nu)}_{Y-,p}(h_n)$ and $\hat{\tau}_{T,\nu,p}(h_n)=\hat{\mu}^{(\nu)}_{T+,p}(h_n)-\hat{\mu}^{(\nu)}_{T-,p}(h_n)$, where $\hat{\mu}^{(\nu)}_{Y+,p}(h_n)=\nu!e_\nu'\hat{\beta}_{Y+,p}(h_n)$, $\hat{\mu}^{(\nu)}_{Y-,p}(h_n)=\nu!e_\nu'\hat{\beta}_{Y-,p}(h_n)$, $\hat{\mu}^{(\nu)}_{T+,p}(h_n)=\nu!e_\nu'\hat{\beta}_{T+,p}(h_n)$, and $\hat{\mu}^{(\nu)}_{T-,p}(h_n)=\nu!e_\nu'\hat{\beta}_{T-,p}(h_n)$ with

$$\begin{aligned}
\hat{\beta}_{Y+,\ell}(h_n)&=H_\ell(h_n)\Gamma^{-1}_{+,\ell}(h_n)X_\ell(h_n)'W_+(h_n)Y/n,\\
\hat{\beta}_{Y-,\ell}(h_n)&=H_\ell(h_n)\Gamma^{-1}_{-,\ell}(h_n)X_\ell(h_n)'W_-(h_n)Y/n,\\
\hat{\beta}_{T+,\ell}(h_n)&=H_\ell(h_n)\Gamma^{-1}_{+,\ell}(h_n)X_\ell(h_n)'W_+(h_n)T/n,\\
\hat{\beta}_{T-,\ell}(h_n)&=H_\ell(h_n)\Gamma^{-1}_{-,\ell}(h_n)X_\ell(h_n)'W_-(h_n)T/n.
\end{aligned}$$

Using the expansion

$$\frac{\hat{a}}{\hat{b}}-\frac{a}{b}=\frac{1}{b}(\hat{a}-a)-\frac{a}{b^2}(\hat{b}-b)+\frac{a}{b^2\hat{b}}(\hat{b}-b)^2-\frac{1}{b\hat{b}}(\hat{a}-a)(\hat{b}-b),$$

we obtain $\hat{\varsigma}_{\nu,p}(h_n)-\varsigma_\nu=\tilde{\varsigma}_{\nu,p}(h_n)+R_n$ with

$$\begin{aligned}
\tilde{\varsigma}_{\nu,p}(h_n)&=\frac{1}{\tau_{T,\nu}}\big(\hat{\tau}_{Y,\nu,p}(h_n)-\tau_{Y,\nu}\big)-\frac{\tau_{Y,\nu}}{\tau^2_{T,\nu}}\big(\hat{\tau}_{T,\nu,p}(h_n)-\tau_{T,\nu}\big),\\
R_n&=\frac{\tau_{Y,\nu}}{\tau^2_{T,\nu}\hat{\tau}_{T,\nu,p}(h_n)}\big(\hat{\tau}_{T,\nu,p}(h_n)-\tau_{T,\nu}\big)^2\\
&\quad-\frac{1}{\tau_{T,\nu}\hat{\tau}_{T,\nu,p}(h_n)}\big(\hat{\tau}_{Y,\nu,p}(h_n)-\tau_{Y,\nu}\big)\big(\hat{\tau}_{T,\nu,p}(h_n)-\tau_{T,\nu}\big).
\end{aligned}$$

We restate Lemma A.2 and discuss its proof next.

LEMMA A.2: *Suppose Assumptions* 1–3 *hold with* $S \geq p+2$, *and* $nh_n \to \infty$. *Let* $r \in \mathbb{N}$.

(R) *If* $h_n \to 0$ *and* $nh_n^{1+2\nu} \to \infty$, *then*

$$R_n = O_p\left(\frac{1}{nh_n^{1+2\nu}} + h_n^{2(p+1-\nu)}\right).$$

(B) *If* $h_n \to 0$, *then*

$$\mathbb{E}\big[\tilde{s}_{\nu,p}(h_n)|\mathcal{X}_n\big] = h_n^{p+1-\nu}\mathsf{B}_{\mathrm{F},\nu,p,p+1}(h_n) + h_n^{p+2-\nu}\mathsf{B}_{\mathrm{F},\nu,p,p+2}(h_n) + o_p\big(h_n^{p+2-\nu}\big),$$

*where*

$$\mathsf{B}_{\mathrm{F},\nu,p,r}(h_n) = \frac{1}{\tau_{T,\nu}}\mathsf{B}_{Y,\nu,p,r}(h_n) - \frac{\tau_{Y,\nu}}{\tau_{T,\nu}^2}\mathsf{B}_{T,\nu,p,r}(h_n),$$

*with*

$$\mathsf{B}_{Y,\nu,p,r}(h_n) = \frac{\mu_{Y+}^{(r)}}{r!}\mathcal{B}_{+,\nu,p,r}(h_n) - \frac{\mu_{Y-}^{(r)}}{r!}\mathcal{B}_{-,\nu,p,r}(h_n),$$

$$\mathsf{B}_{T,\nu,p,r}(h_n) = \frac{\mu_{T+}^{(r)}}{r!}\mathcal{B}_{+,\nu,p,r}(h_n) - \frac{\mu_{T-}^{(r)}}{r!}\mathcal{B}_{-,\nu,p,r}(h_n).$$

(V) *If* $h_n \to 0$, *then* $\mathsf{V}_{\mathrm{F},\nu,p}(h_n) = \mathbb{V}[\tilde{s}_{\nu,p}(h_n)|\mathcal{X}_n] = \mathsf{V}_{\mathrm{F},+,\nu,p}(h_n) + \mathsf{V}_{\mathrm{F},-,\nu,p}(h_n)$, *where*

$$\mathsf{V}_{\mathrm{F},+,\nu,p}(h_n) = \frac{1}{\tau_{T,\nu}^2}\mathcal{V}_{YY+,\nu,p}(h_n) - \frac{2\tau_{Y,\nu}}{\tau_{T,\nu}^3}\mathcal{V}_{YT+,\nu,p}(h_n) + \frac{\tau_{Y,\nu}^2}{\tau_{T,\nu}^4}\mathcal{V}_{TT+,\nu,p}(h_n),$$

*with*

$$\mathcal{V}_{YY+,\nu,p}(h_n) = \frac{1}{nh_n^{2\nu}}\nu!^2 e_\nu'\Gamma_{+,p}^{-1}(h_n)\Psi_{YY+,p}(h_n)\Gamma_{+,p}^{-1}(h_n)e_\nu = \frac{1}{nh_n^{1+2\nu}}\frac{\sigma_{YY+}^2}{f}\nu!^2 e_\nu'\Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\big\{1+o_p(1)\big\},$$

$$\mathcal{V}_{YT+,\nu,p}(h_n) = \frac{1}{nh_n^{2\nu}}\nu!^2 e_\nu'\Gamma_{+,p}^{-1}(h_n)\Psi_{YT+,p}(h_n)\Gamma_{+,p}^{-1}(h_n)e_\nu = \frac{1}{nh_n^{1+2\nu}}\frac{\sigma_{YT+}^2}{f}\nu!^2 e_\nu'\Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\big\{1+o_p(1)\big\},$$

$$\mathcal{V}_{TT+,\nu,p}(h_n) = \frac{1}{nh_n^{2\nu}}\nu!^2 e_\nu' \Gamma_{+,p}^{-1}(h_n)\Psi_{TT+,p}(h_n)\Gamma_{+,p}^{-1}(h_n)e_\nu$$
$$= \frac{1}{nh_n^{1+2\nu}}\frac{\sigma_{TT+}^2}{f}\nu!^2 e_\nu' \Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\{1+o_p(1)\},$$

*and*

$$\mathsf{V}_{\mathrm{F},-,\nu,p}(h_n) = \frac{1}{\tau_{T,\nu}^2}\mathcal{V}_{YY-,\nu,p}(h_n) - \frac{2\tau_{Y,\nu}}{\tau_{T,\nu}^3}\mathcal{V}_{YT-,\nu,p}(h_n) + \frac{\tau_{Y,\nu}^2}{\tau_{T,\nu}^4}\mathcal{V}_{TT-,\nu,p}(h_n),$$

*with*

$$\mathcal{V}_{YY-,\nu,p}(h_n) = \frac{1}{nh_n^{2\nu}}\nu!^2 e_\nu' \Gamma_{-,p}^{-1}(h_n)\Psi_{YY-,p}(h_n)\Gamma_{-,p}^{-1}(h_n)e_\nu$$
$$= \frac{1}{nh_n^{1+2\nu}}\frac{\sigma_{YY-}^2}{f}\nu!^2 e_\nu' \Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\{1+o_p(1)\},$$
$$\mathcal{V}_{YT-,\nu,p}(h_n) = \frac{1}{nh_n^{2\nu}}\nu!^2 e_\nu' \Gamma_{-,p}^{-1}(h_n)\Psi_{YT-,p}(h_n)\Gamma_{-,p}^{-1}(h_n)e_\nu$$
$$= \frac{1}{nh_n^{1+2\nu}}\frac{\sigma_{YT-}^2}{f}\nu!^2 e_\nu' \Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\{1+o_p(1)\},$$
$$\mathcal{V}_{TT-,\nu,p}(h_n) = \frac{1}{nh_n^{2\nu}}\nu!^2 e_\nu' \Gamma_{-,p}^{-1}(h_n)\Psi_{TT-,p}(h_n)\Gamma_{-,p}^{-1}(h_n)e_\nu$$
$$= \frac{1}{nh_n^{1+2\nu}}\frac{\sigma_{TT-}^2}{f}\nu!^2 e_\nu' \Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\{1+o_p(1)\}.$$

(D) *If* $nh_n^{2p+5} \to 0$ *and* $nh_n^{1+2\nu} \to \infty$, *then*

$$\frac{\hat{\varsigma}_{\nu,p}(h_n) - \varsigma_\nu - h_n^{p+1-\nu}\mathsf{B}_{\mathrm{F},\nu,p,p+1}(h_n)}{\sqrt{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)}} \to_d \mathcal{N}(0,1).$$

PROOF: Lemma A.1 implies that

$$\left(\hat{\tau}_{T,\nu,p}(h_n) - \tau_{T,\nu}\right)^2 = O_p\left(\frac{1}{nh_n^{1+2\nu}} + h_n^{2(p+1-\nu)}\right)$$

and

$$\big(\hat{\tau}_{Y,\nu,p}(h_n) - \tau_{Y,\nu}\big)\big(\hat{\tau}_{T,\nu,p}(h_n) - \tau_{T,\nu}\big)$$
$$= O_p\Bigg(\sqrt{\frac{1}{nh_n^{1+2\nu}} + h_n^{2(p+1-\nu)}}\Bigg) O_p\Bigg(\sqrt{\frac{1}{nh_n^{1+2\nu}} + h_n^{2(p+1-\nu)}}\Bigg),$$

provided that $h_n \to 0$ and $nh_n \to \infty$. This gives part (R), provided that $\hat{\tau}_{T,\nu,p}(h_n) \to_p \tau_{T,\nu} > 0$, which follows by $nh_n^{1+2\nu} \to \infty$ (and $h_n \to 0$). Parts (B) and (V) follow directly from Lemma A.1 by computing the conditional moments of $\tilde{\varsigma}_p^{(\nu)}(h_n)$, which are (fixed) linear combinations of $(\hat{\tau}_{Y,\nu,p}(h_n) - \tau_{Y,\nu})$ and $(\hat{\tau}_{T,\nu,p}(h_n) - \tau_{T,\nu})$. Finally, for part (D), note first

$$\frac{\tilde{\varsigma}_{\nu,p}(h_n) - h_n^{p+1-\nu}\mathsf{B}_{\mathrm{F},\nu,p,p+1}(h_n)}{\sqrt{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)}} \to_d \mathcal{N}(0,1),$$

using Lemma A.1 and the Cramér–Wold device, and provided that $nh_n \to \infty$ and $nh_n^{2p+5} \to 0$. Thus, using part (R), we have

$$\frac{\hat{\varsigma}_{\nu,p}(h_n) - \varsigma_\nu - h_n^{p+1-\nu}\mathsf{B}_{\mathrm{F},\nu,p,p+1}(h_n)}{\sqrt{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)}}$$
$$= \frac{\tilde{\varsigma}_{\nu,p}(h_n) - h_n^{p+1-\nu}\mathsf{B}_{\mathrm{F},\nu,p,p+1}(h_n)}{\sqrt{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)}} + \frac{R_n}{\sqrt{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)}} \to_d \mathcal{N}(0,1),$$

because

$$\frac{R_n}{\sqrt{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)}} = O_p\Bigg(\frac{\sqrt{nh_n^{1+2\nu}}}{nh_n^{1+2\nu}} + \sqrt{nh_n^{1+2\nu}}h_n^{2(p+1-\nu)}\Bigg)$$
$$= O_p\Bigg(\frac{1}{\sqrt{nh_n^{1+2\nu}}} + \sqrt{nh_n^{4p-2\nu+5}}\Bigg) = o_p(1),$$

provided that $nh_n^{1+2\nu} \to \infty$. Note that $nh_n^{4p-2\nu+5} = nh_n^{2(p-\nu)+2p+5} \to 0$ for any $p \geq \nu$. *Q.E.D.*

Next, for the proof of Theorem A.2, which gives an analogue of Theorem A.1 for the bias-corrected fuzzy RD estimator, recall that

$$\hat{\varsigma}^{\mathrm{bc}}_{\nu,p,q}(h_n, b_n) = \hat{\varsigma}_{\nu,p}(h_n) - h_n^{p+1-\nu}\hat{\mathsf{B}}_{\mathrm{F},\nu,p,p+1,q}(h_n, b_n),$$

with

$$\hat{\mathsf{B}}_{\mathrm{F},\nu,p,q}(h_n,b_n)=\frac{1}{\hat{\tau}_{T,\nu,p}(h_n)}\big(\big(e'_{p+1}\hat{\beta}_{Y+,q}(b_n)\big)\mathcal{B}_{+,\nu,p,p+1}(h_n)$$
$$-\big(e'_{p+1}\hat{\beta}_{Y-,q}(b_n)\big)\mathcal{B}_{-,\nu,p,p+1}(h_n)\big)$$
$$-\frac{\hat{\tau}_{Y,\nu,p}(h_n)}{\hat{\tau}_{T,\nu,p}(h_n)^2}\big(\big(e'_{p+1}\hat{\beta}_{T+,q}(b_n)\big)\mathcal{B}_{+,\nu,p,p+1}(h_n)$$
$$-\big(e'_{p+1}\hat{\beta}_{T-,q}(b_n)\big)\mathcal{B}_{-,\nu,p,p+1}(h_n)\big).$$

Linearizing the estimator, we obtain

$$\begin{aligned}
&\hat{\varsigma}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)-\varsigma_\nu\\
&\quad=\hat{\varsigma}_{\nu,p}(h_n)-h_n^{p+1-\nu}\hat{\mathsf{B}}_{\mathrm{F},\nu,p,q}(h_n,b_n)-\varsigma_\nu\\
&\quad=\tilde{\varsigma}_{\nu,p}(h_n)+R_n-h_n^{p+1-\nu}\hat{\mathsf{B}}_{\mathrm{F},\nu,p,q}(h_n,b_n)\\
&\quad=\tilde{\varsigma}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)+R_n-h_n^{p+1-\nu}\big(\hat{\mathsf{B}}_{\mathrm{F},\nu,p,q}(h_n,b_n)-\check{\mathsf{B}}_{\mathrm{F},\nu,p,q}(h_n,b_n)\big)\\
&\quad=\tilde{\varsigma}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)+R_n-R_n^{\mathrm{bc}},
\end{aligned}$$

with

$$\tilde{\varsigma}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)=\frac{1}{\tau_{T,\nu}}\big(\hat{\tau}^{\mathrm{bc}}_{Y,\nu,p,q}(h_n,b_n)-\tau_{Y,\nu}\big)$$
$$-\frac{\tau_{Y,\nu}}{\tau^2_{T,\nu}}\big(\hat{\tau}^{\mathrm{bc}}_{T,\nu,p,q}(h_n,b_n)-\tau_{T,\nu}\big),$$

$$R_n=\frac{\tau_{Y,\nu}}{\tau^2_{T,\nu}\hat{\tau}_{T,\nu,p}(h_n)}\big(\hat{\tau}_{T,\nu,p}(h_n)-\tau_{T,\nu}\big)^2$$
$$-\frac{1}{\tau_{T,\nu}\hat{\tau}_{T,\nu,p}(h_n)}\big(\hat{\tau}_{Y,\nu,p}(h_n)-\tau_{Y,\nu}\big)\big(\hat{\tau}_{T,\nu,p}(h_n)-\tau_{T,\nu}\big),$$

$$\check{\mathsf{B}}_{\nu,p,q}(h_n,b_n)=\frac{1}{\tau_{T,\nu}}\big(\big(e'_{p+1}\hat{\beta}_{Y+,q}(b_n)\big)\mathcal{B}_{+,\nu,p,p+1}(h_n)$$
$$-\big(e'_{p+1}\hat{\beta}_{Y-,q}(b_n)\big)\mathcal{B}_{-,\nu,p,p+1}(h_n)\big)$$
$$-\frac{\tau_{Y,\nu}}{\tau^2_{T,\nu}}\big(\big(e'_{p+1}\hat{\beta}_{T+,q}(b_n)\big)\mathcal{B}_{+,\nu,p,p+1}(h_n)$$
$$-\big(e'_{p+1}\hat{\beta}_{T-,q}(b_n)\big)\mathcal{B}_{-,\nu,p,p+1}(h_n)\big),$$

$$R_n^{\mathrm{bc}}=h_n^{p+1-\nu}\big(\hat{\mathsf{B}}_{\mathrm{F},\nu,p,q}(h_n,b_n)-\check{\mathsf{B}}_{\mathrm{F},\nu,p,q}(h_n,b_n)\big).$$

Next, we restate Theorem A.2 and discuss its proof.

THEOREM A.2: *Suppose Assumptions* 1–3 *hold with* $S \geq p+2$, *and* $n\min\{h_n, b_n\} \to \infty$.

($\mathrm{R}^{\mathrm{bc}}$) *If* $h_n \to 0$ *and* $nh_n^{1+2\nu} \to \infty$, *and provided that* $\kappa b_n < \kappa_0$, *then*

$$R_n^{\mathrm{bc}} = O_p\left(\frac{h_n^{p+1-\nu}}{\sqrt{nh_n^{1+2\nu}}} + h_n^{2(p+1-\nu)}\right) O_p\left(1 + \frac{1}{\sqrt{nb_n^{3+2p}}}\right).$$

(B) *If* $\max\{h_n, b_n\} \to 0$, *then*

$$\begin{aligned}\mathbb{E}\big[\tilde{\varsigma}^{\mathrm{bc}}_{\nu,p,q}(h_n, b_n)|\mathcal{X}_n\big] = {} & h_n^{p+2-\nu}\mathsf{B}_{\mathrm{F},\nu,p,p+2}(h_n)\big[1+o_p(1)\big] \\ & + h_n^{p+1-\nu} b_n^{q-p} \mathsf{B}^{\mathrm{bc}}_{\mathrm{F},\nu,p,q}(h_n, b_n)\big[1+o_p(1)\big],\end{aligned}$$

*where*

$$\mathsf{B}^{\mathrm{bc}}_{\mathrm{F},\nu,p,q}(h,b) = \frac{1}{\tau_{T,\nu}}\mathsf{B}^{\mathrm{bc}}_{Y,\nu,p,q}(h_n,b_n) - \frac{\tau_{Y,\nu}}{\tau^2_{T,\nu}}\mathsf{B}^{\mathrm{bc}}_{T,\nu,p,q}(h_n,b_n),$$

*with*

$$\begin{aligned}\mathsf{B}^{\mathrm{bc}}_{Y,\nu,p,q}(h,b) = {} & \frac{\mu^{(q+1)}_{Y+}}{(q+1)!}\mathcal{B}_{+,p+1,q,q+1}(b)\frac{\mathcal{B}_{+,\nu,p,p+1}(h)}{(p+1)!} \\ & - \frac{\mu^{(q+1)}_{Y-}}{(q+1)!}\mathcal{B}_{-,p+1,q,q+1}(b)\frac{\mathcal{B}_{-,\nu,p,p+1}(h)}{(p+1)!}, \\ \mathsf{B}^{\mathrm{bc}}_{T,\nu,p,q}(h,b) = {} & \frac{\mu^{(q+1)}_{T+}}{(q+1)!}\mathcal{B}_{+,p+1,q,q+1}(b)\frac{\mathcal{B}_{+,\nu,p,p+1}(h)}{(p+1)!} \\ & - \frac{\mu^{(q+1)}_{T-}}{(q+1)!}\mathcal{B}_{-,p+1,q,q+1}(b)\frac{\mathcal{B}_{-,\nu,p,p+1}(h)}{(p+1)!}.\end{aligned}$$

(V) $\mathsf{V}^{\mathrm{bc}}_{\mathrm{F},\nu,p,q}(h_n,b_n) = \mathbb{V}[\tilde{\varsigma}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)|\mathcal{X}_n] = \mathsf{V}^{\mathrm{bc}}_{\mathrm{F},+,\nu,p,q}(h_n,b_n) + \mathsf{V}^{\mathrm{bc}}_{\mathrm{F},-,\nu,p,q}(h_n,b_n)$, *where*

$$\begin{aligned}\mathsf{V}^{\mathrm{bc}}_{\mathrm{F},+,\nu,p,q}(h,b) = {} & \mathsf{V}_{\mathrm{F},+,\nu,p}(h) - 2h^{p+1-\nu}\mathcal{C}_{\mathrm{F},+,\nu,p,q}(h,b)\frac{\mathcal{B}_{+,\nu,p,p+1}(h)}{(p+1)!} \\ & + h^{2p+2-2\nu}\mathsf{V}_{\mathrm{F},+,p+1,q}(b)\frac{\mathcal{B}^2_{+,\nu,p,p+1}(h)}{(p+1)!^2}, \\ \mathsf{V}^{\mathrm{bc}}_{\mathrm{F},-,\nu,p,q}(h,b) = {} & \mathsf{V}_{\mathrm{F},-,\nu,p}(h) - 2h^{p+1-\nu}\mathcal{C}_{\mathrm{F},-,\nu,p,q}(h,b)\frac{\mathcal{B}_{-,\nu,p,p+1}(h)}{(p+1)!} \\ & + h^{2p+2-2\nu}\mathsf{V}_{\mathrm{F},-,p+1,q}(b)\frac{\mathcal{B}^2_{-,\nu,p,p+1}(h)}{(p+1)!^2},\end{aligned}$$

$$\mathcal{C}_{\mathtt{F},+,\nu,p,q}(h,b) = \frac{1}{\tau_{T,\nu}^2}\mathcal{C}_{YY+,\nu,p,q}(h,b) - \frac{2\tau_{Y,\nu}}{\tau_{T,\nu}^3}\mathcal{C}_{YT+,\nu,p,q}(h,b) + \frac{\tau_{Y,\nu}^2}{\tau_{T,\nu}^4}\mathcal{C}_{TT+,\nu,p,q}(h,b),$$

$$\mathcal{C}_{\mathtt{F},-,\nu,p,q}(h,b) = \frac{1}{\tau_{T,\nu}^2}\mathcal{C}_{YY-,\nu,p,q}(h,b) - \frac{2\tau_{Y,\nu}}{\tau_{T,\nu}^3}\mathcal{C}_{YT-,\nu,p,q}(h,b) + \frac{\tau_{Y,\nu}^2}{\tau_{T,\nu}^4}\mathcal{C}_{TT-,\nu,p,q}(h,b),$$

*where*

$$\mathcal{C}_{YY+,\nu,p,q}(h,b) = \frac{1}{nh^\nu b^{p+1}}\nu!(p+1)! \times e_\nu' \Gamma_{+,p}^{-1}(h)\Psi_{YY+,p,q}(h,b)\Gamma_{+,q}^{-1}(b)e_{p+1},$$

$$\mathcal{C}_{YT+,\nu,p,q}(h,b) = \frac{1}{nh^\nu b^{p+1}}\nu!(p+1)! \times e_\nu' \Gamma_{+,p}^{-1}(h)\Psi_{YT+,p,q}(h,b)\Gamma_{+,q}^{-1}(b)e_{p+1},$$

$$\mathcal{C}_{TT+,\nu,p,q}(h,b) = \frac{1}{nh^\nu b^{p+1}}\nu!(p+1)! \times e_\nu' \Gamma_{+,p}^{-1}(h)\Psi_{TT+,p,q}(h,b)\Gamma_{+,q}^{-1}(b)e_{p+1},$$

$$\mathcal{C}_{YY-,\nu,p,q}(h,b) = \frac{1}{nh^\nu b^{p+1}}\nu!(p+1)! \times e_\nu' \Gamma_{-,p}^{-1}(h)\Psi_{YY-,p,q}(h,b)\Gamma_{-,q}^{-1}(b)e_{p+1},$$

$$\mathcal{C}_{YT-,\nu,p,q}(h,b) = \frac{1}{nh^\nu b^{p+1}}\nu!(p+1)! \times e_\nu' \Gamma_{-,p}^{-1}(h)\Psi_{YT-,p,q}(h,b)\Gamma_{-,q}^{-1}(b)e_{p+1},$$

$$\mathcal{C}_{TT-,\nu,p,q}(h,b) = \frac{1}{nh^\nu b^{p+1}}\nu!(p+1)! \times e_\nu' \Gamma_{-,p}^{-1}(h)\Psi_{TT-,p,q}(h,b)\Gamma_{-,q}^{-1}(b)e_{p+1}.$$

(D) *If* $n\min\{h_n^{2p+3}, b_n^{2p+3}\}\max\{h_n^2, b_n^{2(q-p)}\} \to 0$ *and* $n\min\{h_n^{1+2\nu}, b_n\} \to \infty$, *and provided that* $h_n \to 0$ *and* $\kappa b_n < \kappa_0$, *then*

$$T_{\mathtt{F},\nu,p,q}^{\mathtt{rbc}}(h_n,b_n) = \frac{\hat{\varsigma}_{\nu,p,q}^{\mathtt{bc}}(h_n,b_n) - \varsigma_\nu}{\sqrt{\mathsf{V}_{\mathtt{F},\nu,p,q}^{\mathtt{bc}}(h_n,b_n)}} \to_d \mathcal{N}(0,1).$$

PROOF: For part ($\mathrm{R}^{\mathrm{bc}}$), note that

$$\begin{aligned}|R_n^{\mathrm{bc}}| &\lesssim h_n^{p+1-\nu}\left|\frac{1}{\tau_{T,\nu}}-\frac{1}{\hat{\tau}_{T,\nu,p}(h_n)}\right|\left(\left|e'_{p+1}\hat{\beta}_{Y+,q}(b_n)\right|+\left|e'_{p+1}\hat{\beta}_{Y-,q}(b_n)\right|\right)\\ &\quad+h_n^{p+1-\nu}\left|\frac{\tau_{Y,\nu}}{\tau_{T,\nu}^2}-\frac{\hat{\tau}_{Y,\nu,p}(h_n)}{\hat{\tau}_{T,\nu,p}(h_n)^2}\right|\\ &\quad\times\left(\left|e'_{p+1}\hat{\beta}_{T+,q}(b_n)\right|+\left|e'_{p+1}\hat{\beta}_{T-,q}(b_n)\right|\right)\\ &=h_n^{p+1-\nu}O_p\left(\frac{1}{\sqrt{nh_n^{1+2\nu}}}+h_n^{p+1-\nu}\right)O_p\left(1+\frac{1}{\sqrt{nb_n^{3+2p}}}\right),\end{aligned}$$

provided that $\hat{\tau}_{T,\nu,p}(h_n)\to_p\tau_{T,\nu}>0$, which follows by $nh_n^{1+2\nu}\to\infty$ (and $h_n\to 0$), and $nb_n\to\infty$ and $\kappa b_n<\kappa_0$.

For part (B), first note that

$$\begin{aligned}\mathbb{E}\left[\tilde{s}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)|\mathcal{X}_n\right]&=\mathbb{E}\left[\frac{1}{\tau_{T,\nu}}\left(\hat{\tau}^{\mathrm{bc}}_{Y,\nu,p,q}(h_n,b_n)-\tau_y^{(\nu)}\right)\right.\\ &\qquad\left.-\frac{\tau_{Y,\nu}}{\tau_{T,\nu}^2}\left(\hat{\tau}^{\mathrm{bc}}_{T,\nu,p,q}(h_n,b_n)-\tau_t^{(\nu)}\right)\middle|\mathcal{X}_n\right]\\ &=B_1-h_n^{p+1-\nu}B_2,\end{aligned}$$

with

$$\begin{aligned}B_1&=\mathbb{E}\left[\tilde{s}_{\nu,p}(h_n)-h_n^{p+1-\nu}\mathsf{B}_{\mathrm{F},\nu,p,p+1}(h_n)|\mathcal{X}_n\right]\\ &=\mathbb{E}\left[\tilde{s}_{\nu,p}(h_n)|\mathcal{X}_n\right]-h_n^{p+1-\nu}\mathsf{B}_{\mathrm{F},\nu,p,p+1}(h_n),\\ B_2&=\mathbb{E}\left[\check{\mathsf{B}}_{\nu,p,q}(h_n,b_n)-\mathsf{B}_{\mathrm{F},\nu,p,p+1}(h_n)|\mathcal{X}_n\right]\\ &=h_n^{p+1-\nu}\left(\mathbb{E}\left[\check{\mathsf{B}}_{\nu,p,q}(h_n,b_n)|\mathcal{X}_n\right]-\mathsf{B}_{\mathrm{F},\nu,p,p+1}(h_n)\right).\end{aligned}$$

Using Lemma S.A.3 and Theorem A.1,

$$B_1=h_n^{p+2-\nu}\mathsf{B}_{\mathrm{F},\nu,p,p+2}(h_n)\left[1+o_p(1)\right]$$

and

$$\begin{aligned}B_2&=\mathbb{E}\left[\check{\mathsf{B}}_{\nu,p,q}(h_n,b_n)-\mathsf{B}_{\mathrm{F},\nu,p,p+1}(h_n)|\mathcal{X}_n\right]\\ &=\left(\frac{1}{\tau_{T,\nu}}\left(\mathbb{E}\left[e'_{p+1}\hat{\beta}_{Y+,q}(b_n)|\mathcal{X}_n\right]-e'_{p+1}\beta_{Y+,q}\right)\mathcal{B}_{+,\nu,p,p+1}(h_n)\right.\\ &\quad\left.-\frac{1}{\tau_{T,\nu}}\left(\mathbb{E}\left[e'_{p+1}\hat{\beta}_{Y-,q}(b_n)|\mathcal{X}_n\right]-e'_{p+1}\beta_{Y-,q}\right)\mathcal{B}_{-,\nu,p,p+1}(h_n)\right)\end{aligned}$$

$$
\begin{aligned}
&- \bigg( \frac{\tau_{Y,\nu}}{\tau_{T,\nu}^2} \big( \mathbb{E}\big[e'_{p+1}\hat{\beta}_{T+,q}(b_n)|\mathcal{X}_n\big] - e'_{p+1}\beta_{T+,q}\big)\mathcal{B}_{+,\nu,p,p+1}(h_n) \\
&- \frac{\tau_{Y,\nu}}{\tau_{T,\nu}^2} \big( \mathbb{E}\big[e'_{p+1}\hat{\beta}_{T-,q}(b_n)|\mathcal{X}_n\big] - e'_{p+1}\beta_{T-,q}\big)\mathcal{B}_{-,\nu,p,p+1}(h_n)\bigg) \\
= {} & \mathsf{B}^{\mathrm{bc}}_{\mathrm{F},\nu,p,p+1,q}(h_n, b_n)
\end{aligned}
$$

because

$$
\begin{aligned}
&\mathbb{E}\big[(p+1)!e'_{p+1}\hat{\beta}_{Y+,q}(b_n)|\mathcal{X}_n\big] \\
&= e'_{p+1}\beta_{Y+,q} + b_n^{q-p}\frac{\mu_+^{(q+1)}}{(q+1)!}\mathcal{B}_{+,p+1,q,q+1}(b_n) + o_p\big(b_n^{q-p}\big),
\end{aligned}
$$

and similarly for $\{Y-, T+, T-\}$. Collecting terms, the result in part (B) follows.

For part (V), first note that $\mathsf{V}^{\mathrm{bc}}_{\mathrm{F},\nu,p,q}(h_n, b_n) = \mathbb{V}[\tilde{\varsigma}^{\mathrm{bc}}_{+,\nu,p,q}(h_n, b_n)|\mathcal{X}_n] + \mathbb{V}[\tilde{\varsigma}^{\mathrm{bc}}_{-,\nu,p,q}(h_n, b_n)|\mathcal{X}_n]$ with

$$
\begin{aligned}
\tilde{\varsigma}^{\mathrm{bc}}_{+,\nu,p,q}(h,b) &= \tilde{\varsigma}_{+,\nu,p}(h) - h_n^{p+1-\nu}\check{\mathsf{B}}_{\mathrm{F},+,\nu,p,q}(h,b) \\
&= \frac{1}{\tau_{T,\nu}}\big(\hat{\mu}^{(\nu)}_{Y+,p}(h) - \mu^{(\nu)}_{Y+}\big) - \frac{\tau_{Y,\nu}}{\tau_{T,\nu}^2}\big(\hat{\mu}^{(\nu)}_{T+,p}(h) - \mu^{(\nu)}_{T+}\big) \\
&\quad - h^{p+1-\nu}\bigg(\frac{1}{\tau_{T,\nu}}e'_{p+1}\hat{\beta}_{Y+,q}(b) - \frac{\tau_{Y,\nu}}{\tau_{T,\nu}^2}e'_{p+1}\hat{\beta}_{T+,q}(b)\bigg) \\
&\quad \times \mathcal{B}_{+,\nu,p,p+1}(h)
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{\varsigma}^{\mathrm{bc}}_{-,\nu,p,q}(h,b) &= \tilde{\varsigma}_{-,\nu,p}(h) - h^{p+1-\nu}\check{\mathsf{B}}_{\mathrm{F},-,\nu,p,q}(h,b) \\
&= \frac{1}{\tau_{T,\nu}}\big(\hat{\mu}^{(\nu)}_{Y-,p}(h) - \mu^{(\nu)}_{Y-}\big) - \frac{\tau_{Y,\nu}}{\tau_{T,\nu}^2}\big(\hat{\mu}^{(\nu)}_{T-,p}(h) - \mu^{(\nu)}_{T-}\big) \\
&\quad - h^{p+1-\nu}\bigg(\frac{1}{\tau_{T,\nu}}e'_{p+1}\hat{\beta}_{Y-,q}(b) - \frac{\tau_{Y,\nu}}{\tau_{T,\nu}^2}e'_{p+1}\hat{\beta}_{T-,q}(b)\bigg) \\
&\quad \times \mathcal{B}_{+,\nu,p,p+1}(h).
\end{aligned}
$$

This decomposition implies that

$$
\mathsf{V}^{\mathrm{bc}}_{\mathrm{F},+,\nu,p,q}(h,b) = \mathbb{V}\big[\tilde{\varsigma}^{\mathrm{bc}}_{+,\nu,p,q}(h,b)|\mathcal{X}_n\big] = V_{+1} + V_{+2} - 2C_{+12},
$$

with

$$V_{+1} = \mathsf{V}_{\mathrm{F},+,\nu,p}(h), \quad V_{+2} = h^{2(p+1-\nu)}\mathsf{V}_{\mathrm{F},+,p+1,q}(b)\frac{\mathcal{B}^2_{+,\nu,p,p+1}(h)}{(p+1)!^2},$$

and

$$\begin{aligned} C_{+12} &= \mathbb{C}\biggl[\frac{1}{\tau_{T,\nu}}\hat{\mu}^{(\nu)}_{Y+,p}(h) - \frac{\tau_{Y,\nu}}{\tau^2_{T,\nu}}\hat{\mu}^{(\nu)}_{T+,p}(h), \\ &\qquad h^{p+1-\nu}\biggl(\frac{1}{\tau_{T,\nu}}e'_{p+1}\hat{\beta}_{Y+,q}(b) - \frac{\tau_{Y,\nu}}{\tau^2_{T,\nu}}e'_{p+1}\hat{\beta}_{T+,q}(b)\biggr) \\ &\qquad \times \mathcal{B}_{+,\nu,p,p+1}(h)\Big|\mathcal{X}_n\biggr] \\ &= h^{p+1-\nu}[C_{+121} - C_{+122} + C_{+124}]\frac{\mathcal{B}_{+,\nu,p,p+1}(h)}{(p+1)!}, \end{aligned}$$

with

$$\begin{aligned} C_{+121} &= \mathbb{C}\biggl[\frac{1}{\tau_{T,\nu}}\nu!e'_{\nu}\hat{\beta}_{Y+,p}(h), \frac{1}{\tau_{T,\nu}}(p+1)!e'_{p+1}\hat{\beta}_{Y+,q}(b)\Big|\mathcal{X}_n\biggr] \\ &= \frac{1}{\tau^2_{T,\nu}}\mathbb{C}\bigl[\nu!e'_{\nu}\hat{\beta}_{Y+,p}(h), (p+1)!e'_{p+1}\hat{\beta}_{Y+,q}(b)|\mathcal{X}_n\bigr] \\ &= \frac{1}{\tau^2_{T,\nu}}\mathcal{C}_{YY+,\nu,p,q}(h,b), \\ C_{+122} &= 2\mathbb{C}\biggl[\frac{1}{\tau_{T,\nu}}\nu!e'_{\nu}\hat{\beta}_{Y+,p}(h), \frac{\tau_{Y,\nu}}{\tau^2_{T,\nu}}(p+1)!e'_{p+1}\hat{\beta}_{T+,q}(b)\Big|\mathcal{X}_n\biggr] \\ &= \frac{2\tau_{Y,\nu}}{\tau^3_{T,\nu}}\mathbb{C}\bigl[\nu!e'_{\nu}\hat{\beta}_{Y+,p}(h), (p+1)!e'_{p+1}\hat{\beta}_{T+,q}(b)|\mathcal{X}_n\bigr] \\ &= \frac{2\tau_{Y,\nu}}{\tau^3_{T,\nu}}\mathcal{C}_{YT+,\nu,p,q}(h,b), \\ C_{+123} &= \mathbb{C}\biggl[\frac{\tau_{Y,\nu}}{\tau^2_{T,\nu}}\nu!e'_{\nu}\hat{\beta}_{T+,p}(h), \frac{\tau_{Y,\nu}}{\tau^2_{T,\nu}}(p+1)!e'_{p+1}\hat{\beta}_{T+,q}(b)\Big|\mathcal{X}_n\biggr] \\ &= \frac{\tau^2_{Y,\nu}}{\tau^4_{T,\nu}}\mathbb{C}\bigl[\nu!e'_{\nu}\hat{\beta}_{T+,p}(h), (p+1)!e'_{p+1}\hat{\beta}_{T+,q}(b)|\mathcal{X}_n\bigr] \\ &= \frac{\tau^2_{Y,\nu}}{\tau^4_{T,\nu}}\mathcal{C}_{TT+,\nu,p,q}(h,b), \end{aligned}$$

because, for example,

$$\begin{aligned}
&\mathbb{C}\big[\nu! e_\nu' \hat{\beta}_{Y+,p}(h), (p+1)! e_{p+1}' \hat{\beta}_{T+,q}(b) | \mathcal{X}_n\big] \\
&\quad = h^{-\nu} \nu!(p+1)! e_\nu' \Gamma_{+,p}^{-1}(h) X_p(h) W_+(h) \mathbb{C}[Y, T | \mathcal{X}_n] \\
&\qquad \times W_+(b) X_q(b) \Gamma_{+,q}^{-1}(b) e_{p+1} b^{-p-1}/n^2 \\
&\quad = \frac{1}{n h^\nu b^{p+1}} \nu!(p+1)! e_\nu' \Gamma_{+,\ell}^{-1}(h) \Psi_{YT+,p,q}(h,b) \Gamma_{+,q}^{-1}(b) e_{p+1},
\end{aligned}$$

and similarly for the other two covariances. Thus,

$$\begin{aligned}
\mathsf{V}^{\mathrm{bc}}_{\mathrm{F},+,\nu,p,q}(h,b) &= \mathsf{V}_{\mathrm{F},+,\nu,p}(h) + h^{2(p+1-\nu)} \mathsf{V}_{\mathrm{F},+,p+1,q}(b) \frac{\mathcal{B}^2_{+,\nu,p,p+1}(h)}{(p+1)!^2} \\
&\quad - 2h^{p+1-\nu} \mathcal{C}_{\mathrm{F},+,\nu,p,q}(h,b) \frac{\mathcal{B}_{+,\nu,p,p+1}(h)}{(p+1)!},
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{C}_{\mathrm{F},+,\nu,p,q}(h,b) &= \frac{1}{\tau^2_{T,\nu}} \mathcal{C}_{YY+,\nu,p,q}(h,b) - \frac{2\tau_{Y,\nu}}{\tau^3_{T,\nu}} \mathcal{C}_{YT+,\nu,p,q}(h,b) \\
&\quad + \frac{\tau^2_{Y,\nu}}{\tau^4_{T,\nu}} \mathcal{C}_{TT+,\nu,p,q}(h,b).
\end{aligned}$$

The term $\mathsf{V}^{\mathrm{bc}}_{\mathrm{F},-,\nu,p,q}(h,b)$ is derived analogously. This completes the derivation of part (V).

Finally, consider part (D). First we show that $R_n^2 = o_p(\mathsf{V}_{\mathrm{F},\nu,p}(h_n))$ and $(R_n^{\mathrm{bc}})^2 = o_p(\mathsf{V}_{\mathrm{F},\nu,p}(h_n))$. Recall that

$$\mathsf{V}_{\mathrm{F},\nu,p}(h_n) = O_p\left(\frac{1}{n h^{1+2\nu}} + \frac{h_n^{2(p+1-\nu)}}{n b_n^{3+2p}}\right).$$

Thus, we have

$$\begin{aligned}
\frac{R_n^2}{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)} &= O_p\left(\min\left\{n h_n^{1+2\nu}, \frac{n b_n^{3+2p}}{h_n^{2(p+1-\nu)}}\right\}\right) \\
&\quad \times O_p\left(\frac{1}{n^2 h_n^{2+4\nu}} + h_n^{4(p+1-\nu)}\right) \\
&= O_p\left(\min\left\{\frac{1}{n h_n^{1+2\nu}}, \frac{b_n^{3+2p}}{n h_n^{2p+4+2\nu}}\right\}\right) \\
&\quad + O_p\left(\min\left\{n h_n^{4p+5-2\nu}, n b_n^{3+2p} h_n^{2(p+1-\nu)}\right\}\right)
\end{aligned}$$

$$= O_p\left(\frac{1}{nh_n^{1+2\nu}} \min\left\{1, \frac{1}{\rho_n^{3+2p}}\right\}\right)$$
$$+ O_p\left(nh_n^{2(p+1-\nu)} \min\left\{h_n^{3+2p}, b_n^{3+2p}\right\}\right)$$
$$= o_p(1),$$

because $nh_n^{1+2\nu} \to \infty$ and $n\min\{h_n^{3+2p}, b_n^{3+2p}\}h_n^{2(p+1-\nu)} = o(n\min\{h_n^{3+2p}, b_n^{3+2p}\} \times h_n^2) \to 0$ for any $p \geq \nu$. Also, we have

$$\frac{(R_n^{\mathrm{bc}})^2}{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)} = O_p\left(\min\left\{nh_n^{1+2\nu}, \frac{nb_n^{3+2p}}{h_n^{2(p+1-\nu)}}\right\}\right)$$
$$\times O_p\left(\frac{h_n^{2p+3}}{nh_n^{2+4\nu}} + h_n^{4(p+1-\nu)}\right) O_p\left(1 + \frac{1}{nb_n^{3+2p}}\right)$$
$$= O_p\left(\frac{h_n^{2p+3}}{nh_n^{1+2\nu}} \min\left\{1, \frac{1}{\rho_n^{3+2p}}\right\}\right.$$
$$\left. + nh_n^{2(p+1-\nu)} \min\left\{h_n^{3+2p}, b_n^{3+2p}\right\}\right)$$
$$+ O_p\left(\frac{h_n^{2p+3}}{nh_n^{1+2\nu}} \min\left\{1, \frac{1}{\rho_n^{3+2p}}\right\}\right.$$
$$\left. + nh_n^{2(p+1-\nu)} \min\left\{h_n^{3+2p}, b_n^{3+2p}\right\}\right) O_p\left(\frac{1}{nb_n^{3+2p}}\right)$$
$$= o_p(1) + O_p\left(\frac{1}{nh_n^{1+2\nu}} \frac{1}{n} \min\left\{\rho_n^{3+2p}, 1\right\}\right.$$
$$\left. + h_n^{2(p+1-\nu)} \min\left\{\rho_n^{3+2p}, 1\right\}\right)$$
$$= o_p(1),$$

using the previous calculations. These results imply

$$\frac{\hat{\mathsf{s}}_{\nu,p,q}^{\mathrm{bc}}(h_n, b_n) - \mathsf{s}_\nu}{\sqrt{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)}} = \frac{\tilde{\mathsf{s}}_{\nu,p}^{\mathrm{bc}}(h_n, b_n)}{\sqrt{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)}} + \frac{R_n + R_n^{\mathrm{bc}}}{\sqrt{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)}}$$
$$= \frac{\tilde{\mathsf{s}}_{\nu,p}^{\mathrm{bc}}(h_n, b_n)}{\sqrt{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)}} + o_p(1),$$

provided that $n\min\{h_n^{2p+3}, b_n^{2p+3}\}\max\{h_n^2, b_n^{2(q-p)}\} \to 0$ and $n\min\{h_n^{1+2\nu}, b_n\} \to \infty$.

Now, proceeding as in Lemma S.A.4 and using the Cramér–Wold device, it can be shown that

$$\frac{\tilde{\varsigma}^{\mathrm{bc}}_{\nu,p}(h_n,b_n)}{\sqrt{\mathsf{V}_{\mathrm{F},\nu,p}(h_n)}} = \frac{\tilde{\varsigma}^{\mathrm{bc}}_{\nu,p}(h_n,b_n)}{\sqrt{\mathbb{V}[\tilde{\varsigma}^{\mathrm{bc}}_{\nu,p}(h_n,b_n)|\mathcal{X}_n]}} \to_d \mathcal{N}(0,1),$$

from which the result in part (D) follows, provided that $n\min\{h_n^{2p+3}, b_n^{2p+3}\} \times \max\{h_n^2, b_n^{2(q-p)}\} \to 0$ and $n\min\{h_n, b_n\} \to \infty$. *Q.E.D.*

## S.2.4. *Consistent Standard Error Estimators*

As explained in Section 5 of CCT, consistent standard errors may be constructed by replacing the matrices

$$\begin{aligned}\Psi_{UV+,p,q}(h_n,b_n) &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i \geq 0)K_{h_n}(X_i)K_{b_n}(X_i)\\ &\quad\times r_p(X_i/h_n)r_q(X_i/b_n)'\sigma^2_{UV+}(X_i),\\ \Psi_{UV-,p,q}(h_n,b_n) &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i < 0)K_{h_n}(X_i)K_{b_n}(X_i)\\ &\quad\times r_p(X_i/h_n)r_q(X_i/b_n)'\sigma^2_{UV-}(X_i),\end{aligned}$$

with appropriate consistent estimators thereof, where $\sigma^2_{UV+}(x) = \mathbb{C}\mathrm{ov}[U(1), V(1)|X = x]$ and $\sigma^2_{UV-}(x) = \mathbb{C}\mathrm{ov}[U(0), V(0)|X = x]$, and $U$ and $V$ are placeholders for either $Y$ or $T$. A natural choice is to employ estimated residuals, leading to

$$\begin{aligned}\check{\Psi}_{UV+,p,q}(h_n,b_n) &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i \geq 0)K_{h_n}(X_i)K_{b_n}(X_i)\\ &\quad\times r_p(X_i/h_n)r_q(X_i/b_n)'\check{\varepsilon}_{U+,i}\check{\varepsilon}_{V+,i},\\ \check{\Psi}_{UV-,p,q}(h_n,b_n) &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i < 0)K_{h_n}(X_i)K_{b_n}(X_i)\\ &\quad\times r_p(X_i/h_n)r_q(X_i/b_n)'\check{\varepsilon}_{U-,i}\check{\varepsilon}_{V-,i},\end{aligned}$$

where $\check{\varepsilon}_{U+,i}$, $\check{\varepsilon}_{V+,i}$, $\check{\varepsilon}_{U-,i}$, and $\check{\varepsilon}_{V-,i}$ are consistent residual estimators of their population counterparts, with $U$ and $V$ placeholders for either $Y$ or $T$. This type of approach can be used to construct standard error estimators using conventional nonparametric techniques. For example, in Theorem A.1, a consistent estimator of $\mathsf{V}_{\nu,p}(h_n)$ using this approach is $\check{\mathsf{V}}_{\nu,p}(h_n) = \check{\mathcal{V}}_{+,\nu,p}(h_n) +$

$\check{\mathcal{V}}_{-,\nu,p}(h_n)$, where

$$\check{\mathcal{V}}_{+,\nu,p}(h_n) = \frac{1}{nh_n^{2\nu}}\nu!^2 e_\nu' \Gamma_{+,p}^{-1}(h_n)\check{\Psi}_{YY+,p,q}(h_n, h_n)\Gamma_{+,p}^{-1}(h_n)e_\nu$$

and

$$\check{\mathcal{V}}_{-,\nu,p}(h_n) = \frac{1}{nh_n^{2\nu}}\nu!^2 e_\nu' \Gamma_{-,p}^{-1}(h_n)\check{\Psi}_{YY-,p,q}(h_n, h_n)\Gamma_{-,p}^{-1}(h_n)e_\nu,$$

where

$$\begin{aligned}\check{\Psi}_{YY+,p,q}(h_n, h_n) &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i \geq 0)K_{h_n}(X_i)K_{h_n}(X_i) \\ &\quad \times r_p(X_i/h_n)r_q(X_i/h_n)'\check{\varepsilon}_{+,i}^2, \\ \check{\Psi}_{YY-,p,q}(h_n, h_n) &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i < 0)K_{h_n}(X_i)K_{h_n}(X_i) \\ &\quad \times r_p(X_i/h_n)r_q(X_i/h_n)'\check{\varepsilon}_{-,i}^2,\end{aligned}$$

where $\check{\varepsilon}_{+,i} = Y_i - \hat{\mu}_{+,p}(h_n)$ and $\check{\varepsilon}_{-,i} = Y_i - \hat{\mu}_{-,p}(h_n)$. In this case, the matrices $\mathcal{C}_{+,\nu,p,q}(h, b)$, $\mathcal{C}_{-,\nu,p,q}(h, b)$, $\mathcal{V}_{+,p+1,q}(b)$, and $\mathcal{V}_{-,p+1,q}(b)$ are estimated using the same logic. This approach leads to a consistent estimator of $\mathsf{V}_{\nu,p,q}^{\text{bc}}(h_n, b_n) = \mathbb{V}[\hat{\tau}_{\nu,p,q}^{\text{bc}}(h_n, b_n)|\mathcal{X}_n]$ based on plug-in estimated residuals, which is the common way of constructing consistent standard errors in nonparametrics in general, and in RD applications in particular. This approach can be implemented directly by using general purpose software for linear regression estimation. It is easy to verify that the resulting standard error estimators are consistent under both conventional asymptotics and our asymptotics, with and without bias correction.

We propose an alternative standard error estimator based on ideas in Abadie and Imbens (2006). Specifically, we consider

$$\begin{aligned}\hat{\Psi}_{UV+,p,q}(h_n, b_n) &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i \geq 0)K_{h_n}(X_i)K_{b_n}(X_i) \\ &\quad \times r_p(X_i/h_n)r_q(X_i/b_n)'\hat{\sigma}_{UV+}^2(X_i), \\ \hat{\Psi}_{UV-,p,q}(h_n, b_n) &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i < 0)K_{h_n}(X_i)K_{b_n}(X_i) \\ &\quad \times r_p(X_i/h_n)r_q(X_i/b_n)'\hat{\sigma}_{UV-}^2(X_i),\end{aligned}$$

with

$$\hat{\sigma}^2_{UV+}(X_i) = \mathbf{1}(X_i \geq 0)\frac{J}{J+1} \times \left(U_i - \sum_{j=1}^{J} U_{\ell_{+,j}(i)}/J\right)\left(V_i - \sum_{j=1}^{J} V_{\ell_{+,j}(i)}/J\right),$$

$$\hat{\sigma}^2_{UV-}(X_i) = \mathbf{1}(X_i < 0)\frac{J}{J+1} \times \left(U_i - \sum_{j=1}^{J} U_{\ell_{-,j}(i)}/J\right)\left(V_i - \sum_{j=1}^{J} V_{\ell_{-,j}(i)}/J\right),$$

where, again, $U$ and $V$ are placeholders for either $Y$ or $T$. The construction of the standard error estimator is exactly as explained above, and is justified by the following theorem.

THEOREM A.3—Fixed NN-Based Standard Error Estimators:
(Sharp RD) *Suppose the conditions in Theorem A.1(D) hold. If, in addition, $\sigma^2_+(x)$ and $\sigma^2_-(x)$ are Lipschitz continuous on $(-\kappa_0, \kappa_0)$, then*

$$\hat{\Psi}_{YY+,p,q}(h_n, b_n) = \Psi_{YY+,p,q}(h_n, b_n) + o_p\big(\min\{h_n^{-1}, b_n^{-1}\}\big),$$
$$\hat{\Psi}_{YY-,p,q}(h_n, b_n) = \Psi_{YY-,p,q}(h_n, b_n) + o_p\big(\min\{h_n^{-1}, b_n^{-1}\}\big).$$

(Fuzzy RD) *Suppose the conditions in Theorem A.2(D) hold. If, in addition, $\sigma^2_{UV+}(x)$ and $\sigma^2_{UV-}(x)$ are Lipschitz continuous on $(-\kappa_0, \kappa_0)$, then*

$$\hat{\Psi}_{UV+,p,q}(h_n, b_n) = \Psi_{UV+,p,q}(h_n, b_n) + o_p\big(\min\{h_n^{-1}, b_n^{-1}\}\big),$$
$$\hat{\Psi}_{UV-,p,q}(h_n, b_n) = \Psi_{UV-,p,q}(h_n, b_n) + o_p\big(\min\{h_n^{-1}, b_n^{-1}\}\big)$$

*for $U = Y, T$ and $V = Y, T$.*

PROOF: We only prove the result for $\hat{\Psi}_{YT+,p,q}(h_n, b_n)$ because the proof of the other cases is analogous. First note that, for all $X_i \geq 0$,

$$\hat{\sigma}^2_{YT+}(X_i) = \frac{J}{J+1}\left(Y_i - \frac{1}{J}\sum_{j=1}^{J} Y_{\ell_{+,j}(i)}\right)\left(T_i - \frac{1}{J}\sum_{j=1}^{J} T_{\ell_{+,j}(i)}\right)$$

$$= \frac{J}{J+1}\left(\varepsilon_{Y,i} - \frac{1}{J}\sum_{j=1}^{J}\varepsilon_{Y,\ell_j^+(i)} + \frac{1}{J}\sum_{j=1}^{J}\bigl(\mu_{Y+}(X_i) - \mu_{Y+}(X_{\ell_j^+(i)})\bigr)\right)$$
$$\times \left(\varepsilon_{T,i} - \frac{1}{J}\sum_{j=1}^{J}\varepsilon_{T,\ell_j^+(i)} + \frac{1}{J}\sum_{j=1}^{J}\bigl(\mu_{T+}(X_i) - \mu_{T+}(X_{\ell_j^+(i)})\bigr)\right),$$

and therefore

$$\hat{\sigma}^2_{YT+}(X_i) = \varepsilon_{Y,i}\varepsilon_{T,i} + \hat{\sigma}^2_{1,i} + \hat{\sigma}^2_{2,i} + \hat{\sigma}^2_{3,i} - \hat{\sigma}^2_{4,i} - \hat{\sigma}^2_{5,i} + \hat{\sigma}^2_{6,i} + \hat{\sigma}^2_{7,i} - \hat{\sigma}^2_{8,i} - \hat{\sigma}^2_{9,i},$$

with

$$\hat{\sigma}^2_{1,i} = \frac{1}{J(J+1)}\sum_{j=1}^{J}(\varepsilon_{Y,\ell_j^+(i)}\varepsilon_{T,\ell_j^+(i)} - \varepsilon_{Y,i}\varepsilon_{T,i}),$$

$$\hat{\sigma}^2_{2,i} = \frac{2}{J(J+1)}\sum_{1\le j<k\le J}\varepsilon_{Y,\ell_j^+(i)}\varepsilon_{T,\ell_k^+(i)},$$

$$\hat{\sigma}^2_{3,i} = \frac{1}{J(J+1)}\left(\sum_{j=1}^{J}\bigl(\mu_{Y+}(X_i) - \mu_{Y+}(X_{\ell_j^+(i)})\bigr)\right) \times \left(\sum_{j=1}^{J}\bigl(\mu_{T+}(X_i) - \mu_{T+}(X_{\ell_j^+(i)})\bigr)\right),$$

$$\hat{\sigma}^2_{4,i} = \varepsilon_{Y,i}\frac{1}{J+1}\sum_{j=1}^{J}\varepsilon_{T,\ell_j^+(i)}, \qquad \hat{\sigma}^2_{5,i} = \varepsilon_{T,i}\frac{1}{J+1}\sum_{j=1}^{J}\varepsilon_{Y,\ell_j^+(i)},$$

$$\hat{\sigma}^2_{6,i} = \varepsilon_{Y,i}\frac{1}{J+1}\sum_{j=1}^{J}\bigl(\mu_{T+}(X_i) - \mu_{T+}(X_{\ell_j^+(i)})\bigr),$$

$$\hat{\sigma}^2_{7,i} = \varepsilon_{T,i}\frac{1}{J+1}\sum_{j=1}^{J}\bigl(\mu_{Y+}(X_i) - \mu_{Y+}(X_{\ell_j^+(i)})\bigr),$$

$$\hat{\sigma}^2_{8,i} = \frac{1}{J(J+1)}\sum_{j=1}^{J}\varepsilon_{Y,\ell_j^+(i)}\sum_{k=1}^{J}\bigl(\mu_{T+}(X_i) - \mu_{T+}(X_{\ell_k^+(i)})\bigr),$$

$$\hat{\sigma}^2_{9,i} = \frac{1}{J(J+1)}\sum_{j=1}^{J}\varepsilon_{T,\ell_j^+(i)}\sum_{k=1}^{J}\bigl(\mu_{Y+}(X_i) - \mu_{Y+}(X_{\ell_k^+(i)})\bigr).$$

Thus, using the expansion of $\hat{\sigma}^2_+(X_i)$, we obtain

$$\hat{\Psi}_{+,p,q}(h_n,b_n) = \Psi_{+,p,q}(h_n,b_n) + \eta_{1,n} + \eta_{2,n} + \eta_{3,n} - \eta_{4,n} + \eta_{5,n} - \eta_{6,n},$$

with

$$\eta_{1,n} = \frac{1}{J(J+1)} \sum_{j=1}^{J} \eta_{1,n,j},$$

$$\eta_{1,n,j} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \geq 0) K_{h_n}(X_i) K_{b_n}(X_i) \times (\varepsilon_{Y,\ell_j^+(i)} \varepsilon_{T,\ell_j^+(i)} - \varepsilon_{Y,i}\varepsilon_{T,i}) r_p(X_i/h_n) r_q(X_i/b_n)',$$

$$\eta_{2,n} = \frac{2}{J(J+1)} \sum_{1\leq j<k\leq J} \eta_{2,n,j,k},$$

$$\eta_{2,n,j,k} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \geq 0) K_{h_n}(X_i) K_{b_n}(X_i) \times [\varepsilon_{Y,\ell_j^+(i)} \varepsilon_{T,\ell_k^+(i)}] r_p(X_i/h_n) r_q(X_i/b_n)',$$

$$\eta_{3,n} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \geq 0) K_{h_n}(X_i) K_{b_n}(X_i) \times [\hat{\sigma}^2_{3,i}] r_p(X_i/h_n) r_q(X_i/b_n)',$$

$$\eta_{4,n} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \geq 0) K_{h_n}(X_i) K_{b_n}(X_i) \times [\hat{\sigma}^2_{4,i} + \hat{\sigma}^2_{5,i}] r_p(X_i/h_n) r_q(X_i/b_n)',$$

$$\eta_{5,n} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \geq 0) K_{h_n}(X_i) K_{b_n}(X_i) \times [\hat{\sigma}^2_{6,i} + \hat{\sigma}^2_{7,i}] r_p(X_i/h_n) r_q(X_i/b_n)',$$

$$\eta_{6,n} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \geq 0) K_{h_n}(X_i) K_{b_n}(X_i) \times [\hat{\sigma}^2_{8,i} + \hat{\sigma}^2_{9,i}] r_p(X_i/h_n) r_q(X_i/b_n)'.$$

Therefore, it suffices to show that $\eta_{l,n} = o_p(\min\{h_n^{-1}, b_n^{-1}\})$ for $l = 1, \ldots, 6$.

The rest of the proof uses the following result in Abadie and Imbens (2006): for any (fixed) $J \in \mathbb{N}_+$,

$$\max_{1\le j\le J}\max_{1\le i\le n:X_i\ge 0}|X_{\ell_{+,j}(i)}-X_i| = o_p(1).$$

For the first reminder, if $n\min\{h_n, b_n\}\to\infty$ and $\kappa\max\{h_n, b_n\} < \kappa_0$,

$$\begin{aligned}
&\mathbb{E}[\eta_{1,n}|\mathcal{X}_n] \\
&\quad = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i\ge 0)K_{h_n}(X_i)K_{b_n}(X_i) \\
&\qquad \times \frac{1}{J(J+1)}\sum_{j=1}^{J}\big(\mathbb{E}[\varepsilon_{Y,\ell_j^+(i)}\varepsilon_{T,\ell_j^+(i)}-\varepsilon_{Y,i}\varepsilon_{T,i}|\mathcal{X}_n]\big) \\
&\qquad \times r_p(X_i/h_n)r_q(X_i/b_n)' \\
&\quad = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i\ge 0)K_{h_n}(X_i)K_{b_n}(X_i) \\
&\qquad \times \frac{1}{J(J+1)}\sum_{j=1}^{J}\big(\sigma_{YT}^2(X_{\ell_{+,j}(i)})-\sigma_{YT}^2(X_i)\big) \\
&\qquad \times r_p(X_i/h_n)r_q(X_i/b_n)' \\
&\quad = o_p\big(\min\{h_n^{-1}, b_n^{-1}\}\big),
\end{aligned}$$

because, by Lemma S.A.2,

$$\begin{aligned}
\big|\mathbb{E}[\eta_{1,n}|\mathcal{X}_n]\big| &\lesssim \max_{1\le j\le J}\max_{1\le i\le n:X_i\ge 0}|X_{\ell_{+,j}(i)}-X_i| \\
&\qquad \times \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i\ge 0)K_{h_n}(X_i)K_{b_n}(X_i) \\
&\qquad \times \big|r_p(X_i/h_n)\big|\big|r_q(X_i/b_n)\big| \\
&= o_p(1)O_p\left(\frac{\min\{h_n, b_n\}}{h_n b_n}\right) = o_p\big(\min\{h_n^{-1}, b_n^{-1}\}\big).
\end{aligned}$$

Thus, $\eta_{1,n} = o_p(\min\{h_n^{-1}, b_n^{-1}\})$ because, by Lemma S.A.2,

$$\begin{aligned}
&\mathbb{E}\big[\big|\eta_{1,n}-\mathbb{E}[\eta_{1,n}|\mathcal{X}_n]\big|^2\big] \\
&\quad \lesssim n^{-1}\int_0^\infty K_{h_n}^2(x)K_{b_n}^2(x)\big|r_p(x)\big|^2\big|r_q(x)\big|^2 f(x)\,\mathrm{d}u
\end{aligned}$$

$$
\begin{aligned}
&= n^{-1} O\left(\frac{\min\{h_n, b_n\}}{h_n^2 b_n^2}\right) = O\left(\frac{\min\{h_n^{-2}, b_n^{-2}\}}{n \min\{h_n, b_n\}}\right) \\
&= o\big(\min\{h_n^{-2}, b_n^{-2}\}\big),
\end{aligned}
$$

provided $n \min\{h_n, b_n\} \to \infty$ and $\kappa \max\{h_n, b_n\} < \kappa_0$.

Similarly, $\mathbb{E}[\eta_{2,n}|\mathcal{X}_n] = 0$ and, proceeding as above, $\mathbb{V}[\eta_{2,n}|\mathcal{X}_n] = o_p(\min\{h_n^{-1}, b_n^{-1}\})$. Therefore, $\eta_{2,n} = o_p(\min\{h_n^{-1}, b_n^{-1}\})$.

Next, for $\eta_{3,n}$, simply note that

$$
\begin{aligned}
|\eta_{3,n}| &\precsim \max_{1\le j\le J} \max_{1\le i\le n: X_i\ge 0} |X_{\ell_{+,j}(i)} - X_i|^2 \\
&\quad \times \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \ge 0) K_{h_n}(X_i) K_{b_n}(X_i) \big|r_p(X_i/h_n)\big| \big|r_q(X_i/b_n)\big| \\
&\le o_p(1) O_p\big(\min\{h_n^{-1}, b_n^{-1}\}\big) = o_p\big(\min\{h_n^{-1}, b_n^{-1}\}\big).
\end{aligned}
$$

The last three (cross-product) terms can be analyzed analogously. *Q.E.D.*

### S.2.5. *Proofs of Lemma 1 and Lemma 2*

First we prove Lemma 1, which we restate next. Recall the definition

$$
\begin{aligned}
\mathsf{MSE}_{\nu,p,s}(\mathsf{h}_n) = \mathbb{E}\Big[\Big(\big(\hat{\mu}_{+,p}^{(\nu)}(\mathsf{h}_n) - (-1)^s \hat{\mu}_{-,p}^{(\nu)}(\mathsf{h}_n)\big) \\
- \big(\mu_{+}^{(\nu)} - (-1)^s \mu_{-}^{(\nu)}\big)\Big)^2 \Big| \mathcal{X}_n\Big],
\end{aligned}
$$

where $\mathsf{h}_n$ is a bandwidth sequence. Also,

$$
\begin{aligned}
\mathsf{B}_{\nu,p,r,s} &= \frac{\mu_{+}^{(r)} - (-1)^{\nu+r+s}\mu_{-}^{(r)}}{r!} \nu! e_\nu' \Gamma_p^{-1} \vartheta_{p,r}, \\
\mathsf{V}_{\nu,p} &= \frac{\sigma_-^2 + \sigma_+^2}{f} \nu!^2 e_\nu' \Gamma_p^{-1} \Psi_p \Gamma_p^{-1} e_\nu.
\end{aligned}
$$

LEMMA 1: *Suppose Assumptions* 1–2 *hold with* $S \ge p+1$, *and* $\nu \le p$. *If* $\mathsf{h}_n \to 0$ *and* $n\mathsf{h}_n \to \infty$, *then*

$$
\mathsf{MSE}_{\nu,p,s}(\mathsf{h}_n) = \mathsf{h}_n^{2(p+1-\nu)}\big[\mathsf{B}_{\nu,p,p+1,s}^2 + o_p(1)\big] + \frac{1}{n\mathsf{h}_n^{1+2\nu}}\big[\mathsf{V}_{\nu,p} + o_p(1)\big].
$$

*If, in addition*, $\mathsf{B}_{\nu,p,p+1,s} \neq 0$, *then the* (*asymptotic*) *MSE-optimal bandwidth is*

$$
\mathsf{h}_{\mathtt{MSE},\nu,p,s} = \mathsf{C}_{\mathtt{MSE},\nu,p,s}^{1/(2p+3)} n^{-1/(2p+3)}, \quad \mathsf{C}_{\mathtt{MSE},\nu,p,s} = \frac{(1+2\nu)\mathsf{V}_{\nu,p}}{2(p+1-\nu)\mathsf{B}_{\nu,p,p+1,s}^2}.
$$

PROOF: We have

$$\begin{aligned}\mathsf{MSE}_{\nu,p,s}(\mathsf{h}_n) &= \mathbb{E}\big[\big(\nu!e_\nu'\big(\hat{\beta}_{+,p}(\mathsf{h}_n)-(-1)^s\hat{\beta}_{-,p}(\mathsf{h}_n)\big)\\ &\quad -\nu!e_\nu'\big(\beta_{+,p}-(-1)^s\beta_{-,p}\big)\big)^2|\mathcal{X}_n\big]\\ &= \nu!^2\mathbb{V}\big[e_\nu'\big(\hat{\beta}_{+,p}(\mathsf{h}_n)-(-1)^s\hat{\beta}_{-,p}(\mathsf{h}_n)\big)|\mathcal{X}_n\big]\\ &\quad +\nu!^2\big(\mathbb{E}\big[e_\nu'\big(\hat{\beta}_{+,p}(\mathsf{h}_n)-(-1)^s\hat{\beta}_{-,p}(\mathsf{h}_n)\big)\\ &\quad -e_\nu'\big(\beta_{+,p}-(-1)^s\beta_{-,p}\big)|\mathcal{X}_n\big]\big)^2,\end{aligned}$$

where, by Lemma S.A.3, we verify

$$\begin{aligned}&\mathbb{V}\big[e_\nu'\big(\hat{\beta}_{+,p}(\mathsf{h}_n)-(-1)^s\hat{\beta}_{-,p}(\mathsf{h}_n)\big)|\mathcal{X}_n\big]\\ &\quad = \mathbb{V}\big[e_\nu'\hat{\beta}_{+,p}(\mathsf{h}_n)|\mathcal{X}_n\big]+\mathbb{V}\big[\hat{\beta}_{-,p}(\mathsf{h}_n)|\mathcal{X}_n\big]\\ &\quad = \frac{1}{n\mathsf{h}_n^{1+2\nu}}\frac{\sigma_-^2+\sigma_+^2}{f}e_\nu'\Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\big\{1+o_p(1)\big\}\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}\big[e_\nu'\big(\hat{\beta}_{+,p}(\mathsf{h}_n)-(-1)^s\hat{\beta}_{-,p}(\mathsf{h}_n)\big)-e_\nu'\big(\beta_{+,p}-(-1)^s\beta_{-,p}\big)|\mathcal{X}_n\big]\\ &\quad = \mathbb{E}\big[e_\nu'\big(\hat{\beta}_{+,p}(\mathsf{h}_n)-\beta_{+,p}\big)|\mathcal{X}_n\big]-(-1)^s\mathbb{E}\big[e_\nu'\big(\hat{\beta}_{-,p}(\mathsf{h}_n)-\beta_{-,p}\big)|\mathcal{X}_n\big]\\ &\quad = \mathsf{h}_n^{p+1-\nu}\frac{\mu_+^{(p+1)}-(-1)^{\nu+p+s}\mu_-^{(p+1)}}{(p+1)!}e_\nu'\Gamma_p^{-1}\vartheta_{p,p+1}\big\{1+o_p(1)\big\}.\end{aligned}$$

This completes the derivation. *Q.E.D.*

Next, we consider Lemma 2. Recall that $\mathsf{MSE}_{\nu,p}(\mathsf{h}_n)=\mathbb{E}[(\tilde{\varsigma}_{\nu,p}(\mathsf{h}_n))^2|\mathcal{X}_n]$, with

$$\tilde{\varsigma}_{\nu,p}(\mathsf{h}_n)=\frac{1}{\tau_{T,\nu}}\big(\hat{\tau}_{Y,\nu,p}(\mathsf{h}_n)-\tau_{Y,\nu}\big)-\frac{\tau_{Y,\nu}}{\tau_{T,\nu}^2}\big(\hat{\tau}_{T,\nu,p}(\mathsf{h}_n)-\tau_{T,\nu}\big).$$

LEMMA 2: *Suppose Assumptions* 1–3 *hold with* $S\geq p+1$, *and* $\nu\leq p$. *If* $\mathsf{h}_n\to 0$ *and* $n\mathsf{h}_n\to\infty$, *then*

$$\begin{aligned}\mathbb{E}\big[\big(\tilde{\varsigma}_{\nu,p}(\mathsf{h}_n)\big)^2|\mathcal{X}_n\big] &= \mathsf{h}_n^{2(p+1-\nu)}\big[\mathsf{B}_{\mathrm{F},\nu,p,p+1}^2+o_p(1)\big]\\ &\quad +\frac{1}{n\mathsf{h}_n^{1+2\nu}}\big[\mathsf{V}_{\mathrm{F},\nu,p}+o_p(1)\big],\end{aligned}$$

*where*

$$\mathsf{B}_{\mathrm{F},\nu,p,r} = \Bigg(\frac{1}{\tau_{T,\nu}}\frac{\mu_{Y+}^{(r)} - (-1)^{\nu+r}\mu_{Y-}^{(r)}}{r!} - \frac{\tau_{Y,\nu}}{\tau_{T,\nu}^2}\frac{\mu_{T+}^{(r)} - (-1)^{\nu+r}\mu_{T-}^{(r)}}{r!}\Bigg)\nu! e_\nu' \Gamma_p^{-1}\vartheta_{p,r},$$

$$\mathsf{V}_{\mathrm{F},\nu,p} = \Bigg(\frac{1}{\tau_{T,\nu}}\frac{\sigma_{YY-}^2 + \sigma_{YY+}^2}{f} - \frac{2\tau_{Y,\nu}}{\tau_{T,\nu}^3}\frac{\sigma_{YT-}^2 + \sigma_{YT+}^2}{f} + \frac{\tau_{Y,\nu}^2}{\tau_{T,\nu}^4}\frac{\sigma_{TT-}^2 + \sigma_{TT+}^2}{f}\Bigg)\nu!^2 e_\nu' \Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu.$$

*If, in addition*, $\mathsf{B}_{\mathrm{F},\nu,p,p+1} \neq 0$, *then the* (*asymptotic*) *MSE-optimal bandwidth is*

$$\mathsf{h}_{\mathrm{MSE,F},\nu,p} = \mathsf{C}_{\mathrm{MSE,F},\nu,p}^{1/(2p+3)} n^{-1/(2p+3)}, \quad \mathsf{C}_{\mathrm{MSE,F},\nu,p} = \frac{(2\nu+1)\mathsf{V}_{\mathrm{F},\nu,p}}{2(p+1-\nu)\mathsf{B}_{\mathrm{F},\nu,p,p+1}^2}.$$

PROOF: Observe that

$$\mathsf{MSE}_{\nu,p}(\mathsf{h}_n) = \mathbb{V}\big[\tilde{\varsigma}_{\nu,p}(\mathsf{h}_n)|\mathcal{X}_n\big] + \big(\mathbb{E}\big[\tilde{\varsigma}_{\nu,p}(\mathsf{h}_n)|\mathcal{X}_n\big]\big)^2.$$

Using Lemma A.2, we obtain

$$\begin{aligned}\mathsf{V}_{\mathrm{F},\nu,p} &= \mathbb{V}\big[\tilde{\varsigma}_{\nu,p}(\mathsf{h}_n)|\mathcal{X}_n\big] \\ &= \frac{\nu!^2}{n\mathsf{h}_n^{1+2\nu}}\Bigg(\frac{1}{\tau_{T,\nu}}\frac{\sigma_{YY-}^2 + \sigma_{YY+}^2}{f} - \frac{2\tau_{Y,\nu}}{\tau_{T,\nu}^3}\frac{\sigma_{YT-}^2 + \sigma_{YT+}^2}{f} \\ &\quad + \frac{\tau_{Y,\nu}^2}{\tau_{T,\nu}^4}\frac{\sigma_{TT-}^2 + \sigma_{TT+}^2}{f}\Bigg) e_\nu' \Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\big\{1 + o_p(1)\big\}\end{aligned}$$

and

$$\begin{aligned}\mathsf{B}_{\mathrm{F},\nu,p,p+1} &= \mathbb{E}\big[\tilde{\varsigma}_{\nu,p}(\mathsf{h}_n)|\mathcal{X}_n\big] = \mathsf{h}_n^{p+1-\nu}\mathsf{B}_{\mathrm{F},\nu,p,p+1}(\mathsf{h}_n)\big\{1 + o_p(1)\big\} \\ &= \mathsf{h}_n^{p+1-\nu}\nu!\Bigg(\frac{1}{\tau_{T,\nu}}\frac{\mu_{Y+}^{(p+1)} - (-1)^{\nu+p+1}\mu_{Y-}^{(p+1)}}{(p+1)!} \\ &\quad - \frac{\tau_{Y,\nu}}{\tau_{T,\nu}^2}\frac{\mu_{T+}^{(p+1)} - (-1)^{\nu+p+1}\mu_{T-}^{(p+1)}}{(p+1)!}\Bigg)\big\{1 + o_p(1)\big\},\end{aligned}$$

which completes the derivation. *Q.E.D.*

S.2.6. *Consistent Bandwidth Selection for Sharp RD Designs*

We propose consistent, MSE-optimal bandwidth selectors for sharp RD designs. Recall that $\nu \leq p < q$. As we explain in the main paper, we construct both an MSE-optimal choice of bandwidth $h_n$ for the RD point estimator $\hat{\tau}_{\nu,p}(h_n)$ and an MSE-optimal choice of bandwidth $b_n$ for the leading bias of RD point estimator, which depends on the preliminary estimator $\hat{\mathsf{b}}_{\nu,p,q}(b_n) = \hat{\mu}^{(p+1)}_{+,q}(b_n) - (-1)^{\nu+p+1}\hat{\mu}^{(p+1)}_{-,q}(b_n)$.

For any $\nu \leq p$, let $\hat{\mathsf{V}}_{\nu,p}(h_n) = \hat{\mathcal{V}}_{+,\nu,p}(h_n) + \hat{\mathcal{V}}_{-,\nu,p}(h_n)$, with

$$\hat{\mathcal{V}}_{+,\nu,p}(h_n) = \nu!^2 e'_\nu \Gamma^{-1}_{+,p}(h_n)\hat{\Psi}_{YY+,p}(h_n)\Gamma^{-1}_{+,p}(h_n)e_\nu/nh_n^{2\nu},$$

$$\hat{\mathcal{V}}_{-,\nu,p}(h_n) = \nu!^2 e'_\nu \Gamma^{-1}_{-,p}(h_n)\hat{\Psi}_{YY-,p}(h_n)\Gamma^{-1}_{-,p}(h_n)e_\nu/nh_n^{2\nu},$$

where $\hat{\Psi}_{YY+,p}(h_n)$ and $\hat{\Psi}_{YY-,p}(h_n)$ are constructed using the nearest-neighbor approach described in Section 5 of CCT and in Section S.2.4 above.

*Plug-in Bandwidths Selectors*

Fix $\nu, p, q \in \mathbb{N}$ with $\nu \leq p < q$. Let $\mathcal{B}_{\nu,p} = \nu! e'_\nu \Gamma_p^{-1} \vartheta_{p,p+1}$.

*Step 0: Initial bandwidths* $(v_n, c_n)$. (i) Suppose $v_n \to_p 0$ and $nv_n \to_p \infty$. In particular, let

$$v_n = 2.58 \cdot \min\{S_X, IQR_X/1.349\} \cdot n^{-1/5},$$

where $S_X^2$ and $IQR_X$ denote, respectively, the sample variance and interquartile range of $\{X_i : 1 \leq i \leq n\}$.

(ii) Suppose $c_n \to_p 0$ and $nc_n^{2q+3} \to_p \infty$. In particular, let

$$c_n = \check{C}^{1/(2q+5)}_{\nu,p,q} n^{-1/(2q+5)},$$

$$\check{C}_{\nu,p,q} = \frac{(2q+3)nv_n^{2q+3}\hat{\mathsf{V}}_{q+1,q+1}(v_n)}{2\mathcal{B}^2_{q+1,q+1}(e'_{q+2}\hat{\gamma}_{+,q+2} - (-1)^{\nu+q}e'_{q+2}\hat{\gamma}_{-,q+2})^2},$$

where $\hat{\gamma}_{+,m}$ and $\hat{\gamma}_{-,m}$ denote the estimated coefficients of an $(m+1)$th-order global polynomial fit at either side of the threshold:

$$\hat{\gamma}_{+,m} = \arg\min_{\gamma\in\mathbb{R}^{m+1}} \sum_{i=1}^{n} \mathbf{1}(X_i \geq 0)\big(Y_i - r_m(X_i)'\gamma\big)^2,$$

$$\hat{\gamma}_{-,m} = \arg\min_{\gamma\in\mathbb{R}^{m+1}} \sum_{i=1}^{n} \mathbf{1}(X_i < 0)\big(Y_i - r_m(X_i)'\gamma\big)^2,$$

with $r_m(x) = (1, x, x^2, \dots, x^m)'$.

*Step 1: Pilot bandwidth $b_n$*. We estimate $b_n = \mathsf{h}_{\texttt{MSE},p+1,q,\nu+p+1}$. Compute

$$\begin{aligned}
\hat{b}_{\nu,p,q} &= \hat{C}_{\nu,p,q}^{1/(2q+3)} n^{-1/(2q+3)}, \\
\hat{C}_{\nu,p,q} &= (2p+3) n v_n^{2p+3} \hat{\mathsf{V}}_{p+1,q}(v_n) \\
&\quad \Big/ \Big( 2(q-p) \\
&\quad \times \mathcal{B}_{p+1,q}^2 \big[ \big( e'_{q+1} \hat{\beta}_{+,q+1}(c_n) - (-1)^{\nu+q+1} e'_{q+1} \hat{\beta}_{-,q+1}(c_n) \big)^2 \\
&\quad + 3 \hat{\mathsf{V}}_{q+1,q+1}(c_n) \big] \Big).
\end{aligned}$$

*Step 2: Main bandwidth $h_n$*. We estimate $h_n = \mathsf{h}_{\texttt{MSE},\nu,p,0}$. Set $b_n = \hat{b}_{\nu,p,q}$, and compute

$$\begin{aligned}
\hat{h}_{\nu,p} &= \hat{C}_{\nu,p}^{1/(2p+3)} n^{-1/(2p+3)}, \\
\hat{C}_{\nu,p} &= (2\nu+1) n v_n^{1+2\nu} \hat{\mathsf{V}}_{\nu,p}(v_n) \\
&\quad \Big/ \Big( 2(p+1-\nu) \\
&\quad \times \mathcal{B}_{\nu,p}^2 \big[ \big( e'_{p+1} \hat{\beta}_{+,p+1}(b_n) - (-1)^{\nu+p+1} e'_{p+1} \hat{\beta}_{-,p+1}(b_n) \big)^2 \\
&\quad + 3 \hat{\mathsf{V}}_{p+1,q}(b_n) \big] \Big).
\end{aligned}$$

The first step (Step 0) constructs preliminary bandwidths ($v_n$ and $c_n$) to estimate the asymptotic variance terms and preliminary bias term entering the plug-in rules. The choice of constant in $v_n$ is a modified Silverman's rule of thumb: because we employ triangular kernels in our estimates, we modify the constant in front accordingly. Specifically, recall that Silverman's rule of thumb is

$$\begin{aligned}
\mathsf{h}_{\texttt{IMSE}} &= \sigma \left( \frac{8\sqrt{\pi} \int K(u)^2 \,\mathrm{d}u}{3 \left( \int u^2 K(u) \,\mathrm{d}u \right)^2} \frac{1}{n} \right)^{1/5} \\
&= \sigma \left( \frac{8\sqrt{\pi} \int K(u)^2 \,\mathrm{d}u}{3 \left( \int u^2 K(u) \,\mathrm{d}u \right)^2} \right)^{1/5} n^{-1/5}.
\end{aligned}$$

For $K(u) = \exp(-u^2/2)/\sqrt{2\pi}$, we have

$$\int K(u)^2\,\mathrm{d}u = \int_{-\infty}^{\infty} \bigl(\exp(-u^2/2)/\sqrt{2\pi}\bigr)^2\,\mathrm{d}u = \frac{1}{2\sqrt{\pi}},$$
$$\int u^2 K(u)\,\mathrm{d}u = \int_{-\infty}^{\infty} u^2 \exp(-u^2/2)/\sqrt{2\pi}\,\mathrm{d}u = 1,$$

and hence $\mathsf{h}_{\mathtt{IMSE}} = 1.0592 \cdot \sigma \cdot n^{-1/5}$. For $K(u) = \mathbf{1}(|u| \le 1)$, we have

$$\int K(u)^2\,\mathrm{d}u = \int_{-1}^{1} \mathrm{d}u = 2, \quad \int u^2 K(u)\,\mathrm{d}u = \int_{-1}^{1} u^2\,\mathrm{d}u = \frac{2}{3},$$

and hence $\mathsf{h}_{\mathtt{IMSE}} = 1.8431 \cdot \sigma \cdot n^{-1/5}$. For $K(u) = (1-|u|)\mathbf{1}(|u| \le 1)$, we obtain

$$\int K(u)^2\,\mathrm{d}u = \int_{-1}^{1} \bigl(1-|u|\bigr)^2\,\mathrm{d}u = \frac{2}{3},$$
$$\int u^2 K(u)\,\mathrm{d}u = \int_{-1}^{1} u^2\bigl(1-|u|\bigr)\,\mathrm{d}u = \frac{1}{6},$$

and hence $\mathsf{h}_{\mathtt{IMSE}} = 2.576 \cdot \sigma \cdot n^{-1/5}$. Step 0 also constructs a preliminary, possibly inconsistent bandwidth ($c_n$) to estimate the bias term entering the rule of thumb of $b_n$. The bandwidth choices in Steps 1 and 2 follow directly from Lemma 1 applied to the key component of the bias estimate ($\hat{\mathsf{b}}_{\nu,p,q}(b_n)$) and the main RD estimator ($\hat{\tau}_{\nu,p}(h_n)$), respectively. Our proposed bandwidths include regularization terms; see Imbens and Kalyanaraman (2012) and the discussion on their implementation further below for more details.

The following theorem establishes consistency and MSE-optimality of these bandwidth selectors.

THEOREM A.4—Consistency of Plug-in Bandwidth Selectors: *Let* $\nu \le p < q$. *Suppose Assumptions* 1–2 *hold with* $S \ge q+2$. *In addition, suppose* $e'_{q+2}\hat{\gamma}_{+,q+2} - (-1)^{\nu+q+1} e'_{q+2}\hat{\gamma}_{-,q+2} \to_p c \ne 0$.

*Step 1. If* $\mathsf{B}_{p+1,q,q+1,\nu+p+1} \ne 0$, *then*

$$\frac{\hat{b}_{\nu,p,q}}{\mathsf{h}_{\mathtt{MSE},p+1,q,\nu+p+1}} \to_p 1 \quad \textit{and} \quad \frac{\mathsf{MSE}_{p+1,q,\nu+p+1}(\hat{b}_{\nu,p,q})}{\mathsf{MSE}_{p+1,q,\nu+p+1}(\mathsf{h}_{\mathtt{MSE},p+1,q,\nu+p+1})} \to_p 1.$$

*Step 2. If* $\mathsf{B}_{\nu,p,p+1,0} \ne 0$, *then*

$$\frac{\hat{h}_{\nu,p}}{\mathsf{h}_{\mathtt{MSE},\nu,p,0}} \to_p 1 \quad \textit{and} \quad \frac{\mathsf{MSE}_{\nu,p,0}(\hat{h}_{\nu,p})}{\mathsf{MSE}_{\nu,p,0}(\mathsf{h}_{\mathtt{MSE},\nu,p,0})} \to_p 1.$$

PROOF: Recall that, if $c_n \to_p 0$ and $nc_n \to_p \infty$, Lemma S.A.3 and Theorem A.3 imply

$$\hat{\mathsf{V}}_{s,p}(c_n) = \frac{1}{nc_n^{1+2s}}\mathsf{V}_{s,p}\{1 + o_p(1)\} \quad \text{for all } s \leq p,$$

with $\hat{\mathsf{V}}_{s,p}(c_n)$ constructed using the NN-based estimators introduced in Section 5 of the paper. This implies consistency of the numerators of $\hat{C}_{\nu,p}$, $\hat{C}_{\nu,p,q}$, and $\check{C}_{\nu,p,q}$, for any $\nu \leq p < q$. In addition, for the regularization terms, we have

$$\begin{aligned}\hat{\mathsf{V}}_{q+2,q+2}(v_n) &= \frac{1}{nv_n^{1+2q+4}}\mathsf{V}_{q+2,q+2}\{1 + o_p(1)\} \\ &= \frac{1}{n^{2q/5}}\mathsf{V}_{q+2,q+2}\{1 + o_p(1)\}\end{aligned}$$

because $v_n = 2.58 \cdot \omega \cdot n^{-1/5}$,

$$\begin{aligned}\hat{\mathsf{V}}_{q+1,q+1}(c_n) &= \frac{1}{nc_n^{1+2q+2}}\mathsf{V}_{q+1,q+1}\{1 + o_p(1)\} \\ &= \frac{1}{n^{2/(2q+5)}}\mathsf{V}_{q+1,q+1}\{1 + o_p(1)\}\end{aligned}$$

because $c_n = \hat{C}_{q+1,q+1}n^{-1/(2q+5)}$, and

$$\begin{aligned}\hat{\mathsf{V}}_{p+1,q}(b_n) &= \frac{1}{nb_n^{1+2p+2}}\mathsf{V}_{p+1,q}\{1 + o_p(1)\} \\ &= \frac{1}{n^{2(q-p)/(2q+3)}}\mathsf{V}_{p+1,q}\{1 + o_p(1)\}\end{aligned}$$

because $b_n = \hat{C}_{p+1,q}n^{-1/(2q+3)}$.

Next, by Lemma S.A.3, if $nh_n \to_p \infty$ and $h_n \to_p 0$, and $\mu_+^{(p+1)} \neq \mu_-^{(p+1)}$,

$$e_s'\hat{\beta}_{+,p}(h_n) \pm e_s'\beta_{+,p} = O_p\left(h_n^{1+p-s} + \frac{1}{\sqrt{nh_n^{1+2s}}}\right) \quad \text{for all } s \leq p.$$

Now we consider each step.

*Step 0*: Assuming $e_{q+2}'\hat{\gamma}_{+,q+2} - (-1)^{\nu+q+1}e_{q+2}'\hat{\gamma}_{-,q+2} \to_p c \neq 0$, we verify that

$$c_n = \check{C}_{\nu,p,q}^{1/(2q+5)}n^{-1/(2q+5)} = O_p\big(n^{-1/(2q+5)}\big).$$

*Step 1*: We have $c_n = O_p(n^{-1/(2q+5)})$, and

$$\begin{aligned}
e'_{q+1}\hat{\beta}_{+,q+1}(c_n) &= e'_{q+1}\beta_{+,q+1} + O_p\left(c_n + \frac{1}{\sqrt{nc_n^{2q+3}}}\right) \\
&= e'_{q+1}\beta_{+,q+1} + O_p\big(n^{-1/(2q+5)}\big), \\
e'_{q+1}\hat{\beta}_{-,q+1}(c_n) &= e'_{q+1}\beta_{-,q+1} + O_p\left(c_n + \frac{1}{\sqrt{nc_n^{2q+3}}}\right) \\
&= e'_{q+1}\beta_{-,q+1} + O_p\big(n^{-1/(2q+5)}\big).
\end{aligned}$$

Thus, if $\mathsf{B}_{p+1,q,q+1,\nu+p+1} \neq 0$, then, using previous results, we have

$$\frac{\hat{C}_{\nu,p,q}}{\mathsf{C}_{\mathtt{MSE},p+1,q,\nu+p+1}} \to_p 1, \quad \frac{\hat{b}_{\nu,p,q}}{\mathsf{h}_{\mathtt{MSE},p+1,q,\nu+p+1}} \to_p 1,$$

$$\hat{b}_{\nu,p,q} = O_p\big(n^{-1/(2q+3)}\big), \quad \frac{\mathsf{MSE}_{p+1,q,\nu+p+1}(\hat{b}_{\nu,p,q})}{\mathsf{MSE}_{p+1,q,\nu+p+1}(\mathsf{h}_{\mathtt{MSE},p+1,q,\nu+p+1})} \to_p 1.$$

*Step 2*: We have $b_n = \hat{b}_{\nu,p,q} = O_p(n^{-1/(2q+3)})$, and

$$\begin{aligned}
e'_{p+1}\hat{\beta}_{+,q}(b_n) &= e'_{p+1}\beta_{+,q} + O_p\left(b_n^{q-p} + \frac{1}{\sqrt{nb_n^{2p+3}}}\right) \\
&= e'_{p+1}\beta_{+,q} + O_p\big(n^{-(q-p)/(2q+3)}\big), \\
e'_{p+1}\hat{\beta}_{-,q}(b_n) &= e'_{p+1}\beta_{-,q} + O_p\left(b_n^{q-p} + \frac{1}{\sqrt{nb_n^{2p+3}}}\right) \\
&= e'_{p+1}\beta_{-,q} + O_p\big(n^{-(q-p)/(2q+3)}\big),
\end{aligned}$$

and therefore

$$\frac{\hat{C}_{\nu,p}}{\mathsf{C}_{\mathtt{MSE},\nu,p,0}} \to_p 1, \quad \frac{\hat{h}_{\nu,p}}{\mathsf{h}_{\mathtt{MSE},\nu,p,0}} \to_p 1,$$

$$\hat{h}_{\nu,p} = O_p\big(n^{-1/(2p+3)}\big), \quad \frac{\mathsf{MSE}_{\nu,p,0}(\hat{h}_{\nu,p})}{\mathsf{MSE}_{\nu,p,0}(\mathsf{h}_{\mathtt{MSE},\nu,p,0})} \to_p 1.$$

This completes the proof. *Q.E.D.*

S.2.7. *Details on Remark 3 and Remark 7*

Remark 3 in the paper generalizes as follows for sharp RD designs. Recall that $\nu \leq p < q$.

REMARK S.A.3: Three main limiting cases are obtained depending on the limit $\rho_n \to \rho \in [0, \infty]$.

*Case 1*: $\rho = 0$. In this case, $h_n = o(b_n)$ and

$$\begin{aligned}\mathbb{V}\big[\hat{\tau}^{\mathrm{bc}}_{\nu,p,q}(h_n, b_n)|\mathcal{X}_n\big] &= \mathbb{V}\big[\hat{\tau}_{\nu,p}(h_n)|\mathcal{X}_n\big]\big\{1 + o_p(1)\big\} \\ &= \frac{1}{nh_n^{1+2\nu}}\frac{\sigma^2_+ + \sigma^2_-}{f}\nu!^2\big(e'_\nu \Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\big) \\ &\quad \times \big\{1 + o_p(1)\big\},\end{aligned}$$

which is the classical approach to bias correction.

*Case 2*: $\rho \in (0, \infty)$. In this case, $h_n = \rho b_n$ and

$$\begin{aligned}&\mathbb{V}\big[\hat{\tau}^{\mathrm{bc}}_{\nu,p,q}(h_n, b_n)|\mathcal{X}_n\big] \\ &\quad = \frac{\nu!^2}{nh_n^{1+2\nu}}\bigg[\frac{\sigma^2_+ + \sigma^2_-}{f}\big(e'_\nu\Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\big) \\ &\quad + \rho^{2p+3}\frac{\sigma^2_+ + \sigma^2_-}{f}\big(e'_{p+1}\Gamma_q^{-1}\Psi_q\Gamma_q^{-1}e_{p+1}\big)\big(e'_\nu\Gamma_p^{-1}\vartheta_{p,p+1}\big)^2 \\ &\quad - \rho^{p+2}\bigg(e'_\nu\Gamma_p^{-1}\bigg(\frac{\sigma^2_+}{f}\Psi_{p,q}(\rho) + \frac{\sigma^2_-}{f}\Psi_{p,q}(-\rho)\bigg)\Gamma_q^{-1}e_{p+1}\bigg) \\ &\quad \times \big(e'_\nu\Gamma_p^{-1}\vartheta_{p+1}\big)\bigg] \\ &\quad \times \big\{1 + o_p(1)\big\},\end{aligned}$$

with $\Psi_{p,q}(\rho) = \int_0^\infty K(u)K(\rho u)r_p(u)r_q(\rho u)'\,\mathrm{d}u$. For conventional choices of kernel $K(\cdot)$, the limiting variance is increasing in $\rho$.

*Case 3*: $\rho = \infty$. In this case, $b_n = o(h_n)$ and

$$\begin{aligned}&\mathbb{V}\big[\hat{\tau}^{\mathrm{bc}}_{\nu,p,q}(h_n, b_n)|\mathcal{X}_n\big] \\ &\quad = h_n^{2(p+1-\nu)}\mathbb{V}\big[\hat{\mathsf{B}}_{\nu,p,q}(h_n, b_n)|\mathcal{X}_n\big]\big\{1 + o_p(1)\big\} \\ &\quad = \frac{\rho_n^{2(p+1-\nu)}}{nb_n^{1+2\nu}}\frac{\sigma^2_+ + \sigma^2_-}{f}\nu!^2\big(e'_{p+1}\Gamma_q^{-1}\Psi_q\Gamma_q^{-1}e_{p+1}\big)\big(e'_\nu\Gamma_p^{-1}\vartheta_{p,p+1}\big)^2 \\ &\quad \times \big\{1 + o_p(1)\big\},\end{aligned}$$

which implies that the bias estimate is first order while the actual estimator $\hat{\tau}_p(h_n)$ is of smaller order.

This remark is established by noting that

$$\mathbb{V}\big[\hat{\tau}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)|\mathcal{X}_n\big] = \mathbb{V}\big[\hat{\tau}_{\nu,p}(h_n)|\mathcal{X}_n\big] + h_n^{2(p+1-\nu)}\mathbb{V}\big[\hat{\mathsf{B}}_{\nu,p,q}(h_n,b_n)|\mathcal{X}_n\big] - 2h_n^{p+1-\nu}\mathbb{C}\big[\hat{\tau}_{\nu,p}(h_n),\hat{\mathsf{B}}_{\nu,p,q}(h_n,b_n)|\mathcal{X}_n\big],$$

where these terms are given in Theorem A.1. The rest is obtained as follows.

*Case 1*: $\rho = 0$. In this case, $h_n = o(b_n)$. Using the previous calculations,

$$\begin{aligned}
&\mathbb{V}\big[\hat{\tau}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)|\mathcal{X}_n\big] \\
&\quad = \mathbb{V}\big[\hat{\tau}_{\nu,p}(h_n)|\mathcal{X}_n\big] + O_p\bigg(\frac{\rho_n^{p+2}}{nh_n^{1+2\nu}} + \frac{\rho_n^{2p+3}}{nh_n^{1+2\nu}}\bigg) \\
&\quad = \frac{1}{nh_n^{1+2\nu}}\frac{\sigma_+^2+\sigma_-^2}{f}\nu!^2\big(e_\nu'\Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\big)\big\{1+o_p(1)\big\}.
\end{aligned}$$

*Case 2*: $\rho \in (0,\infty)$. In this case, $h_n = \rho b_n$. By previous calculations,

$$\begin{aligned}
&\mathbb{V}\big[\hat{\tau}_{\nu,p}(h_n)|\mathcal{X}_n\big] = \frac{1}{nh_n^{1+2\nu}}\frac{\sigma_+^2+\sigma_-^2}{f}\nu!^2\big(e_\nu'\Gamma_p^{-1}\Psi_p\Gamma_p^{-1}e_\nu\big)\big\{1+o_p(1)\big\}, \\
&h_n^{p+1-\nu}\mathbb{C}\big[\hat{\tau}_{\nu,p}(h_n),\hat{\mathsf{B}}_{\nu,p,q}(h_n,b_n)|\mathcal{X}_n\big] \\
&\quad = \frac{\rho^{p+2}}{nh_n^{1+2\nu}}\nu!^2\bigg(e_\nu'\Gamma_p^{-1}\bigg(\frac{\sigma_+^2}{f}\Psi_{p,q}(\rho)+\frac{\sigma_-^2}{f}\Psi_{p,q}(-\rho)\bigg)\Gamma_q^{-1}e_{p+1}\bigg) \\
&\qquad \times \big(e_\nu'\Gamma_p^{-1}\vartheta_{p+1}\big)\big\{1+o_p(1)\big\}, \\
&h_n^{2(p+1-\nu)}\mathbb{V}\big[\hat{\mathsf{B}}_{\nu,p,q}(h_n,b_n)|\mathcal{X}_n\big] \\
&\quad = \frac{\rho^{2p+3}}{nh_n^{1+2\nu}}\frac{\sigma_+^2+\sigma_-^2}{f}\nu!^2\big(e_{p+1}'\Gamma_q^{-1}\Psi_q\Gamma_q^{-1}e_{p+1}\big)\big(e_\nu'\Gamma_p^{-1}\vartheta_{p,p+1}\big)^2 \\
&\qquad \times \big\{1+o_p(1)\big\}.
\end{aligned}$$

*Case 3*: $\rho = \infty$. In this case, $b_n = o(h_n)$. By previous calculations,

$$\begin{aligned}
&\mathbb{V}\big[\hat{\tau}^{\mathrm{bc}}_{\nu,p,q}(h_n,b_n)|\mathcal{X}_n\big] \\
&\quad = h_n^{2(p+1-\nu)}\mathbb{V}\big[\hat{\mathsf{B}}_{\nu,p,p+1,q}(h_n,b_n)|\mathcal{X}_n\big] + O_p\bigg(\frac{1}{nh_n^{1+2\nu}} + \frac{\rho_n^{p+2}}{nh_n^{1+2\nu}}\bigg) \\
&\quad = \frac{\rho_n^{2(p+1-\nu)}}{nb_n^{1+2\nu}}\frac{\sigma_+^2+\sigma_-^2}{f}\nu!^2\big(e_{p+1}'\Gamma_q^{-1}\Psi_q\Gamma_q^{-1}e_{p+1}\big)\big(e_\nu'\Gamma_p^{-1}\vartheta_{p,p+1}\big)^2 \\
&\qquad \times \big\{1+o_p(1)\big\},
\end{aligned}$$

which implies that the bias estimate is first order while the actual estimator $\hat{\tau}_{\nu,p}(h_n)$ is of smaller order.

Similar, but more cumbersome, expressions may be derived for fuzzy RD designs.

Next, we have the following generalization of Remark 7.

REMARK S.A.7: If $h_n = b_n$ (and the same kernel function $K(\cdot)$ is used), then

$$\hat{\tau}^{\mathtt{bc}}_{\nu,p,p+1}(h_n, h_n) = \hat{\tau}_{\nu,p+1}(h_n) \quad \text{and} \quad T^{\mathtt{rbc}}_{\nu,p,p+1}(h_n, h_n) = T_{\nu,p+1}(h_n),$$

which gives a simple relationship between local polynomial estimators of order $p$ and $p+1$, and their relation to manual bias correction. The result extends to $\hat{\tau}^{\mathtt{bc}}_{\nu,p,p+r}(h_n, h_n) = \hat{\tau}_{\nu,p+r}(h_n)$ and $T^{\mathtt{rbc}}_{\nu,p,p+r}(h_n, h_n) = T_{\nu,p+r}(h_n)$, $r \geq 1$, when the natural generalization of the bias-correction estimate is used.

This equivalence follows by properties of linear models. Consider the right-side estimators $\hat{\beta}_{+,p}(h_n)$ and $\hat{\beta}_{+,p+1}(h_n)$, and define

$$\begin{aligned}\hat{\beta}_{+,p+1}(h) &= H_{p+1}(h)\Gamma^{-1}_{+,p+1}(h)X_{p+1}(h)'W_+(h)Y/n \\ &= \left[\tilde{\beta}_{+,p+1}(h)', e'_{p+1}\hat{\beta}_{+,p+1}(h)\right]',\end{aligned}$$

where

$$\tilde{\beta}_{+,p+1}(h) = \left[e'_0\hat{\beta}_{+,p+1}(h), \dots, e'_p\hat{\beta}_{+,p+1}(h)\right]'.$$

The normal equations for the estimator $\hat{\beta}_{+,p+1}(h)$ are

$$\begin{aligned}&\begin{bmatrix} X_p(h)'W_+(h)X_p(h) & X_p(h)'W_+(h)S_{p+1}(h) \\ S_{p+1}(h)'W_+(h)X_p(h) & X_p(h)'W_+(h)S_{p+1}(h)\end{bmatrix}\begin{bmatrix}\tilde{\beta}_{+,p}(h) \\ e'_{p+1}\hat{\beta}_{+,p+1}(h)\end{bmatrix} \\ &\quad = H_{p+1}(h)\begin{bmatrix} X_p(h)'W_+(h)Y \\ S_{p+1}(h)'W_+(h)Y\end{bmatrix},\end{aligned}$$

and therefore, after some algebra,

$$\begin{aligned}\tilde{\beta}_{+,p}(h) &= H_{p+1}(h)\left[X_p(h)'W_+(h)X_p(h)\right]^{-1}X_{p+1}(h)'W_+(h)Y \\ &\quad - H_{p+1}(h)\left[X_p(h)'W_+(h)X_p(h)\right]^{-1} \\ &\quad \times X_p(h)'W_+(h)S_{p+1}(h)\left(e'_{p+1}\hat{\beta}_{+,p+1}(h)\right) \\ &= H_{p+1}(h)\Gamma^{-1}_{+,p}(h)X_p(h)'W_+(h)Y/n \\ &\quad - H_{p+1}(h)\Gamma^{-1}_{+,p}\vartheta_{+,p,p+1}(h)\left(e'_{p+1}\hat{\beta}_{+,p+1}(h)\right) \\ &= \hat{\beta}_{+,p}(h) - H_{p+1}(h)\Gamma^{-1}_{+,p}\vartheta_{+,p,p+1}(h)\left(e'_{p+1}\hat{\beta}_{+,p+1}(h)\right).\end{aligned}$$

This result immediately gives, for any $\nu \leq p$,

$$\begin{aligned} e_\nu' \hat{\beta}_{+,p+1}(h) &= e_\nu' \tilde{\beta}_{+,p}(h) \\ &= e_\nu' \hat{\beta}_{+,p}(h) - e_\nu' \Gamma_{+,p}^{-1} \vartheta_{+,p,p+1}(h) \big( e_{p+1}' \hat{\beta}_{+,p+1}(h) \big), \end{aligned}$$

as claimed. It follows that $\hat{\tau}^{\mathrm{bc}}_{\nu,p,p+1}(h_n, h_n) = \hat{\tau}_{\nu,p+1}(h_n)$, and $T^{\mathrm{rbc}}_{\nu,p,p+1}(h_n, h_n) = T_{\nu,p+1}(h_n)$.

To generalize this result to multiple levels of bias correction, note that the same argument gives

$$\begin{aligned} \tilde{\beta}_{+,p}(h) &= \hat{\beta}_{+,p}(h) - \Gamma_{+,p}^{-1} X_p(h)' W_+(h) \\ &\quad \times \big[ S_{p+1}(h), S_{p+2}(h), \ldots, S_{p+r}(h) \big] \begin{bmatrix} e_{p+1}' \hat{\beta}_{+,p+r}(h) \\ e_{p+2}' \hat{\beta}_{+,p+r}(h) \\ \vdots \\ e_{p+r}' \hat{\beta}_{+,p+r}(h) \end{bmatrix}. \end{aligned}$$

## S.3. SIMULATIONS

We provide further details on each data generating process (DGP) employed in our simulation study, and on the implementation of the alternative bandwidth selectors described in the text. We also include further numerical results not presented in the paper.

### S.3.1. *Data Generating Processes*

All DGPs employ the same simulation setup, with the only exception of the functional form of the regression function. Specifically, for each replication, the data are generated as i.i.d. draws, $i = 1, 2, \ldots, n$ with $n = 500$, as follows:

$$\begin{aligned} &Y_i = \mu_j(X_i) + \varepsilon_i, \quad X_i \sim \big( 2\mathcal{B}(2,4) - 1 \big), \\ &\varepsilon_i \sim \mathcal{N}\big(0, \sigma_\varepsilon^2\big), \quad j = 1, 2, 3, \end{aligned}$$

where $\mathcal{B}(\alpha, \beta)$ denotes a beta distribution with parameters $\alpha$ and $\beta$, $\varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ with $\sigma_\varepsilon = 0.1295$, and $\mu_j(X_i)$ with $j = 1, 2, 3$ as discussed below. Up to the regression function form, this setup coincides exactly with the one employed in Imbens and Kalyanaraman (2012).

#### S.3.1.1. *Model 1: Lee (2008) Data*

This model employs the regression function form described in Imbens and Kalyanaraman (2012), which was generated using data from Lee (2008). Lee studied the incumbency advantage in elections, and thus his identification strategy was based on the discontinuity generated by the rule that the party with a majority vote share wins. The forcing variable is the difference in vote share between the Democratic candidate and her strongest opponent in a given election, with the threshold level set at $c = 0$. The outcome variable is the Democratic vote share in the following election.

The regression function is obtained by fitting a fifth-order global polynomial with different coefficients for $X_i < 0$ and $X_i > 0$. The resulting coefficients estimated on the Lee (2008) data, after discarding observations with past vote share differences greater than 0.99 and less than −0.99, leads to the following functional form:

$$\mu_1(x) = \begin{cases} 0.48 + 1.27x + 7.18x^2 & \\ \quad + 20.21x^3 + 21.54x^4 + 7.33x^5 & \text{if } x < 0, \\ 0.52 + 0.84x - 3.00x^2 & \\ \quad + 7.99x^3 - 9.01x^4 + 3.56x^5 & \text{if } x \geq 0. \end{cases}$$

S.3.1.2. *Model 2: Ludwig and Miller (2007) Data*

Ludwig and Miller (2007) studied the effect of Head Start funding to identify the program's effects on health and schooling. For each county, eligibility is based on the county's poverty rate, inducing a natural RD design. For each county $i = 1, 2, \dots, n$, the forcing variable is the county's 1960 poverty rate with treatment assignment given by $T_i = \mathbf{1}(X_i \geq \bar{x})$, where $X_i$ represents the county's poverty rate in 1960 and $\bar{x}$ is the fixed threshold level. The cutoff is set to the poverty rate value of the 300th poorest county in 1960, which in this data set is given by $\bar{x} = 59.198$. We consider as outcome variable the mortality rates per 100,000 for children between 5 and 9 years old, with Head Start-related causes, for 1973–1983 (see Panel A, Figure IV in Ludwig and Miller (2007)).

As above, we estimate the regression function using a fifth-order polynomial, with separate coefficients for $X_i < 0$ and $X_i > 0$ (after discarding observations with differences greater than 0.99 and less than −0.99 of the rescaled running variable), leading to

$$\mu_2(x) = \begin{cases} 3.71 + 2.30x + 3.28x^2 & \\ \quad + 1.45x^3 + 0.23x^4 + 0.03x^5 & \text{if } x < 0, \\ 0.26 + 18.49x - 54.81x^2 & \\ \quad + 74.30x^3 - 45.02x^4 + 9.83x^5 & \text{if } x \geq 0. \end{cases}$$

S.3.1.3. *Model 3: An Alternative DGP*

We also explored other regression function specifications, and in all cases we obtained qualitatively similar results to those reported in the main text.

One such specification is given by

$$\mu_3(x) = \begin{cases} 0.48 + 1.27x - \mathbf{0.5} \cdot 7.18x^2 \\ \quad + \mathbf{0.7} \cdot 20.21x^3 + \mathbf{1.1} \cdot 21.54x^4 + \mathbf{1.5} \cdot 7.33x^5 & \text{if } x < 0, \\ 0.52 + 0.84x - \mathbf{0.1} \cdot 3.00x^2 \\ \quad - \mathbf{0.3} \cdot 7.99x^3 - \mathbf{0.1} \cdot 9.01x^4 + 3.56x^5 & \text{if } x \geq 0. \end{cases}$$

This specification was motivated by altering some of the coefficients in Model 1 (in bold). Our goal was to increase the overall "curvature" of the regression function while roughly preserving its monotonicity. Our main goal was to generate a plausible model with substantial size distortion when the theoretical, MSE-optimal bandwidths were employed, an important feature not present in the previous two models.

### S.3.2. *Bandwidths Selection*

We consider the following choices of MSE-optimal bandwidths $h_n$ and $b_n$:

(i) Imbens and Kalyanaraman (2012): denoted $\hat{h}_{\mathtt{IK}}$ and $\hat{b}_{\mathtt{IK}}$.

(ii) DesJardins and McCall (2009): denoted $\hat{h}_{\mathtt{DM}}$ and $\hat{b}_{\mathtt{DM}}$.

(iii) Ludwig and Miller (2007): denoted $\hat{h}_{\mathtt{CV}}$.

(iv) Second generation approach (CCT): denoted $\hat{h}_{\mathtt{CCT}}$ and $\hat{b}_{\mathtt{CCT}}$.

All the procedures are implemented for $p = 1$ and $q = 2$. For future reference, we define the constant:

$$C_{\nu,p}(K) = \frac{(2\nu+1)e_\nu' \Gamma_p^{-1} \Psi_p \Gamma_p^{-1} e_\nu}{2(p+1-\nu)(e_\nu' \Gamma_p^{-1} \vartheta_{p,p+1})^2},$$

which depends on the kernel employed. We denote $K_U$ and $K_T$ the uniform and triangular kernels, respectively.

#### S.3.2.1. *Imbens and Kalyanaraman (2012)*

We follow as closely as possible their implementation for $h_n$. We also extend their method to construct a plug-in, consistent estimator for $b_n$. Note that their preliminary estimates cannot be used as valid estimates of $b_{\mathtt{MSE},p+1,q}$ because those estimates are not consistent.

*Step 1: Estimation of density and conditional variances*. We employ their modified Silverman rule of thumb to obtain a pilot bandwidth for calculating the density and variances. That is, we use the formula $\hat{h}_1 = 1.84 \cdot S_X \cdot n^{-1/5}$, where $S_X^2$ is the sample variance of the forcing variable $X_i$, and the constant 1.84 corresponds to the uniform kernel (see CCT procedures below).

We estimate the density at $X_i = 0$ and the conditional variances of $Y_i$ given $X_i = 0$ from the left and from the right separately as follows (replacing $h_1$ by $\hat{h}_1$):

$$\hat{f}(h_1) = \frac{N_{h_1,-} + N_{h_1,+}}{2nh_1},$$

$$\hat{\sigma}_-^2(h_1) = \frac{1}{N_{h_1,-} - 1} \sum_{i:-h_1 \le X_i < 0} (Y_i - \bar{Y}_{h_1,-})^2,$$

$$\hat{\sigma}_+^2(h_1) = \frac{1}{N_{h_1,+} - 1} \sum_{i:0 \le X_i \le h_1} (Y_i - \bar{Y}_{h_1,+})^2,$$

where the components are simply the number of units and the average outcomes on either side of the threshold:

$$N_{h,-} = \sum_{i=1}^{n} \mathbf{1}(-h \le X_i \le 0), \quad N_{h,+} = \sum_{i=1}^{n} \mathbf{1}(0 \le X_i \le h),$$

$$\bar{Y}_{h,-} = \frac{1}{N_{h,-}} \sum_{i:-h \le X_i < 0} Y_i, \quad \bar{Y}_{h,+} = \frac{1}{N_{h,+}} \sum_{i:0 \le X_i \le h} Y_i.$$

This step is common for both estimators $\hat{h}_{\mathtt{IK}}$ and $\hat{b}_{\mathtt{IK}}$.

*Step 2: Estimation of bandwidth $h_n$.* We first discuss how we estimate the main bandwidth $h_n$, following a procedure proposed by IK. We employ our notation whenever possible to make clear how this procedure is extended to the case of selecting $b_n$. The IK bandwidth estimator is given by

$$\hat{h}_{\mathtt{IK}} = \left( C_{0,1}(K_T) \frac{(\hat{\sigma}_+^2(\hat{h}_1) + \hat{\sigma}_-^2(\hat{h}_1))/\hat{f}(\hat{h}_1)}{((\hat{\mu}_+^{(2)}(\hat{h}_{2+}) - \hat{\mu}_-^{(2)}(\hat{h}_{2-}))/2!)^2 + (1/2!^2)\hat{r}_h} \right)^{1/5} n^{-1/5},$$

where this procedure requires constructing the estimates $\hat{h}_{2+}$, $\hat{h}_{2-}$, $\hat{\mu}_+^{(2)}(\hat{h}_{2+})$, $\hat{\mu}_-^{(2)}(\hat{h}_{2-})$, and $\hat{r}_h$. Note that, as in IK, we have $(C_{0,1}(K_T)2!^2)^{1/5} \approx 3.4375$.

First, consider selecting the preliminary bandwidths $\hat{h}_{2+}$ and $\hat{h}_{2-}$. Following closely IK, we use a possibly inconsistent estimator of the derivatives $\mu_-^{(3)}(0)$ and $\mu_+^{(3)}(0)$ to construct a plug-in rule for these preliminary bandwidths. Specifically, employing only the 50% observations closest to the cutoff from either side, we fit a global cubic polynomial to the data at either side of the threshold separately, and then use the corresponding derivative estimates to construct the plug-in bandwidth selector. We denote the models

$$Y_i = \varpi \mathbf{1}(X_i \ge 0) + \gamma_0 + \gamma_1 X_i + \gamma_2 X_i^2 + \gamma_3 X_i^3 + \varepsilon_i$$
$$\text{for all } i : X_{-,[1/2]} \le X_i < X_{+,[1/2]},$$

where $X_{-,[1/2]}$ and $X_{+,[1/2]}$ denote, respectively, the median of the data for units below and above the cutoff. We denote the associated least-squares estimates by $(\hat{\varpi}, \hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3)'$. Thus, we implement

$$\hat{h}_{2-} = \left( C_{2,2}(K_U) \frac{\hat{\sigma}_-^2(\hat{h}_1)/\hat{f}(\hat{h}_1)}{N_-(\hat{\mu}^{(3)}/3!)^2} \right)^{1/7}, \quad \hat{\mu}^{(3)} = (3!)\hat{\gamma}_3,$$

$$\hat{h}_{2+} = \left( C_{2,2}(K_U) \frac{\hat{\sigma}_+^2(\hat{h}_1)/\hat{f}(\hat{h}_1)}{N_+(\hat{\mu}^{(3)}/3!)^2} \right)^{1/7}, \quad \hat{\mu}^{(3)} = (3!)\hat{\gamma}_3,$$

and note that $(C_{2,2}(K_U)3!^2)^{1/7} \approx 3.56$, as in IK.

Second, using $\hat{h}_{2-}$ and $\hat{h}_{2+}$, we construct consistent estimators $\hat{\mu}_+^{(2)}(\hat{h}_{2+})$ and $\hat{\mu}_-^{(2)}(\hat{h}_{2-})$ of, respectively, $\mu_+^{(2)}(0)$ and $\mu_-^{(2)}(0)$. Specifically, we let $\hat{\mu}_+^{(2)}(h_{2+})$ and $\hat{\mu}_-^{(2)}(h_{2-})$ denote unweighted local-quadratic fits only employing the observations in $0 \leq X_i \leq h_{2+}$ and $-h_{2-} \leq X_i < 0$, respectively; that is, we estimate the models

$$Y_{-,i} = \lambda_{0,-} + \lambda_{1,-}X_{-,i} + \lambda_{2,-}X_{-,i}^2 + \varepsilon_{-,i} \quad \text{for all } i : -h_{2-} \leq X_i < 0,$$

$$Y_{+,i} = \lambda_{0,+} + \lambda_{1,+}X_{+,i} + \lambda_{2,+}X_{+,i}^2 + \varepsilon_{+,i} \quad \text{for all } i : 0 \leq X_i \leq h_{2+},$$

and denote the associated least-squares estimates by $(\hat{\lambda}_{0,-}, \hat{\lambda}_{1,-}, \hat{\lambda}_{2,-})'$ and $(\hat{\lambda}_{0,+}, \hat{\lambda}_{1,+}, \hat{\lambda}_{2,+})'$, respectively. Therefore, the final estimates are given by

$$\hat{\mu}_-^{(2)}(\hat{h}_{2-}) = (2!)\hat{\lambda}_{2,-}, \quad \hat{\mu}_+^{(2)}(\hat{h}_{2+}) = (2!)\hat{\lambda}_{2,+}.$$

Finally, in IK implementation the regularization term is given by

$$\hat{r}_h = \hat{\sigma}_+^2(\hat{h}_1) \frac{2160}{N_{\hat{h}_{2,+}} \hat{h}_{2,+}^4} + \hat{\sigma}_-^2(\hat{h}_1) \frac{2160}{N_{\hat{h}_{2,-}} \hat{h}_{2,-}^4}.$$

*Step 3: Estimation of bandwidth $b_n$.* We follow the logic described in Step 2 to construct a consistent plug-in estimate for the pilot bandwidth $b_n$ that follows the approach of IK. The resulting estimator is

$$\hat{b}_{\mathtt{IK}} = \left( C_{2,2}(K_T) \frac{(\hat{\sigma}_+^2(\hat{h}_1) + \hat{\sigma}_-^2(\hat{h}_1))/\hat{f}(\hat{h}_1)}{((\hat{\mu}_+^{(3)}(\hat{b}_{3+}) + \hat{\mu}_-^{(3)}(\hat{b}_{3-}))/3!)^2 + (1/3!^2)\hat{r}_b} \right)^{1/7} n^{-1/7},$$

where this procedure now requires constructing the estimates $\hat{b}_{3+}$, $\hat{b}_{3-}$, $\hat{\mu}_+^{(3)}(\hat{b}_{3+})$, $\hat{\mu}_-^{(3)}(\hat{b}_{3-})$, and $\hat{r}_b$.

As before, we first select the preliminary bandwidths $\hat{b}_{3+}$ and $\hat{b}_{3-}$, for which we use a possibly inconsistent estimator of the derivatives $\mu_-^{(4)}(0)$ and $\mu_+^{(4)}(0)$ to construct a plug-in rule. We now fit a global fourth-order polynomial to the data, and then use the corresponding derivative estimates to construct the plug-in bandwidth selector. The model is

$$Y_i = \varpi \mathbf{1}(X_i \geq 0) + \gamma_0 + \gamma_1 X_i + \gamma_2 X_i^2 + \gamma_3 X_i^3 + \gamma_4 X_i^4 + \varepsilon_i$$
$$\text{for all } i : X_{-,[1/2]} \leq X_i < X_{+,[1/2]},$$

where $X_{-,[1/2]}$ and $X_{+,[1/2]}$ denote, respectively, the median of the data for units below and above the cutoff. We denote the associated least-squares estimates by $(\hat{\varpi}, \hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3, \hat{\gamma}_4)'$. Thus, we implement

$$\hat{b}_{3-} = \left( C_{3,3}(K_U) \frac{\hat{\sigma}^2_-(\hat{h}_1)/\hat{f}(\hat{h}_1)}{N_-(\hat{\mu}^{(4)}_-/4!)^2} \right)^{1/9}, \quad \hat{\mu}^{(4)}_- = (4!)\hat{\gamma}_{4,-},$$

$$\hat{b}_{3+} = \left( C_{3,3}(K_U) \frac{\hat{\sigma}^2_+(\hat{h}_1)/\hat{f}(\hat{h}_1)}{N_+(\hat{\mu}^{(4)}_+/4!)^2} \right)^{1/9}, \quad \hat{\mu}^{(4)}_+ = (4!)\hat{\gamma}_{4,+}.$$

Next, using $\hat{b}_{3-}$ and $\hat{b}_{3+}$, and proceeding as before, we construct consistent estimators $\hat{\mu}^{(3)}_+(\hat{b}_{3+})$ and $\hat{\mu}^{(3)}_-(\hat{b}_{3-})$ of $\mu^{(3)}_+(0)$ and $\mu^{(3)}_-(0)$, respectively, by fitting an unweighted local-quadratic polynomial regression employing the observations in $0 \leq X_i \leq h_{2,+}$ and $-h_{2,-} \leq X_i < 0$ separately: we estimate the models

$$Y_{-,i} = \lambda_{0,-} + \lambda_{1,-} X_{-,i} + \lambda_{2,-} X^2_{-,i} + \lambda_{3,-} X^3_{-,i} + \varepsilon_{-,i}$$
$$\text{for all } i : -\hat{b}_{3-} \leq X_i < 0,$$
$$Y_{+,i} = \lambda_{0,+} + \lambda_{1,+} X_{+,i} + \lambda_{2,+} X^2_{+,i} + \lambda_{3,+} X^3_{+,i} + \varepsilon_{+,i}$$
$$\text{for all } i : 0 \leq X_i \leq \hat{b}_{3+},$$

and denote the associated least-squares estimates by $(\hat{\lambda}_{0,-}, \hat{\lambda}_{1,-}, \hat{\lambda}_{2,-}, \hat{\lambda}_{3,-})'$ and $(\hat{\lambda}_{0,+}, \hat{\lambda}_{1,+}, \hat{\lambda}_{2,+}, \hat{\lambda}_{3,+})'$, respectively. Therefore, the final estimates are given by

$$\hat{\mu}^{(3)}_-(\hat{b}_{3-}) = (3!)\hat{\lambda}_{3,-}, \quad \hat{\mu}^{(3)}_+(\hat{b}_{3+}) = (3!)\hat{\lambda}_{3,+}.$$

Finally, for the corresponding regularization term, we employ

$$\hat{r}_h = \hat{\sigma}^2_+(\hat{h}_1) \frac{3!^2 \cdot 3 \cdot 2800}{N_{\hat{b}_{3+},+} \hat{b}^6_{3+}} + \hat{\sigma}^2_-(\hat{h}_1) \frac{3!^2 \cdot 3 \cdot 2800}{N_{\hat{b}_{3-},-} \hat{b}^6_{3-}}.$$

*Details on Regularization Terms*. As discussed in IK, the regularization terms are introduced to account for cases where the denominators may be small in finite samples (e.g., because of the lack of curvature of the underlying regression function). The regularization term is derived by first linearizing the denominator and then computing its expectation. Specifically, under appropriate regularity conditions and using

$$\frac{1}{\hat{a}}-\frac{1}{a}=-\frac{\hat{a}-a}{a^2}+\frac{(\hat{a}-a)^2}{a^3}-\frac{1}{\hat{a}}\frac{(\hat{a}-a)^3}{a^3},$$

we have (let $\hat{a}=(e_\nu'\hat{\beta}_{+,p}(h_n)-(-1)^{\nu+p+1}e_\nu'\hat{\beta}_{-,p}(h_n))^2$ and $a=(e_\nu'\beta_{+,p}-(-1)^{\nu+p+1}e_\nu'\beta_{-,p})^2$):

$$\begin{aligned}
&\frac{1}{(e_\nu'\hat{\beta}_{+,p}(h_n)-(-1)^{\nu+p+1}e_\nu'\hat{\beta}_{-,p}(h_n))^2}\\
&\qquad-\frac{1}{(e_\nu'\beta_{+,p}-(-1)^{\nu+p+1}e_\nu'\beta_{-,p})^2}\\
&\quad=3\cdot\Big(e_\nu'\big(\hat{\beta}_{+,p}(h_n)-(-1)^{\nu+p+1}\beta_{+,p}\big)\\
&\qquad-e_\nu'\big(\hat{\beta}_{-,p}(h_n)-(-1)^{\nu+p+1}\beta_{-,p}\big)\Big)^2\\
&\qquad/\big(e_\nu'\beta_{+,p}-(-1)^{\nu+p+1}e_\nu'\beta_{-,p}\big)^4\\
&\qquad+o_p(1).
\end{aligned}$$

Therefore, the regularization term can be shown to be equal to

$$3\cdot\big[\mathbb{V}\big[e_\nu'\hat{\beta}_{+,p}(h_n)\big]+\mathbb{V}\big[e_\nu'\hat{\beta}_{-,p}(h_n)\big]\big]$$

for any $\nu$ and any $p$ with $\nu\leq p$.

IK proposed an approximation based on the simplified formula for the variances for the particular case of homoskedasticity with a uniform kernel:

$$\mathbb{V}\big[e_\nu'\hat{\beta}_{+,p}(h_n)\big]=\sigma_+^2\big(X_+'X_+\big)^{-1}$$

and

$$\mathbb{V}\big[e_\nu'\hat{\beta}_{-,p}(h_n)\big]=\sigma_-^2\big(X_-'X_-\big)^{-1},$$

which they further approximated (for small $h_n$). This gives the asymptotic formulas

$$\begin{aligned}
\mathbb{V}\big[e_2'\hat{\beta}_{+,2}(h_n)\big]&\approx\sigma_+^2\frac{180}{N_{h_n,+}h_n^4}, &\quad \mathbb{V}\big[e_2'\hat{\beta}_{-,2}(h_n)\big]&\approx\sigma_-^2\frac{180}{N_{h_n,-}h_n^4},\\
\mathbb{V}\big[e_3'\hat{\beta}_{+,3}(h_n)\big]&\approx\sigma_+^2\frac{2800}{N_{h_n,+}h_n^6}, &\quad \mathbb{V}\big[e_3'\hat{\beta}_{-,3}(h_n)\big]&\approx\sigma_-^2\frac{2800}{N_{h_n,-}h_n^6},
\end{aligned}$$

which are then used above to construct the regularization terms.

An alternative approach employs the pre-asymptotic approximation to these variance terms, thus avoiding several approximations. This approach leads to the alternative regularization terms that we employ for our bandwidth selectors discussed further below:

$$\mathbb{V}\big[e_\nu' \hat{\beta}_{+,p}(h_n)\big] = \frac{1}{nh_n^{2\nu}} e_\nu' \Gamma_{+,p}^{-1}(h_n) \Psi_{+,p}(h_n) \Gamma_{+,p}^{-1}(h_n) e_\nu,$$
$$\mathbb{V}\big[e_\nu' \hat{\beta}_{-,p}(h_n)\big] = \frac{1}{nh_n^{2\nu}} e_\nu' \Gamma_{-,p}^{-1}(h_n) \Psi_{-,p}(h_n) \Gamma_{-,p}^{-1}(h_n) e_\nu$$

for any appropriate values of $\nu$ and $p$.

S.3.2.2. *DesJardins and McCall (2009)*

DesJardins and McCall (2009) used an alternative bandwidth choice method, which minimizes an objective criterion based on the sum of squared differences. This leads to the optimal choice

$$h_{\text{DM}} = \left( C_{\nu,p}(K) \frac{\sigma_+^2/f + \sigma_-^2/f}{(\mu_+^{(p+1)}/(p+1)!)^2 + (\mu_-^{(p+1)}/(p+1)!)^2} \right)^{1/(2p+3)} \times n^{-1/(2p+3)}, \quad p = 1, \nu = 0,$$

and

$$b_{\text{DM}} = \left( C_{\nu,q}(K) \frac{\sigma_+^2/f + \sigma_-^2/f}{(\mu_+^{(q+1)}/(q+1)!)^2 + (\mu_-^{(q+1)}/(q+1)!)^2} \right)^{1/(2q+3)} \times n^{-1/(2q+3)}, \quad q = 2, \nu = 2$$

for a choice of kernel function $K(\cdot)$.

Thus, we implement these bandwidth selectors exactly the same way as discussed above. Observe that this implementation does not include regularization, although in some cases such correction may be appropriate (e.g., see Models 1 and 2 in our simulation results).

S.3.2.3. *Ludwig and Miller (2007)*

Ludwig and Miller (2007) proposed a cross-validation approach to select the main bandwidth $h_n$, specifically aimed at the regression discontinuity setting, which we denote as $\hat{h}_{\text{CV}}$. Following Imbens and Kalyanaraman (2012, Section 4.5), we construct this bandwidth estimate as follows: bandwidth choice

$\hat{h}_{\mathrm{CV}}$ employs the cross-validation criterion of the form

$$\hat{h}_{\mathrm{CV}} = \arg\min_{h>0} \mathrm{CV}_\delta(h),$$

$$\mathrm{CV}_\delta(h) = \sum_{i=1}^{n} \mathbf{1}(X_{-,[\delta]} \leq X_i \leq X_{+,[\delta]})\big(Y_i - \hat{\mu}(X_i; h)\big)^2,$$

where

$$\hat{\mu}(x; h) = \begin{cases} e_0'\hat{\beta}_{+,p}(x, h) & \text{if } x \geq 0, \\ e_0'\hat{\beta}_{-,p}(x, h) & \text{if } x < 0, \end{cases}$$

$$\hat{\beta}_{+,p}(x, h_n) = \arg\min_{\beta\in\mathbb{R}^p} \sum_{i=1}^{n} \mathbf{1}(X_i \geq x)\big(Y_i - r_p(X_i - x)'\beta\big)^2 K_{h_n}(X_i - x),$$

$$\hat{\beta}_{-,p}(x, h_n) = \arg\min_{\beta\in\mathbb{R}^p} \sum_{i=1}^{n} \mathbf{1}(X_i < x)\big(Y_i - r_p(X_i - x)'\beta\big)^2 K_{h_n}(X_i - x),$$

and, for $\delta \in (0, 1)$, $X_{-,[\delta]}$ and $X_{+,[\delta]}$ denote the $\delta$th quantile of $\{X_i : X_i < 0\}$ and $\{X_i : X_i \geq 0\}$, respectively. In our implementation we use $\delta = 0.5$.

S.3.2.4. *CCT Procedures*

We employ the general construction given in Section 2.6. Here we provide more details on our proposed bandwidth selectors used in the simulations. Recall that, in this case, $(\nu, p, q) = (0, 1, 2)$.

*Step 0: Initial bandwidths* $(v_n, c_n)$. First, we construct a preliminary bandwidth to estimate the asymptotic variance terms, denoted $\hat{v}_n$:

$$\hat{v}_n = 2.58 \cdot \omega \cdot n^{-1/5}, \quad \omega = \min\left\{S_X, \frac{IQR_X}{1.349}\right\},$$

where $S_X^2$ and $IQR_X$ denote, respectively, the sample variance and interquartile range of $\{X_i : 1 \leq i \leq n\}$.

Second, we construct a preliminary, possibly inconsistent bandwidth to estimate the bias term, denoted $\hat{c}_n$:

$$\hat{c}_n = \check{C}_{\nu,p,q}^{1/9} n^{-1/9}, \quad \check{C}_{0,1,2} = \frac{7n\hat{v}_n^7 \hat{\mathsf{V}}_{3,3}(\hat{v}_n)}{2\mathcal{B}_{3,3}^2 (e_4'\hat{\gamma}_{+,4} - e_4'\hat{\gamma}_{-,4})^2},$$

where $\hat{\gamma}_{+,4}$ and $\hat{\gamma}_{-,4}$ denote the estimated coefficients of a $(p+1)$th-order global polynomial fit at either side of the threshold.

*Step 1: Pilot bandwidth* $\hat{b}_{\text{CCT}}$. We compute the pilot bandwidth $\hat{b}_{\text{CCT}} = \hat{b}_{0,1,2}$:

$$\hat{b}_{\text{CCT}} = \hat{C}_{0,1,2}^{1/7} n^{-1/7},$$

$$\hat{C}_{0,1,2} = \frac{5n\hat{v}_n^5 \hat{\mathsf{V}}_{2,2}(\hat{v}_n)}{2\mathcal{B}_{2,2}^2\{(e_3'\hat{\beta}_{+,3}(\hat{c}_n) + e_3'\hat{\beta}_{-,3}(\hat{c}_n))^2 + 3\hat{\mathsf{V}}_{3,3}(\hat{c}_n)\}}.$$

*Step 2: Main bandwidth* $\hat{h}_{\text{CCT}}$. We compute the main bandwidth $\hat{h}_{\text{CCT}} = \hat{h}_{0,1}$:

$$\hat{h}_{\text{CCT}} = \hat{C}_{0,1,0}^{1/5} n^{-1/5},$$

$$\hat{C}_{0,1,0} = \frac{n\hat{v}_n \hat{\mathsf{V}}_{0,1}(\hat{v}_n)}{4\mathcal{B}_{0,1}^2\{(e_2'\hat{\beta}_{+,2}(\hat{b}_{\text{CCT}}) - e_2'\hat{\beta}_{-,2}(\hat{b}_{\text{CCT}}))^2 + 3\hat{\mathsf{V}}_{2,2}(\hat{b}_{\text{CCT}})\}}.$$

### S.3.3. *Additional Simulation Results*

Tables S.A.I through S.A.VIII present additional simulation results not reported in CCT. Using the same Monte Carlo simulation setup, each table now also includes the empirical coverage and average length for $I_{\text{SRD}}^{\text{bc}}(h_n, b_n)$, the confidence interval obtained using the bias-corrected statistic $T_{\text{SRD}}^{\text{bc}}(h_n, b_n)$. We also report results using the alternative bandwidth selection procedure proposed in DesJardins and McCall (2009), denoted $h_{\text{DM}}$, as discussed above.

Table S.A.I employs the infeasible standard errors based on $\mathsf{V}_{\text{SRD}}(h_n)$ and $\mathsf{V}_{\text{SRD}}^{\text{bc}}(h_n, b_n)$, while Tables S.A.II and S.A.III use the fully data-driven standard errors $\hat{\mathsf{V}}_{\text{SRD}}(h_n)$ and $\hat{\mathsf{V}}_{\text{SRD}}^{\text{bc}}(h_n, b_n)$ with $J = 3$ and $J = 1$, respectively. The simulation results across these tables are qualitatively very similar, with the feasible versions of the confidence intervals showing slightly more empirical coverage distortion and longer intervals on average. In Table S.A.IV, we also report results employing the traditional standard error estimators constructed using plug-in estimated residuals, which lead to even more undercoverage in our simulations. Overall, the robust standard error estimators lead to important improvements in empirical coverage with only moderate increments in the average empirical length of the resulting confidence intervals.

Finally, Tables S.A.V through S.A.VIII present the same set of results, but now using an ad hoc undersmoothing approach that implements each of the methods considered using the corresponding bandwidth divided by 2 (i.e., replacing $h_n$ and $b_n$ by $h_n/2$ and $b_n/2$, respectively). This undersmoothing approach led to some numerical instability in the case of Model 2, but otherwise performed reasonably well in our simulation study. The improvements in coverage rates in this case are associated with longer confidence intervals, which suggests that, in our simulation setup, the undersmoothing approach performs in general on par with our robust bias-correction approach. This numerical evidence, however, may be dependent on the particular data generating process

TABLE S.A.I

EMPIRICAL COVERAGE AND AVERAGE INTERVAL LENGTH OF DIFFERENT 95% CONFIDENCE INTERVALS USING INFEASIBLE ASYMPTOTIC VARIANCE[a]

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 1 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 93.5 | 0.225 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 88.8 | 0.225 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 94.5 | 0.273 | 0.166 | 0.251 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 93.2 | 0.213 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 88.4 | 0.213 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 94.3 | 0.262 | 0.184 | 0.271 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 83.4 | 0.153 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 70.9 | 0.153 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 92.4 | 0.270 | 0.375 | 0.350 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 78.5 | 0.136 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 61.6 | 0.136 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 88.9 | 0.269 | 0.501 | 0.429 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 80.8 | 0.145 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 76.0 | 0.145 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 91.8 | 0.213 | 0.428 | 0.428 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 90.7 | 0.206 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 87.6 | 0.206 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 92.7 | 0.243 | 0.204 | 0.332 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 81.0 | 0.225 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 94.7 | 0.339 | 0.166 | 0.166 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 81.0 | 0.213 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 94.8 | 0.319 | 0.184 | 0.184 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 78.5 | 0.153 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 92.8 | 0.226 | 0.375 | 0.375 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 71.5 | 0.136 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 89.4 | 0.198 | 0.501 | 0.501 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 81.3 | 0.206 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 94.8 | 0.308 | 0.204 | 0.204 |
| Model 2 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 92.7 | 0.327 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 92.7 | 0.327 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 94.8 | 0.355 | 0.082 | 0.189 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 92.4 | 0.323 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 92.6 | 0.323 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 94.8 | 0.352 | 0.084 | 0.190 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 27.2 | 0.214 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 83.4 | 0.214 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 89.3 | 0.247 | 0.184 | 0.325 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 14.3 | 0.206 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 80.7 | 0.206 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 86.9 | 0.238 | 0.198 | 0.347 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 76.8 | 0.264 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 80.2 | 0.264 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 94.6 | 0.401 | 0.124 | 0.124 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 87.4 | 0.300 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 91.7 | 0.300 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 94.1 | 0.326 | 0.097 | 0.223 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 79.5 | 0.327 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 95.2 | 0.513 | 0.082 | 0.082 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 79.7 | 0.323 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 95.2 | 0.506 | 0.084 | 0.084 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 80.3 | 0.214 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 94.1 | 0.320 | 0.184 | 0.184 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 80.4 | 0.206 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 94.3 | 0.308 | 0.198 | 0.198 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 80.0 | 0.300 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 94.7 | 0.465 | 0.097 | 0.097 |

(*Continues*)

TABLE S.A.I—*Continued*

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 3 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 85.8 | 0.179 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 86.2 | 0.179 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 94.7 | 0.235 | 0.260 | 0.322 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 87.3 | 0.182 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 85.8 | 0.182 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 94.7 | 0.242 | 0.251 | 0.305 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 85.7 | 0.187 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 87.7 | 0.187 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 94.8 | 0.234 | 0.241 | 0.352 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 90.7 | 0.197 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 90.8 | 0.197 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 94.7 | 0.223 | 0.215 | 0.437 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 93.1 | 0.219 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 81.2 | 0.219 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 94.9 | 0.329 | 0.177 | 0.177 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 91.4 | 0.216 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 90.9 | 0.216 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 95.0 | 0.249 | 0.183 | 0.329 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 81.3 | 0.179 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 94.9 | 0.266 | 0.260 | 0.260 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 81.3 | 0.182 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 94.9 | 0.271 | 0.251 | 0.251 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 81.0 | 0.187 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 94.9 | 0.278 | 0.241 | 0.241 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 81.1 | 0.197 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 94.8 | 0.295 | 0.215 | 0.215 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 82.0 | 0.216 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 95.4 | 0.324 | 0.183 | 0.183 |

[a]Notes: (i) EC = Empirical Coverage in percentage points, (ii) IL = Average Interval Length, (iii) columns under “Bandwidths” report the population and average estimated bandwidths choices, as appropriate, for main bandwidth $h_n$ and pilot bandwidth $b_n$.

TABLE S.A.II

EMPIRICAL COVERAGE AND AVERAGE INTERVAL LENGTH OF DIFFERENT 95% CONFIDENCE INTERVALS USING ESTIMATED ASYMPTOTIC VARIANCE WITH $J = 3$ NEAREST-NEIGHBORS[a]

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 1 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 92.0 | 0.223 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 87.4 | 0.223 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 93.0 | 0.270 | 0.166 | 0.251 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 91.7 | 0.211 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 86.8 | 0.211 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 93.1 | 0.259 | 0.184 | 0.271 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 82.3 | 0.152 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 70.0 | 0.152 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 91.4 | 0.267 | 0.375 | 0.350 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 78.0 | 0.135 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 61.1 | 0.135 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 87.6 | 0.266 | 0.501 | 0.429 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 79.7 | 0.144 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 75.2 | 0.144 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 90.5 | 0.211 | 0.428 | 0.428 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 89.4 | 0.203 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 86.1 | 0.203 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 91.6 | 0.239 | 0.204 | 0.332 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 79.7 | 0.223 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 92.4 | 0.332 | 0.166 | 0.166 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 79.7 | 0.211 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 92.6 | 0.313 | 0.184 | 0.184 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 77.5 | 0.152 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 91.5 | 0.223 | 0.375 | 0.375 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 71.0 | 0.135 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 88.4 | 0.197 | 0.501 | 0.501 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 79.8 | 0.203 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 92.8 | 0.300 | 0.204 | 0.204 |
| Model 2 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 91.3 | 0.355 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 91.5 | 0.355 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 93.6 | 0.386 | 0.082 | 0.189 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 91.2 | 0.350 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 91.5 | 0.350 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 93.7 | 0.382 | 0.084 | 0.190 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 30.3 | 0.225 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 84.0 | 0.225 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 89.5 | 0.261 | 0.184 | 0.325 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 16.9 | 0.216 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 81.3 | 0.216 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 87.5 | 0.251 | 0.198 | 0.347 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 77.3 | 0.281 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 81.0 | 0.281 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 93.5 | 0.439 | 0.124 | 0.124 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 87.3 | 0.319 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 91.2 | 0.319 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 93.2 | 0.347 | 0.097 | 0.223 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 80.5 | 0.355 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 93.3 | 0.569 | 0.082 | 0.082 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 80.6 | 0.350 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 93.2 | 0.561 | 0.084 | 0.084 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 81.0 | 0.225 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 93.6 | 0.345 | 0.184 | 0.184 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 81.3 | 0.216 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 93.9 | 0.331 | 0.198 | 0.198 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 80.8 | 0.319 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 93.2 | 0.508 | 0.097 | 0.097 |

(*Continues*)

TABLE S.A.II—*Continued*

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 3 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 84.6 | 0.178 | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 85.0 | 0.178 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 93.5 | 0.233 | 0.260 | 0.322 |
| $I_{\mathtt{SRD}}(h_{\mathrm{DM}})$ | 86.0 | 0.181 | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, b_{\mathrm{DM}})$ | 84.6 | 0.181 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, b_{\mathrm{DM}})$ | 93.4 | 0.239 | 0.251 | 0.305 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 84.2 | 0.185 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 86.7 | 0.185 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 93.6 | 0.231 | 0.241 | 0.352 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}})$ | 89.4 | 0.195 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{b}_{\mathrm{DM}})$ | 89.7 | 0.195 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{b}_{\mathrm{DM}})$ | 93.4 | 0.221 | 0.215 | 0.437 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathrm{CV}})$ | 91.6 | 0.217 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{CV}}, \hat{h}_{\mathrm{CV}})$ | 80.0 | 0.217 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{CV}}, \hat{h}_{\mathrm{CV}})$ | 92.6 | 0.322 | 0.177 | 0.177 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 89.8 | 0.213 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 89.4 | 0.213 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 93.3 | 0.245 | 0.183 | 0.329 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 79.8 | 0.178 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 93.2 | 0.262 | 0.260 | 0.260 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, h_{\mathrm{DM}})$ | 79.8 | 0.181 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, h_{\mathrm{DM}})$ | 93.2 | 0.267 | 0.251 | 0.251 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 79.8 | 0.185 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 93.2 | 0.274 | 0.241 | 0.241 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{h}_{\mathrm{DM}})$ | 79.9 | 0.195 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{h}_{\mathrm{DM}})$ | 93.0 | 0.290 | 0.215 | 0.215 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 80.3 | 0.213 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 93.1 | 0.316 | 0.183 | 0.183 |

[a](i) EC = Empirical Coverage in percentage points, (ii) IL = Average Interval Length, (iii) columns under "Bandwidths" report the population and average estimated bandwidths choices, as appropriate, for main bandwidth $h_n$ and pilot bandwidth $b_n$.

TABLE S.A.III

EMPIRICAL COVERAGE AND AVERAGE INTERVAL LENGTH OF DIFFERENT 95% CONFIDENCE INTERVALS USING ESTIMATED ASYMPTOTIC VARIANCE WITH $J = 1$ NEAREST-NEIGHBORS[a]

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 1 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 91.4 | 0.221 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 86.9 | 0.221 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 92.5 | 0.268 | 0.166 | 0.251 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 91.1 | 0.209 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 86.2 | 0.209 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 92.4 | 0.257 | 0.184 | 0.271 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 82.0 | 0.152 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 69.5 | 0.152 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 90.6 | 0.265 | 0.375 | 0.350 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 77.5 | 0.135 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 60.8 | 0.135 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 87.2 | 0.265 | 0.501 | 0.429 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 79.6 | 0.144 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 74.9 | 0.144 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 90.0 | 0.209 | 0.428 | 0.428 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 88.8 | 0.201 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 85.6 | 0.201 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 90.8 | 0.237 | 0.204 | 0.332 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 78.9 | 0.221 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 91.5 | 0.328 | 0.166 | 0.166 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 79.2 | 0.209 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 91.7 | 0.310 | 0.184 | 0.184 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 77.3 | 0.152 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 91.1 | 0.222 | 0.375 | 0.375 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 70.6 | 0.135 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 87.9 | 0.196 | 0.501 | 0.501 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 79.5 | 0.201 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 91.8 | 0.297 | 0.204 | 0.204 |
| Model 2 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 88.9 | 0.328 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 89.2 | 0.328 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 91.7 | 0.356 | 0.082 | 0.189 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 88.8 | 0.324 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 89.3 | 0.324 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 91.7 | 0.353 | 0.084 | 0.190 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 28.3 | 0.215 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 81.8 | 0.215 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 87.6 | 0.249 | 0.184 | 0.325 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 15.7 | 0.207 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 79.0 | 0.207 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 85.6 | 0.240 | 0.198 | 0.347 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 74.8 | 0.264 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 78.8 | 0.264 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 91.0 | 0.401 | 0.124 | 0.124 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 84.9 | 0.300 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 88.9 | 0.300 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 91.2 | 0.326 | 0.095 | 0.220 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 77.0 | 0.328 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 90.1 | 0.510 | 0.082 | 0.082 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 77.1 | 0.324 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 90.1 | 0.503 | 0.084 | 0.084 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 78.9 | 0.215 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 91.9 | 0.322 | 0.184 | 0.184 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 79.4 | 0.207 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 92.1 | 0.309 | 0.198 | 0.198 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 77.8 | 0.300 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 90.4 | 0.463 | 0.095 | 0.095 |

(*Continues*)

TABLE S.A.III—*Continued*

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 3 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 84.0 | 0.177 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 84.7 | 0.177 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 93.0 | 0.231 | 0.260 | 0.322 |
| $I_{\mathtt{SRD}}(h_{\mathrm{DM}})$ | 85.8 | 0.180 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, b_{\mathrm{DM}})$ | 84.3 | 0.180 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, b_{\mathrm{DM}})$ | 92.8 | 0.237 | 0.251 | 0.305 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 83.9 | 0.184 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 86.0 | 0.184 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 93.1 | 0.230 | 0.241 | 0.352 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}})$ | 89.2 | 0.194 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{b}_{\mathrm{DM}})$ | 89.2 | 0.194 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{b}_{\mathrm{DM}})$ | 92.9 | 0.219 | 0.215 | 0.437 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathrm{CV}})$ | 91.1 | 0.215 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{CV}}, \hat{h}_{\mathrm{CV}})$ | 79.1 | 0.215 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{CV}}, \hat{h}_{\mathrm{CV}})$ | 91.8 | 0.319 | 0.177 | 0.177 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 89.4 | 0.211 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 88.7 | 0.211 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 92.7 | 0.243 | 0.182 | 0.328 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 79.4 | 0.177 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 92.7 | 0.260 | 0.260 | 0.260 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, h_{\mathrm{DM}})$ | 79.5 | 0.180 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, h_{\mathrm{DM}})$ | 92.6 | 0.265 | 0.251 | 0.251 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 79.1 | 0.184 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 92.4 | 0.272 | 0.241 | 0.241 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{h}_{\mathrm{DM}})$ | 79.3 | 0.194 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{h}_{\mathrm{DM}})$ | 92.1 | 0.287 | 0.215 | 0.215 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 79.7 | 0.211 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 92.2 | 0.312 | 0.182 | 0.182 |

[a](i) EC = Empirical Coverage in percentage points, (ii) IL = Average Interval Length, (iii) columns under "Bandwidths" report the population and average estimated bandwidths choices, as appropriate, for main bandwidth $h_n$ and pilot bandwidth $b_n$.

TABLE S.A.IV

EMPIRICAL COVERAGE AND AVERAGE INTERVAL LENGTH OF DIFFERENT 95% CONFIDENCE INTERVALS USING ESTIMATED ASYMPTOTIC VARIANCE WITH PLUG-IN RESIDUALS ESTIMATES[a]

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 1 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 91.0 | 0.213 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 86.0 | 0.213 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 92.2 | 0.258 | 0.166 | 0.251 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 90.7 | 0.203 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 85.6 | 0.203 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 92.2 | 0.248 | 0.184 | 0.271 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 81.5 | 0.149 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 69.1 | 0.149 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 91.1 | 0.262 | 0.375 | 0.350 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 77.0 | 0.133 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 60.2 | 0.133 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 87.5 | 0.263 | 0.501 | 0.429 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 79.0 | 0.141 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 74.5 | 0.141 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 90.0 | 0.206 | 0.428 | 0.428 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 88.4 | 0.195 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 84.7 | 0.195 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 90.7 | 0.231 | 0.203 | 0.332 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 78.4 | 0.213 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 92.0 | 0.315 | 0.166 | 0.166 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 78.4 | 0.203 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 92.2 | 0.299 | 0.184 | 0.184 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 76.8 | 0.149 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 91.2 | 0.219 | 0.375 | 0.375 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 70.1 | 0.133 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 87.6 | 0.193 | 0.501 | 0.501 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 79.0 | 0.195 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 92.4 | 0.288 | 0.203 | 0.203 |
| Model 2 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 86.4 | 0.290 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 87.1 | 0.290 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 89.9 | 0.315 | 0.082 | 0.189 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 86.3 | 0.287 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 87.1 | 0.287 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 90.0 | 0.314 | 0.084 | 0.190 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 30.1 | 0.223 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 84.2 | 0.223 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 90.1 | 0.262 | 0.184 | 0.325 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 18.1 | 0.221 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 82.6 | 0.221 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 88.8 | 0.260 | 0.198 | 0.347 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 72.8 | 0.249 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 77.2 | 0.249 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 91.6 | 0.376 | 0.124 | 0.124 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 80.8 | 0.265 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 87.7 | 0.265 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 90.5 | 0.289 | 0.104 | 0.237 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 73.4 | 0.290 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 89.3 | 0.441 | 0.082 | 0.082 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 73.7 | 0.287 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 89.4 | 0.437 | 0.084 | 0.084 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 81.3 | 0.223 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 94.4 | 0.344 | 0.184 | 0.184 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 82.7 | 0.221 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 95.2 | 0.342 | 0.198 | 0.198 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 75.9 | 0.265 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 90.5 | 0.399 | 0.104 | 0.104 |

(*Continues*)

TABLE S.A.IV—*Continued*

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 3 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 84.0 | 0.175 | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 84.6 | 0.175 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 93.6 | 0.229 | 0.260 | 0.322 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 85.6 | 0.177 | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 84.2 | 0.177 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 93.4 | 0.234 | 0.251 | 0.305 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 83.6 | 0.181 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 86.1 | 0.181 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 93.5 | 0.227 | 0.241 | 0.352 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 88.8 | 0.190 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 89.0 | 0.190 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 92.9 | 0.214 | 0.215 | 0.437 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 90.8 | 0.207 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 78.5 | 0.207 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 92.2 | 0.307 | 0.177 | 0.177 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 89.1 | 0.205 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 88.4 | 0.205 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 92.6 | 0.236 | 0.182 | 0.328 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 79.7 | 0.175 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 93.4 | 0.258 | 0.260 | 0.260 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 79.5 | 0.177 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 93.2 | 0.262 | 0.251 | 0.251 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 79.1 | 0.181 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 93.2 | 0.268 | 0.241 | 0.241 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 79.1 | 0.190 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 92.9 | 0.280 | 0.215 | 0.215 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 79.1 | 0.205 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 92.5 | 0.302 | 0.182 | 0.182 |

[a](i) EC = Empirical Coverage in percentage points, (ii) IL = Average Interval Length, (iii) columns under "Bandwidths" report the population and average estimated bandwidths choices, as appropriate, for main bandwidth $h_n$ and pilot bandwidth $b_n$.

TABLE S.A.V

EMPIRICAL COVERAGE AND AVERAGE INTERVAL LENGTH OF DIFFERENT 95% CONFIDENCE INTERVALS USING INFEASIBLE ASYMPTOTIC VARIANCE AND AD-HOC “UNDERSMOOTHING” (ALL BANDWIDTHS DIVIDED BY 2)[a]

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 1 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 94.9 | 0.326 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 89.0 | 0.326 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 94.7 | 0.401 | 0.083 | 0.126 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 94.9 | 0.308 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 88.5 | 0.308 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 94.9 | 0.382 | 0.092 | 0.135 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 93.0 | 0.216 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 74.3 | 0.216 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 95.2 | 0.397 | 0.187 | 0.175 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 90.4 | 0.188 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 69.5 | 0.188 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 94.6 | 0.408 | 0.250 | 0.215 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 91.8 | 0.203 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 81.6 | 0.203 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 94.8 | 0.303 | 0.214 | 0.214 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 95.0 | 0.297 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 90.7 | 0.297 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 95.2 | 0.353 | 0.102 | 0.166 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 79.5 | 0.326 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 95.2 | 0.512 | 0.083 | 0.083 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 80.2 | 0.308 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 95.1 | 0.478 | 0.092 | 0.092 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 81.0 | 0.216 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 94.8 | 0.325 | 0.187 | 0.187 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 81.7 | 0.188 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 95.0 | 0.281 | 0.250 | 0.250 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 81.1 | 0.297 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 95.5 | 0.460 | 0.102 | 0.102 |
| Model 2 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | NA | NA | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | NA | NA | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | NA | NA | 0.041 | 0.095 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | NA | NA | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | NA | NA | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | NA | NA | 0.042 | 0.095 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 90.4 | 0.309 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 90.6 | 0.309 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 94.7 | 0.359 | 0.092 | 0.162 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | NA | NA | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | NA | NA | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | NA | NA | 0.099 | 0.173 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 93.7 | 0.387 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 78.5 | 0.387 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 95.2 | 0.636 | 0.062 | 0.062 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 94.3 | 0.449 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 92.6 | 0.449 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 94.9 | 0.490 | 0.048 | 0.111 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | NA | NA | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | NA | NA | 0.041 | 0.041 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | NA | NA | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | NA | NA | 0.042 | 0.042 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 79.9 | 0.309 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 94.9 | 0.480 | 0.092 | 0.092 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | NA | NA | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | NA | NA | 0.099 | 0.099 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 76.9 | 0.449 | $I^{\mathrm{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 95.5 | 0.793 | 0.048 | 0.048 |

(*Continues*)

TABLE S.A.V—*Continued*

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 3 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 94.8 | 0.256 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 86.2 | 0.256 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 94.7 | 0.339 | 0.130 | 0.161 |
| $I_{\mathtt{SRD}}(h_{\mathrm{DM}})$ | 94.8 | 0.261 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, b_{\mathrm{DM}})$ | 85.7 | 0.261 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, b_{\mathrm{DM}})$ | 94.9 | 0.350 | 0.125 | 0.152 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 94.5 | 0.268 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 88.5 | 0.268 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 94.9 | 0.338 | 0.120 | 0.176 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}})$ | 94.8 | 0.284 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{b}_{\mathrm{DM}})$ | 91.8 | 0.284 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{b}_{\mathrm{DM}})$ | 94.9 | 0.320 | 0.108 | 0.218 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathrm{CV}})$ | 95.0 | 0.317 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{CV}}, \hat{h}_{\mathrm{CV}})$ | 79.8 | 0.317 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{CV}}, \hat{h}_{\mathrm{CV}})$ | 95.2 | 0.495 | 0.088 | 0.088 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 94.9 | 0.312 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 91.5 | 0.312 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 95.1 | 0.362 | 0.091 | 0.164 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 81.0 | 0.256 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 94.9 | 0.388 | 0.130 | 0.130 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, h_{\mathrm{DM}})$ | 80.8 | 0.261 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, h_{\mathrm{DM}})$ | 95.0 | 0.396 | 0.125 | 0.125 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 80.5 | 0.268 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 94.7 | 0.409 | 0.120 | 0.120 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{h}_{\mathrm{DM}})$ | 80.5 | 0.284 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{h}_{\mathrm{DM}})$ | 95.0 | 0.436 | 0.108 | 0.108 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 80.4 | 0.312 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 95.5 | 0.487 | 0.091 | 0.091 |

[a](i) EC = Empirical Coverage in percentage points, (ii) IL = Average Interval Length, (iii) columns under "Bandwidths" report the population and average estimated bandwidths choices, as appropriate, for main bandwidth $h_n$ and pilot bandwidth $b_n$ (all divided by 2 to perform ad hoc "undersmoothing"), (iv) NA = not available due to numerical instability.

TABLE S.A.VI

EMPIRICAL COVERAGE AND AVERAGE INTERVAL LENGTH OF DIFFERENT 95% CONFIDENCE INTERVALS USING ESTIMATED ASYMPTOTIC VARIANCE WITH $J = 3$ NEAREST-NEIGHBORS AND AD-HOC "UNDERSMOOTHING" (ALL BANDWIDTHS DIVIDED BY 2)[a]

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 1 | | | | | | | | | | |
| $I_{\text{SRD}}(h_{\text{MSE}})$ | 91.9 | 0.319 | $I^{\text{bc}}_{\text{SRD}}(h_{\text{MSE}}, b_{\text{MSE}})$ | 85.9 | 0.319 | $I^{\text{rbc}}_{\text{SRD}}(h_{\text{MSE}}, b_{\text{MSE}})$ | 92.2 | 0.391 | 0.083 | 0.126 |
| $I_{\text{SRD}}(h_{\text{DM}})$ | 92.2 | 0.301 | $I^{\text{bc}}_{\text{SRD}}(h_{\text{DM}}, b_{\text{DM}})$ | 85.8 | 0.301 | $I^{\text{rbc}}_{\text{SRD}}(h_{\text{DM}}, b_{\text{DM}})$ | 92.3 | 0.373 | 0.092 | 0.135 |
| $I_{\text{SRD}}(\hat{h}_{\text{IK}})$ | 91.4 | 0.214 | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{IK}}, \hat{b}_{\text{IK}})$ | 72.6 | 0.214 | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{IK}}, \hat{b}_{\text{IK}})$ | 92.8 | 0.389 | 0.187 | 0.175 |
| $I_{\text{SRD}}(\hat{h}_{\text{DM}})$ | 89.0 | 0.186 | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{DM}}, \hat{b}_{\text{DM}})$ | 68.5 | 0.186 | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{DM}}, \hat{b}_{\text{DM}})$ | 92.6 | 0.401 | 0.250 | 0.215 |
| $I_{\text{SRD}}(\hat{h}_{\text{CV}})$ | 90.4 | 0.201 | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{CV}}, \hat{h}_{\text{CV}})$ | 80.1 | 0.201 | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{CV}}, \hat{h}_{\text{CV}})$ | 92.9 | 0.298 | 0.214 | 0.214 |
| $I_{\text{SRD}}(\hat{h}_{\text{CCT}})$ | 92.1 | 0.288 | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{CCT}}, \hat{b}_{\text{CCT}})$ | 87.7 | 0.288 | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{CCT}}, \hat{b}_{\text{CCT}})$ | 92.6 | 0.343 | 0.102 | 0.166 |
| | | | $I^{\text{bc}}_{\text{SRD}}(h_{\text{MSE}}, h_{\text{MSE}})$ | 77.2 | 0.319 | $I^{\text{rbc}}_{\text{SRD}}(h_{\text{MSE}}, h_{\text{MSE}})$ | 91.4 | 0.492 | 0.083 | 0.083 |
| | | | $I^{\text{bc}}_{\text{SRD}}(h_{\text{DM}}, h_{\text{DM}})$ | 77.9 | 0.301 | $I^{\text{rbc}}_{\text{SRD}}(h_{\text{DM}}, h_{\text{DM}})$ | 91.3 | 0.460 | 0.092 | 0.092 |
| | | | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{IK}}, \hat{h}_{\text{IK}})$ | 79.7 | 0.214 | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{IK}}, \hat{h}_{\text{IK}})$ | 92.7 | 0.318 | 0.187 | 0.187 |
| | | | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{DM}}, \hat{h}_{\text{DM}})$ | 80.5 | 0.186 | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{DM}}, \hat{h}_{\text{DM}})$ | 93.0 | 0.276 | 0.250 | 0.250 |
| | | | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{CCT}}, \hat{h}_{\text{CCT}})$ | 78.7 | 0.288 | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{CCT}}, \hat{h}_{\text{CCT}})$ | 92.1 | 0.440 | 0.102 | 0.102 |
| Model 2 | | | | | | | | | | |
| $I_{\text{SRD}}(h_{\text{MSE}})$ | NA | NA | $I^{\text{bc}}_{\text{SRD}}(h_{\text{MSE}}, b_{\text{MSE}})$ | NA | NA | $I^{\text{rbc}}_{\text{SRD}}(h_{\text{MSE}}, b_{\text{MSE}})$ | NA | NA | 0.041 | 0.095 |
| $I_{\text{SRD}}(h_{\text{DM}})$ | NA | NA | $I^{\text{bc}}_{\text{SRD}}(h_{\text{DM}}, b_{\text{DM}})$ | NA | NA | $I^{\text{rbc}}_{\text{SRD}}(h_{\text{DM}}, b_{\text{DM}})$ | NA | NA | 0.042 | 0.095 |
| $I_{\text{SRD}}(\hat{h}_{\text{IK}})$ | 89.4 | 0.334 | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{IK}}, \hat{b}_{\text{IK}})$ | 90.0 | 0.334 | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{IK}}, \hat{b}_{\text{IK}})$ | 93.8 | 0.390 | 0.092 | 0.162 |
| $I_{\text{SRD}}(\hat{h}_{\text{DM}})$ | NA | NA | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{DM}}, \hat{b}_{\text{DM}})$ | NA | NA | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{DM}}, \hat{b}_{\text{DM}})$ | NA | NA | 0.099 | 0.173 |
| $I_{\text{SRD}}(\hat{h}_{\text{CV}})$ | 91.9 | 0.425 | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{CV}}, \hat{h}_{\text{CV}})$ | 79.0 | 0.425 | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{CV}}, \hat{h}_{\text{CV}})$ | 93.4 | 0.704 | 0.062 | 0.062 |
| $I_{\text{SRD}}(\hat{h}_{\text{CCT}})$ | 92.7 | 0.492 | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{CCT}}, \hat{b}_{\text{CCT}})$ | 91.3 | 0.492 | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{CCT}}, \hat{b}_{\text{CCT}})$ | 93.4 | 0.537 | 0.048 | 0.111 |
| | | | $I^{\text{bc}}_{\text{SRD}}(h_{\text{MSE}}, h_{\text{MSE}})$ | NA | NA | $I^{\text{rbc}}_{\text{SRD}}(h_{\text{MSE}}, h_{\text{MSE}})$ | NA | NA | 0.041 | 0.041 |
| | | | $I^{\text{bc}}_{\text{SRD}}(h_{\text{DM}}, h_{\text{DM}})$ | NA | NA | $I^{\text{rbc}}_{\text{SRD}}(h_{\text{DM}}, h_{\text{DM}})$ | NA | NA | 0.042 | 0.042 |
| | | | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{IK}}, \hat{h}_{\text{IK}})$ | 80.8 | 0.334 | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{IK}}, \hat{h}_{\text{IK}})$ | 93.2 | 0.531 | 0.092 | 0.092 |
| | | | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{DM}}, \hat{h}_{\text{DM}})$ | NA | NA | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{DM}}, \hat{h}_{\text{DM}})$ | NA | NA | 0.099 | 0.099 |
| | | | $I^{\text{bc}}_{\text{SRD}}(\hat{h}_{\text{CCT}}, \hat{h}_{\text{CCT}})$ | 77.6 | 0.492 | $I^{\text{rbc}}_{\text{SRD}}(\hat{h}_{\text{CCT}}, \hat{h}_{\text{CCT}})$ | 93.4 | 0.862 | 0.048 | 0.048 |

(*Continues*)

TABLE S.A.VI—*Continued*

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 3 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 92.9 | 0.252 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 84.2 | 0.252 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 92.5 | 0.333 | 0.130 | 0.161 |
| $I_{\mathtt{SRD}}(h_{\mathrm{DM}})$ | 92.7 | 0.257 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, b_{\mathrm{DM}})$ | 83.7 | 0.257 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, b_{\mathrm{DM}})$ | 92.5 | 0.343 | 0.125 | 0.152 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 92.6 | 0.263 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 86.2 | 0.263 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 93.0 | 0.331 | 0.120 | 0.176 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}})$ | 92.5 | 0.278 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{b}_{\mathrm{DM}})$ | 89.3 | 0.278 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{b}_{\mathrm{DM}})$ | 92.8 | 0.314 | 0.108 | 0.218 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathrm{CV}})$ | 92.3 | 0.310 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{CV}}, \hat{h}_{\mathrm{CV}})$ | 77.5 | 0.310 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{CV}}, \hat{h}_{\mathrm{CV}})$ | 91.7 | 0.476 | 0.088 | 0.088 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 92.2 | 0.303 | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 88.8 | 0.303 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 92.7 | 0.351 | 0.091 | 0.164 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 79.1 | 0.252 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 92.2 | 0.378 | 0.130 | 0.130 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, h_{\mathrm{DM}})$ | 79.2 | 0.257 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(h_{\mathrm{DM}}, h_{\mathrm{DM}})$ | 92.1 | 0.386 | 0.125 | 0.125 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 78.8 | 0.263 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 91.8 | 0.397 | 0.120 | 0.120 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{h}_{\mathrm{DM}})$ | 78.3 | 0.278 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathrm{DM}}, \hat{h}_{\mathrm{DM}})$ | 91.6 | 0.422 | 0.108 | 0.108 |
| | | | $I^{\mathrm{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 77.8 | 0.303 | $I^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 91.6 | 0.466 | 0.091 | 0.091 |

[a](i) EC = Empirical Coverage in percentage points, (ii) IL = Average Interval Length, (iii) columns under "Bandwidths" report the population and average estimated bandwidths choices, as appropriate, for main bandwidth $h_n$ and pilot bandwidth $b_n$ (all divided by 2 to perform ad hoc "undersmoothing"), (iv) NA = not available due to numerical instability.

TABLE S.A.VII

EMPIRICAL COVERAGE AND AVERAGE INTERVAL LENGTH OF DIFFERENT 95% CONFIDENCE INTERVALS USING ESTIMATED ASYMPTOTIC VARIANCE WITH $J = 1$ NEAREST-NEIGHBORS AND AD-HOC "UNDERSMOOTHING" (ALL BANDWIDTHS DIVIDED BY 2)[a]

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 1 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 91.0 | 0.315 | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 85.1 | 0.315 | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 91.0 | 0.386 | 0.083 | 0.126 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 91.3 | 0.298 | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 84.9 | 0.298 | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 91.3 | 0.369 | 0.092 | 0.135 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 90.7 | 0.212 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 72.3 | 0.212 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 91.8 | 0.385 | 0.187 | 0.175 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 88.4 | 0.185 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 68.3 | 0.185 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 92.0 | 0.397 | 0.250 | 0.215 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 89.8 | 0.200 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 79.5 | 0.200 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 92.1 | 0.295 | 0.214 | 0.214 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 91.0 | 0.284 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 86.9 | 0.284 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 91.6 | 0.338 | 0.102 | 0.166 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 75.9 | 0.315 | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 89.1 | 0.484 | 0.083 | 0.083 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 76.6 | 0.298 | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 89.4 | 0.453 | 0.092 | 0.092 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 79.0 | 0.212 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 91.8 | 0.315 | 0.187 | 0.187 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 80.4 | 0.185 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 92.3 | 0.274 | 0.250 | 0.250 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 77.4 | 0.284 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 90.1 | 0.433 | 0.102 | 0.102 |
| Model 2 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | NA | NA | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | NA | NA | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | NA | NA | 0.041 | 0.095 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | NA | NA | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | NA | NA | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | NA | NA | 0.042 | 0.095 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 87.2 | 0.310 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 87.7 | 0.310 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 92.0 | 0.361 | 0.092 | 0.162 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | NA | NA | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | NA | NA | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | NA | NA | 0.099 | 0.173 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 89.3 | 0.386 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 74.6 | 0.386 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 89.5 | 0.627 | 0.062 | 0.062 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 89.2 | 0.444 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 87.7 | 0.444 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 89.9 | 0.485 | 0.048 | 0.110 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | NA | NA | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | NA | NA | 0.041 | 0.041 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | NA | NA | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | NA | NA | 0.042 | 0.042 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 77.7 | 0.310 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 90.4 | 0.478 | 0.092 | 0.092 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | NA | NA | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | NA | NA | 0.099 | 0.099 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 72.0 | 0.444 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 88.9 | 0.778 | 0.048 | 0.048 |

(*Continues*)

TABLE S.A.VII—*Continued*

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 3 | | | | | | | | | | |
| $I_{\texttt{SRD}}(h_{\texttt{MSE}})$ | 92.2 | 0.250 | $I^{\text{bc}}_{\texttt{SRD}}(h_{\texttt{MSE}}, b_{\texttt{MSE}})$ | 83.4 | 0.250 | $I^{\text{rbc}}_{\texttt{SRD}}(h_{\texttt{MSE}}, b_{\texttt{MSE}})$ | 91.8 | 0.329 | 0.130 | 0.161 |
| $I_{\texttt{SRD}}(h_{\text{DM}})$ | 92.1 | 0.254 | $I^{\text{bc}}_{\texttt{SRD}}(h_{\text{DM}}, b_{\text{DM}})$ | 82.9 | 0.254 | $I^{\text{rbc}}_{\texttt{SRD}}(h_{\text{DM}}, b_{\text{DM}})$ | 91.6 | 0.339 | 0.125 | 0.152 |
| $I_{\texttt{SRD}}(\hat{h}_{\texttt{IK}})$ | 92.0 | 0.261 | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\texttt{IK}}, \hat{b}_{\texttt{IK}})$ | 85.1 | 0.261 | $I^{\text{rbc}}_{\texttt{SRD}}(\hat{h}_{\texttt{IK}}, \hat{b}_{\texttt{IK}})$ | 92.1 | 0.327 | 0.120 | 0.176 |
| $I_{\texttt{SRD}}(\hat{h}_{\text{DM}})$ | 91.8 | 0.275 | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\text{DM}}, \hat{b}_{\text{DM}})$ | 88.5 | 0.275 | $I^{\text{rbc}}_{\texttt{SRD}}(\hat{h}_{\text{DM}}, \hat{b}_{\text{DM}})$ | 92.1 | 0.311 | 0.108 | 0.218 |
| $I_{\texttt{SRD}}(\hat{h}_{\text{CV}})$ | 91.3 | 0.306 | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\text{CV}}, \hat{h}_{\text{CV}})$ | 76.4 | 0.306 | $I^{\text{rbc}}_{\texttt{SRD}}(\hat{h}_{\text{CV}}, \hat{h}_{\text{CV}})$ | 89.6 | 0.468 | 0.088 | 0.088 |
| $I_{\texttt{SRD}}(\hat{h}_{\texttt{CCT}})$ | 91.2 | 0.299 | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\texttt{CCT}}, \hat{b}_{\texttt{CCT}})$ | 87.6 | 0.299 | $I^{\text{rbc}}_{\texttt{SRD}}(\hat{h}_{\texttt{CCT}}, \hat{b}_{\texttt{CCT}})$ | 91.8 | 0.346 | 0.091 | 0.164 |
| | | | $I^{\text{bc}}_{\texttt{SRD}}(h_{\texttt{MSE}}, h_{\texttt{MSE}})$ | 78.3 | 0.250 | $I^{\text{rbc}}_{\texttt{SRD}}(h_{\texttt{MSE}}, h_{\texttt{MSE}})$ | 90.8 | 0.373 | 0.130 | 0.130 |
| | | | $I^{\text{bc}}_{\texttt{SRD}}(h_{\text{DM}}, h_{\text{DM}})$ | 78.0 | 0.254 | $I^{\text{rbc}}_{\texttt{SRD}}(h_{\text{DM}}, h_{\text{DM}})$ | 90.8 | 0.381 | 0.125 | 0.125 |
| | | | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\texttt{IK}}, \hat{h}_{\texttt{IK}})$ | 77.8 | 0.261 | $I^{\text{rbc}}_{\texttt{SRD}}(\hat{h}_{\texttt{IK}}, \hat{h}_{\texttt{IK}})$ | 90.4 | 0.391 | 0.120 | 0.120 |
| | | | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\text{DM}}, \hat{h}_{\text{DM}})$ | 77.4 | 0.275 | $I^{\text{rbc}}_{\texttt{SRD}}(\hat{h}_{\text{DM}}, \hat{h}_{\text{DM}})$ | 90.0 | 0.415 | 0.108 | 0.108 |
| | | | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\texttt{CCT}}, \hat{h}_{\texttt{CCT}})$ | 76.5 | 0.299 | $I^{\text{rbc}}_{\texttt{SRD}}(\hat{h}_{\texttt{CCT}}, \hat{h}_{\texttt{CCT}})$ | 89.6 | 0.458 | 0.091 | 0.091 |

[a](i) EC = Empirical Coverage in percentage points, (ii) IL = Average Interval Length, (iii) columns under "Bandwidths" report the population and average estimated bandwidths choices, as appropriate, for main bandwidth $h_n$ and pilot bandwidth $b_n$ (all divided by 2 to perform ad hoc "undersmoothing"), (iv) NA = not available due to numerical instability.

TABLE S.A.VIII

EMPIRICAL COVERAGE AND AVERAGE INTERVAL LENGTH OF DIFFERENT 95% CONFIDENCE INTERVALS USING ESTIMATED ASYMPTOTIC VARIANCE WITH PLUG-IN RESIDUALS ESTIMATES AND AD-HOC "UNDERSMOOTHING" (ALL BANDWIDTHS DIVIDED BY 2)[a]

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 1 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | 89.5 | 0.287 | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 82.7 | 0.287 | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | 89.6 | 0.352 | 0.083 | 0.126 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | 90.3 | 0.275 | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 82.9 | 0.275 | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | 90.2 | 0.340 | 0.092 | 0.135 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 90.4 | 0.205 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 71.1 | 0.205 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 92.5 | 0.371 | 0.187 | 0.175 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | 88.1 | 0.181 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 67.4 | 0.181 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | 92.6 | 0.387 | 0.250 | 0.215 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 89.6 | 0.193 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 78.8 | 0.193 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 92.6 | 0.285 | 0.214 | 0.214 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 90.2 | 0.265 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 85.3 | 0.265 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 90.6 | 0.315 | 0.102 | 0.166 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 73.1 | 0.287 | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | 89.0 | 0.436 | 0.083 | 0.083 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 74.4 | 0.275 | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | 89.7 | 0.413 | 0.092 | 0.092 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 78.3 | 0.205 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 92.3 | 0.303 | 0.187 | 0.187 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 79.6 | 0.181 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | 93.0 | 0.266 | 0.250 | 0.250 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 75.6 | 0.265 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 90.1 | 0.399 | 0.102 | 0.102 |
| Model 2 | | | | | | | | | | |
| $I_{\mathtt{SRD}}(h_{\mathtt{MSE}})$ | NA | NA | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | NA | NA | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, b_{\mathtt{MSE}})$ | NA | NA | 0.041 | 0.095 |
| $I_{\mathtt{SRD}}(h_{\mathtt{DM}})$ | NA | NA | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | NA | NA | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, b_{\mathtt{DM}})$ | NA | NA | 0.042 | 0.095 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}})$ | 84.7 | 0.279 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 85.5 | 0.279 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{b}_{\mathtt{IK}})$ | 90.5 | 0.325 | 0.092 | 0.162 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}})$ | NA | NA | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | NA | NA | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{b}_{\mathtt{DM}})$ | NA | NA | 0.099 | 0.173 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}})$ | 85.8 | 0.322 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 69.2 | 0.322 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CV}}, \hat{h}_{\mathtt{CV}})$ | 86.9 | 0.511 | 0.062 | 0.062 |
| $I_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}})$ | 85.0 | 0.341 | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 82.9 | 0.341 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{b}_{\mathtt{CCT}})$ | 85.9 | 0.375 | 0.052 | 0.119 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | NA | NA | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{MSE}}, h_{\mathtt{MSE}})$ | NA | NA | 0.041 | 0.041 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | NA | NA | $I^{\text{rbc}}_{\mathtt{SRD}}(h_{\mathtt{DM}}, h_{\mathtt{DM}})$ | NA | NA | 0.042 | 0.042 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 74.2 | 0.279 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{IK}}, \hat{h}_{\mathtt{IK}})$ | 89.7 | 0.421 | 0.092 | 0.092 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | NA | NA | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{DM}}, \hat{h}_{\mathtt{DM}})$ | NA | NA | 0.099 | 0.099 |
| | | | $I^{\text{bc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 65.8 | 0.341 | $I^{\text{rbc}}_{\mathtt{SRD}}(\hat{h}_{\mathtt{CCT}}, \hat{h}_{\mathtt{CCT}})$ | 85.7 | 0.562 | 0.052 | 0.052 |

(*Continues*)

TABLE S.A.VIII—*Continued*

| Conventional | | | Bias-Corrected | | | Robust Approach | | | Bandwidths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | EC (%) | IL | | EC (%) | IL | | EC (%) | IL | $h_n$ | $b_n$ |
| Model 3 | | | | | | | | | | |
| $I_{\texttt{SRD}}(h_{\texttt{MSE}})$ | 91.6 | 0.237 | $I^{\text{bc}}_{\texttt{SRD}}(h_{\texttt{MSE}}, b_{\texttt{MSE}})$ | 82.2 | 0.237 | $I^{\texttt{rbc}}_{\texttt{SRD}}(h_{\texttt{MSE}}, b_{\texttt{MSE}})$ | 91.5 | 0.312 | 0.130 | 0.161 |
| $I_{\texttt{SRD}}(h_{\text{DM}})$ | 91.4 | 0.241 | $I^{\text{bc}}_{\texttt{SRD}}(h_{\text{DM}}, b_{\text{DM}})$ | 81.6 | 0.241 | $I^{\texttt{rbc}}_{\texttt{SRD}}(h_{\text{DM}}, b_{\text{DM}})$ | 91.3 | 0.320 | 0.125 | 0.152 |
| $I_{\texttt{SRD}}(\hat{h}_{\texttt{IK}})$ | 91.1 | 0.246 | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\texttt{IK}}, \hat{b}_{\texttt{IK}})$ | 84.1 | 0.246 | $I^{\texttt{rbc}}_{\texttt{SRD}}(\hat{h}_{\texttt{IK}}, \hat{b}_{\texttt{IK}})$ | 91.6 | 0.308 | 0.120 | 0.176 |
| $I_{\texttt{SRD}}(\hat{h}_{\text{DM}})$ | 90.8 | 0.258 | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\text{DM}}, \hat{b}_{\text{DM}})$ | 87.4 | 0.258 | $I^{\texttt{rbc}}_{\texttt{SRD}}(\hat{h}_{\text{DM}}, \hat{b}_{\text{DM}})$ | 91.3 | 0.291 | 0.108 | 0.218 |
| $I_{\texttt{SRD}}(\hat{h}_{\texttt{CV}})$ | 90.1 | 0.280 | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\texttt{CV}}, \hat{h}_{\texttt{CV}})$ | 73.4 | 0.280 | $I^{\texttt{rbc}}_{\texttt{SRD}}(\hat{h}_{\texttt{CV}}, \hat{h}_{\texttt{CV}})$ | 89.4 | 0.424 | 0.088 | 0.088 |
| $I_{\texttt{SRD}}(\hat{h}_{\texttt{CCT}})$ | 90.2 | 0.276 | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\texttt{CCT}}, \hat{b}_{\texttt{CCT}})$ | 86.1 | 0.276 | $I^{\texttt{rbc}}_{\texttt{SRD}}(\hat{h}_{\texttt{CCT}}, \hat{b}_{\texttt{CCT}})$ | 90.7 | 0.320 | 0.091 | 0.164 |
| | | | $I^{\text{bc}}_{\texttt{SRD}}(h_{\texttt{MSE}}, h_{\texttt{MSE}})$ | 77.1 | 0.237 | $I^{\texttt{rbc}}_{\texttt{SRD}}(h_{\texttt{MSE}}, h_{\texttt{MSE}})$ | 91.1 | 0.352 | 0.130 | 0.130 |
| | | | $I^{\text{bc}}_{\texttt{SRD}}(h_{\text{DM}}, h_{\text{DM}})$ | 76.7 | 0.241 | $I^{\texttt{rbc}}_{\texttt{SRD}}(h_{\text{DM}}, h_{\text{DM}})$ | 90.9 | 0.358 | 0.125 | 0.125 |
| | | | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\texttt{IK}}, \hat{h}_{\texttt{IK}})$ | 76.5 | 0.246 | $I^{\texttt{rbc}}_{\texttt{SRD}}(\hat{h}_{\texttt{IK}}, \hat{h}_{\texttt{IK}})$ | 90.6 | 0.366 | 0.120 | 0.120 |
| | | | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\text{DM}}, \hat{h}_{\text{DM}})$ | 75.4 | 0.258 | $I^{\texttt{rbc}}_{\texttt{SRD}}(\hat{h}_{\text{DM}}, \hat{h}_{\text{DM}})$ | 90.2 | 0.385 | 0.108 | 0.108 |
| | | | $I^{\text{bc}}_{\texttt{SRD}}(\hat{h}_{\texttt{CCT}}, \hat{h}_{\texttt{CCT}})$ | 74.3 | 0.276 | $I^{\texttt{rbc}}_{\texttt{SRD}}(\hat{h}_{\texttt{CCT}}, \hat{h}_{\texttt{CCT}})$ | 89.5 | 0.418 | 0.091 | 0.091 |

[a](i) EC = Empirical Coverage in percentage points, (ii) IL = Average Interval Length, (iii) columns under “Bandwidths” report the population and average estimated bandwidths choices, as appropriate, for main bandwidth $h_n$ and pilot bandwidth $b_n$ (all divided by 2 to perform ad hoc “undersmoothing”), (iv) NA = not available due to numerical instability.

used: because shrinking the bandwidth reduces the effective sample size, the Gaussian approximation may fail if a non-Gaussian data generating processes is used (notice that here, $\varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ in all cases). In any case, our simulation evidence indicates that our approach performs as well as, if not better than, the conventional one in all cases considered. In particular, the conventional and robust confidence intervals exhibit similar coverage rates and interval lengths when the bandwidth is “small” and, in addition, our robust confidence intervals continue to perform well when the bandwidth is “large”.

## S.4. EMPIRICAL ILLUSTRATION

In this section, we illustrate the performance of our methods and compare them to other conventional alternatives employing household data from Oportunidades (formerly known as Progresa), a well-known large-scale anti-poverty conditional cash transfer program in Mexico. Our goal is to show how the different methods perform in a substantive, realistic empirical application: the impact of the program on household consumption. All estimates and figures were constructed using the STATA package described in Calonico, Cattaneo, and Titiunik (2014b). See Calonico, Cattaneo, and Titiunik (2014d) for a companion R package.

### S.4.1. *The Program*

Oportunidades is a conditional cash transfer program that targets poor households in rural and urban areas in Mexico. The program started in 1998 under the name of Progresa in rural areas. The most important elements of the program are the nutrition, health, and education components. The nutrition component consists of a cash grant for all treated households and an additional supplement for households with young children and pregnant or lactating mothers. The educational grant is linked to regular attendance in school and starts on the third grade of primary school and continues until the last grade of secondary school. The transfer constituted a significant contribution to the income of eligible families.

This social program is best known for its experimental design: treatment was initially randomly assigned at the locality level in rural areas. Indeed, its experimental features have spiked a huge body of work focusing on a variety of economic, health, and related outcomes.[2] It was successfully implemented with a take-up rate of around 97%. Progresa was expanded to urban areas in 2003. Unlike the rural program, the allocation across treatment and control

[2]Recent examples include Attanasio, Meghir, and Santiago (2011), Behrman, Gallardo-García, Parker, Todd, and Vélez-Grajales (2012), Djebbari and Smith (2008), Dubois, de Janvry, and Sadoulet (2012), Fernald, Gertler, and Neufeld (2009), among many others. These papers also include references to early reviews and research work.

areas was not random. Instead, it was first offered in blocks with the highest density of poor households.[3] Still, in order to accurately target the program to poor households, in both rural and urban areas Mexican officials constructed a pre-intervention (at baseline) household poverty-index that determined each household's eligibility. In rural communities, seven distinct poverty cutoffs were used depending on the geographic area, while one common cutoff was used in all urban localities. Thus, Progresa/Oportunidades' eligibility assignment rule naturally leads to eight sharp (intention-to-treat) regression-discontinuity designs.[4] Intention-to-treat is a useful policy parameter because it measures the average program effect on the households who are offered the treatment, that is, regardless of whether they participate in the program or not. By ignoring the determinants of participation, it requires less restrictive assumptions than other common parameters such as the average treatment on the treated. Angelucci, Attanasio, and Di Maro (2012) discussed this issue in more detail, and compared different estimates of the impact of Oportunidades on household consumption, savings, and transfers.

We first illustrate our methods employing data from the urban and one of the seven rural RD designs (the one corresponding to the median household population size, Region 3, Sierra-Negra-Zongolica-Mazateca).[5] We do not pool the RD designs, nor do we compare them with each other or to the experimental estimates from the rural areas, since without further (strong) assumptions the associated estimands need not to coincide with each other. Instead, we treat the RD designs as different examples, which vary in observable, and possibly unobservable, characteristics.

Our empirical exercise investigates the program treatment effect on three measures of household consumption: food, non-food, and total consumption expenditures. Studying the effects on consumption is important for several reasons. First, consumption is a measure of household wellbeing and, therefore, changes in consumption reflect more accurately the effectiveness of the program in reducing poverty than other variables. Also, consumption dynamics can capture both the perception that individual households have of the program and its sustainability, by reflecting other changes in behavior and sources of income induced by the program that take time to adjust. We can do this by looking at the effects up to two years after the program started.

Looking at the allocation between food and the rest of consumption is also very relevant. First, one of the main justifications for cash transfers is that poor households might have a better notion of their needs and might target the resources offered by the program more effectively than alternative sources, such as in-kind transfers. It is therefore important to consider how the transfer is

[3]Initial take-up was also much lower (around 50%).

[4]Buddelmeyer and Skoufias (2004) were the first to note the RD features of this social program, and used it to study its effects on child school attendance and child work.

[5]The remaining regions are analyzed at the end of this section.

spent. For example, one could expect the share of food consumption to total expenditures to decrease with an increase in total consumption or, more generally, with living standards. Finally, since the transfers are mainly targeted to women, they could change the balance of power within the household and shift expenditure shares to reflect the increased influence of women and their preferences.

Related literature on this topic includes Hoddinott and Skoufias (2004), Angelucci and Attanasio (2009), Angelucci and De Giorgi (2009), Gertler, Martinez, and Rubio-Codina (2012), and Angelucci and Attanasio (2013), who have also investigated the effect of Oportunidades/Progresa on consumption using experimental methods (in rural areas) and non-experimental matching methods (in urban areas). Our illustrative results therefore contribute to this literature by presenting new empirical evidence based on non-experimental RD estimates.

### S.4.2. *Data*

Our databases correspond to household data for both rural and urban communities in Progresa/Oportunidades. With the exception of the poverty-index and region identifier in rural areas, all the data used in our empirical illustration are publicly available in the following location:

- http://www.oportunidades.gob.mx/EVALUACION/es/eval_cuant/bases_cuanti.php.

The households' poverty-index at baseline and the region identifier for rural areas were obtained with the help of Habiba Djebbari (Université Laval), Paul Gertler (UC-Berkeley), and Jeff Smith (University of Michigan).

In this application, $X_i$ denotes the household's poverty-index, $\bar{x} = 0$ denotes the centered cutoff for each RD design, and $Y_i$ denotes the two different measures of household consumption. Our final database contains 691 control households ($X_i < 0$) and 2,118 intention-to-treat households ($X_i \geq 0$) in the urban RD design ($n = 2,809$, $X_i \in [-2.25, 4.11]$), and 315 control households ($X_i < 0$) and 618 intention-to-treat households ($X_i \geq 0$) in the rural RD design ($n = 933$, $X_i \in [-456.6, 338.4]$). We address the empirical validity of these RD designs by conducting standard balance and falsification tests on pre-intervention covariates. These results give empirical support for the RD assumptions. Figures S.A.1 and S.A.2 present, respectively, the usual RD plots for the urban and rural areas (cf. Figure 1). In these figures, the solid lines correspond to distinct fourth-order global polynomial fits for control and treatments units, and the solid dots correspond to sample averages of the outcome variable for each bin (or partition) of the running variable. The number of bins was chosen as discussed in Calonico, Cattaneo, and Titiunik (2014a).

The final sample sizes and other related information are given in Table S.A.IX.

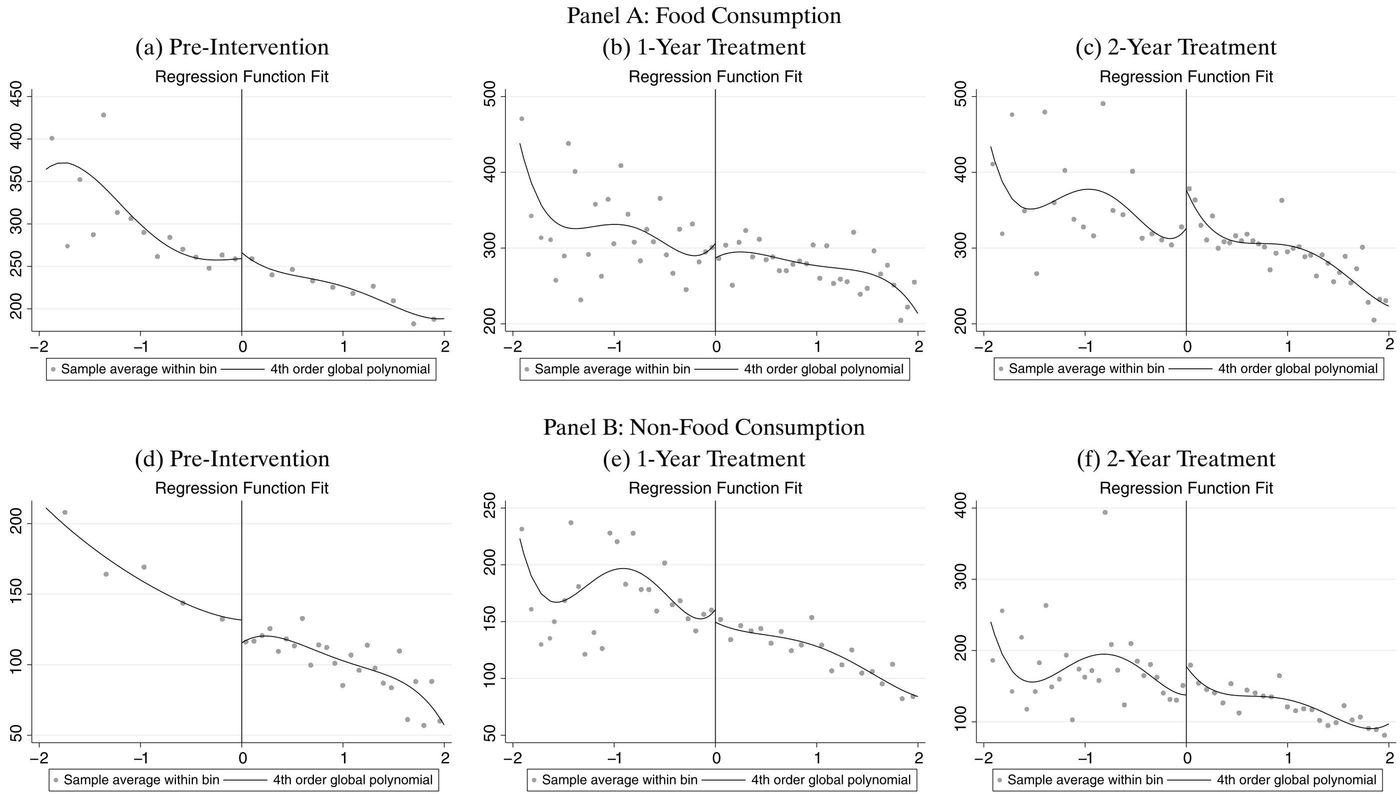


FIGURE S.A.1.—RD plots of Progresa/Oportunidades on food and non-food consumption, urban localities.

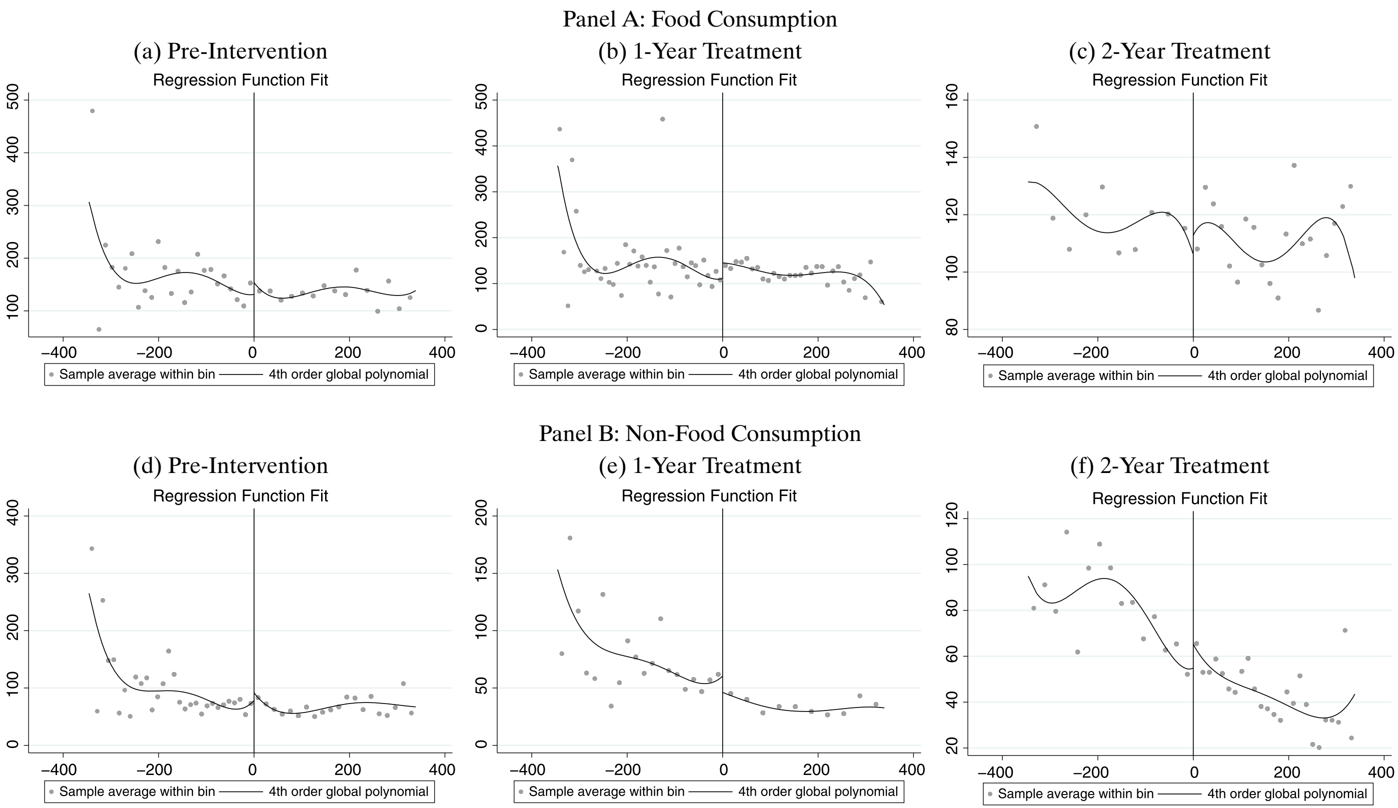


FIGURE S.A.2.—RD plots of Progresa/Oportunidades on food and non-food consumption, rural localities.

TABLE S.A.IX

SAMPLE SIZES AND RD CUTOFFS FOR PROGRESA/OPORTUNIDADES REGIONS

| | All Sample | | Treatment | | Control | |
|---|---|---|---|---|---|---|
| Region | $N$ | Cutoff | $N$ | Range | $N$ | Range |
| Urban | 2809 | 0.69 | 2118 | [0.69, 4.80] | 691 | [−1.56, 0.69] |
| Region 3 | 933 | 759.4 | 618 | [421.0, 759.4] | 315 | [760.0, 1216.0] |
| Region 4 | 1189 | 753.0 | 810 | [365.0, 753.0] | 379 | [754.0, 1184.1] |
| Region 5 | 3116 | 751.5 | 2003 | [354.0, 751.5] | 1113 | [752.0, 1346.0] |
| Region 6 | 541 | 751.0 | 441 | [394.0, 751.0] | 100 | [756.3, 994.0] |
| Region 12 | 78 | 569.0 | 40 | [298.0, 569.0] | 38 | [573.0, 701.0] |
| Region 27 | 828 | 691.0 | 614 | [232.0, 691.0] | 214 | [691.5, 1008.5] |
| Region 28 | 175 | 853.3 | 157 | [309.0, 853.3] | 18 | [863.5, 1006.0] |

### S.4.3. *Main Results*

Our main empirical results are reported in Table S.A.X. Panels A and B correspond, respectively, to the urban and rural RD designs. We consider three time periods: pre-intervention (as a falsification test), one year after the program started (1-Year Treatment), and two years after the program started (2-Year Treatment). Thus, each panel reports six groups of RD estimates (i.e., 2 outcomes $\times$ 3 periods). For each combination of outcome and time period, we conduct RD estimation and inference employing the same setup as in our simulation study: local-linear estimator of $\tau_{\mathtt{SRD}}$, conventional confidence interval, and robust confidence interval (with local-quadratic bias correction), each implemented with the three different data-driven bandwidth choices $\hat{h}_{\mathtt{CCT}}$, $\hat{h}_{\mathtt{IK}}$, and $\hat{h}_{\mathtt{CV}}$. To be specific, for each panel, outcome, period, and bandwidth selection method, we report $\hat{\tau}_{\mathtt{SRD}}(\hat{h}_n)$, $\hat{I}_{\mathtt{SRD}}(\hat{h}_n)$, $\hat{I}^{\mathtt{rbc}}_{\mathtt{SRD}}(\hat{h}_n, \hat{b}_n)$, $\hat{h}_n$, and $\hat{b}_n$.

This empirical exercise offers an array of interesting examples to discuss the performance of our proposed methods. First of all, using the pre-intervention data (columns 1–3, Panels A and B), we find no effects of the program in any case (i.e., food or non-food consumption in urban or rural localities).[6] This result gives additional evidence in favor of the validity of the RD designs, since households in control and treatment areas exhibit, on average, the same levels of pre-intervention consumption. In the 1-year after treatment data, we find statistically significant effects of the program on food consumption in rural

[6]In rural areas, pre- and post-intervention food consumption data differ in two main aspects. First, the pre-intervention survey only provides information on expenditures (i.e., it omits home production). Second, it reports expenditures only by food groups rather than asking detailed item-by-item questions, as in later waves. See, for example, Angelucci and De Giorgi (2009) for further details.

TABLE S.A.X

SHARP RD TREATMENT EFFECT ESTIMATES OF PROGRESA/OPORTUNIDADES ON CONSUMPTION[a]

| | Pre-Intervention | | | 1-Year Treatment | | | 2-Year Treatment | | |
|---|---|---|---|---|---|---|---|---|---|
| | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV |
| | | | | Panel A: Urban Localities | | | | | |
| Food | | | | | | | | | |
| | 3.4 | 6.4 | 6.9 | −10.9 | 2.4 | 6.1 | 48.4 | 48.6 | 49.0 |
| | (−29.7, 36.6) | (−19.0, 31.7) | (−17.1, 30.9) | (−44.9, 23.2) | (−22.9, 27.6) | (−17.7, 29.8) | (−8.1, 105.0)* | (1.9, 95.3)** | (0.9, 97.0)** |
| | [−49.3, 44.3] | [−35.2, 37.5] | [−32.1, 37.0] | [−36.0, 55.5] | [−49.9, 20.6] | [−40.5, 24.9] | [−16.3, 131.4] | [−18.4, 117.3] | [−20.1, 118.1] |
| | {−37.8, 40.3} | {−31.2, 37.0} | | {−55.1, 22.7} | {−45.7, 76.8} | | {−20.0, 116.8} | {−22.7, 125.5} | |
| | $\hat{h}_{\texttt{CCT}} = 0.57$ | $\hat{h}_{\texttt{IK}} = 1.09$ | $\hat{h}_{\text{CV}} = 1.25$ | $\hat{h}_{\texttt{CCT}} = 0.43$ | $\hat{h}_{\texttt{IK}} = 0.89$ | $\hat{h}_{\text{CV}} = 1.13$ | $\hat{h}_{\texttt{CCT}} = 0.47$ | $\hat{h}_{\texttt{IK}} = 0.67$ | $\hat{h}_{\text{CV}} = 0.64$ |
| | $\hat{b}_{\texttt{CCT}} = 0.90$ | $\hat{b}_{\texttt{IK}} = 1.26$ | | $\hat{b}_{\texttt{CCT}} = 0.77$ | $\hat{b}_{\texttt{IK}} = 0.56$ | | $\hat{b}_{\texttt{CCT}} = 0.67$ | $\hat{b}_{\texttt{IK}} = 0.58$ | |
| Non-Food | | | | | | | | | |
| | −10.1 | −10.2 | −10.5 | −8.9 | −1.3 | −0.8 | 41.2 | 38.1 | 36.2 |
| | (−34.6, 14.4) | (−32.3, 11.9) | (−31.9, 10.8) | (−38.4, 20.6) | (−19.4, 16.8) | (−22.1, 20.5) | (2.0, 80.3)** | (6.0, 70.3)** | (5.2, 67.2)** |
| | [−55.5, 10.5] | [−42.6, 16.1] | [−40.3, 16.7] | [−40.0, 34.8] | [−25.0, 23.1] | [−36.4, 21.2] | [−16.8, 82.8] | [−6.2, 86.3]* | [−2.2, 87.9]* |
| | {−38.3, 18.6} | {−60.0, 10.5} | | {−46.7, 21.4} | {−153.3, 58.2} | | {−5.4, 88.3}* | {3.4, 89.1}** | |
| | $\hat{h}_{\texttt{CCT}} = 0.56$ | $\hat{h}_{\texttt{IK}} = 0.76$ | $\hat{h}_{\text{CV}} = 0.84$ | $\hat{h}_{\texttt{CCT}} = 0.37$ | $\hat{h}_{\texttt{IK}} = 1.66$ | $\hat{h}_{\text{CV}} = 0.91$ | $\hat{h}_{\texttt{CCT}} = 0.44$ | $\hat{h}_{\texttt{IK}} = 0.64$ | $\hat{h}_{\text{CV}} = 0.68$ |
| | $\hat{b}_{\texttt{CCT}} = 0.91$ | $\hat{b}_{\texttt{IK}} = 0.62$ | | $\hat{b}_{\texttt{CCT}} = 0.63$ | $\hat{b}_{\texttt{IK}} = 0.62$ | | $\hat{b}_{\texttt{CCT}} = 0.63$ | $\hat{b}_{\texttt{IK}} = 0.75$ | |

(*Continues*)

TABLE S.A.X—*Continued*

| | Pre-Intervention | | | 1-Year Treatment | | | 2-Year Treatment | | |
|---|---|---|---|---|---|---|---|---|---|
| | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV |
| **Panel B: Rural Localities** | | | | | | | | | |
| Food | | | | | | | | | |
| | 16.5 | 15.6 | 6.6 | 33.7 | 41.7 | 38.2 | 8.3 | 3.1 | 3.5 |
| | (−24.6, 57.6) | (−16.0, 47.3) | (−21.1, 34.2) | (3.8, 63.6)** | (15.7, 67.8)*** | (14.5, 62.0)*** | (−15.1, 31.7) | (−24.8, 31.0) | (−24.5, 31.6) |
| | [−47.1, 61.1] | [−21.5, 66.3] | [−15.2, 64.2] | [−4.1, 70.4]* | [−1.4, 66.1]* | [19.2, 76.6]*** | [−7.7, 59.4] | [−15.6, 36.5] | [−14.1, 36.7] |
| | {−35.0, 61.2} | {−57.2, 59.0} | | {−3.7, 65.4}* | {0.3, 67.1}** | | {−13.8, 35.7} | {−33.5, 79.3} | |
| | $\hat{h}_{\text{CCT}} = 80.73$ | $\hat{h}_{\text{IK}} = 155.83$ | $\hat{h}_{\text{CV}} = 200.00$ | $\hat{h}_{\text{CCT}} = 63.21$ | $\hat{h}_{\text{IK}} = 109.40$ | $\hat{h}_{\text{CV}} = 200.00$ | $\hat{h}_{\text{CCT}} = 80.26$ | $\hat{h}_{\text{IK}} = 173.63$ | $\hat{h}_{\text{CV}} = 165.00$ |
| | $\hat{b}_{\text{CCT}} = 132.40$ | $\hat{b}_{\text{IK}} = 113.73$ | | $\hat{b}_{\text{CCT}} = 110.66$ | $\hat{b}_{\text{IK}} = 112.67$ | | $\hat{b}_{\text{CCT}} = 148.35$ | $\hat{b}_{\text{IK}} = 106.72$ | |
| Non-Food | | | | | | | | | |
| | 19.1 | 24.0 | 17.5 | −11.0 | −9.5 | −8.4 | 14.6 | 10.5 | 10.3 |
| | (−8.0, 46.1) | (−8.3, 56.2) | (−8.5, 43.5) | (−35.5, 13.5) | (−31.6, 12.7) | (−24.8, 8.0) | (−2.7, 31.9)* | (−4.4, 25.4) | (−2.2, 22.9) |
| | [−11.5, 68.4] | [−16.4, 75.0] | [−10.0, 66.7] | [−52.9, 20.8] | [−48.5, 17.4] | [−34.0, 16.2] | [−7.2, 44.9] | [−2.3, 39.9]* | [−5.9, 28.8] |
| | {−8.2, 54.0} | {−8.3, 62.9} | | {−41.6, 16.0} | {−64.6, 25.9} | | {−3.5, 36.3} | {−16.4, 50.2} | |
| | $\hat{h}_{\text{CCT}} = 105.39$ | $\hat{h}_{\text{IK}} = 75.63$ | $\hat{h}_{\text{CV}} = 115.00$ | $\hat{h}_{\text{CCT}} = 108.89$ | $\hat{h}_{\text{IK}} = 138.88$ | $\hat{h}_{\text{CV}} = 245.00$ | $\hat{h}_{\text{CCT}} = 97.18$ | $\hat{h}_{\text{IK}} = 145.92$ | $\hat{h}_{\text{CV}} = 235.00$ |
| | $\hat{b}_{\text{CCT}} = 179.96$ | $\hat{b}_{\text{IK}} = 156.64$ | | $\hat{b}_{\text{CCT}} = 175.75$ | $\hat{b}_{\text{IK}} = 101.21$ | | $\hat{b}_{\text{CCT}} = 170.88$ | $\hat{b}_{\text{IK}} = 100.73$ | |

[a](i) BW-CCT, BW-IK, and BW-CV correspond to estimation methods using, respectively, CCT, IK, and cross-validation bandwidth selectors. (ii) For each bandwidth selection method and outcome, the table reports RD local-linear point estimator, conventional 95% confidence intervals in parentheses, robust 95% confidence intervals in square-brackets using estimated $h_n$ and $\rho_n = 1$, robust 95% confidence intervals in curly-brackets using estimated $h_n$ and $b_n$, and estimated bandwidths values. All confidence intervals are also robust to heteroskedasticity. (iii) For each confidence interval, accompanying stars denote associated null hypothesis of no-treatment effect rejected at: *statistically significant at 10% level, **statistically significant at 5% level, and ***statistically significant at 1% level.

areas (columns 4–6, Panel B). This result is present in all cases when using both the conventional as well as the robust confidence intervals. On the other hand, in the same period, we find no statistically significant effects on non-food consumption in rural areas (columns 4–6, Panel B) nor on any of the outcomes in urban areas (columns 4–6, Panel A). These results are consistent across inference procedures.

The results from the 2-year after treatment data are the most interesting. In this case, for food consumption in urban areas (columns 7–9, Panel A), we find statistically significant results when using the conventional confidence intervals, but these results are not statistically significant when using the robust confidence intervals proposed in this paper. This empirical example offers an instance where the conventional inference approach suggests the presence of a strong positive treatment effect, but our methods cast doubt on such a conclusion. On the other hand, when examining non-food consumption in urban areas (still columns 7–9, Panel A), the results appear to be more robust, as they are statistically significant at standard levels when using both the conventional and the robust confidence intervals. Finally, in the case of the rural RD design (columns 7–9, Panel B), we find no statistically significant effects on food consumption using either method, but we find a statistically significant (10-percent level) treatment effect on non-food consumption when using conventional confidence intervals. The latter result, however, is not particularly robust based on our proposed confidence intervals.

To summarize, the empirical findings suggest that the program Progresa/Oportunidades had (i) a positive, significant effect on non-food consumption in urban areas two years after its introduction, and (ii) a positive, significant effect on food consumption in rural areas one year after its introduction. Both results appear to be robust according to our proposed methods. In addition, the empirical findings using conventional methods suggest that the program had positive, significant effects on food consumption in urban areas and on non-food consumption in rural areas two years after its introduction, but these findings are not robust according to our proposed inference procedures.

### S.4.4. *Falsification Tests and Additional Empirical Results*

Tables S.A.XI and S.A.XII report difference-in-means tests for pre-intervention covariates for urban localities and rural localities in Region 3, respectively. In each table, we present results for the full sample and for a window near the cutoff. These tables show that most covariates are balanced near the cutoff; that is, households near the cutoff in control areas have, on average, the same observable characteristics as households near the cutoff in treatment areas.

TABLE S.A.XI

BALANCE TESTS FOR PROGRESA/OPORTUNIDADES, URBAN LOCALITIES[a]

| | (a) All Observations | | | | | (b) Observations With Index in $(-0.40, 0.40)$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Treatment | | Control | | Difference | Treatment | | Control | | Difference |
| Variables | *N* | Mean | *N* | Mean | Mean | *N* | Mean | *N* | Mean | Mean |
| Household size | 2118 | 5.83 | 691 | 5.28 | −0.55 | 460 | 5.42 | 351 | 5.40 | −0.02 |
| | | (0.10) | | (0.17) | (0.19)*** | | (0.13) | | (0.15) | (0.20) |
| Age of head | 2118 | 38.53 | 691 | 40.02 | 1.50 | 460 | 38.29 | 351 | 38.93 | 0.64 |
| | | (0.51) | | (0.86) | (1.00) | | (0.55) | | (0.64) | (0.84) |
| Male head of household | 2118 | 1.10 | 691 | 1.07 | −0.03 | 460 | 1.10 | 351 | 1.07 | −0.04 |
| | | (0.01) | | (0.02) | (0.02) | | (0.02) | | (0.03) | (0.04) |
| Head's years of education | 1729 | 5.83 | 628 | 6.58 | 0.75 | 402 | 6.22 | 323 | 6.54 | 0.32 |
| | | (0.17) | | (0.28) | (0.33)** | | (0.21) | | (0.24) | (0.32) |
| Head is employed | 2118 | 0.88 | 691 | 0.90 | 0.02 | 460 | 0.90 | 351 | 0.92 | 0.02 |
| | | (0.01) | | (0.02) | (0.02) | | (0.01) | | (0.02) | (0.02) |
| Spouse's age | 2008 | 34.45 | 677 | 36.87 | 2.42 | 448 | 34.12 | 344 | 35.45 | 1.32 |
| | | (0.54) | | (0.89) | (1.04)** | | (0.69) | | (0.81) | (1.06) |
| Spouse's years of education | 1744 | 2.37 | 612 | 2.45 | 0.09 | 410 | 2.48 | 308 | 2.44 | −0.04 |
| | | (0.03) | | (0.06) | (0.07) | | (0.05) | | (0.06) | (0.08) |
| Number of children 0–5 years old | 2118 | 1.05 | 691 | 0.56 | −0.49 | 460 | 0.80 | 351 | 0.67 | −0.13 |
| | | (0.07) | | (0.12) | (0.14)*** | | (0.05) | | (0.06) | (0.08)* |
| Number of boys 0–5 years old | 2118 | 0.54 | 691 | 0.32 | −0.23 | 460 | 0.40 | 351 | 0.38 | −0.01 |
| | | (0.03) | | (0.05) | (0.05)*** | | (0.03) | | (0.04) | (0.05) |

(*Continues*)

TABLE S.A.XI—*Continued*

| | (a) All Observations | | | | | (b) Observations With Index in $(-0.40, 0.40)$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Treatment | | Control | | Difference | Treatment | | Control | | Difference |
| Variables | *N* | Mean | *N* | Mean | Mean | *N* | Mean | *N* | Mean | Mean |
| Owns a house | 2109 | 0.76 | 688 | 0.80 | 0.04 | 458 | 0.78 | 348 | 0.77 | −0.00 |
| | | (0.04) | | (0.06) | (0.07) | | (0.04) | | (0.05) | (0.07) |
| Cement floors | 2110 | 0.45 | 689 | 0.84 | 0.40 | 458 | 0.69 | 349 | 0.78 | 0.09 |
| | | (0.07) | | (0.12) | (0.14)*** | | (0.06) | | (0.07) | (0.09) |
| Number of rooms | 2110 | 1.27 | 688 | 1.64 | 0.37 | 458 | 1.37 | 349 | 1.54 | 0.17 |
| | | (0.06) | | (0.10) | (0.11)*** | | (0.08) | | (0.09) | (0.12) |
| Water connection | 2111 | 0.60 | 689 | 0.76 | 0.16 | 458 | 0.69 | 349 | 0.76 | 0.06 |
| | | (0.07) | | (0.11) | (0.13) | | (0.07) | | (0.08) | (0.10) |
| Water connection inside the house | 2100 | 0.23 | 681 | 0.40 | 0.17 | 454 | 0.31 | 345 | 0.36 | 0.05 |
| | | (0.05) | | (0.09) | (0.10) | | (0.06) | | (0.07) | (0.09) |
| Has a bathroom | 2110 | 0.56 | 688 | 0.66 | 0.10 | 458 | 0.62 | 348 | 0.64 | 0.01 |
| | | (0.07) | | (0.11) | (0.13) | | (0.07) | | (0.09) | (0.12) |
| Has electricity | 2118 | 0.94 | 691 | 0.97 | 0.03 | 460 | 0.96 | 351 | 0.98 | 0.02 |
| | | (0.02) | | (0.04) | (0.04) | | (0.02) | | (0.03) | (0.04) |

[a](i) Table reports sample size, sample mean, and standard errors for treatment and control units. It also reports difference in means with heteroskedasticity-robust standard errors. (ii) Significance levels for difference-in-means tests: *statistically significant at 10% level, **statistically significant at 5% level, and ***statistically significant at 1% level.

TABLE S.A.XII
BALANCE TESTS FOR PROGRESA/OPORTUNIDADES, RURAL LOCALITIES (REGION 3)[a]

| | (a) All Observations | | | | | (b) Observations With Index in $(-0.40, 0.40)$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Treatment | | Control | | Difference | Treatment | | Control | | Difference |
| Variables | *N* | Mean | *N* | Mean | Mean | *N* | Mean | *N* | Mean | Mean |
| Household size | 618 | 6.39 | 315 | 6.04 | −0.35 | 156 | 6.25 | 89 | 6.25 | −0.00 |
| | | (0.12) | | (0.17) | (0.21)* | | (0.20) | | (0.27) | (0.34) |
| Age of head | 618 | 40.06 | 315 | 48.77 | 8.71 | 156 | 41.00 | 89 | 47.34 | 6.34 |
| | | (1.08) | | (1.57) | (1.91)*** | | (1.54) | | (2.09) | (2.59)** |
| Male head of household | 618 | 0.91 | 315 | 0.93 | 0.01 | 156 | 0.94 | 89 | 0.90 | −0.04 |
| | | (0.01) | | (0.02) | (0.02) | | (0.02) | | (0.03) | (0.03) |
| Head's years of education | 615 | 2.29 | 315 | 3.14 | 0.85 | 155 | 2.91 | 89 | 3.13 | 0.23 |
| | | (0.22) | | (0.32) | (0.39)** | | (0.24) | | (0.32) | (0.40) |
| Head is employed | 617 | 0.90 | 311 | 0.93 | 0.03 | 156 | 0.92 | 88 | 0.94 | 0.03 |
| | | (0.02) | | (0.02) | (0.03) | | (0.02) | | (0.03) | (0.04) |
| Spouse's age | 547 | 34.45 | 282 | 43.56 | 9.11 | 143 | 35.43 | 72 | 40.03 | 4.59 |
| | | (1.04) | | (1.51) | (1.83)*** | | (1.25) | | (1.80) | (2.19)** |
| Spouse's years of education | 547 | 2.03 | 283 | 3.05 | 1.02 | 143 | 2.90 | 73 | 2.70 | −0.20 |
| | | (0.24) | | (0.34) | (0.42)** | | (0.31) | | (0.45) | (0.55) |
| Number of children 0–5 years old | 618 | 1.38 | 315 | 0.77 | −0.62 | 156 | 1.21 | 89 | 1.04 | −0.16 |
| | | (0.08) | | (0.12) | (0.15)*** | | (0.11) | | (0.15) | (0.18) |
| Number of boys 0–5 years old | 618 | 0.73 | 315 | 0.41 | −0.32 | 156 | 0.62 | 89 | 0.60 | −0.02 |
| | | (0.05) | | (0.08) | (0.09)*** | | (0.06) | | (0.08) | (0.11) |

(*Continues*)

TABLE S.A.XII—*Continued*

| | (a) All Observations | | | | | (b) Observations With Index in $(-0.40, 0.40)$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Treatment | | Control | | Difference | Treatment | | Control | | Difference |
| Variables | $N$ | Mean | $N$ | Mean | Mean | $N$ | Mean | $N$ | Mean | Mean |
| Owns a house | 618 | 0.95<br>(0.01) | 315 | 0.98<br>(0.01) | 0.03<br>(0.02) | 156 | 0.94<br>(0.02) | 89 | 0.98<br>(0.03) | 0.04<br>(0.04) |
| Cement floors | 617 | 0.24<br>(0.07) | 314 | 0.69<br>(0.11) | 0.45<br>(0.13)*** | 155 | 0.44<br>(0.08) | 89 | 0.54<br>(0.12) | 0.10<br>(0.15) |
| Number of rooms | 615 | 1.45<br>(0.12) | 315 | 2.38<br>(0.17) | 0.93<br>(0.21)*** | 155 | 1.76<br>(0.10) | 89 | 1.88<br>(0.14) | 0.12<br>(0.17) |
| Water connection | 618 | 0.42<br>(0.08) | 315 | 0.71<br>(0.12) | 0.30<br>(0.15)** | 156 | 0.74<br>(0.10) | 89 | 0.75<br>(0.13) | 0.02<br>(0.17) |
| Water connection inside the house | 616 | 0.07<br>(0.03) | 312 | 0.18<br>(0.05) | 0.11<br>(0.06)* | 155 | 0.13<br>(0.04) | 88 | 0.09<br>(0.05) | −0.04<br>(0.07) |
| Has a bathroom | 618 | 0.59<br>(0.07) | 314 | 0.62<br>(0.10) | 0.03<br>(0.12) | 156 | 0.58<br>(0.09) | 88 | 0.56<br>(0.13) | −0.03<br>(0.16) |
| Has electricity | 618 | 0.54<br>(0.09) | 315 | 0.93<br>(0.14) | 0.39<br>(0.17)** | 156 | 0.84<br>(0.09) | 89 | 0.90<br>(0.12) | 0.06<br>(0.15) |

[a](i) Table reports sample size, sample mean, and standard errors for treatment and control units. It also reports difference in means with heteroskedasticity-robust standard errors. (ii) Significance levels for difference-in-means tests: *statistically significant at 10% level, **statistically significant at 5% level, and ***statistically significant at 1% level.

For completeness, Tables S.A.XIII through S.A.XX report RD estimates for food, non-food, and total consumption for urban localities and rural localities in all seven regions. Each table considers one geographic region. The tables have the same structure as Tables S.A.XI and S.A.XII, with the additional results corresponding to total consumption.

### S.4.5. *Implementation Details*

The data cleaning and data preparation were done as follows:

• *Urban Communities*. All data are publicly available. We employ the following databases: baseline or pre-intervention (2002), and two follow-ups (2003 and 2004). We constructed our final database as follows:

– Matched household in the three periods.
– Drop households in the control areas.
– Drop households in the treatment areas who did not apply to the program.
– Drop households with missing information for the consumption variables.
– Drop household with invalid entries in some key pre-intervention variables (household size and age of household head).
– Using the baseline database (2002), we constructed the pre-intervention variables used in Table S.A.XI.
– Using all three databases (2002, 2003, 2004), we constructed the following outcome variables (averaged in the household over all its members, expressed as monthly expenses):

∗ *Food Consumption*: household level data on food outlays made in the seven days preceding the interview for 36 food items. It also includes the value of food consumed from own production in that same period of time, valued using household self-reported information.

∗ *Non-Food Consumption*: expenses reported on a weekly, monthly, and quarterly basis. Non-food expenses reported on a weekly basis include transportation and tobacco. Monthly outlays include school tuition, health-related expenses, home cleaning, electricity, and home fuel expenditures. Expenditures reported on a quarterly basis include home and school supplies, clothes, shoes, toys, and payments for special events.

∗ *Total Consumption*: computed as the sum of non-food expenditures and the value of food consumption.

• *Rural Communities*. We merged the publicly available data with the households' poverty-index at baseline and the region identifier. We employ the following databases: baseline or pre-intervention (1997 and March 1998), and two follow-ups (November 1998 and November 1999). We constructed our final database as follows:

– Matched household in the three periods.
– Drop households in the control areas.

TABLE S.A.XIII

SHARP RD TREATMENT EFFECT ESTIMATES OF PROGRESA/OPORTUNIDADES ON CONSUMPTION, URBAN LOCALITIES[a]

| | Pre-Intervention | | | 1-Year Treatment | | | 2-Year Treatment | | |
|---|---|---|---|---|---|---|---|---|---|
| | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV |
| Food | | | | | | | | | |
| | 3.4 | 6.4 | 6.9 | −10.9 | 2.4 | 6.1 | 48.4 | 48.6 | 49.0 |
| | (−29.7, 36.6) | (−19.0, 31.7) | (−17.1, 30.9) | (−44.9, 23.2) | (−22.9, 27.6) | (−17.7, 29.8) | (−8.1, 105.0)$^{*}$ | (1.9, 95.3)$^{**}$ | (0.9, 97.0)$^{**}$ |
| | [−49.3, 44.3] | [−35.2, 37.5] | [−32.1, 37.0] | [−36.0, 55.5] | [−49.9, 20.6] | [−40.5, 24.9] | [−16.3, 131.4] | [−18.4, 117.3] | [−20.1, 118.1] |
| | {−37.8, 40.3} | {−31.2, 37.0} | | {−55.1, 22.7} | {−45.7, 76.8} | | {−20.0, 116.8} | {−22.7, 125.5} | |
| | $\hat{h}_{\mathtt{CCT}} = 0.57$ | $\hat{h}_{\mathtt{IK}} = 1.09$ | $\hat{h}_{\mathtt{CV}} = 1.25$ | $\hat{h}_{\mathtt{CCT}} = 0.43$ | $\hat{h}_{\mathtt{IK}} = 0.89$ | $\hat{h}_{\mathtt{CV}} = 1.13$ | $\hat{h}_{\mathtt{CCT}} = 0.47$ | $\hat{h}_{\mathtt{IK}} = 0.67$ | $\hat{h}_{\mathtt{CV}} = 0.64$ |
| | $\hat{b}_{\mathtt{CCT}} = 0.90$ | $\hat{b}_{\mathtt{IK}} = 1.26$ | | $\hat{b}_{\mathtt{CCT}} = 0.77$ | $\hat{b}_{\mathtt{IK}} = 0.56$ | | $\hat{b}_{\mathtt{CCT}} = 0.67$ | $\hat{b}_{\mathtt{IK}} = 0.58$ | |
| Non-Food | | | | | | | | | |
| | −10.1 | −10.2 | −10.5 | −8.9 | −1.3 | −0.8 | 41.2 | 38.1 | 36.2 |
| | (−34.6, 14.4) | (−32.3, 11.9) | (−31.9, 10.8) | (−38.4, 20.6) | (−19.4, 16.8) | (−22.1, 20.5) | (2.0, 80.3)$^{**}$ | (6.0, 70.3)$^{**}$ | (5.2, 67.2)$^{**}$ |
| | [−55.5, 10.5] | [−42.6, 16.1] | [−40.3, 16.7] | [−40.0, 34.8] | [−25.0, 23.1] | [−36.4, 21.2] | [−16.8, 82.8] | [−6.2, 86.3]$^{*}$ | [−2.2, 87.9]$^{*}$ |
| | {−38.3, 18.6} | {−60.0, 10.5} | | {−46.7, 21.4} | {−153.3, 58.2} | | {−5.4, 88.3}$^{*}$ | {3.4, 89.1}$^{**}$ | |
| | $\hat{h}_{\mathtt{CCT}} = 0.56$ | $\hat{h}_{\mathtt{IK}} = 0.76$ | $\hat{h}_{\mathtt{CV}} = 0.84$ | $\hat{h}_{\mathtt{CCT}} = 0.37$ | $\hat{h}_{\mathtt{IK}} = 1.66$ | $\hat{h}_{\mathtt{CV}} = 0.91$ | $\hat{h}_{\mathtt{CCT}} = 0.44$ | $\hat{h}_{\mathtt{IK}} = 0.64$ | $\hat{h}_{\mathtt{CV}} = 0.68$ |
| | $\hat{b}_{\mathtt{CCT}} = 0.91$ | $\hat{b}_{\mathtt{IK}} = 0.62$ | | $\hat{b}_{\mathtt{CCT}} = 0.63$ | $\hat{b}_{\mathtt{IK}} = 0.62$ | | $\hat{b}_{\mathtt{CCT}} = 0.63$ | $\hat{b}_{\mathtt{IK}} = 0.75$ | |
| Total | | | | | | | | | |
| | −6.6 | −5.3 | −0.3 | −17.7 | 4.0 | 4.2 | 90.3 | 87.8 | 87.0 |
| | (−57.3, 44.0) | (−46.7, 36.2) | (−37.7, 37.0) | (−72.6, 37.1) | (−34.4, 42.4) | (−34.0, 42.4) | (−0.6, 181.2)$^{*}$ | (8.7, 166.9)$^{**}$ | (12.5, 161.5)$^{**}$ |
| | [−94.9, 45.4] | [−68.3, 47.6] | [−63.1, 41.5] | [−59.3, 84.5] | [−71.6, 34.5] | [−70.4, 35.2] | [−23.6, 206.6] | [−21.8, 200.6] | [−18.6, 196.8] |
| | {−68.2, 50.9} | {−90.0, 58.3} | | {−88.0, 39.4} | {−117.1, 90.8} | | {−19.6, 199.4} | {−21.8, 199.7} | |
| | $\hat{h}_{\mathtt{CCT}} = 0.56$ | $\hat{h}_{\mathtt{IK}} = 0.96$ | $\hat{h}_{\mathtt{CV}} = 1.29$ | $\hat{h}_{\mathtt{CCT}} = 0.38$ | $\hat{h}_{\mathtt{IK}} = 0.99$ | $\hat{h}_{\mathtt{CV}} = 1.01$ | $\hat{h}_{\mathtt{CCT}} = 0.45$ | $\hat{h}_{\mathtt{IK}} = 0.57$ | $\hat{h}_{\mathtt{CV}} = 0.64$ |
| | $\hat{b}_{\mathtt{CCT}} = 0.89$ | $\hat{b}_{\mathtt{IK}} = 0.72$ | | $\hat{b}_{\mathtt{CCT}} = 0.64$ | $\hat{b}_{\mathtt{IK}} = 0.58$ | | $\hat{b}_{\mathtt{CCT}} = 0.62$ | $\hat{b}_{\mathtt{IK}} = 0.58$ | |

[a](i) BW-CCT, BW-IK, and BW-CV correspond to estimation methods using, respectively, CCT, IK, and cross-validation bandwidth selectors. (ii) For each bandwidth selection method and outcome, the table reports RD local-linear point estimator, conventional 95% confidence intervals in parentheses, robust 95% confidence intervals in square-brackets using estimated $h_n$ and $\rho_n = 1$, robust 95% confidence intervals in curly-brackets using estimated $h_n$ and $b_n$, and estimated bandwidths values. All confidence intervals are also robust to heteroskedasticity. (iii) For each confidence interval, accompanying stars denote associated null hypothesis of no-treatment effect rejected at: $^{*}$statistically significant at 10% level, $^{**}$statistically significant at 5% level, and $^{***}$statistically significant at 1% level.

TABLE S.A.XIV

SHARP RD TREATMENT EFFECT ESTIMATES OF PROGRESA/OPORTUNIDADES ON CONSUMPTION, RURAL LOCALITIES (REGION 3)[a]

| | Pre-Intervention | | | 1-Year Treatment | | | 2-Year Treatment | | |
|---|---|---|---|---|---|---|---|---|---|
| | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV |
| Food | | | | | | | | | |
| | 16.5 | 15.6 | 6.6 | 33.7 | 41.7 | 38.2 | 8.3 | 3.1 | 3.5 |
| | (−24.6, 57.6) | (−16.0, 47.3) | (−21.1, 34.2) | (3.8, 63.6)** | (15.7, 67.8)*** | (14.5, 62.0)*** | (−15.1, 31.7) | (−24.8, 31.0) | (−24.5, 31.6) |
| | [−47.1, 61.1] | [−21.5, 66.3] | [−15.2, 64.2] | [−4.1, 70.4]* | [−1.4, 66.1]* | [19.2, 76.6]*** | [−7.7, 59.4] | [−15.6, 36.5] | [−14.1, 36.7] |
| | {−35.0, 61.2} | {−57.2, 59.0} | | {−3.7, 65.4}* | {0.3, 67.1}** | | {−13.8, 35.7} | {−33.5, 79.3} | |
| | $\hat{h}_{CCT} = 80.73$ | $\hat{h}_{IK} = 155.83$ | $\hat{h}_{CV} = 200.00$ | $\hat{h}_{CCT} = 63.21$ | $\hat{h}_{IK} = 109.40$ | $\hat{h}_{CV} = 200.00$ | $\hat{h}_{CCT} = 80.26$ | $\hat{h}_{IK} = 173.63$ | $\hat{h}_{CV} = 165.00$ |
| | $\hat{b}_{CCT} = 132.40$ | $\hat{b}_{IK} = 113.73$ | | $\hat{b}_{CCT} = 110.66$ | $\hat{b}_{IK} = 112.67$ | | $\hat{b}_{CCT} = 148.35$ | $\hat{b}_{IK} = 106.72$ | |
| Non-Food | | | | | | | | | |
| | 19.1 | 24.0 | 17.5 | −11.0 | −9.5 | −8.4 | 14.6 | 10.5 | 10.3 |
| | (−8.0, 46.1) | (−8.3, 56.2) | (−8.5, 43.5) | (−35.5, 13.5) | (−31.6, 12.7) | (−24.8, 8.0) | (−2.7, 31.9)* | (−4.4, 25.4) | (−2.2, 22.9) |
| | [−11.5, 68.4] | [−16.4, 75.0] | [−10.0, 66.7] | [−52.9, 20.8] | [−48.5, 17.4] | [−34.0, 16.2] | [−7.2, 44.9] | [−2.3, 39.9]* | [−5.9, 28.8] |
| | {−8.2, 54.0} | {−8.3, 62.9} | | {−41.6, 16.0} | {−64.6, 25.9} | | {−3.5, 36.3} | {−16.4, 50.2} | |
| | $\hat{h}_{CCT} = 105.39$ | $\hat{h}_{IK} = 75.63$ | $\hat{h}_{CV} = 115.00$ | $\hat{h}_{CCT} = 108.89$ | $\hat{h}_{IK} = 138.88$ | $\hat{h}_{CV} = 245.00$ | $\hat{h}_{CCT} = 97.18$ | $\hat{h}_{IK} = 145.92$ | $\hat{h}_{CV} = 235.00$ |
| | $\hat{b}_{CCT} = 179.96$ | $\hat{b}_{IK} = 156.64$ | | $\hat{b}_{CCT} = 175.75$ | $\hat{b}_{IK} = 101.21$ | | $\hat{b}_{CCT} = 170.88$ | $\hat{b}_{IK} = 100.73$ | |
| Total | | | | | | | | | |
| | 38.5 | 38.6 | 37.9 | 21.0 | 30.8 | 30.1 | 22.5 | 9.3 | 13.5 |
| | (−21.0, 98.0) | (−21.0, 98.3) | (−16.5, 92.4) | (−25.8, 67.7) | (−8.6, 70.1) | (−3.4, 63.6)* | (−12.0, 57.0) | (−22.6, 41.2) | (−22.4, 49.3) |
| | [−43.3, 119.8] | [−43.6, 119.7] | [−39.3, 114.7] | [−48.8, 81.9] | [−38.4, 71.4] | [−5.9, 80.5]* | [−1.0, 102.1]* | [−18.2, 58.2] | [−9.3, 64.2] |
| | {−28.4, 112.0} | {−37.0, 113.0} | | {−37.0, 72.3} | {−21.9, 72.2} | | {−10.1, 65.0} | {−108.2, 244.9} | |
| | $\hat{h}_{CCT} = 86.49$ | $\hat{h}_{IK} = 85.92$ | $\hat{h}_{CV} = 105.00$ | $\hat{h}_{CCT} = 69.59$ | $\hat{h}_{IK} = 109.89$ | $\hat{h}_{CV} = 200.00$ | $\hat{h}_{CCT} = 80.23$ | $\hat{h}_{IK} = 283.28$ | $\hat{h}_{CV} = 165.00$ |
| | $\hat{b}_{CCT} = 140.32$ | $\hat{b}_{IK} = 111.21$ | | $\hat{b}_{CCT} = 118.76$ | $\hat{b}_{IK} = 160.11$ | | $\hat{b}_{CCT} = 150.33$ | $\hat{b}_{IK} = 113.02$ | |

[a](i) BW-CCT, BW-IK, and BW-CV correspond to estimation methods using, respectively, CCT, IK, and cross-validation bandwidth selectors. (ii) For each bandwidth selection method and outcome, the table reports RD local-linear point estimator, conventional 95% confidence intervals in parentheses, robust 95% confidence intervals in square-brackets using estimated $h_n$ and $\rho_n = 1$, robust 95% confidence intervals in curly-brackets using estimated $h_n$ and $b_n$, and estimated bandwidths values. All confidence intervals are also robust to heteroskedasticity. (iii) For each confidence interval, accompanying stars denote associated null hypothesis of no-treatment effect rejected at: *statistically significant at 10% level, **statistically significant at 5% level, and ***statistically significant at 1% level.

TABLE S.A.XV

SHARP RD TREATMENT EFFECT ESTIMATES OF PROGRESA/OPORTUNIDADES ON CONSUMPTION, RURAL LOCALITIES (REGION 4)[a]

| | Pre-Intervention | | | 1-Year Treatment | | | 2-Year Treatment | | |
|---|---|---|---|---|---|---|---|---|---|
| | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV |
| Food | | | | | | | | | |
| | −11.8 | −11.7 | −12.2 | 21.5 | 15.1 | 16.3 | 0.8 | 0.5 | −1.9 |
| | (−34.2, 10.5) | (−34.3, 10.8) | (−26.6, 2.2)* | (−16.7, 59.8) | (−12.3, 42.6) | (−7.7, 40.3) | (−20.5, 22.2) | (−21.0, 22.1) | (−15.6, 11.8) |
| | [−54.4, 16.8] | [−55.4, 16.5] | [−35.6, 6.1] | [−28.2, 79.8] | [−22.5, 57.3] | [−20.7, 49.1] | [−36.2, 24.0] | [−36.4, 24.3] | [−23.0, 16.4] |
| | {−38.7, 15.6} | {−58.2, 15.6} | | {−19.5, 71.3} | {−99.6, 176.7} | | {−23.3, 27.8} | {−45.5, 25.5} | |
| | $\hat{h}_{\mathtt{CCT}} = 87.25$ | $\hat{h}_{\mathtt{IK}} = 85.72$ | $\hat{h}_{\mathtt{CV}} = 254.30$ | $\hat{h}_{\mathtt{CCT}} = 117.11$ | $\hat{h}_{\mathtt{IK}} = 257.18$ | $\hat{h}_{\mathtt{CV}} = 388.00$ | $\hat{h}_{\mathtt{CCT}} = 114.89$ | $\hat{h}_{\mathtt{IK}} = 111.99$ | $\hat{h}_{\mathtt{CV}} = 388.00$ |
| | $\hat{b}_{\mathtt{CCT}} = 143.31$ | $\hat{b}_{\mathtt{IK}} = 83.52$ | | $\hat{b}_{\mathtt{CCT}} = 183.80$ | $\hat{b}_{\mathtt{IK}} = 108.90$ | | $\hat{b}_{\mathtt{CCT}} = 182.93$ | $\hat{b}_{\mathtt{IK}} = 91.66$ | |
| Non-Food | | | | | | | | | |
| | 3.6 | 3.2 | −4.8 | 5.8 | 5.5 | 1.1 | 7.4 | 5.1 | 0.6 |
| | (−12.8, 20.0) | (−14.6, 20.9) | (−16.0, 6.4) | (−10.5, 22.0) | (−8.6, 19.6) | (−9.8, 11.9) | (−12.1, 26.8) | (−12.1, 22.3) | (−12.0, 13.2) |
| | [−25.5, 25.9] | [−29.8, 27.5] | [−17.0, 10.7] | [−14.0, 25.5] | [−12.7, 24.9] | [−14.0, 13.2] | [−28.4, 36.7] | [−18.6, 36.7] | [−14.7, 21.9] |
| | {−14.3, 25.1} | {−22.0, 25.6} | | {−13.9, 24.4} | {−13.9, 25.7} | | {−12.8, 34.2} | {−51.2, 45.5} | |
| | $\hat{h}_{\mathtt{CCT}} = 85.39$ | $\hat{h}_{\mathtt{IK}} = 73.93$ | $\hat{h}_{\mathtt{CV}} = 388.00$ | $\hat{h}_{\mathtt{CCT}} = 64.81$ | $\hat{h}_{\mathtt{IK}} = 98.19$ | $\hat{h}_{\mathtt{CV}} = 388.00$ | $\hat{h}_{\mathtt{CCT}} = 113.59$ | $\hat{h}_{\mathtt{IK}} = 142.66$ | $\hat{h}_{\mathtt{CV}} = 311.60$ |
| | $\hat{b}_{\mathtt{CCT}} = 152.00$ | $\hat{b}_{\mathtt{IK}} = 95.75$ | | $\hat{b}_{\mathtt{CCT}} = 111.13$ | $\hat{b}_{\mathtt{IK}} = 86.53$ | | $\hat{b}_{\mathtt{CCT}} = 190.26$ | $\hat{b}_{\mathtt{IK}} = 96.80$ | |
| Total | | | | | | | | | |
| | −7.9 | −8.3 | −12.6 | 27.2 | 21.4 | 17.4 | 8.1 | 7.9 | −2.8 |
| | (−41.1, 25.2) | (−42.3, 25.7) | (−33.1, 7.9) | (−19.6, 74.0) | (−17.6, 60.5) | (−10.3, 45.1) | (−27.2, 43.4) | (−27.9, 43.7) | (−24.2, 18.6) |
| | [−74.4, 30.9] | [−77.9, 30.8] | [−46.9, 6.0] | [−28.3, 102.2] | [−16.9, 90.3] | [−25.5, 53.1] | [−56.0, 52.5] | [−57.5, 52.6] | [−29.8, 32.7] |
| | {−45.5, 34.5} | {−66.4, 32.9} | | {−21.2, 88.1} | {−47.9, 128.7} | | {−29.1, 55.5} | {−77.2, 60.9} | |
| | $\hat{h}_{\mathtt{CCT}} = 81.87$ | $\hat{h}_{\mathtt{IK}} = 78.32$ | $\hat{h}_{\mathtt{CV}} = 388.00$ | $\hat{h}_{\mathtt{CCT}} = 91.34$ | $\hat{h}_{\mathtt{IK}} = 150.38$ | $\hat{h}_{\mathtt{CV}} = 388.00$ | $\hat{h}_{\mathtt{CCT}} = 114.39$ | $\hat{h}_{\mathtt{IK}} = 111.10$ | $\hat{h}_{\mathtt{CV}} = 388.00$ |
| | $\hat{b}_{\mathtt{CCT}} = 140.80$ | $\hat{b}_{\mathtt{IK}} = 86.76$ | | $\hat{b}_{\mathtt{CCT}} = 145.27$ | $\hat{b}_{\mathtt{IK}} = 97.08$ | | $\hat{b}_{\mathtt{CCT}} = 188.06$ | $\hat{b}_{\mathtt{IK}} = 89.30$ | |

[a](i) BW-CCT, BW-IK, and BW-CV correspond to estimation methods using, respectively, CCT, IK, and cross-validation bandwidth selectors. (ii) For each bandwidth selection method and outcome, the table reports RD local-linear point estimator, conventional 95% confidence intervals in parentheses, robust 95% confidence intervals in square-brackets using estimated $h_n$ and $\rho_n = 1$, robust 95% confidence intervals in curly-brackets using estimated $h_n$ and $b_n$, and estimated bandwidths values. All confidence intervals are also robust to heteroskedasticity. (iii) For each confidence interval, accompanying stars denote associated null hypothesis of no-treatment effect rejected at: *statistically significant at 10% level, **statistically significant at 5% level, and ***statistically significant at 1% level.

TABLE S.A.XVI

SHARP RD TREATMENT EFFECT ESTIMATES OF PROGRESA/OPORTUNIDADES ON CONSUMPTION, RURAL LOCALITIES (REGION 5)[a]

| | Pre-Intervention | | | 1-Year Treatment | | | 2-Year Treatment | | |
|---|---|---|---|---|---|---|---|---|---|
| | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV |
| Food | | | | | | | | | |
| | 8.0 | 4.3 | 9.2 | 30.9 | 30.8 | 25.3 | 18.0 | 17.7 | 22.7 |
| | (−8.1, 24.1) | (−16.1, 24.8) | (−1.1, 19.4)* | (12.5, 49.2)*** | (13.5, 48.0)*** | (12.5, 38.2)*** | (0.3, 35.7)** | (−0.3, 35.7)* | (14.1, 31.4)*** |
| | [−23.5, 24.1] | [−32.6, 24.9] | [−5.5, 23.0] | [3.7, 53.7]** | [5.8, 53.6]** | [14.3, 51.2]*** | [−14.4, 34.3] | [−15.0, 34.1] | [13.5, 38.4]*** |
| | {−12.0, 26.5} | {−25.1, 25.8} | | {11.7, 54.1}*** | {−9.6, 53.6} | | {−5.5, 34.8} | {−13.1, 34.6} | |
| | $\hat{h}_{\texttt{CCT}} = 121.99$ | $\hat{h}_{\texttt{IK}} = 75.82$ | $\hat{h}_{\texttt{CV}} = 397.50$ | $\hat{h}_{\texttt{CCT}} = 118.77$ | $\hat{h}_{\texttt{IK}} = 136.05$ | $\hat{h}_{\texttt{CV}} = 278.85$ | $\hat{h}_{\texttt{CCT}} = 79.25$ | $\hat{h}_{\texttt{IK}} = 76.79$ | $\hat{h}_{\texttt{CV}} = 397.50$ |
| | $\hat{b}_{\texttt{CCT}} = 202.28$ | $\hat{b}_{\texttt{IK}} = 104.46$ | | $\hat{b}_{\texttt{CCT}} = 203.25$ | $\hat{b}_{\texttt{IK}} = 98.18$ | | $\hat{b}_{\texttt{CCT}} = 143.62$ | $\hat{b}_{\texttt{IK}} = 83.53$ | |
| Non-Food | | | | | | | | | |
| | −3.1 | −3.7 | −3.7 | 5.3 | 5.1 | 3.3 | 17.2 | 17.1 | 8.7 |
| | (−15.4, 9.2) | (−16.6, 9.3) | (−11.1, 3.7) | (−4.2, 14.7) | (−4.1, 14.3) | (−2.0, 8.6) | (5.1, 29.4)*** | (4.9, 29.4)*** | (1.8, 15.5)** |
| | [−24.7, 10.6] | [−25.8, 10.7] | [−13.8, 6.8] | [−8.4, 17.8] | [−7.7, 17.8] | [−0.7, 14.2]* | [4.2, 40.7]** | [4.7, 41.6]** | [5.6, 24.0]*** |
| | {−18.1, 11.0} | {−20.2, 10.7} | | {−6.7, 15.1} | {−7.5, 17.7} | | {4.3, 33.3}** | {2.7, 38.2}** | |
| | $\hat{h}_{\texttt{CCT}} = 105.22$ | $\hat{h}_{\texttt{IK}} = 94.57$ | $\hat{h}_{\texttt{CV}} = 397.50$ | $\hat{h}_{\texttt{CCT}} = 86.94$ | $\hat{h}_{\texttt{IK}} = 91.87$ | $\hat{h}_{\texttt{CV}} = 397.50$ | $\hat{h}_{\texttt{CCT}} = 91.40$ | $\hat{h}_{\texttt{IK}} = 89.65$ | $\hat{h}_{\texttt{CV}} = 397.50$ |
| | $\hat{b}_{\texttt{CCT}} = 170.87$ | $\hat{b}_{\texttt{IK}} = 148.20$ | | $\hat{b}_{\texttt{CCT}} = 149.82$ | $\hat{b}_{\texttt{IK}} = 93.84$ | | $\hat{b}_{\texttt{CCT}} = 150.49$ | $\hat{b}_{\texttt{IK}} = 96.29$ | |
| Total | | | | | | | | | |
| | 5.1 | 0.0 | 5.5 | 36.0 | 37.1 | 32.2 | 40.1 | 36.5 | 31.4 |
| | (−19.6, 29.8) | (−30.0, 30.1) | (−9.6, 20.6) | (13.6, 58.4)*** | (16.4, 57.7)*** | (15.8, 48.6)*** | (19.9, 60.3)*** | (13.3, 59.7)*** | (19.0, 43.9)*** |
| | [−42.6, 29.8] | [−54.3, 30.4] | [−16.1, 26.7] | [2.5, 64.2]** | [4.5, 62.5]** | [17.2, 64.1]*** | [2.3, 62.4]** | [0.3, 67.9]** | [23.2, 58.1]*** |
| | {−25.1, 33.9} | {−39.8, 31.2} | | {11.8, 64.0}*** | {−12.7, 68.9} | | {17.7, 65.8}*** | {5.7, 61.3}** | |
| | $\hat{h}_{\texttt{CCT}} = 117.67$ | $\hat{h}_{\texttt{IK}} = 79.41$ | $\hat{h}_{\texttt{CV}} = 397.50$ | $\hat{h}_{\texttt{CCT}} = 115.67$ | $\hat{h}_{\texttt{IK}} = 138.63$ | $\hat{h}_{\texttt{CV}} = 239.30$ | $\hat{h}_{\texttt{CCT}} = 117.49$ | $\hat{h}_{\texttt{IK}} = 89.89$ | $\hat{h}_{\texttt{CV}} = 397.50$ |
| | $\hat{b}_{\texttt{CCT}} = 190.93$ | $\hat{b}_{\texttt{IK}} = 127.57$ | | $\hat{b}_{\texttt{CCT}} = 194.03$ | $\hat{b}_{\texttt{IK}} = 97.42$ | | $\hat{b}_{\texttt{CCT}} = 199.06$ | $\hat{b}_{\texttt{IK}} = 140.82$ | |

[a](i) BW-CCT, BW-IK, and BW-CV correspond to estimation methods using, respectively, CCT, IK, and cross-validation bandwidth selectors. (ii) For each bandwidth selection method and outcome, the table reports RD local-linear point estimator, conventional 95% confidence intervals in parentheses, robust 95% confidence intervals in square-brackets using estimated $h_n$ and $\rho_n = 1$, robust 95% confidence intervals in curly-brackets using estimated $h_n$ and $b_n$, and estimated bandwidths values. All confidence intervals are also robust to heteroskedasticity. (iii) For each confidence interval, accompanying stars denote associated null hypothesis of no-treatment effect rejected at: *statistically significant at 10% level, **statistically significant at 5% level, and ***statistically significant at 1% level.

TABLE S.A.XVII

SHARP RD TREATMENT EFFECT ESTIMATES OF PROGRESA/OPORTUNIDADES ON CONSUMPTION, RURAL LOCALITIES (REGION 6)[a]

| | Pre-Intervention | | | 1-Year Treatment | | | 2-Year Treatment | | |
|---|---|---|---|---|---|---|---|---|---|
| | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV |
| Food | | | | | | | | | |
| | 1.5 | 8.3 | 5.5 | 53.4 | 28.3 | 18.3 | −19.9 | −17.4 | −7.0 |
| | (−44.3, 47.2) | (−32.1, 48.6) | (−22.2, 33.1) | (−14.3, 121.1) | (−24.9, 81.4) | (−35.1, 71.8) | (−58.4, 18.5) | (−56.8, 22.1) | (−33.9, 20.0) |
| | [−59.8, 73.4] | [−74.3, 40.1] | [−29.8, 53.9] | [−35.9, 115.2] | [−37.2, 121.9] | [−20.9, 121.5] | [−51.7, 54.5] | [−52.4, 55.6] | [−54.0, 17.3] |
| | {−62.8, 47.7} | {−56.3, 74.2} | | {−15.9, 146.9} | {−21.3, 169.1} | | {−69.1, 26.0} | {−57.3, 61.7} | |
| | $\hat{h}_{\mathtt{CCT}} = 52.24$ | $\hat{h}_{\mathtt{IK}} = 89.07$ | $\hat{h}_{\mathtt{CV}} = 237.67$ | $\hat{h}_{\mathtt{CCT}} = 46.13$ | $\hat{h}_{\mathtt{IK}} = 96.44$ | $\hat{h}_{\mathtt{CV}} = 202.92$ | $\hat{h}_{\mathtt{CCT}} = 74.50$ | $\hat{h}_{\mathtt{IK}} = 68.37$ | $\hat{h}_{\mathtt{CV}} = 237.67$ |
| | $\hat{b}_{\mathtt{CCT}} = 92.06$ | $\hat{b}_{\mathtt{IK}} = 72.61$ | | $\hat{b}_{\mathtt{CCT}} = 66.44$ | $\hat{b}_{\mathtt{IK}} = 69.92$ | | $\hat{b}_{\mathtt{CCT}} = 121.40$ | $\hat{b}_{\mathtt{IK}} = 57.13$ | |
| Non-Food | | | | | | | | | |
| | −6.4 | −6.0 | 9.0 | 1.7 | −19.5 | −21.9 | −12.6 | 0.8 | 10.6 |
| | (−38.4, 25.7) | (−38.0, 26.0) | (−14.2, 32.3) | (−34.6, 38.0) | (−53.0, 14.0) | (−48.9, 5.1) | (−36.6, 11.4) | (−16.0, 17.6) | (−5.4, 26.7) |
| | [−19.1, 64.0] | [−23.3, 60.6] | [−25.1, 38.1] | [8.3, 110.1]** | [−34.0, 55.1] | [−66.6, 8.3] | [−73.3, 14.0] | [−46.5, 8.8] | [−29.9, 17.6] |
| | {−48.2, 30.4} | {33.3, 120.3}*** | | {−35.9, 56.4} | {27.5, 185.2}*** | | {−49.2, 7.2} | {−57.2, 9.9} | |
| | $\hat{h}_{\mathtt{CCT}} = 70.33$ | $\hat{h}_{\mathtt{IK}} = 72.87$ | $\hat{h}_{\mathtt{CV}} = 237.67$ | $\hat{h}_{\mathtt{CCT}} = 44.67$ | $\hat{h}_{\mathtt{IK}} = 73.92$ | $\hat{h}_{\mathtt{CV}} = 214.50$ | $\hat{h}_{\mathtt{CCT}} = 33.88$ | $\hat{h}_{\mathtt{IK}} = 60.43$ | $\hat{h}_{\mathtt{CV}} = 87.08$ |
| | $\hat{b}_{\mathtt{CCT}} = 118.65$ | $\hat{b}_{\mathtt{IK}} = 57.45$ | | $\hat{b}_{\mathtt{CCT}} = 79.91$ | $\hat{b}_{\mathtt{IK}} = 48.55$ | | $\hat{b}_{\mathtt{CCT}} = 69.93$ | $\hat{b}_{\mathtt{IK}} = 51.11$ | |
| Total | | | | | | | | | |
| | −7.1 | 15.6 | 14.5 | 51.5 | 6.2 | −8.9 | −10.2 | −10.6 | −1.4 |
| | (−68.7, 54.5) | (−37.1, 68.3) | (−27.4, 56.4) | (−30.5, 133.5) | (−65.6, 78.0) | (−68.6, 50.9) | (−59.1, 38.7) | (−59.7, 38.6) | (−32.8, 29.9) |
| | [−21.5, 142.2] | [−73.5, 69.0] | [−42.0, 79.2] | [2.6, 194.7]** | [−46.9, 146.5] | [−72.7, 110.4] | [−89.9, 47.4] | [−89.7, 47.4] | [−47.6, 38.6] |
| | {−91.7, 56.2} | {−52.8, 271.7} | | {−27.9, 174.2} | {−24.7, 306.1}* | | {−72.3, 51.5} | {−81.5, 54.9} | |
| | $\hat{h}_{\mathtt{CCT}} = 58.95$ | $\hat{h}_{\mathtt{IK}} = 123.22$ | $\hat{h}_{\mathtt{CV}} = 237.67$ | $\hat{h}_{\mathtt{CCT}} = 46.90$ | $\hat{h}_{\mathtt{IK}} = 90.98$ | $\hat{h}_{\mathtt{CV}} = 237.67$ | $\hat{h}_{\mathtt{CCT}} = 65.01$ | $\hat{h}_{\mathtt{IK}} = 64.02$ | $\hat{h}_{\mathtt{CV}} = 226.09$ |
| | $\hat{b}_{\mathtt{CCT}} = 101.75$ | $\hat{b}_{\mathtt{IK}} = 66.01$ | | $\hat{b}_{\mathtt{CCT}} = 79.91$ | $\hat{b}_{\mathtt{IK}} = 55.39$ | | $\hat{b}_{\mathtt{CCT}} = 97.16$ | $\hat{b}_{\mathtt{IK}} = 58.81$ | |

[a](i) BW-CCT, BW-IK, and BW-CV correspond to estimation methods using, respectively, CCT, IK, and cross-validation bandwidth selectors. (ii) For each bandwidth selection method and outcome, the table reports RD local-linear point estimator, conventional 95% confidence intervals in parentheses, robust 95% confidence intervals in square-brackets using estimated $h_n$ and $\rho_n = 1$, robust 95% confidence intervals in curly-brackets using estimated $h_n$ and $b_n$, and estimated bandwidths values. All confidence intervals are also robust to heteroskedasticity. (iii) For each confidence interval, accompanying stars denote associated null hypothesis of no-treatment effect rejected at: *statistically significant at 10% level, **statistically significant at 5% level, and ***statistically significant at 1% level.

TABLE S.A.XVIII

SHARP RD TREATMENT EFFECT ESTIMATES OF PROGRESA/OPORTUNIDADES ON CONSUMPTION, RURAL LOCALITIES (REGION 12)[a]

| | Pre-Intervention | | | 1-Year Treatment | | | 2-Year Treatment | | |
|---|---|---|---|---|---|---|---|---|---|
| | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV |
| Food | | | | | | | | | |
| | 31.3 | 22.2 | 11.6 | −21.6 | −17.2 | −18.8 | 67.1 | 66.6 | 16.0 |
| | (−89.6, 152.1) | (−75.5, 119.9) | (−67.8, 91.1) | (−123.5, 80.4) | (−109.9, 75.4) | (−78.4, 40.9) | (−4.6, 138.8)* | (−5.1, 138.3)* | (−37.4, 69.4) |
| | [−225.7, 142.8] | [−127.0, 173.3] | [−78.2, 155.8] | [−167.2, 181.9] | [−148.1, 148.0] | [−113.8, 67.6] | [−45.7, 166.2] | [−44.8, 165.2] | [−38.5, 97.2] |
| | {−123.8, 178.8} | {−267.7, 146.6} | | {−136.2, 111.6} | {−147.5, 145.9} | | {3.2, 168.5}** | {−20.6, 166.9} | |
| | $\hat{h}_{CCT} = 48.03$ | $\hat{h}_{IK} = 70.36$ | $\hat{h}_{CV} = 128.00$ | $\hat{h}_{CCT} = 41.45$ | $\hat{h}_{IK} = 47.70$ | $\hat{h}_{CV} = 128.00$ | $\hat{h}_{CCT} = 39.64$ | $\hat{h}_{IK} = 41.26$ | $\hat{h}_{CV} = 128.00$ |
| | $\hat{b}_{CCT} = 67.17$ | $\hat{b}_{IK} = 52.35$ | | $\hat{b}_{CCT} = 64.21$ | $\hat{b}_{IK} = 48.07$ | | $\hat{b}_{CCT} = 72.77$ | $\hat{b}_{IK} = 43.54$ | |
| Non-Food | | | | | | | | | |
| | −7.1 | −15.4 | −19.1 | −43.1 | −28.9 | −19.2 | 0.1 | −0.2 | −3.7 |
| | (−55.0, 40.7) | (−54.1, 23.4) | (−55.6, 17.4) | (−85.6, −0.6)** | (−61.3, 3.5)* | (−46.9, 8.5) | (−36.4, 36.6) | (−27.9, 27.6) | (−29.3, 22.0) |
| | [−115.6, 45.2] | [−57.3, 53.4] | [−56.8, 43.5] | [−108.9, 21.3] | [−98.1, −4.5]** | [−72.7, 0.9]* | [−66.1, 48.1] | [−38.0, 44.7] | [−31.6, 42.2] |
| | {−62.5, 55.3} | {−137.9, 62.7} | | {−101.2, 4.7}* | {−96.1, −4.4}** | | {−43.7, 45.1} | {−115.9, 72.6} | |
| | $\hat{h}_{CCT} = 48.35$ | $\hat{h}_{IK} = 101.21$ | $\hat{h}_{CV} = 123.90$ | $\hat{h}_{CCT} = 39.29$ | $\hat{h}_{IK} = 73.18$ | $\hat{h}_{CV} = 123.90$ | $\hat{h}_{CCT} = 45.57$ | $\hat{h}_{IK} = 98.69$ | $\hat{h}_{CV} = 128.00$ |
| | $\hat{b}_{CCT} = 90.45$ | $\hat{b}_{IK} = 63.30$ | | $\hat{b}_{CCT} = 64.45$ | $\hat{b}_{IK} = 77.14$ | | $\hat{b}_{CCT} = 71.11$ | $\hat{b}_{IK} = 61.78$ | |
| Total | | | | | | | | | |
| | 23.0 | 10.3 | −7.6 | −61.0 | −54.3 | −39.4 | 66.2 | 45.4 | 12.3 |
| | (−149.0, 195.1) | (−109.2, 129.8) | (−109.8, 94.6) | (−185.8, 63.8) | (−156.8, 48.1) | (−107.6, 28.8) | (−37.1, 169.6) | (−45.3, 136.1) | (−61.4, 86.1) |
| | [−426.4, 158.8] | [−177.6, 217.4] | [−120.9, 183.6] | [−250.9, 197.9] | [−221.2, 127.7] | [−163.0, 45.3] | [−129.1, 203.7] | [−51.9, 210.2] | [−61.4, 130.7] |
| | {−188.9, 228.9} | {−253.9, 204.3} | | {−206.7, 104.1} | {−218.6, 122.8} | | {−31.6, 207.4} | {−73.1, 220.4} | |
| | $\hat{h}_{CCT} = 44.66$ | $\hat{h}_{IK} = 78.29$ | $\hat{h}_{CV} = 128.00$ | $\hat{h}_{CCT} = 38.13$ | $\hat{h}_{IK} = 50.19$ | $\hat{h}_{CV} = 119.80$ | $\hat{h}_{CCT} = 41.88$ | $\hat{h}_{IK} = 52.17$ | $\hat{h}_{CV} = 128.00$ |
| | $\hat{b}_{CCT} = 69.75$ | $\hat{b}_{IK} = 63.06$ | | $\hat{b}_{CCT} = 61.70$ | $\hat{b}_{IK} = 51.65$ | | $\hat{b}_{CCT} = 72.24$ | $\hat{b}_{IK} = 46.03$ | |

[a] (i) BW-CCT, BW-IK, and BW-CV correspond to estimation methods using, respectively, CCT, IK, and cross-validation bandwidth selectors. (ii) For each bandwidth selection method and outcome, the table reports RD local-linear point estimator, conventional 95% confidence intervals in parentheses, robust 95% confidence intervals in square-brackets using estimated $h_n$ and $\rho_n = 1$, robust 95% confidence intervals in curly-brackets using estimated $h_n$ and $b_n$, and estimated bandwidths values. All confidence intervals are also robust to heteroskedasticity. (iii) For each confidence interval, accompanying stars denote associated null hypothesis of no-treatment effect rejected at: *statistically significant at 10% level, **statistically significant at 5% level, and ***statistically significant at 1% level.

## TABLE S.A.XIX

SHARP RD TREATMENT EFFECT ESTIMATES OF PROGRESA/OPORTUNIDADES ON CONSUMPTION, RURAL LOCALITIES (REGION 27)[a]

| | Pre-Intervention | | | 1-Year Treatment | | | 2-Year Treatment | | |
|---|---|---|---|---|---|---|---|---|---|
| | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV |
| Food | | | | | | | | | |
| | 0.6 | 7.3 | 3.7 | 22.8 | 22.1 | 15.0 | 9.7 | 19.1 | 33.9 |
| | (−39.6, 40.9) | (−27.6, 42.2) | (−22.8, 30.1) | (−10.2, 55.7) | (−9.3, 53.5) | (−6.3, 36.2) | (−28.5, 47.9) | (−14.1, 52.3) | (6.6, 61.3)** |
| | [−54.2, 59.9] | [−54.2, 47.8] | [−22.6, 45.6] | [−22.7, 65.0] | [−20.5, 65.3] | [−11.1, 45.8] | [−73.0, 45.9] | [−43.3, 46.1] | [−8.8, 59.2] |
| | {−49.9, 46.7} | {−57.9, 69.6} | | {−15.4, 63.3} | {−11.0, 61.2} | | {−40.4, 47.2} | {−62.9, 57.2} | |
| | $\hat{h}_{\mathrm{CCT}} = 71.68$ | $\hat{h}_{\mathrm{IK}} = 99.10$ | $\hat{h}_{\mathrm{CV}} = 317.00$ | $\hat{h}_{\mathrm{CCT}} = 73.24$ | $\hat{h}_{\mathrm{IK}} = 82.84$ | $\hat{h}_{\mathrm{CV}} = 317.00$ | $\hat{h}_{\mathrm{CCT}} = 62.89$ | $\hat{h}_{\mathrm{IK}} = 104.64$ | $\hat{h}_{\mathrm{CV}} = 317.00$ |
| | $\hat{b}_{\mathrm{CCT}} = 117.27$ | $\hat{b}_{\mathrm{IK}} = 77.47$ | | $\hat{b}_{\mathrm{CCT}} = 116.26$ | $\hat{b}_{\mathrm{IK}} = 167.58$ | | $\hat{b}_{\mathrm{CCT}} = 117.81$ | $\hat{b}_{\mathrm{IK}} = 82.05$ | |
| Non-Food | | | | | | | | | |
| | 0.4 | 1.0 | 2.8 | 5.4 | 4.4 | 7.8 | 17.1 | 13.1 | 13.8 |
| | (−33.4, 34.2) | (−31.7, 33.8) | (−18.8, 24.4) | (−24.6, 35.3) | (−20.3, 29.2) | (−10.0, 25.6) | (−12.2, 46.4) | (−9.7, 35.9) | (−7.5, 35.1) |
| | [−50.0, 40.3] | [−48.0, 39.6] | [−28.7, 27.6] | [−50.9, 37.8] | [−28.5, 41.1] | [−24.2, 23.7] | [−42.1, 42.4] | [−20.2, 54.6] | [−9.1, 37.1] |
| | {−42.2, 37.5} | {−46.7, 39.0} | | {−29.9, 41.1} | {−40.3, 40.0} | | {−16.5, 54.1} | {−31.4, 53.5} | |
| | $\hat{h}_{\mathrm{CCT}} = 62.36$ | $\hat{h}_{\mathrm{IK}} = 68.36$ | $\hat{h}_{\mathrm{CV}} = 317.00$ | $\hat{h}_{\mathrm{CCT}} = 62.23$ | $\hat{h}_{\mathrm{IK}} = 101.58$ | $\hat{h}_{\mathrm{CV}} = 317.00$ | $\hat{h}_{\mathrm{CCT}} = 61.22$ | $\hat{h}_{\mathrm{IK}} = 92.17$ | $\hat{h}_{\mathrm{CV}} = 317.00$ |
| | $\hat{b}_{\mathrm{CCT}} = 98.44$ | $\hat{b}_{\mathrm{IK}} = 70.61$ | | $\hat{b}_{\mathrm{CCT}} = 93.62$ | $\hat{b}_{\mathrm{IK}} = 83.86$ | | $\hat{b}_{\mathrm{CCT}} = 118.87$ | $\hat{b}_{\mathrm{IK}} = 79.06$ | |
| Total | | | | | | | | | |
| | 2.0 | 10.8 | 6.5 | 27.6 | 27.0 | 22.8 | 28.1 | 30.3 | 47.8 |
| | (−67.9, 71.9) | (−47.2, 68.7) | (−35.9, 48.9) | (−26.7, 81.8) | (−19.5, 73.4) | (−9.4, 54.9) | (−23.2, 79.4) | (−14.0, 74.7) | (8.5, 87.0)** |
| | [−99.3, 94.3] | [−91.6, 77.5] | [−44.5, 66.4] | [−59.2, 87.1] | [−36.7, 91.9] | [−25.7, 59.9] | [−78.8, 74.5] | [−42.8, 84.0] | [−6.0, 84.5]* |
| | {−87.8, 78.5} | {−112.2, 123.8} | | {−36.7, 92.2} | {−33.8, 92.5} | | {−37.7, 81.9} | {−89.1, 90.1} | |
| | $\hat{h}_{\mathrm{CCT}} = 63.59$ | $\hat{h}_{\mathrm{IK}} = 93.72$ | $\hat{h}_{\mathrm{CV}} = 317.00$ | $\hat{h}_{\mathrm{CCT}} = 61.25$ | $\hat{h}_{\mathrm{IK}} = 89.50$ | $\hat{h}_{\mathrm{CV}} = 317.00$ | $\hat{h}_{\mathrm{CCT}} = 68.58$ | $\hat{h}_{\mathrm{IK}} = 105.08$ | $\hat{h}_{\mathrm{CV}} = 317.00$ |
| | $\hat{b}_{\mathrm{CCT}} = 102.39$ | $\hat{b}_{\mathrm{IK}} = 69.16$ | | $\hat{b}_{\mathrm{CCT}} = 89.83$ | $\hat{b}_{\mathrm{IK}} = 95.31$ | | $\hat{b}_{\mathrm{CCT}} = 136.03$ | $\hat{b}_{\mathrm{IK}} = 78.16$ | |

[a] (i) BW-CCT, BW-IK, and BW-CV correspond to estimation methods using, respectively, CCT, IK, and cross-validation bandwidth selectors. (ii) For each bandwidth selection method and outcome, the table reports RD local-linear point estimator, conventional 95% confidence intervals in parentheses, robust 95% confidence intervals in square-brackets using estimated $h_n$ and $\rho_n = 1$, robust 95% confidence intervals in curly-brackets using estimated $h_n$ and $b_n$, and estimated bandwidths values. All confidence intervals are also robust to heteroskedasticity. (iii) For each confidence interval, accompanying stars denote associated null hypothesis of no-treatment effect rejected at: *statistically significant at 10% level, **statistically significant at 5% level, and ***statistically significant at 1% level.

TABLE S.A.XX

SHARP RD TREATMENT EFFECT ESTIMATES OF PROGRESA/OPORTUNIDADES ON CONSUMPTION, RURAL LOCALITIES (REGION 28)[a]

| | Pre-Intervention | | | 1-Year Treatment | | | 2-Year Treatment | | |
|---|---|---|---|---|---|---|---|---|---|
| | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV | BW-CCT | BW-IK | BW-CV |
| Food | | | | | | | | | |
| | 71.1 | 18.0 | −3.4 | 245.9 | −25.8 | −16.3 | 288.5 | 18.3 | −30.5 |
| | (−44.4, 186.7) | (−34.7, 70.7) | (−58.6, 51.7) | (−86.7, 578.6) | (−100.6, 48.9) | (−99.3, 66.6) | (−42.2, 619.2)* | (−100.3, 137.0) | (−131.3, 70.2) |
| | [−184.2, 306.0] | [−26.9, 180.4] | [−15.0, 127.6] | [−606.1, 1129.2] | [−105.2, 127.7] | [−107.0, 178.0] | [−328.6, 992.0] | [−74.7, 488.1] | [−80.6, 285.2] |
| | {−52.5, 221.1} | {−44.2, 201.5} | | {−79.8, 683.2} | {−134.4, 549.5} | | {−37.4, 711.7}* | {−33.8, 1169.6}* | |
| | $\hat{h}_{CCT} = 49.02$ | $\hat{h}_{IK} = 104.93$ | $\hat{h}_{CV} = 142.53$ | $\hat{h}_{CCT} = 31.18$ | $\hat{h}_{IK} = 182.13$ | $\hat{h}_{CV} = 142.53$ | $\hat{h}_{CCT} = 43.56$ | $\hat{h}_{IK} = 102.51$ | $\hat{h}_{CV} = 142.53$ |
| | $\hat{b}_{CCT} = 95.81$ | $\hat{b}_{IK} = 95.06$ | | $\hat{b}_{CCT} = 65.59$ | $\hat{b}_{IK} = 104.59$ | | $\hat{b}_{CCT} = 90.79$ | $\hat{b}_{IK} = 75.45$ | |
| Non-Food | | | | | | | | | |
| | 215.8 | −40.0 | 54.6 | −57.8 | −43.6 | −45.1 | −85.1 | −123.3 | −128.7 |
| | (−168.3, 599.9) | (−191.0, 111.0) | (−149.6, 258.7) | (−141.3, 25.6) | (−96.6, 9.4) | (−99.3, 9.0) | (−271.7, 101.4) | (−225.6, −21.1)** | (−242.4, −15.0)** |
| | [−138.4, 1579.4] | [−133.7, 416.1] | [−162.8, 667.9] | [−211.9, 105.0] | [−107.5, 16.3] | [−118.4, 17.7] | [−326.6, 233.0] | [−293.7, 27.1] | [−307.1, 61.0] |
| | {−174.6, 714.9} | {−355.1, 1643.5} | | {−167.3, 54.5} | {−147.1, 26.1} | | {−298.6, 158.8} | {−312.4, 111.3} | |
| | $\hat{h}_{CCT} = 46.44$ | $\hat{h}_{IK} = 125.13$ | $\hat{h}_{CV} = 85.52$ | $\hat{h}_{CCT} = 44.95$ | $\hat{h}_{IK} = 176.70$ | $\hat{h}_{CV} = 142.53$ | $\hat{h}_{CCT} = 57.79$ | $\hat{h}_{IK} = 210.57$ | $\hat{h}_{CV} = 142.53$ |
| | $\hat{b}_{CCT} = 89.46$ | $\hat{b}_{IK} = 71.31$ | | $\hat{b}_{CCT} = 73.66$ | $\hat{b}_{IK} = 130.72$ | | $\hat{b}_{CCT} = 93.13$ | $\hat{b}_{IK} = 136.69$ | |
| Total | | | | | | | | | |
| | 313.7 | 6.1 | 90.8 | 194.9 | −66.4 | −61.5 | 90.2 | −152.4 | −159.2 |
| | (−185.7, 813.1) | (−172.0, 184.1) | (−141.3, 322.8) | (−174.1, 563.9) | (−170.4, 37.5) | (−169.2, 46.3) | (−257.6, 438.0) | (−337.3, 32.5) | (−338.5, 20.1)* |
| | [−216.6, 1826.7] | [−122.0, 659.3] | [−168.2, 843.0] | [−1626.4, 1171.7] | [−171.7, 119.1] | [−179.4, 149.8] | [−230.2, 1054.6] | [−290.5, 342.5] | [−312.8, 271.3] |
| | {−193.6, 959.2} | {−343.6, 1721.0} | | {−174.0, 695.7} | {−200.0, 229.1} | | {−266.1, 563.8} | {−215.5, 1336.3} | |
| | $\hat{h}_{CCT} = 42.49$ | $\hat{h}_{IK} = 106.03$ | $\hat{h}_{CV} = 85.52$ | $\hat{h}_{CCT} = 30.99$ | $\hat{h}_{IK} = 162.84$ | $\hat{h}_{CV} = 142.53$ | $\hat{h}_{CCT} = 58.07$ | $\hat{h}_{IK} = 130.30$ | $\hat{h}_{CV} = 142.53$ |
| | $\hat{b}_{CCT} = 86.82$ | $\hat{b}_{IK} = 70.63$ | | $\hat{b}_{CCT} = 63.01$ | $\hat{b}_{IK} = 129.66$ | | $\hat{b}_{CCT} = 106.65$ | $\hat{b}_{IK} = 82.31$ | |

[a](i) BW-CCT, BW-IK, and BW-CV correspond to estimation methods using, respectively, CCT, IK, and cross-validation bandwidth selectors. (ii) For each bandwidth selection method and outcome, the table reports RD local-linear point estimator, conventional 95% confidence intervals in parentheses, robust 95% confidence intervals in square-brackets using estimated $h_n$ and $\rho_n = 1$, robust 95% confidence intervals in curly-brackets using estimated $h_n$ and $b_n$, and estimated bandwidths values. All confidence intervals are also robust to heteroskedasticity. (iii) For each confidence interval, accompanying stars denote associated null hypothesis of no-treatment effect rejected at: *statistically significant at 10% level, **statistically significant at 5% level, and ***statistically significant at 1% level.

– Drop households originally classified as ineligibles, but reclassified in 1999, who received positive transfers that year.

– Drop households with missing information for the consumption variables.

– Drop household with invalid entries in some key pre-intervention variables (household size and age of household head).

– Using the baseline database (1997), we constructed the pre-intervention variables used in Table S.A.XII.

– Using all three databases (March 1998, November 1998, and November 1999), we constructed the following outcome variables (averaged in the household over all its members, expressed as monthly expenses):

∗ *Food Consumption*: household level data on food outlays made in the seven days preceding the interview for 36 food items. It also includes the value of food consumed from own production in that same period of time, valued by imputing a locality level price (based on interviews with local leaders in each village).

∗ *Non-Food Consumption*: expenses reported on a weekly, monthly, and semi-annual basis. Non-food expenses reported on a weekly basis include transportation and tobacco. Monthly outlays include school tuition, health-related expenses, home cleaning, electricity, and home fuel expenditures. Expenditures reported on a semi-annual basis include home and school supplies, clothes, shoes, toys, and payments for special events.

∗ *Total Consumption*: computed as the sum of non-food expenditures and the value of food consumption.

All empirical results were obtained using the `STATA` package `rdrobust`. Here we briefly describe the main underlying implementation details for each part of our empirical illustration.

• *Balance Tests*. We conducted difference-in-means tests of pre-intervention covariates for the full sample and for observations near the cutoffs, as is common in empirical studies employing RD designs.

• *Figures S.A.1 and S.A.2*. The RD plots were obtained using the `STATA` command `rdbinselect` with a scale factor of 5.

• *MSE-Optimal Bandwidth Selection*. Estimation results using CCT and IK methods were obtained using the default options in the command `rdrobust`.

• *Cross-Validation Bandwidth Selection*. Figures S.A.3 and S.A.4 present plots of the cross-validation objective functions for each case analyzed in Table S.A.X. These figures were constructed using the option `cvplot` in the command `rdbwselect`, after selecting the CV tuning parameters as appropriate. In particular, we restricted the range of the running variable, and also adjusted the parameter $\delta \in \{0.1, 0.15, 0.2, 0.25\}$ in the CV cross-validation objective function.

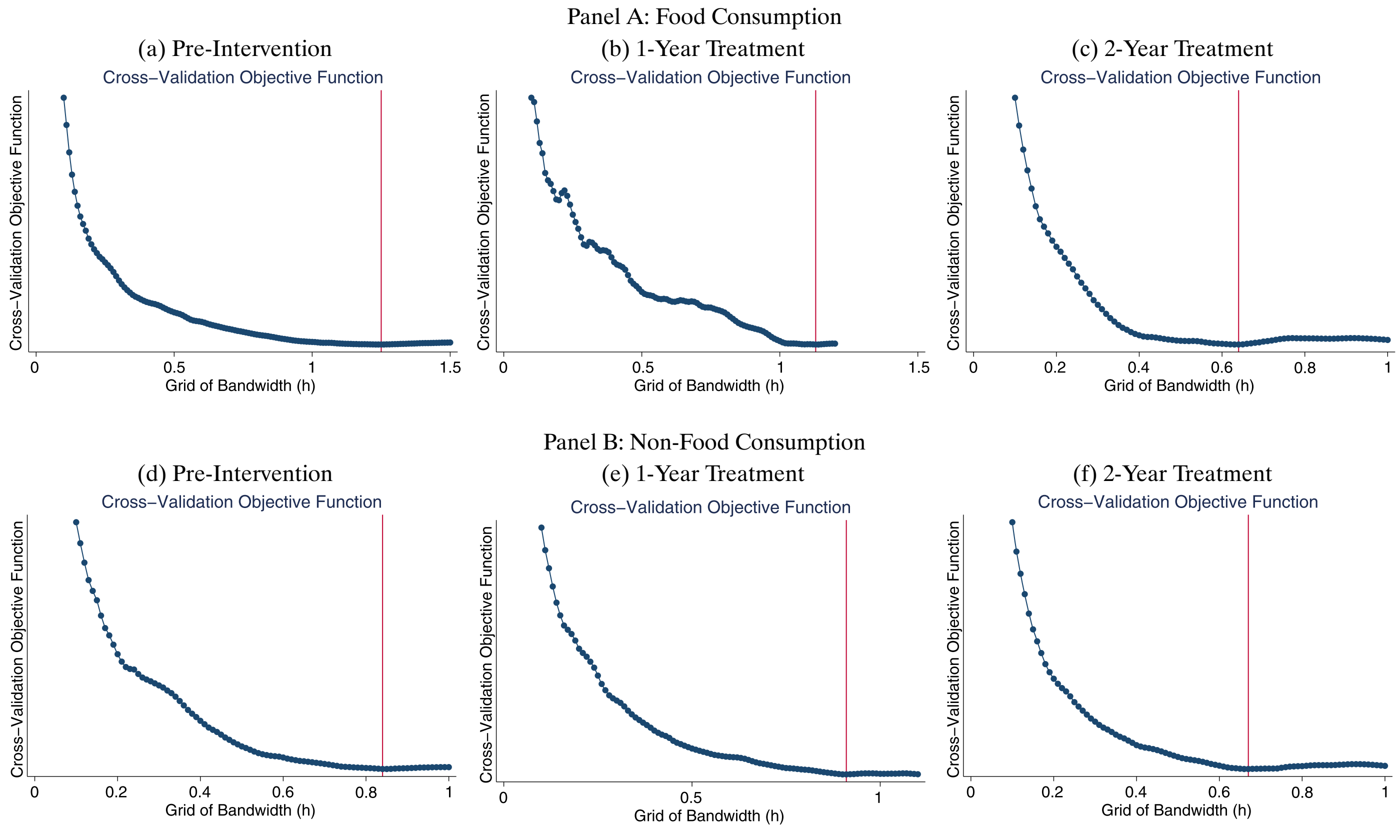


FIGURE S.A.3.—Cross-validation objective functions for bandwidth selection, urban localities.

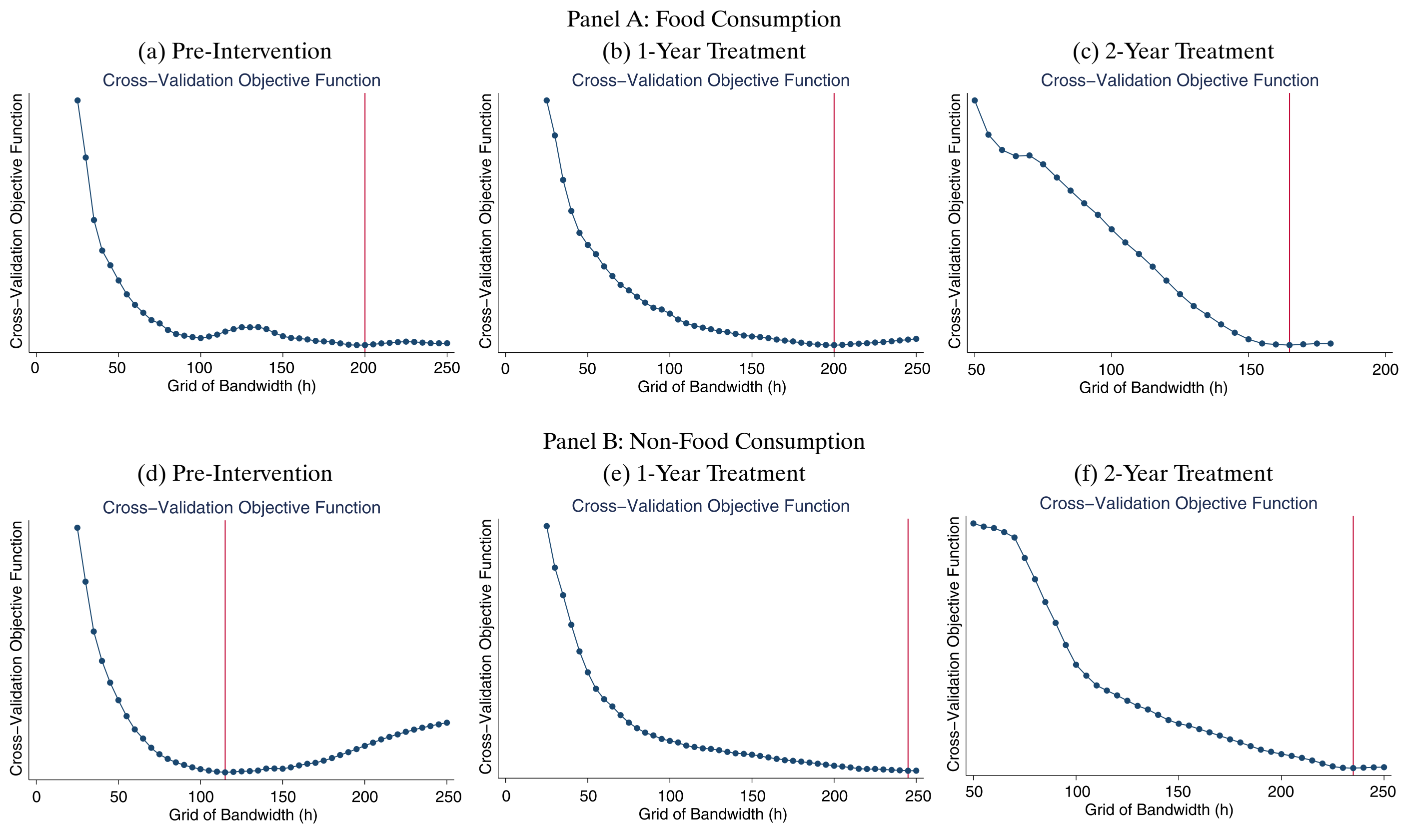


FIGURE S.A.4.—Cross-validation objective functions for bandwidth selection, rural localities.

## REFERENCES


ABADIE, A., AND G. W. IMBENS (2006): "Large Sample Properties of Matching Estimators for Average Treatment Effects," *Econometrica*, 74 (1), 235–267. [33,37]

ANGELUCCI, M., AND O. ATTANASIO (2009): "Program Effect on Consumption, Low Participation, and Methodological Issues," *Economic Development and Cultural Change*, 57 (3), 479–506. [77]

——— (2013): "The Demand for Food of Poor Urban Mexican Households: Understanding Policy Impacts Using Structural Models," *American Economic Journal: Economic Policy*, 5 (1), 146–178. [77]

ANGELUCCI, M., AND G. DE GIORGI (2009): "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?" *American Economic Review*, 99 (1), 486–508. [77, 80]

ANGELUCCI, M., O. ATTANASIO, AND V. DI MARO (2012): "The Impact of Oportunidades on Consumption, Savings and Transfers," *Fiscal Studies*, 33 (3), 305–334. [76]

ATTANASIO, O., C. MEGHIR, AND A. SANTIAGO (2011): "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate Progresa," *Review of Economic Studies*, 79 (1), 37–66. [75]

BEHRMAN, J. R., J. GALLARDO-GARCÍA, S. W. PARKER, P. E. TODD, AND V. VÉLEZ-GRAJALES (2012): "Are Conditional Cash Transfers Effective in Urban Areas? Evidence From Mexico," *Education Economics*, 20 (3), 233–259. [75]

BUDDELMEYER, H., AND E. SKOUFIAS (2004): "An Evaluation of the Performance of Regression Discontinuity Design on PROGRESA," Policy Research Working Paper 3386, World Bank. [76]

CALONICO, S., M. D. CATTANEO, AND M. H. FARRELL (2014): "On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Estimation," Working Paper, University of Michigan. [2]

CALONICO, S., M. D. CATTANEO, AND R. TITIUNIK (2014a): "Optimal Data-Driven Regression Discontinuity Plots," Working Paper, University of Michigan. [77]

——— (2014b): "Robust Data-Driven Inference in the Regression-Discontinuity Design," *Stata Journal* (forthcoming). [1,75]

——— (2014c): "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 82 (6), 2295–2326. [1,2]

——— (2014d): "`rdrobust`: An R Package for Robust Inference in Regression-Discontinuity Designs," Working Paper, University of Michigan. [1,75]

DESJARDINS, S. L., AND B. P. MCCALL (2009): "The Impact of the Gates Millennium Scholars Program on the Retention, College Finance- and Work-Related Choices, and Future Educational Aspirations of Low-Income Students," Working Paper, University of Michigan. [1,51,56, 58]

DJEBBARI, H., AND J. SMITH (2008): "Heterogeneous Impacts in PROGRESA," *Journal of Econometrics*, 145 (1–2), 64–80. [75]

DUBOIS, P., A. DE JANVRY, AND E. SADOULET (2012): "Effects on School Enrollment and Performance of a Conditional Cash Transfer Program in Mexico," *Journal of Labor Economics*, 30 (3), 555–589. [75]

FAN, J., AND I. GIJBELS (1996): *Local Polynomial Modelling and Its Applications*. New York: Chapman & Hall/CRC. [2]

FERNALD, L. C. H., P. J. GERTLER, AND L. M. NEUFELD (2009): "10-Year Effect of Oportunidades, Mexico's Conditional Cash Transfer Programme, on Child Growth, Cognition, Language, and Behaviour: A Longitudinal Follow-Up Study," *Lancet*, 374, 1997–2005. [75]

GERTLER, P. J., S. W. MARTINEZ, AND M. RUBIO-CODINA (2012): "Investing Cash Transfers to Raise Long-Term Living Standards," *American Economic Journal: Applied Economics*, 4 (1), 164–192. [77]

HODDINOTT, J., AND E. SKOUFIAS (2004): "The Impact of PROGRESA on Food Consumption," *Economic Development and Cultural Change*, 53 (1), 37–61. [77]

IMBENS, G. W., AND K. KALYANARAMAN (2012): “Optimal Bandwidth Choice for the Regression Discontinuity Estimator,” *Review of Economic Studies*, 79 (3), 933–959. [1,43,49,51,56]
LEE, D. S. (2008): “Randomized Experiments From Non-Random Selection in U.S. House Elections,” *Journal of Econometrics*, 142 (2), 675–697. [1,49,50]
LUDWIG, J., AND D. L. MILLER (2007): “Does Head Start Improve Children’s Life Chances? Evidence From a Regression Discontinuity Design,” *Quarterly Journal of Economics*, 122 (1), 159–208. [1,50,51,56]

*Dept. of Economics, University of Miami, 5250 University Dr., Coral Gables, FL 33124, U.S.A.; scalonico@bus.miami.edu,*

*Dept. of Economics, University of Michigan, 611 Tappan Ave., Ann Arbor, MI 48109, U.S.A.; cattaneo@umich.edu,*

*and*

*Dept. of Political Science, University of Michigan, 505 S. State St., Ann Arbor, MI 48109, U.S.A.; titiunik@umich.edu.*

*Manuscript received July, 2013; final revision received June, 2014.*