Regression Discontinuity Design under Self-selection: Supplementary Materials

1 Additional Results

When $\mathbb{E}(Z_i(1)|X_i = c) \neq \mathbb{E}(Z_i(0)|X_i = c)$ (i.e., the covariates are unbalanced at the cutoff) and $\gamma \neq 0$, we extend the weighted average treatment effect (WATE) using the framework below.

$$\begin{aligned}
\tau^{w}_{SRD} &= \mathbb{E}\{Y(1)w_1(Z(1))|X = c\} \\
&\quad - \mathbb{E}\{Y(0)w_0(Z(0))|X = c\} \\
&= \int \Big[\mathbb{E}(Y(1)|X = c, Z(1) = z)w_1(z)f_{Z(1)|X}(z|c) \\
&\quad - \mathbb{E}(Y(0)|X = c, Z(1) = z)w_0(z)f_{Z(0)|X}(z|c)\Big]dz, \qquad (1.1)
\end{aligned}$$

where $w_1(\cdot)$ and $w_0(\cdot)$ denote different choices of weights to form the estimand, and $f_{Z(1)|X}(\cdot|\cdot)$ and $f_{Z(0)|X}(\cdot|\cdot)$ are the conditional density of $Z(1)$ and $Z(0)$ given X. In order to interpret (1.1) as the WATE, we require the following normalization condition for $w_1(\cdot)$ and $w_0(\cdot)$:

$$\int w_1(z)f_{Z(1)|X}(z|c)dz = \int w_0(z)f_{Z(0)|X}(z|c)dz = 1.$$

In particular, by choosing appropriate $w_1(\cdot)$ and $w_0(\cdot)$, (1.1) can be interpreted as the average of the difference of the conditional mean functions corresponding to a target population. It is easy to see that WATE defined in (1.1) recovers τ^{w1}_{SRD} by taking

$$w_1(z) = \frac{f_Z(z)}{f_{Z(1)|X}(z|c)}, \quad \text{and} \quad w_0(z) = \frac{f_Z(z)}{f_{Z(0)|X}(z|c)}. \qquad (1.2)$$

- Average treatment effect over locally untreated population:

$$\tau^{w2}_{SRD} = \int \Delta(c, z)f_{Z(0)|X}(z|c)dz.$$

  In this causal parameter, we average the conditional mean difference over the untreated population right below the threshold whose covariates follow from the *conditional distribution* $f_{Z(0)|X}(z|c)$. Similarly, we obtain τ^{w2}_{SRD} by taking

$$w_1(z) = \frac{f_{Z(0)|X}(z|c)}{f_{Z(1)|X}(z|c)} \quad \text{and} \quad w_0(z) = 1.$$

- Average treatment effect over locally randomized population:

$$\tau_{SRD}^{w3} = \int \Delta(c, z) \frac{f_{Z(0)|X}(z|c) + f_{Z(1)|X}(z|c)}{2} dz.$$

  This is the estimand studied by Frölich and Huber (2018) in the sharp RD case. Under the proposed WATE framework, τ_{SRD}^{w3} can be viewed as the average treatment effect over the population around the threshold which is randomized so that their covariates follow from $f_{Z(0)|X}(z|c)$ and $f_{Z(1)|X}(z|c)$ with equal probability. Similarly, we obtain τ_{SRD}^{w3} by taking

$$w_1(z) = \frac{f_{Z(1)|X}(z|c) + f_{Z(0)|X}(z|c)}{2f_{Z(1)|X}(z|c)} \text{ and } w_0(z) = \frac{f_{Z(1)|X}(z|c) + f_{Z(0)|X}(z|c)}{2f_{Z(0)|X}(z|c)}.$$

- Average treatment effect via classical RD estimand: We note that the proposed WATE reduces to the classical RD estimand τ_{SRD} (2.1) by taking $w_1(z) = w_0(z) = 1$. However, we note that unlike the previous three examples, τ_{SRD} may not be written as the average treatment effect over one well defined population. To see this, recall that when there exists self-selection, the conditional distributions $f_{Z(1)|X}(z|c)$ and $f_{Z(0)|X}(z|c)$ usually differ from each other. Then τ_{SRD} in (2.1) can be written as the difference of the average of $\mathbb{E}(Y_i(t)|X_i = c, Z_i(t) = z)$ over two populations (i.e., the population right below and above the threshold) with covariate distributions $f_{Z(1)|X}(z|c)$ and $f_{Z(0)|X}(z|c)$ respectively. This is the reason for which τ_{SRD} is confounded by the unbalanced covariates.

Estimand	$w_1(z)$	$w_0(z)$	$\pi_1(z)$	$\pi_0(z)$
$\int \Delta(c,z) f_Z(z) dz$	$\frac{f_Z(z)}{f_{Z(1)\|X}(z\|c)}$	$\frac{f_Z(z)}{f_{Z(0)\|X}(z\|c)}$	$\frac{f_{X,Z(1)}(c,z)}{2f_Z(z)}$	$\frac{f_{X,Z(0)}(c,z)}{2f_Z(z)}$
$\int \Delta(c,z) f_{Z(0)\|X}(z\|c) dz$	$\frac{f_{Z(0)\|X}(z\|c)}{f_{Z(1)\|X}(z\|c)}$	1	$\frac{f_{X,Z(1)}(c,z)}{2f_{Z(0)\|X}(z\|c)}$	$\frac{f_X(c)}{2}$
$\int \Delta(c,z) \frac{f_{Z(0)\|X}(z\|c)+f_{Z(1)\|X}(z\|c)}{2} dz$	$\frac{f_{Z(1)\|X}(z\|c)+f_{Z(0)\|X}(z\|c)}{2f_{Z(1)\|X}(z\|c)}$	$\frac{f_{Z(1)\|X}(z\|c)+f_{Z(0)\|X}(z\|c)}{2f_{Z(0)\|X}(z\|c)}$	$\frac{f_{X,Z(1)}(c,z)}{f_{Z(1)\|X}(z\|c)+f_{Z(0)\|X}(z\|c)}$	$\frac{f_{X,Z(0)}(c,z)}{f_{Z(1)\|X}(z\|c)+f_{Z(0)\|X}(z\|c)}$

Table 1: Three examples of the weighted average treatment effect τ^w_{SRD}, where $\Delta(c,z) = \mathbb{E}(Y_i(1)|X_i = c, Z_i(1) = z) - \mathbb{E}(Y_i(0)|X_i = c, Z_i(0) = z)$.

2 Additional Simulation and Empirical Results

In this setting, consider the following data generating process:

$$y_i(1) = 2 + x_i + \beta z_i + \epsilon_i,$$

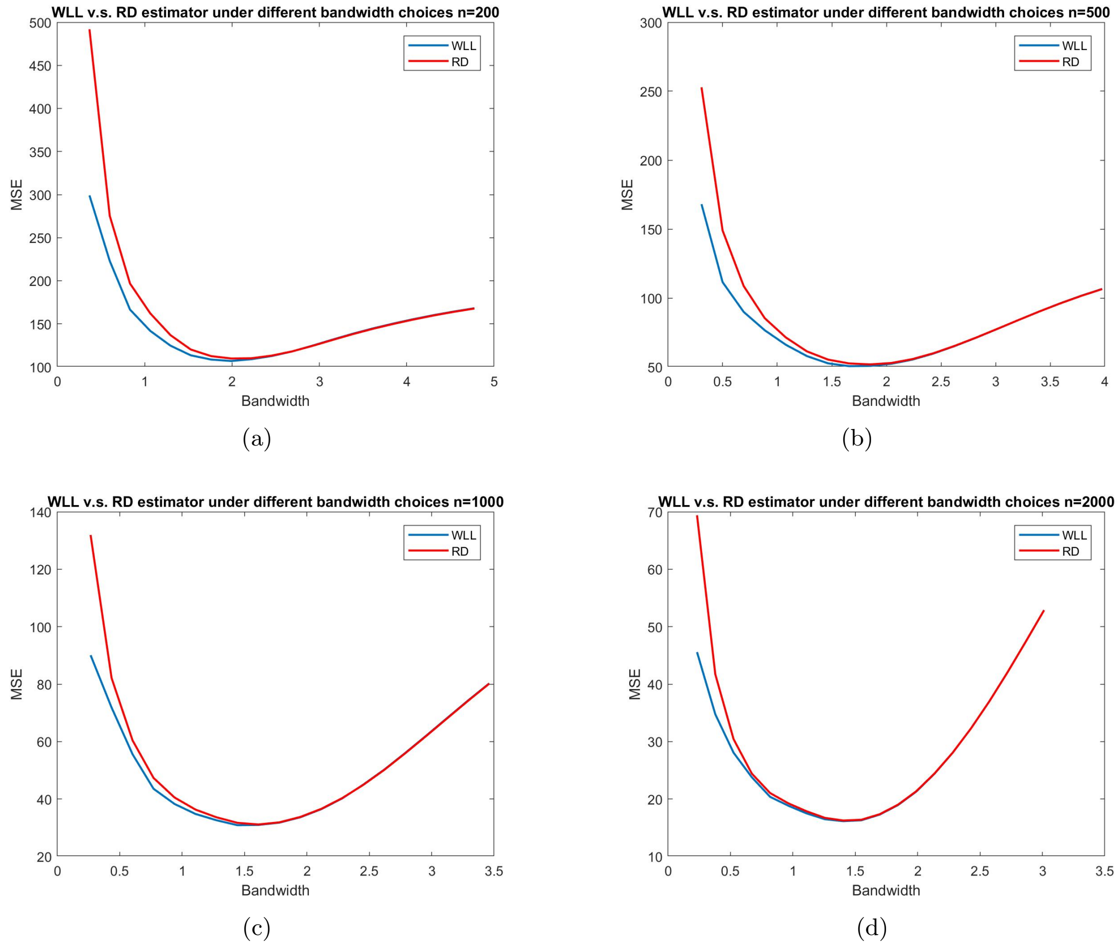
$$y_i(0) = 1 + x_i + \beta z_i + \epsilon_i,$$

where x_i, z_i, and ϵ_i are generated independently from $N(0, 1)$ distribution. The treatment T_i is assigned at the cutoff 0: $T_i = \mathbf{1}(x_i > 0)$. In this case, there is no discontinuity of the conditional distribution of z_i given $x_i = 0$. Thus, our estimaind τ_{SRD}^{w1} equals the standard RD estimand τ_{SRD} (both are equal to 1). We vary β from 0 to 5 and compare weighted local linear (WLL) estimator with the standard RD estimator (Imbens and Lemieux (2008)) in terms of bias, variance, root-mean-squared error (MSE), coverage probability of 95% confidence intervals (Coverage) and its length (CI length). When implementing both methods, we set the bandwidth parameter for standard RD estimator using cross-validation and then use the same bandwidth for WLL. The results based on 500 simulations are shown in Table 2. When $\beta = 0$, there is no covariates involved in the outcome function. Standard RD estimator performs neck to neck with our estimator. When $\beta \neq 0$, the standard RD estimator performs slightly better in terms of bias, however, our estimator consistently has smaller variance and MSE.

		bias		variance		MSE		Coverage		CI length	
n	β	RD	WLL	RD	WLL	RD	WLL	RD	WLL	RD	WLL
	0	-0.0106	**-0.0102**	0.3731	**0.3767**	0.3733	**0.3768**	0.9650	0.9750	1.6666	1.6853
	1	0.0236	**0.0226**	0.6090	**0.5423**	0.6094,	**0.5427**	0.9250	0.9400	2.2335	2.1753
500	2	-0.0583	**-0.0548**	0.9004	**0.7886**	0.9023	**0.7905**	0.9600	0.9600	3.6103	3.2369
	5	-0.0522	**-0.0423**	2.1757	**1.8264**	2.1763	**1.8268**	0.8800	0.9000	7.1005	6.1439
	0	**0.0021**	0.0034	0.2853	**0.2841**	0.2854	**0.2841**	0.9900	0.9900	1.3447	1.4083
	1	**-0.0088**	-0.0118	0.4379	**0.4387**	0.4740	**0.4388**	0.9200	0.9450	1.6833	1.6661
1000	2	**0.0380**	0.0431	0.6046	**0.5118**	0.6058	**0.5136**	0.9400	0.9550	2.2516	2.2635
	5	**0.1106**	0.1129	1.4676	**1.2770**	1.4718	**1.2820**	0.9600	0.9850	6.3293	6.9139
	0	**-0.0107**	-0.0111	0.2098	**0.2096**	0.2101	**0.2099**	0.9300	0.9200	0.7400	0.7229
	1	-0.0320	**-0.0225**	0.3151	**0.2920**	0.3167	**0.2929**	0.9850	0.9850	1.3572	1.3237
2000	2	-0.0177	**-0.0151**	0.5234	**0.4595**	0.5237	**0.4598**	0.9500	0.9650	2.0911	1.9680
	5	**0.1331**	0.1360	1.1707	**1.0019**	1.1783	**1.0111**	0.9150	0.9200	4.2082	3.7723
	0	-0.0149	**-0.0136**	**0.1606**	0.1616	**0.1613**	0.1622	0.9200	0.9250	0.5699	0.5725
	1	**0.0034**	0.0049	0.2451	**0.2275**	0.2451	**0.2275**	0.9350	0.9400	0.9012	0.8793
5000	2	**-0.0130**	-0.0137	0.3301	**0.2940**	0.3304	**0.2943**	0.9450	0.9600	1.3654	1.2931
	5	**0.0053**	0.0139	0.8355	**0.7016**	0.8355	**0.7018**	0.9250	0.9750	3.1578	3.1535

Table 2: Comparison of the standard RD estimator and the proposed weighted local linear estimator (WLL) in the first setting.

Figure 1 compares the MSE of the standard RD estimator with our WLL estimator across different bandwidth choices. The figure is consistent with corollary 1 as our estimator is asymptotically more efficient through including additional covariates into the estimation. Moreover, the advantage of WLL estimator is greater when perform under-smoothing and the difference of the two estimators becomes smaller as the bandwidth increases.



(a) (b) (c) (d)

Figure 1: Comparing MSE of WLL and RD estimator under the same bandwidth choices

Figure 2 plots the connection between search score and click-ability for the second position in the sponsored advertisement area. Once the score passed the cut-off at 15.17, the customers' links will be placed at the sponsored advertisement area and a significant increase in the click traffic can be observed. Figure 3 plots the conditional density of the covariate before and after the cut-off. This indicates a violation of the standard RD assumption. Figure 2 plots the connection between search score and click-ability for the second position in the sponsored advertisement area.

3 Proofs

Figure 3 plots the conditional density of the covariate before and after the cut-off.

In the proof, we use C to denote a generic constant which may change from line to line. Define the following notations:

$$\kappa_\iota = \int_{u>0} K(u)u^\iota du \quad \text{and} \quad \kappa_{2\iota} = \int_{u>0} K(u)^2 u^\iota du, \quad \text{for } \iota = 0, 1, 2, \cdots$$

For the ease of presentation, we introduce the following notations. Let $f'_{X|Z(1)}(c^+|z_i)$, $f''_{X|Z(1)}(c^+|z_i)$ and $f'''_{X|Z(1)}(c^+|z_i)$ denote the right derivatives $\frac{\partial f_{X|Z(1)}(c^+|z_i)}{\partial x}$, $\frac{\partial^2 f_{X|Z(1)}(c^+|z_i)}{\partial x^2}$ and $\frac{\partial^3 f_{X|Z(1)}(c^+|z_i)}{\partial x^3}$. Denote

$$m_1(x, z) = \mathbb{E}(Y(1)|X = x, Z(1) = z),$$

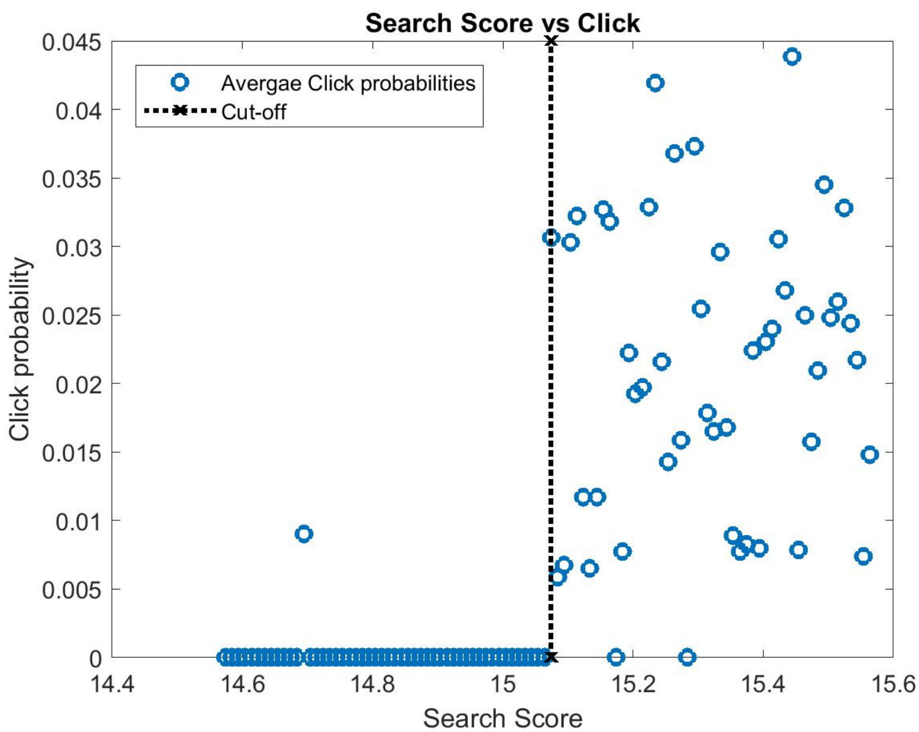


Figure 2: Plot of search score and click-ability. If the search score is above 15.17, the bidder's advertisement will be displayed at the second position. Otherwise, the advertisement will be displayed based on its regular search orders.

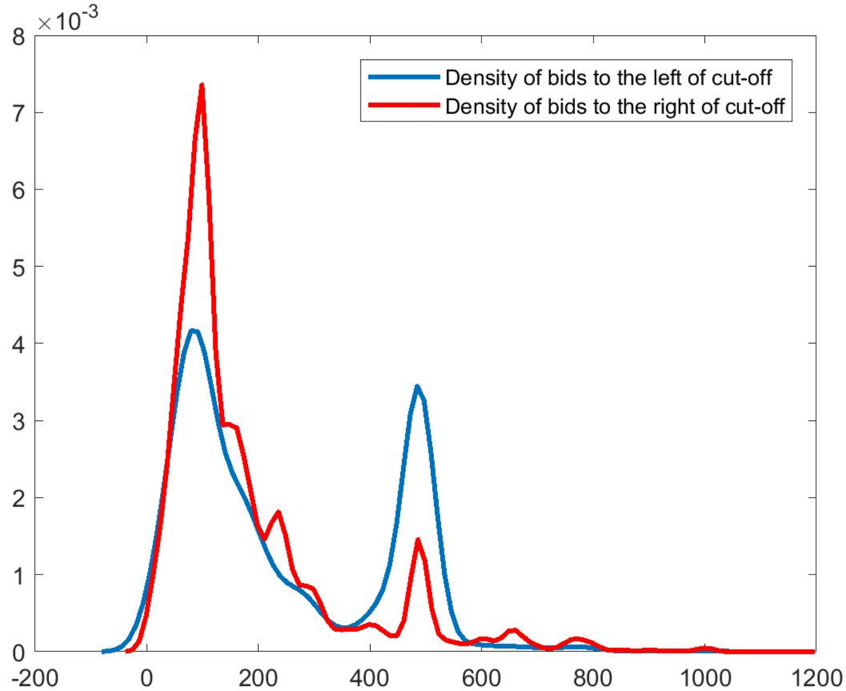


Figure 3: Density plot for bids before and after the cut-off

$$\alpha_1 = \mathbb{E}\{Y(1)w_1(Z(1))|X = c^+\} = \int \frac{f_{X|Z(1)}(c^+|z)}{\pi_1(z)} m_1(c^+, z) f_{Z(1)}(z)dz,$$

and

$$\beta_1 = \int \frac{f'_{X|Z(1)}(c^+|z)}{\pi_1(z)} m_1(c^+, z) f_{Z(1)}(z)dz + \int \frac{f_{X|Z(1)}(c^+|z)}{\pi_1(z)} m_1'(c^+, z) f_{Z(1)}(z)dz,$$

where $m_1'(c^+, z) = \frac{\partial m_1(c^+,z)}{\partial x}$. Define $R_i = y_i - \alpha_1 - \beta_1(x_i - c)$,

$$\begin{aligned}
A_n &= \sum_{i=1}^n \frac{T_i}{\pi_1(z_i)} K\Big(\frac{x_i - c}{h}\Big)(x_i - c)^2, \\
B_n &= -\sum_{i=1}^n \frac{T_i}{\pi_1(z_i)} K\Big(\frac{x_i - c}{h}\Big)(x_i - c), \\
C_n &= \sum_{i=1}^n \frac{T_i}{\pi_1(z_i)} K\Big(\frac{x_i - c}{h}\Big), \\
D_n &= A_n C_n - B_n^2.
\end{aligned}$$

Since A_n, B_n, C_n, D_n all depend on $\pi_t(z)$, we denote the corresponding version with $\widehat{\pi}_1(z)$ as $\widehat{A}_n, \widehat{B}_n, \widehat{C}_n, \widehat{D}_n$.

3.1 Proof of Lemma 3

Proof. By the continuity assumption on $\mathbb{E}(Z_i(t)|X_i = x)$,

$$\mathbb{E}(Z_i(1)|X_i = c) = \lim_{x \to c^+} \mathbb{E}(Z_i(1)|X_i = x) = \lim_{x \to c^+} \mathbb{E}(Z_i|X_i = x),$$

and similarly $\mathbb{E}(Z_i(0)|X_i = c) = \lim_{x \to c^-} \mathbb{E}(Z_i|X_i = x)$. The lemma holds. ∎

3.2 Proof of Theorem 5

We proof an extension version of theorem 5, which includes τ_{SRD}^{w2} and τ_{SRD}^{w3} .

Theorem 6 (Nonparametric Identification)**.** Under Assumption 4, τ_{SRD}^{w1} is identifiable:

$$\tau_{SRD}^{w1} = \int \big[\mathbb{E}(Y|X = c^+, Z = z) - \mathbb{E}(Y|X = c^-, Z = z)\big] f_Z(z)dz,$$

where $\mathbb{E}(Y|X = c^+, Z) = \lim_{x\to c^+} \mathbb{E}(Y|X = x, Z)$ and $\mathbb{E}(Y|X = c^-, Z) = \lim_{x\to c^-} \mathbb{E}(Y|X = x, Z)$. In addition, if $f_{Z(0)|X}(z|x)$ is left continuous in x at $x = c$ for any $z \in \mathcal{Z}$, then τ_{SRD}^{w2} is identifiable:

$$\tau_{SRD}^{w2} = \int \big[\mathbb{E}(Y|X = c^+, Z = z) - \mathbb{E}(Y|X = c^-, Z = z)\big] f_{Z|X}(z|c^-)dz,$$

where $f_{Z|X}(z|c^-) = \lim_{x\to c^-} f_{Z|X}(z|x)$. Furthermore, if $f_{Z(1)|X}(z|x)$ is right continuous in x at $x = c$ for any $z \in \mathcal{Z}$, then τ_{SRD}^{w3} is identifiable:

$$\tau_{SRD}^{w3} = \int \big[\mathbb{E}(Y|X = c^+, Z = z) - \mathbb{E}(Y|X = c^-, Z = z)\big] \frac{f_{Z|X}(z|c^-) + f_{Z|X}(z|c^+)}{2} dz.$$

Proof. To show the identifiability of τ_{SRD}^{w1}, we note that Assumption 4 implies

$$\int \mathbb{E}(Y(1)|X = c, Z(1) = z) f_Z(z)dz = \int \lim_{\delta\to 0^+} \mathbb{E}(Y(1)|X = c + \delta, Z(1) = z) f_Z(z)dz,$$

By the definition of the potential outcome, we have $\mathbb{E}(Y(1)|X = c + \delta, Z(1) = z) = \mathbb{E}(Y|X = c + \delta, Z = z)$ for any $\delta > 0$. Thus,

$$\int \mathbb{E}(Y(1)|X = c, Z(1) = z) f_Z(z)dz = \int \lim_{\delta\to 0^+} \mathbb{E}(Y|X = c + \delta, Z = z) f_Z(z)dz.$$

Following the same step, we can show that

$$\int \mathbb{E}(Y(0)|X = c, Z(0) = z) f_Z(z)dz = \int \lim_{\delta\to 0^-} \mathbb{E}(Y|X = c + \delta, Z = z) f_Z(z)dz.$$

This implies τ_{SRD}^{w1} is identifiable. To show τ_{SRD}^{w2} is identifiable, similarly we have $\mathbb{E}(Y(1)|X = c, Z(1) = z) = \mathbb{E}(Y|X = c^+, Z = z)$. In addition, by the left continuity of $f_{Z(0)|X}(z|x)$, we further have $f_{Z(0)|X}(z|c) = \lim_{\delta\to 0^-} f_{Z(0)|X}(z|c + \delta) = f_{Z|X}(z|c^-)$. Thus, we have

$$\int \mathbb{E}(Y(1)|X = c, Z(1) = z) f_{Z(0)|X}(z|c)dz = \int \mathbb{E}(Y|X = c^+, Z = z) f_{Z|X}(z|c^-)dz.$$

This implies τ_{SRD}^{w2} is identifiable. The identifiability of τ_{SRD}^{w3} follows from the same argument. This completes the proof.

Finally, we show that an alternative identification for τ_{SRD}^{w1} can be written as

$$\tau_{SRD}^{w1} = \lim_{\delta\to 0^+} E\Big(\frac{YT}{\pi_1(Z)}|X = c + \delta\Big) - \lim_{\delta\to 0^+} E\Big(\frac{Y(1-T)}{\pi_0(Z)}|X = c - \delta\Big),$$

Notice that

$$E\Big(\frac{YT}{\pi_1(Z)}|X = c + \delta\Big) = \int \mathbb{E}(Y(1)|X = c^+, Z = z) f_Z(z)dz$$

and

$$E\Big(\frac{YT}{\pi_0(Z)}|X = c - \delta\Big) = \int \mathbb{E}(Y(0)|X = c^-, Z = z) f_Z(z)dz$$

By taking the limit and with the assumption 4, the result holds.

□

Lemma 7. Under assumption 4, 6 and 7, we have

$$\Big|\frac{\widehat{A}_n}{n} - h^3\kappa_2 - h^4\mathbb{E}\left(\frac{f'_{X|Z(1)}(c^+|z_i)}{\pi_1(z_i)}\right)\kappa_3\Big| = O_p(h^5 + rh^3 + \frac{h^{5/2}}{n^{1/2}}),$$

$$\Big|\frac{\widehat{B}_n}{n} + h^2\kappa_1 + h^3\mathbb{E}\left(\frac{f'_{X|Z(1)}(c^+|z_i)}{\pi_1(z_i)}\right)\kappa_2\Big| = O(h^4 + rh^2 + \frac{h^{3/2}}{n^{1/2}}),$$

$$\Big|\frac{\widehat{C}_n}{n} - h/2\Big| = O_p(h^2 + rh + \frac{h^{1/2}}{n^{1/2}}), \tag{3.1}$$

$$\Big|\frac{\widehat{D}_n}{n^2} - h^4\left(\frac{1}{2}\kappa_2 - \kappa_1^2\right)\Big| = O_p(h^5 + rh^4 + \frac{h^{7/2}}{n^{1/2}}),$$

where r satisfies $\sup_{z\in\mathcal{Z}}|\widehat{\pi}_1(z) - \pi_1(z)| = O_p(r)$.

Proof. We will focus on the proof of (3.1). The remain results can be shown following the similar steps. By triangle inequality,

$$\begin{aligned}&\Big|\frac{\widehat{C}_n}{n} - \frac{h}{2}\Big| \\ &\leq \Big|\frac{1}{n}\sum_{i=1}^n \frac{T_i}{\pi_1(z_i)}K\Big(\frac{x_i - c}{h}\Big) - \frac{h}{2}\Big| + \Big|\frac{1}{n}\sum_{i=1}^n \frac{T_i(\pi_1(z_i) - \widehat{\pi}_1(z_i))}{\pi_1(z_i)\widehat{\pi}_1(z_i)}K\Big(\frac{x_i - c}{h}\Big)\Big| := I_1 + I_2.\end{aligned}$$

For term I_1, we further decompose it into two terms,

$$I_1 \leq \Big|\frac{1}{n}\sum_{i=1}^n \frac{T_i}{\pi_1(z_i)}K\Big(\frac{x_i - c}{h}\Big) - \mathbb{E}\Big(\frac{T_i}{\pi_1(z_i)}K\Big(\frac{x_i - c}{h}\Big)\Big)\Big| + \Big|\mathbb{E}\Big(\frac{T_i}{\pi_1(z_i)}K\Big(\frac{x_i - c}{h}\Big)\Big) - h/2\Big| := I_{11} + I_{12}. \tag{3.2}$$

The bias term I_{12} is computed as

$$\begin{aligned}I_{12} &= \Big|\int\int_{x>c}\frac{1}{\pi_1(z)}K\Big(\frac{x_i - c}{h}\Big)f_{X,Z(1)}(x,z)dxdz - \frac{h}{2}\Big| \\ &= h\Big|\int\int_{u>0}\frac{1}{\pi_1(z)}K(u)f_{X,Z(1)}(uh + c, z)dudz - \frac{1}{2}\Big| \\ &\leq h\Big|\int\int_{u>0}\frac{1}{\pi_1(z)}K(u)f_{X,Z(1)}(c^+, z)dudz - \frac{1}{2}\Big| \\ &\quad + h\Big|\int\int_{u>0}\frac{1}{\pi_1(z)}K(u)f'_{X|Z(1)}(\widetilde{u}|z)uhf_{Z(1)}(z)dudz\Big|,\end{aligned}$$

where the last step follows from the mean value theorem for some intermediate value $\widetilde{u}$. Our assumption implies that $|f'_{X|Z(1)}(\widetilde{u}|z)|$ is bounded. In addition, by assumption $\pi_1(z)$ is bounded away from 0 by a constant, thus the second term is of order h^2. For the first term, by the choice of $\pi_1(z)$ we get

$$\int\frac{1}{\pi_1(z)}f_{X|Z(1)}(c^+|z)f_{Z(1)}(z)dz = 1, \tag{3.3}$$

which implies

$$\int\int_{u>0}\frac{1}{\pi_1(z)}K(u)f_{X|Z(1)}(c^+|z)f_{Z(1)}(z)dudz = \int\frac{1}{\pi_1(z)}f_{X|Z(1)}(c^+|z)f_{Z(1)}(z)dz\int_{u>0}K(u)du = \frac{1}{2}.$$

Thus, we have $I_{12} = O(h^2)$. Now we consider I_{11}. By the Markov inequality, $I_{11} \lesssim (\mathbb{E}I_{11}^2)^{1/2}$. Thus, it suffices to compute $\mathbb{E}I_{11}^2$,

$$\begin{aligned}
\mathbb{E}I_{11}^2 &= \frac{1}{n}\mathbb{E}\Big(\frac{T_i}{\pi_1^2(z_i)}K^2\Big(\frac{x_i - c}{h}\Big)\Big) - \frac{1}{n}\Big[\mathbb{E}\Big(\frac{T_i}{\pi_1(z_i)}K\Big(\frac{x_i - c}{h}\Big)\Big)\Big]^2 \\
&\le \frac{1}{n}\int\int_{x>c}\frac{1}{\pi_1^2(z)}K^2\Big(\frac{x-c}{h}\Big)f_{X,Z(1)}(x,z)dxdz \\
&= \frac{h}{n}\int\int_{u>0}\frac{1}{\pi_1^2(z)}K^2(u)f_{X,Z(1)}(uh+c,z)dudz \\
&= O(\frac{h}{n}),
\end{aligned}$$

where the last step follows from the same argument in I_{21}. Putting them together into (3.2), we have $I_1 = O_p(h^2 + \frac{h^{1/2}}{n^{1/2}})$. For the last term I_2, we have

$$\begin{aligned}
I_2 &\le \sup_{z\in\mathcal{Z}}|\widehat{\pi}_1(z) - \pi_1(z)|\frac{1}{n}\sum_{i=1}^n \frac{T_i}{\pi_1(z_i)\widehat{\pi}_1(z_i)}K\Big(\frac{x_i - c}{h}\Big) \\
&\le O_p(r)\frac{1}{n}\sum_{i=1}^n \frac{T_i}{\pi_1(z_i)}K\Big(\frac{x_i - c}{h}\Big) = O_p(rh),
\end{aligned}$$

where the last step holds by the bound for the I_1 term. This implies (3.1). □

Lemma 8. Under assumption 4, 6 and 7, we have

$$\mathbb{E}\left((h_2)^{-1}K(\frac{z_i - z_j}{h_2})\Big|z_i\right) = f_Z(z_i) + O_p(h_2^2),$$

$$\mathbb{E}\left(2(h_1^2)^{-1}T_jK_1(\frac{c - x_j}{h_1}, \frac{z_i - z_j}{h_1})\Big|z_i, x_i\right) = f_{X,Z(1)}(c^+, z_i) + 2h_1\kappa_1\frac{\partial f_{X,Z(1)}(c^+, z_i)}{\partial x_i} + O_p(h_1^2),$$

$$\sup_{z_i}\Big|(nh_2)^{-1}\sum_{j=1}^n K(\frac{z_i - z_j}{h_2}) - f_Z(z_i)\Big| = O_p\left(\sqrt{\frac{\log n}{nh_2}} + h_2^2\right),$$

$$\sup_{z_i}\Big|2(nh_1^2)^{-1}\sum_{j=1}^n T_jK(\frac{c - x_j}{h_1}, \frac{z_i - z_j}{h_1}) - f_{X,Z(1)}(c^+, z_i)\Big| = O_p\left(\sqrt{\frac{\log n}{nh_1^2}} + h_1\right).$$

In case when $Z = Z(1) = Z(0)$,

$$\mathbb{E}\left((h_1^2)^{-1}T_jK_1(\frac{c - x_j}{h_1}, \frac{z_i - z_j}{h_1})\Big|z_i, x_i\right) = f_{X,Z}(c, z_i) + O_p(h_1^2),$$

$$\sup_{z_i}\Big|(nh_1^2)^{-1}\sum_{j=1}^n K(\frac{c - x_j}{h_1}, \frac{z_i - z_j}{h_1}) - f_{X,Z}(c, z_i)\Big| = O_p\left(\sqrt{\frac{\log n}{nh_1^2}} + h_1^2\right).$$

Proof. First, note that

$$\mathbb{P}(Z = z) = \mathbb{P}(Z = z, T = 1) + \mathbb{P}(Z = z, T = 0) = \mathbb{P}(Z(1) = z, X > c) + \mathbb{P}(Z(0) = z, X < c).$$

Then, we have

$$f_Z(z) = \int_{x>c} f_{X,Z(1)}(x,z)dx + \int_{x<c} f_{X,Z(0)}(x,z)dx,$$

which implies that $f_Z(z)$ is second order continuously differentiable by the continuously differentiable property of $f_{X,Z(1)}(x,z)$ and $f_{X,Z(0)}(x,z)$ in assumption 6. Thus, the standard calculation in nonparametric density estimation yields

$$\begin{aligned}\mathbb{E}\left((h_2)^{-1}K(\frac{z_i-z_j}{h_2})\Big|z_i\right) &= \frac{1}{h_2}\int K\left(\frac{z_i-z_j}{h_2}\right)f_Z(z_j)dz_j = \int K(u)\, f_Z(z_i+uh_2)du \\ &= f_Z(z_i)+O_p(h_2^2).\end{aligned}$$

To show the second result, following the similar argument, we get

$$\begin{aligned}&\mathbb{E}\left(2(h_{11}h_{12})^{-1}T_jK_1(\frac{c-x_j}{h_{11}},\frac{z_i-z_j}{h_{12}})\Big|z_i,x_i\right) = \frac{2}{h_{11}h_{12}}\int\int T_jK_1(\frac{c-x_j}{h_{11}},\frac{z_i-z_j}{h_{12}})f_{X,Z(1)}(x_j,z_j)dz_jdx_j \\ &= \frac{2}{h_{11}}\int T_jK(\frac{c-x_j}{h_{11}})\int K(v)f_{X,Z(1)}(x_j,z_i+vh_{12})dvdx_j \\ &= \frac{2}{h_{11}}\int T_jK(\frac{c-x_j}{h_{11}})\left(f_{X,Z(1)}(x_j,z_i)+Ch_{12}^2\right)dx_j \\ &= 2\int_{u>0}K(u)\left(f_{X,Z(1)}(c+uh_{11},z_i)+Ch_{12}^2\right)du \\ &= f_{X,Z(1)}(c^+,z_i)+2h_{11}\kappa_1\frac{\partial f_{X,Z(1)}(c^+,z_i)}{\partial x_i}+O(h_{12}^2+h_{11}^2),\end{aligned}$$

where C is a generic constant. The last two results follow from Fan (1993) together with the above bias calculation. □

3.3 Proof of Theorem 1

Since the local linear estimator is invariant to the scale of $\pi_1(z)$, we can simply take $\frac{1}{\pi_1(z)} = \frac{f_Z(z)}{f_{X,Z(1)}(c,z)}$ in the rest of the proof. It can be estimated by the following kernel estimator:

$$\frac{1}{\widehat{\pi}_1(z)} = \frac{\widehat{f}_Z(z)}{\widehat{f}_{X,Z(1)}(c,z)} = \frac{(nh_2)^{-1}\sum_{i=1}^n K(\frac{z-z_i}{h_2})}{2\cdot(nh_1^2)^{-1}\sum_{x_i>c}K_1(\frac{c-x_i}{h_1},\frac{z-z_i}{h_1})}.$$

Start with the following minimization problem:

$$\left(\widehat{\alpha}_1,\widehat{\beta}_1\right) = \arg\min_{\alpha,\beta}\sum_i \frac{T_i}{\widehat{\pi}_1(z_i)}\Big(y_i-\alpha-(x_i-c)\beta\Big)^2K\Big(\frac{x_i-c}{h}\Big). \tag{3.4}$$

Recall that for any kernel estimates $\widehat{f}=\widehat{f}(x)$ and $\widehat{g}=\widehat{g}(x)$, if f is bounded away from 0, then

$$\frac{1}{\widehat{f}} = \frac{1}{f}-\frac{1}{f^2}(\widehat{f}-f)+O_p(r^2), \tag{3.5}$$

where $\|\widehat{f}-f\|_\infty = O_p(r)$ and $\|\widehat{g}-g\|_\infty = O_p(s)$. Thus, if g is bounded from above, then

$$\begin{aligned}\frac{\widehat{g}}{\widehat{f}} &= \frac{g+(\widehat{g}-g)}{f}-\frac{g+(\widehat{g}-g)}{f^2}(\widehat{f}-f)+O_p(r^2) \\ &= \frac{g}{f}-\frac{g}{f^2}(\widehat{f}-f)+\frac{\widehat{g}-g}{f}+O_p(r^2+rs) \\ &= \frac{g}{f}-\frac{g\widehat{f}}{f^2}+\frac{\widehat{g}}{f}+O_p(r^2+rs).\end{aligned}$$

Following the above discussion, we can show that

$$
\begin{aligned}
&\frac{1}{2\cdot(nh_1^2)^{-1}\sum_{j=1}^n T_jK_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})}\\
&= -\frac{1}{f_{X,Z(1)}(c^+,z_i)^2}\left(2(nh_1^2)^{-1}\sum_{j=1}^n T_jK_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1}) - f_{X,Z(1)}(c^+,z_i)\right)\\
&+\frac{1}{f_{X,Z(1)}(c^+,z_i)} + O_p\left(r^2\right),
\end{aligned}
$$

where Lemma 8 implies

$$
r = \sqrt{\frac{\log n}{nh_1^2}} + h_1.
$$

The gradient of (3.4) can be written as the following U-statistic:

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \frac{T_i\widehat{f}_Z(z_i)}{\widehat{f}_{X,Z(1)}(c,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i = \frac{1}{n}\sum_{i=1}^n\frac{T_i(nh_2)^{-1}\sum_{j=1}^n K(\frac{z_i-z_j}{h_2})}{2\cdot(nh_1^2)^{-1}\sum_{j=1}^n T_jK_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})}K\Big(\frac{x_i-c}{h}\Big)R_i\\
&= \frac{1}{n}\sum_{i=1}^n \frac{T_if_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i + \frac{T_i}{f_{X,Z(1)}(c^+,z_i)}\left((nh_2)^{-1}\sum_{j=1}^n K(\frac{z_i-z_j}{h_2})\right)K\Big(\frac{x_i-c}{h}\Big)R_i\\
&- \frac{T_if_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)^2}\left(2\cdot(nh_1^2)^{-1}\sum_{j=1}^n T_jK_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})\right)K\Big(\frac{x_i-c}{h}\Big)R_i + O_p\left(r^2+rs\right)\\
&= \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\Big\{\frac{T_i}{f_{X,Z(1)}(c^+,z_i)}(h_2)^{-1}K(\frac{z_i-z_j}{h_2})K\Big(\frac{x_i-c}{h}\Big)R_i\\
&- 2\cdot\frac{T_iT_jf_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)^2}(h_1^2)^{-1}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K\Big(\frac{x_i-c}{h}\Big)R_i + \frac{T_if_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i\Big\} + O_p\left(r^2+rs\right),
\end{aligned}
$$

where

$$
s = \sqrt{\frac{\log n}{nh_2}} + h_2^2,
$$

implied by Lemma 8. Define

$$
\begin{aligned}
\phi_{i,j} &= \frac{T_i}{f_{X,Z(1)}(c^+,z_i)}(h_2)^{-1}K(\frac{z_i-z_j}{h_2})K\Big(\frac{x_i-c}{h}\Big)R_i\\
&\quad - 2\cdot\frac{T_iT_jf_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)^2}(h_1^2)^{-1}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K\Big(\frac{x_i-c}{h}\Big)R_i + \frac{T_if_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^n\frac{T_i\widehat{f}_Z(z)}{\widehat{f}_{X|Z(1)}(c,z)}K\Big(\frac{x_i-c}{h}\Big)R_i &= \frac{1}{n^2}\sum_i\sum_j\phi_{i,j} + O_p\left(r^2+rs\right)\\
&= \frac{1}{2n^2}\sum_i\sum_j(\phi_{i,j}+\phi_{j,i}) + O_p\left(r^2+rs\right) \qquad (3.6)\\
&= \frac{1}{n^2}\sum_{i<j}(\phi_{i,j}+\phi_{j,i}) + \frac{1}{n^2}\sum_i\phi_{i,i} + O_p\left(r^2+rs\right),
\end{aligned}
$$

where the first (leading) term is a U-statistic after rescaling. By lemma 12 and Theorem 12.3 in Van der Vaart (2000), we have

$$
\frac{n^{1/2}}{h^{1/2}}\Big(\frac{1}{n(n-1)}\sum_{i<j}(\phi_{i,j}+\phi_{j,i}) - \delta/2 - \frac{1}{n}\sum_{i=1}^n\{\mathbb{E}(\phi_{i,j}+\phi_{j,i}|i)-\delta\}\Big) = o_p(1), \qquad (3.7)
$$

where $\delta = \mathbb{E}(\phi_{i,j} + \phi_{j,i})$, and we use $\mathbb{E}(\cdot|i)$ to denote the conditional expectation given the ith sample. In the following, we approximate $\mathbb{E}(\phi_{i,j} + \phi_{j,i}|i)$. Define $d(x_i, z_i) = m(x_i, z_i) - \alpha_1$. By Lemma 10, we have

$$\mathbb{E}(\phi_{i,j}) = O_p(h^3).$$

The central limit theorem implies

$$\frac{1}{(nh)^{1/2}} \sum_{i=1}^{n} \{\mathbb{E}(\phi_{i,j} + \phi_{j,i}|i) - \delta\} + O_p(\frac{n^{1/2}}{h^{1/2}}(h^3)) \to_d N(0, \xi^2/h), \tag{3.8}$$

where $\delta = O_p(h^3)$ and $\xi^2 = \mathbb{E}\{(\mathbb{E}(\phi_{i,j} + \phi_{j,i}|i) - \delta)^2\} = \mathbb{E}\left\{\mathbb{E}(\phi_{ij}|i)^2 + \mathbb{E}(\phi_{ji}|i)^2 + 2\mathbb{E}(\phi_{ij}|i)\mathbb{E}(\phi_{ji}|i)\right\} - \delta^2$. We now calculate the asymptotic variance as follows. Since $\mathbb{E}(\phi_{j,i}|i) = \mathbb{E}(\phi_{i,j}|j)$, from lemma 10, we have

$$\mathbb{E}(\phi_{ij}|j) = \frac{h}{2} d_1(c^+, z_j) - T_j \frac{h}{h_1} K(\frac{c - x_j}{h_1}) \frac{f_Z(z_j)}{f_{X,Z(1)}(c^+, z_j)} d_1(c^+, z_j) + O_p(h^2 + hh_1 + hh_2).$$

Similarly, we can show that

$$\mathbb{E}(\phi_{i,j}|i) = \frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)} K\Big(\frac{x_i - c}{h}\Big) R_i + O_p(h_2^2 + h_1).$$

Recall that $\sigma^2 = \mathbb{E}(Y(1) - m_1(X, Z(1)))^2$. Since $h \asymp \sqrt{h_1} \asymp h_2$, after some tedious calculation we can show that

$$\begin{aligned}
\frac{1}{n}\mathbb{E}\left(\mathbb{E}(\phi_{ji}|i)^2\right) &= \frac{1}{n}\frac{h^2}{h_1^2} \int_{z_i}\int_{x_i} T_i K(\frac{c - x_i}{h_1})^2 \frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+, z_i)^2} d(c^+, z_i)^2 f_{X,Z(1)}(x_i, z_i) dx_i dz_i + O\left(\frac{h^2}{n}\right) \\
&= \frac{h^2}{h_1}\frac{\kappa_{20}}{n} \int_{z_i} \frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+, z_i)} d(c^+, z_i)^2 dz_i + O\left(\frac{h^2}{n}\right) \\
&= \frac{h^2}{h_1}\frac{\kappa_{20}}{n} \mathbb{E}_Z\left(\frac{f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)} d(c^+, z_i)^2\right) + O\left(\frac{h^2}{n}\right),
\end{aligned}$$

And

$$\xi^2 = \kappa_{20} \underbrace{\mathbb{E}_Z\left(\frac{f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)} d_1(c^+, z_i)^2\right)}_{\omega} + O(h).$$

Combining (3.7) and (3.9),

$$n^{1/2}\Big(\frac{1}{n(n-1)} \sum_{i<j} (\phi_{i,j} + \phi_{j,i}) - \delta/2\Big) + O\Big(n^{1/2} h^3\Big) \to_d N(0, \xi^2).$$

Finally, note that in (3.6),

$$\frac{1}{n^2} \sum_{i=1}^{n} \phi_{i,i} \lesssim \frac{1}{n}\mathbb{E}(\phi_{i,j}) = O_p\left(\frac{h^2}{n}\right),$$

and therefore we obtain that

$$\frac{1}{n^{1/2}} \sum_{i=1}^{n} \frac{T_i \widehat{f}_Z(z)}{\widehat{f}_{X,Z(1)}(c, z)} K\Big(\frac{x_i - c}{h}\Big) R_i + \phi \to_d N(n^{1/2}\delta/2, \xi^2),$$

where

$$\phi = O_p\Big(n^{1/2}(h^3 + r^2 + rs)\Big).$$

Following the similar argument, we can show the joint convergence

$$\frac{1}{n^{1/2}} \sum_{i=1}^{n} \frac{T_i \widehat{f}_Z(z)}{\widehat{f}_{X,Z(1)}(c, z)} K\Big(\frac{x_i - c}{h}\Big) R_i [1, (x_i - c)]^T \to_d N\left(n^{1/2}\begin{pmatrix} O_p(h^3) \\ O_p(h^4)\end{pmatrix}, \omega \begin{pmatrix} \kappa_{20} & h\kappa_{21} \\ h\kappa_{21} & h^2\kappa_{22}\end{pmatrix}\right).$$

By the least squared formulation, the estimator $\widehat{\alpha}_1$ satisfies

$$\sqrt{nh^2}(\widehat{\alpha}_1-\alpha_1) = -e_1^T \begin{pmatrix} \widehat{C}_n/(nh) & -\widehat{B}_n/(nh) \\ -\widehat{B}_n/(nh) & \widehat{A}_n/(nh) \end{pmatrix}^{-1} \frac{1}{n^{1/2}} \sum_{i=1}^{n} \frac{T_i \widehat{f}_Z(z)}{\widehat{f}_{X,Z(1)}(c,z)} K\Big(\frac{x_i-c}{h}\Big) R_i[1,(x_i-c)]^T,$$

where $e_1^T=(1,0)$. From lemma 7 and the matrix inversion formula,

$$\begin{pmatrix} \widehat{C}_n/(nh) & -\widehat{B}_n/(nh) \\ -\widehat{B}_n/(nh) & \widehat{A}_n/(nh) \end{pmatrix}^{-1} = \frac{1}{\widehat{D}_n/(nh)^2} \begin{pmatrix} \widehat{A}_n/(nh) & \widehat{B}_n/(nh) \\ \widehat{B}_n/(nh) & \widehat{C}_n/(nh) \end{pmatrix} \to_p \frac{1}{h^2(\kappa_2/2-\kappa_1^2)} \begin{pmatrix} h^2\kappa_2 & -h\kappa_1 \\ -h\kappa_1 & \frac{1}{2} \end{pmatrix}.$$

Thus, the asymptotic bias of $\sqrt{nh^2}(\widehat{\alpha}_1-\alpha_1)$ is

$$\frac{-e_1^T}{h^2(\kappa_2/2-\kappa_1^2)} \begin{pmatrix} h^2\kappa_2 & -h\kappa_1 \\ -h\kappa_1 & \frac{1}{2} \end{pmatrix} n^{1/2} \begin{pmatrix} O_p(h^3) \\ O_p(h^4) \end{pmatrix} = O(n^{1/2}h^3) = o(1).$$

Similarly, the asymptotic variance of $\sqrt{nh^2}(\widehat{\alpha}_1-\alpha_1)$ is

$$\begin{aligned} &\frac{\omega}{h^4(\kappa_2/2-\kappa_1^2)^2} e_1^T \begin{pmatrix} h^2\kappa_2 & -h\kappa_1 \\ -h\kappa_1 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \kappa_{20} & h\kappa_{21} \\ h\kappa_{21} & h^2\kappa_{22} \end{pmatrix} \begin{pmatrix} h^2\kappa_2 & -h\kappa_1 \\ -h\kappa_1 & \frac{1}{2} \end{pmatrix} e_1 \\ &= \mathbb{E}_Z\left(\frac{f_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)} d(c^+,z_i)^2 \right) \cdot C_v, \end{aligned}$$

where

$$C_v = \frac{\kappa_2^2\kappa_{20}+\kappa_1^2\kappa_{22}-2\kappa_1\kappa_2\kappa_{21}}{\left(\frac{1}{2}\kappa_2-\kappa_1^2\right)^2}.$$

This completes the proof.

Lemma 9. Under the same condition as in Theorem 1,

$$\begin{aligned} M(c,z_i) &:= \frac{1}{h}\int T_i K\left(\frac{x_i-c}{h}\right) \mathbb{E}(R_i|x_i,z_i) f_{X,Z(1)}(x_i,z_i)dx_i \\ &= \frac{1}{2} m_1(c^+,z_i) f_{X,Z(1)}(c^+,z_i) + h\kappa_1 \frac{\partial m_1(c^+,z_i)}{\partial x_i} f_{X,Z(1)}(c^+,z_i) + h\kappa_1 m_1(c^+,z_i) \frac{\partial f_{X,Z(1)}(c^+,z_i)}{\partial x_i} \\ &- \frac{1}{2}\alpha_1 f_{X,Z(1)}(c^+,z_i) - h\kappa_1\beta_1 f_{X,Z(1)}(c^+,z_i) - h\kappa_1\alpha_1 \frac{\partial f_{X,Z(1)}(c^+,z_i)}{\partial x_i} + O_p(h^2), \end{aligned}$$

where O_p terms are valid uniformly over i.

Proof. Following the standard Taylor expansion, we can show that

$$\begin{aligned} M(c,z_i) &= \frac{1}{h}\int T_i K\left(\frac{x_i-c}{h}\right) \mathbb{E}(R_i|x_i,z_i) f_{X,Z(1)}(x_i,z_i)dx_i \\ &= \int_{u>0} K(u)\left(m_1(c+uh,z_i)-\alpha_1-uh\beta_1\right) f_{X,Z(1)}(c+uh,z_i)du \\ &= \int_{u>0} K(u)\left(m_1(c^+,z_i) + \frac{\partial m_1(c^+,z_i)}{\partial x_i}uh + \frac{\partial^2 m_1(c^+,z_i)}{\partial x_i^2}\frac{u^2h^2}{2} + \frac{\partial^3 m_1(\widetilde{c}^{(2)},z_i)}{\partial x_i^3}\frac{u^3h^3}{3!} - \alpha_1 - uh\beta_1 \right) \\ &\left(f_{X,Z(1)}(c^+,z_i) + \frac{\partial f_{X,Z(1)}(c^+,z_i)}{\partial x_i}uh + \frac{\partial^2 f_{X,Z(1)}(c^+,z_i)}{\partial x_i^2}\frac{u^2h^2}{2} + \frac{\partial^3 f_{X,Z(1)}(\widetilde{c}^{(3)},z_i)}{\partial x_i^3}\frac{u^3h^3}{3!} \right) du \\ &= \frac{1}{2}(m_1(c^+,z_i)-\alpha_1) f_{X,Z(1)}(c^+,z_i) + h\kappa_1\left(\frac{\partial m_1(c^+,z_i)}{\partial x_i} - \beta_1\right) f_{X,Z(1)}(c^+,z_i) \\ &+ h\kappa_1(m_1(c^+,z_i)-\alpha_1)\frac{\partial f_{X,Z(1)}(c^+,z_i)}{\partial x_i} + O_p(h^2), \end{aligned}$$

where O_p terms are valid uniformly over i as the (mixed) third derivatives of $f_{X,Z(1)}(x_i,z_i)$ are all bounded. □

Lemma 10. Recall that

$$\phi_{i,j} = \frac{T_i}{f_{X,Z(1)}(c^+, z_i)}(h_2)^{-1}K(\frac{z_i - z_j}{h_2})K\Big(\frac{x_i - c}{h}\Big)R_i$$
$$- 2 \cdot \frac{T_i T_j f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)^2}(h_1^2)^{-1}K_1(\frac{c - x_j}{h_1}, \frac{z_i - z_j}{h_1})K\Big(\frac{x_i - c}{h}\Big)R_i + \frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)}K\Big(\frac{x_i - c}{h}\Big)R_i.$$

Under the same condition as in Theorem 1, and when $h_1 = h^2$

$$\mathbb{E}(\phi_{i,j}|j) = \frac{h}{2}d(c^+, z_j) - T_j \frac{h}{h_1}K(\frac{c - x_j}{h_1})\frac{f_Z(z_j)}{f_{X,Z(1)}(c^+, z_j)}d_1(c^+, z_j) + O_p(h^2 + hh_1 + hh_2),$$
$$\mathbb{E}(\phi_{i,j}|i) = \frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)}K\Big(\frac{x_i - c}{h}\Big)R_i + O_p(h_2^2 + h_1),$$
$$\mathbb{E}(\phi_{i,j}) = O_p(h^3)$$

where O_p terms are valid uniformly over i or j.

Proof.

$$\mathbb{E}(\phi_{i,j}|j) = \underbrace{\mathbb{E}\left(\frac{T_i}{f_{X,Z(1)}(c^+, z_i)}(h_2)^{-1}K(\frac{z_i - z_j}{h_2})K\Big(\frac{x_i - c}{h}\Big)R_i\Big|z_j\right)}_{\text{(Part.BI)}}$$
$$- 2T_j \cdot \underbrace{\mathbb{E}\left(\frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)^2}(h_1^2)^{-1}K_1(\frac{c - x_j}{h_1}, \frac{z_i - z_j}{h_1})K\Big(\frac{x_i - c}{h}\Big)R_i\Big|z_j, x_j\right)}_{\text{(Part.BII)}}$$
$$+ \underbrace{\mathbb{E}\left(\frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)}K\Big(\frac{x_i - c}{h}\Big)R_i\Big|z_j, x_j\right)}_{\text{(Part.BIII)}}.$$

From lemma 9,

$$\begin{aligned}
\text{(Part.BI)} &= \mathbb{E}\left(\frac{T_i}{f_{X,Z(1)}(c^+, z_i)}(h_2)^{-1}K(\frac{z_i - z_j}{h_2})K\Big(\frac{x_i - c}{h}\Big)R_i\Big|z_j\right) \\
&= h\int_{z_i} \frac{1}{f_{X,Z(1)}(c^+, z_i)}(h_2)^{-1}K(\frac{z_i - z_j}{h_2})M(c, z_i)dz_i \\
&= h\int_{z_i} (h_2)^{-1}K(\frac{z_i - z_j}{h_2})\frac{1}{2}(m_1(c^+, z_i) - \alpha_1)dz_i + O_p(h^2) \\
&= \frac{h}{2}(m_1(c^+, z_j) - \alpha_1) + O_p(h^2 + h_2 h),
\end{aligned}$$

where O_p terms are valid uniformly over j. Similarly, for Part.BII and Part.BIII, we can show that

$$\text{(Part.BII)} = \mathbb{E}\left(\frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)^2}(h_1^2)^{-1}K_1(\frac{c - x_j}{h_1}, \frac{z_i - z_j}{h_1})K\Big(\frac{x_i - c}{h}\Big)R_i\Big|z_j, x_j\right)$$
$$= h\int_{z_i} \frac{f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)^2}(h_1^2)^{-1}K_1(\frac{c - x_j}{h_1}, \frac{z_i - z_j}{h_1})M(c, z_i)dz_i$$
$$= \frac{h}{2}\int_{z_i} \frac{f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)}(h_1^2)^{-1}K_1(\frac{c - x_j}{h_1}, \frac{z_i - z_j}{h_1})(m_1(c^+, z_i) - \alpha_1)dz_i + O_p(h^2)$$
$$= \frac{h}{2h_1}K(\frac{c - x_j}{h_1})\frac{f_Z(z_j)}{f_{X,Z(1)}(c^+, z_j)}(m_1(c^+, z_j) - \alpha_1) + O_p(hh_1) + O_p(h^2),$$

$$\text{(Part.BIII)} = \mathbb{E}\left(\frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)}K\Big(\frac{x_i - c}{h}\Big)R_i\Big|z_j, x_j\right) = \frac{h}{2}\int_{z_i} (m_1(c^+, z_i) - \alpha_1)f_Z(z_i)dz_i + O_p(h^2) = O_p(h^2),$$

where the last step follows from the definition of α_1. Define $d_1(x_i, z_i) = m_1(x_i, z_i) - \alpha_1$. Combining the Part BI, BII and BIII, we obtain

$$\mathbb{E}(\phi_{ij}|j) = \frac{h}{2} d_1(c^+, z_j) - T_j \frac{h}{h_1} K(\frac{c - x_j}{h_1}) \frac{f_Z(z_j)}{f_{X,Z(1)}(c^+, z_j)} d_1(c^+, z_j) + O_p(h^2 + hh_1 + hh_2).$$

Following the similar calculation, by lemma 8 we have

$$\begin{aligned}
\mathbb{E}(\phi_{i,j}|i) = & \underbrace{\frac{T_i}{f_{X,Z(1)}(c^+, z_i)} \mathbb{E}\left((h_2)^{-1} K(\frac{z_i - z_j}{h_2}) \Big| z_i \right) K\Big(\frac{x_i - c}{h}\Big) R_i}_{\text{Part. AI}} \\
& \underbrace{- 2 \cdot \frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)^2} \mathbb{E}\left((h_1^2)^{-1} T_j K_1(\frac{c - x_j}{h_1}, \frac{z_i - z_j}{h_1}) \Big| z_i, x_i \right) K\Big(\frac{x_i - c}{h}\Big) R_i}_{\text{Part. AII}} \\
& \underbrace{+ \frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)} K\Big(\frac{x_i - c}{h}\Big) R_i}_{\text{Part. AIII}} \\
& = \frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)} K\Big(\frac{x_i - c}{h}\Big) R_i + O_p(h_1 + h_2^2).
\end{aligned}$$

Finally, we calculate $\mathbb{E}(\phi_{i,j})$ using Lemma 9,

$$\begin{aligned}
\mathbb{E}(\text{ Part.AIII}) &= h \int \frac{f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)} M(c, z_i) dz_i \\
&= \frac{h}{2} \int (m_1(c^+, z_i) - \alpha_1) f_Z(z_i) dz_i - h^2 \kappa_1 \beta_1 - h^2 \kappa_1 \alpha_1 \int \frac{\partial f_{X,Z(1)}(c^+, z_i)}{\partial x_i} \frac{f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)} dz_i \\
&+ h^2 \kappa_1 \int \left(\frac{\partial m_1(c^+, z_i)}{\partial x_i} f_Z(z_i) + m_1(c^+, z_i) \frac{\partial f_{X,Z(1)}(c^+, z_i)}{\partial x_i} \frac{f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)} \right) dz_i + O_p(h^3) \\
&= -h^2 \kappa_1 \alpha_1 \int \frac{\partial f_{X,Z(1)}(c^+, z_i)}{\partial x_i} \frac{f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)} dz_i + O_p(h^3)
\end{aligned}$$

where the O_p terms are valid uniformly over i and the last equality follows as

$$\alpha_1 = \int m_1(c^+, z) f_Z(z) dz,$$

and

$$\beta_1 = \int \frac{\partial m_1(c^+, z_i)}{\partial x_i} f_Z(z_i) dz_i + \int m_1(c^+, z_i) \frac{\partial f_{X,Z(1)}(c^+, z_i)}{\partial x_i} \frac{f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)} dz_i.$$

From Lemma 8, some tedious calculation implies

$$\mathbb{E}(\text{Part.AI}) = \mathbb{E}(\text{Part.AIII}) + O_p(hh_2^2),$$

and

$$\mathbb{E}(\text{Part.AII}) = \mathbb{E}(\text{Part.AIII}) + 2h_1 \kappa_1 \mathbb{E}\left(\frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)^2} \frac{\partial f_{X,Z(1)}(c^+, z_i)}{\partial x_i} K\Big(\frac{x_i - c}{h}\Big) R_i \right) + O_p(hh_1^2).$$

From Lemma 8 and Lemma 9,

$$\begin{aligned}
& h_1 \kappa_1 \mathbb{E}\left(\frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)^2} \frac{\partial f_{X,Z(1)}(c^+, z_i)}{\partial x_i} K\Big(\frac{x_i - c}{h}\Big) R_i \right) \\
&= h_1 \kappa_1 \int \int \frac{T_i f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)^2} \frac{\partial f_{X,Z(1)}(c^+, z_i)}{\partial x_i} K\Big(\frac{x_i - c}{h}\Big) \mathbb{E}(R_i | x_i, z_i) f_{X,Z(1)}(x_i, z_i) dx_i dz_i \\
&= \frac{hh_1 \kappa_1}{2} \int \frac{f_Z(z_i)}{f_{X,Z(1)}(c^+, z_i)} \frac{\partial f_{X,Z(1)}(c^+, z_i)}{\partial x_i} (m_1(c^+, z_i) - \alpha_1) dz_i + O_p(h_1 h^2).
\end{aligned}$$

Thus when $h_1 \asymp h^2$

$$\mathbb{E}(\phi_{i,j}) = O_p(h^3).$$

This completes the proof. $\square$

3.4 Proof of Corollary 1

In the case when $Z = Z(1) = Z(0)$, we can estimate $\frac{1}{\pi_1(z)} = \frac{f_Z(z)}{f_{X,Z}(c,z)}$ by the following kernel estimator:

$$\frac{1}{\widehat{\pi}_1(z)} = \frac{\widehat{f}_Z(z)}{\widehat{f}_{X,Z}(c,z)} = \frac{(nh_2)^{-1}\sum_{i=1}^n K(\frac{z-z_i}{h_2})}{(nh_1^2)^{-1}\sum_{i=1}^n K_1(\frac{c-x_i}{h_1}, \frac{z-z_i}{h_1})}.$$

Similar as in theorem 1,

$$\begin{aligned}
&\frac{1}{(nh_1^2)^{-1}\sum_{j=1}^n K_1(\frac{c-x_j}{h_1}, \frac{z_i-z_j}{h_1})} \\
&= -\frac{1}{f_{X,Z}(c,z_i)^2}\left((nh_1^2)^{-1}\sum_{j=1}^n K_1(\frac{c-x_j}{h_1}, \frac{z_i-z_j}{h_1}) - f_{X,Z}(c,z_i)\right) \\
&+ \frac{1}{f_{X,Z}(c,z_i)} + O_p\left(r^2\right),
\end{aligned}$$

where Lemma 8 implies

$$r = \sqrt{\frac{\log n}{nh_1^2}} + h_1^2.$$

The gradient in (3.4) can be written as the following U-statistic:

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \frac{T_i\widehat{f}_Z(z_i)}{\widehat{f}_{X,Z}(c,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i = \frac{1}{n}\sum_{i=1}^n \frac{T_i(nh_2)^{-1}\sum_{j=1}^n K(\frac{z_i-z_j}{h_2})}{(nh_1^2)^{-1}\sum_{j=1}^n K_1(\frac{c-x_j}{h_1}, \frac{z_i-z_j}{h_1})}K\Big(\frac{x_i-c}{h}\Big)R_i \\
&= \frac{1}{n}\sum_{i=1}^n \frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i + \frac{T_i}{f_{X,Z}(c,z_i)}\left((nh_2)^{-1}\sum_{j=1}^n K(\frac{z_i-z_j}{h_2})\right)K\Big(\frac{x_i-c}{h}\Big)R_i \\
&- \frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)^2}\left((nh_1^2)^{-1}\sum_{j=1}^n K_1(\frac{c-x_j}{h_1}, \frac{z_i-z_j}{h_1})\right)K\Big(\frac{x_i-c}{h}\Big)R_i + O_p\left(r^2+rs\right) \\
&= \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \Big\{\frac{T_i}{f_{X,Z}(c,z_i)}(h_2)^{-1}K(\frac{z_i-z_j}{h_2})K\Big(\frac{x_i-c}{h}\Big)R_i \\
&- \frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)^2}(h_1^2)^{-1}K_1(\frac{c-x_j}{h_1}, \frac{z_i-z_j}{h_1})K\Big(\frac{x_i-c}{h}\Big)R_i + \frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i\Big\} + O_p\left(r^2+rs\right),
\end{aligned}$$

where

$$s = \sqrt{\frac{\log n}{nh_2}} + h_2^2,$$

Define

$$\begin{aligned}
\phi_{i,j}^{rd} &= \frac{T_i}{f_{X,Z}(c,z_i)}(h_2)^{-1}K(\frac{z_i-z_j}{h_2})K\Big(\frac{x_i-c}{h}\Big)R_i \\
&- \frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)^2}(h_1^2)^{-1}K_1(\frac{c-x_j}{h_1}, \frac{z_i-z_j}{h_1})K\Big(\frac{x_i-c}{h}\Big)R_i + \frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i.
\end{aligned}$$

By Lemma 11, we have

$$\mathbb{E}(\phi_{i,j}^{rd}) = O_p(h^3).$$

The central limit theorem implies

$$\frac{1}{(nh)^{1/2}}\sum_{i=1}^{n}\{\mathbb{E}(\phi_{i,j}^{rd}+\phi_{j,i}^{rd}|i)-\delta^{rd}\}+O_p(\frac{n^{1/2}}{h^{1/2}}(h^3+h^2h_1))\to_d N(0,\xi^2/h), \tag{3.9}$$

where $\delta^{rd}=O_p(h^3)$ and $\xi^2=\mathbb{E}\{(\mathbb{E}(\phi_{i,j}^{rd}+\phi_{j,i}^{rd}|i)-\delta)^2\}=\mathbb{E}\left\{\mathbb{E}(\phi_{ij}|i)^2+\mathbb{E}(\phi_{ji}|i)^2+2\mathbb{E}(\phi_{ij}|i)\mathbb{E}(\phi_{ji}|i)\right\}-\delta^2$. We now calculate the asymptotic variance as follows. Since $\mathbb{E}(\phi_{j,i}|i)=\mathbb{E}(\phi_{i,j}|j)$, from lemma 10, we have

$$\mathbb{E}(\phi_{ij}|j)=\frac{h}{2}d_1(c^+,z_j)-T_j\frac{h}{h_1}K(\frac{c-x_j}{h_1})\frac{f_Z(z_j)}{f_{X,Z}(c,z_j)}d_1(c^+,z_j)+O_p(h^2+hh_1+hh_2).$$

Similarly, we can show that

$$\mathbb{E}(\phi_{i,j}|i)=\frac{T_if_Z(z_i)}{f_{X,Z}(c,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i+O_p(h_2^2+h_1).$$

Recall that $\sigma^2=\mathbb{E}(Y(1)-m_1(X,Z))^2$. Since $h\asymp h_1\asymp h_2$, after some tedious calculation we can show that

$$\begin{aligned}
\frac{1}{n}\mathbb{E}\left(\mathbb{E}(\phi_{ij}|i)^2\right)&=\frac{\sigma^2}{n}\int_{z_i}\int_{x_i}\frac{T_if_Z(z_i)^2}{f_{X,Z}(c,z_i)^2}K\Big(\frac{x_i-c}{h}\Big)^2f_{X,Z}(x_i,z_i)dx_idz_i\\
&\quad+\frac{1}{n}\int_{z_i}\int_{x_i}\frac{T_if_Z(z_i)^2}{f_{X,Z}(c,z_i)^2}K\Big(\frac{x_i-c}{h}\Big)^2f_{X,Z}(x_i,z_i)(m_1(x_i,z_i)-\alpha_1)^2dx_idz_i+O\left(\frac{h^2}{n}\right)\\
&=\frac{\sigma^2}{n}h\kappa_{20}\int_{z_i}\frac{f_Z(z_i)^2}{f_{X,Z}(c,z_i)}dz_i+\frac{h}{n}\kappa_{20}\int_{z_i}\frac{f_Z(z_i)^2}{f_{X,Z}(c,z_i)}d_1(c^+,z_i)^2dz_i+O\left(\frac{h^2}{n}\right)\\
&=\kappa_{20}\frac{h}{n}\sigma^2\mathbb{E}_Z\left(\frac{f_Z(z_i)}{f_{X,Z}(c,z_i)}\right)+\kappa_{20}\frac{h}{n}\mathbb{E}_Z\left(\frac{f_Z(z_i)}{f_{X,Z}(c,z_i)}d_1(c^+,z_i)^2\right)+O\left(\frac{h^2}{n}\right),\\
\frac{1}{n}\mathbb{E}\left(\mathbb{E}(\phi_{ji}|i)^2\right)&=\frac{1}{n}\frac{h^2}{h_1^2}\int_{z_i}\int_{x_i}T_iK(\frac{c-x_i}{h_1})^2\frac{f_Z(z_i)^2}{f_{X,Z}(c^+,z_i)^2}d_1(c^+,z_i)^2f_{X,Z}(x_i,z_i)dx_idz_i+O\left(\frac{h^2}{n}\right)\\
&=\frac{h^2}{h_1}\frac{\kappa_{20}}{n}\int_{z_i}\frac{f_Z(z_i)^2}{f_{X,Z}(c,z_i)}d_1(c^+,z_i)^2dz_i+O\left(\frac{h^2}{n}\right)\\
&=\frac{h^2}{h_1}\frac{\kappa_{20}}{n}\mathbb{E}_Z\left(\frac{f_Z(z_i)}{f_{X,Z}(c,z_i)}d_1(c^+,z_i)^2\right)+O\left(\frac{h^2}{n}\right),\\
\frac{2}{n}\mathbb{E}\left(\mathbb{E}(\phi_{ij}|i)\mathbb{E}(\phi_{ji}|i)\right)&=-2\cdot\frac{h}{h_1n}\int_{z_i}\int_{x_i}T_iK(\frac{c-x_j}{h_1})^2\frac{f_Z(z_j)^2}{f_{X,Z}(c,z_j)^2}d_1(c^+,z_j)^2f_{X,Z}(x_i,z_i)dx_idz_i+O\left(\frac{h^2}{n}\right)\\
&=-2\cdot\kappa_{20}\frac{h}{n}\int_{z_i}\frac{f_Z(z_i)^2}{f_{X,Z}(c,z_i)}d_1(c^+,z_j)^2dz_i+O\left(\frac{h^2}{n}\right)\\
&=-2\cdot\kappa_{20}\frac{h}{n}\mathbb{E}_Z\left(\frac{f_Z(z_i)}{f_{X,Z}(c|z_i)}d_1(c^+,z_i)^2\right)+O\left(\frac{h^2}{n}\right).
\end{aligned}$$

Thus, choosing $h_1=h$, we have

$$\xi^2=h\kappa_{20}\underbrace{\sigma^2\mathbb{E}_Z\left(\frac{f_Z(z_i)}{f_{X,Z}(c,z_i)}\right)}_{\omega}+O(h^2).$$

Combining (3.7) and (3.9),

$$\frac{n^{1/2}}{h^{1/2}}\Big(\frac{1}{n(n-1)}\sum_{i<j}(\phi_{i,j}+\phi_{j,i})-\delta/2\Big)+O\Big(\frac{n^{1/2}}{h^{1/2}}(h^3+h^2h_1)\Big)\to_d N(0,\xi^2/h).$$

Finally, note that in (3.6),

$$\frac{1}{n^2}\sum_{i=1}^{n}\phi_{i,i}\lesssim\frac{1}{n}\mathbb{E}(\phi_{i,j})=O_p\left(\frac{h^2}{n}\right),$$

and therefore we obtain that

$$\frac{1}{(nh)^{1/2}}\sum_{i=1}^{n}\frac{T_i\widehat{f}_Z(z)}{\widehat{f}_{X,Z}(c,z)}K\Big(\frac{x_i-c}{h}\Big)R_i+\phi\rightarrow_d N(\frac{n^{1/2}\delta}{2h^{1/2}},\xi^2/h),$$

where

$$\phi=O_p\Big(\frac{n^{1/2}}{h^{1/2}}(h^3+r^2+rs)\Big).$$

Following the similar argument, we can show the joint convergence

$$\frac{1}{(nh)^{1/2}}\sum_{i=1}^{n}\frac{T_i\widehat{f}_Z(z)}{\widehat{f}_{X,Z}(c,z)}K\Big(\frac{x_i-c}{h}\Big)R_i[1,(x_i-c)]^T\rightarrow_d N\left(\frac{n^{1/2}}{h^{1/2}}\begin{pmatrix}O_p(h^3)\\O_p(h^4)\end{pmatrix},\omega\begin{pmatrix}\kappa_{20}&h\kappa_{21}\\h\kappa_{21}&h^2\kappa_{22}\end{pmatrix}\right).$$

By the least squared formulation, the estimator $\widehat{\alpha}_1$ satisfies

$$\sqrt{nh}(\widehat{\alpha}_1-\alpha_1)=-e_1^T\begin{pmatrix}\widehat{C}_n/(nh)&-\widehat{B}_n/(nh)\\-\widehat{B}_n/(nh)&\widehat{A}_n/(nh)\end{pmatrix}^{-1}\frac{1}{(nh)^{1/2}}\sum_{i=1}^{n}\frac{T_i\widehat{f}_Z(z)}{\widehat{f}_{X,Z(1)}(c,z)}K\Big(\frac{x_i-c}{h}\Big)R_i[1,(x_i-c)]^T,$$

where $e_1^T=(1,0)$. From lemma 7 and the matrix inversion formula,

$$\begin{pmatrix}\widehat{C}_n/(nh)&-\widehat{B}_n/(nh)\\-\widehat{B}_n/(nh)&\widehat{A}_n/(nh)\end{pmatrix}^{-1}=\frac{1}{\widehat{D}_n/(nh)^2}\begin{pmatrix}\widehat{A}_n/(nh)&\widehat{B}_n/(nh)\\\widehat{B}_n/(nh)&\widehat{C}_n/(nh)\end{pmatrix}\rightarrow_p\frac{1}{h^2(\kappa_2/2-\kappa_1^2)}\begin{pmatrix}h^2\kappa_2&-h\kappa_1\\-h\kappa_1&\frac{1}{2}\end{pmatrix}.$$

Thus, the asymptotic bias of $\sqrt{nh}(\widehat{\alpha}_1-\alpha_1)$ is

$$\frac{-e_1^T}{h^2(\kappa_2/2-\kappa_1^2)}\begin{pmatrix}h^2\kappa_2&-h\kappa_1\\-h\kappa_1&\frac{1}{2}\end{pmatrix}\frac{n^{1/2}}{h^{1/2}}\begin{pmatrix}O_p(h^3)\\O_p(h^4)\end{pmatrix}=O(n^{1/2}h^{5/2})=o(1).$$

Similarly, the asymptotic variance of $\sqrt{nh}(\widehat{\alpha}_1-\alpha_1)$ is

$$\begin{aligned}&\frac{\omega}{h^4(\kappa_2/2-\kappa_1^2)^2}e_1^T\begin{pmatrix}h^2\kappa_2&-h\kappa_1\\-h\kappa_1&\frac{1}{2}\end{pmatrix}\begin{pmatrix}\kappa_{20}&h\kappa_{21}\\h\kappa_{21}&h^2\kappa_{22}\end{pmatrix}\begin{pmatrix}h^2\kappa_2&-h\kappa_1\\-h\kappa_1&\frac{1}{2}\end{pmatrix}e_1\\&=\sigma^2\int\frac{f_Z(z_i)^2}{f_{X,Z}(c,z_i)}dz_i\cdot C_v,\end{aligned}$$

where

$$C_v=\frac{\kappa_2^2\kappa_{20}+\kappa_1^2\kappa_{22}-2\kappa_1\kappa_2\kappa_{21}}{\left(\frac{1}{2}\kappa_2-\kappa_1^2\right)^2}.$$

This completes the proof.

Lemma 11. Recall that

$$\begin{aligned}\phi_{i,j}^{rd}=&\frac{T_i}{f_{X,Z}(c,z_i)}(h_2)^{-1}K(\frac{z_i-z_j}{h_2})K\Big(\frac{x_i-c}{h}\Big)R_i\\&-\frac{T_if_Z(z_i)}{f_{X,Z}(c,z_i)^2}(h_1^2)^{-1}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K\Big(\frac{x_i-c}{h}\Big)R_i+\frac{T_if_Z(z_i)}{f_{X,Z}(c,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i.\end{aligned}$$

Under the same condition as in Theorem 1, and choose $h_1\asymp h$

$$\begin{aligned}&\mathbb{E}(\phi_{i,j}^{rd}|j)=\frac{h}{2}d_1(c^+,z_j)-\frac{h}{2h_1}K(\frac{c-x_j}{h_1})\frac{f_Z(z_j)}{f_{X,Z}(c,z_j)}d_1(c^+,z_j)+O_p(h^2+hh_1+hh_2),\\&\mathbb{E}(\phi_{i,j}^{rd}|i)=\frac{T_if_Z(z_i)}{f_{X,Z}(c,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i+O_p(h_2^2+h_1^2),\\&\mathbb{E}(\phi_{i,j}^{rd})=O_p(h^3),\end{aligned}$$

where O_p terms are valid uniformly over i or j.

Proof.

$$
\begin{aligned}
\mathbb{E}(\phi_{i,j}^{rd}|j) =& \underbrace{\mathbb{E}\left(\frac{T_i}{f_{X,Z}(c,z_i)}(h_2)^{-1}K(\frac{z_i-z_j}{h_2})K\Big(\frac{x_i-c}{h}\Big)R_i\Big|z_j\right)}_{\text{(Part.EI)}} \\
&- \underbrace{\mathbb{E}\left(\frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)^2}(h_1^2)^{-1}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K\Big(\frac{x_i-c}{h}\Big)R_i\Big|z_j,x_j\right)}_{\text{(Part.EII)}} \\
&+ \underbrace{\mathbb{E}\left(\frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i\Big|z_j,x_j\right)}_{\text{(Part.EIII)}}.
\end{aligned}
$$

From lemma 9,

$$
\begin{aligned}
\text{(Part.EI)} &= \mathbb{E}\left(\frac{T_i}{f_{X,Z}(c,z_i)}(h_2)^{-1}K(\frac{z_i-z_j}{h_2})K\Big(\frac{x_i-c}{h}\Big)R_i\Big|z_j\right) \\
&= h\int_{z_i}\frac{1}{f_{X,Z}(c,z_i)}(h_2)^{-1}K(\frac{z_i-z_j}{h_2})M(c,z_i)dz_i \\
&= h\int_{z_i}(h_2)^{-1}K(\frac{z_i-z_j}{h_2})\frac{1}{2}(m_1(c^+,z_i)-\alpha_1)dz_i + O_p(h^2) \\
&= \frac{h}{2}(m_1(c^+,z_j)-\alpha_1) + O_p(h^2+h_2h),
\end{aligned}
$$

where O_p terms are valid uniformly over j. Similarly, for Part.EII and Part.EIII, we can show that

$$
\begin{aligned}
&\text{(Part.EII)} = \mathbb{E}\left(\frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)^2}(h_1^2)^{-1}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K\Big(\frac{x_i-c}{h}\Big)R_i\Big|z_j,x_j\right) \\
&= h\int_{z_i}\frac{f_Z(z_i)}{f_{X,Z}(c,z_i)^2}(h_1^2)^{-1}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})M(c,z_i)dz_i \\
&= \frac{h}{2}\int_{z_i}\frac{f_Z(z_i)}{f_{X,Z}(c,z_i)}(h_1^2)^{-1}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})(m_1(c^+,z_i)-\alpha_1)dz_i + O_p(h^2) \\
&= \frac{h}{2h_1}K(\frac{c-x_j}{h_1})\frac{f_Z(z_j)}{f_{X,Z}(c,z_j)}(m_1(c^+,z_j)-\alpha_1) + O_p\left(hh_1\right) + O_p(h^2),
\end{aligned}
$$

$$
\text{(Part.EIII)} = \mathbb{E}\left(\frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i\Big|z_j,x_j\right) = \frac{h}{2}\int_{z_i}(m_1(c^+,z_i)-\alpha_1)f_Z(z_i)dz_i + O_p(h^2) = O_p(h^2),
$$

where the last step follows from the definition of α_1. Define $d_1(x_i,z_i) = m_1(x_i,z_i)-\alpha_1$. Combining the Part EI, EII and EIII, we obtain

$$
\mathbb{E}(\phi_{ij}^{rd}|j) = \frac{h}{2}d_1(c^+,z_j) - \frac{h}{2h_1}K(\frac{c-x_j}{h_1})\frac{f_Z(z_j)}{f_{X,Z}(c,z_j)}d_1(c^+,z_j) + O_p(h^2+hh_1+hh_2).
$$

Following the similar calculation, by lemma 8 we have

$$
\begin{aligned}
\mathbb{E}(\phi_{i,j}^{rd}|i) =& \underbrace{\frac{T_i}{f_{X,Z}(c,z_i)}\mathbb{E}\left((h_2)^{-1}K(\frac{z_i-z_j}{h_2})\Big|z_i\right)K\Big(\frac{x_i-c}{h}\Big)R_i}_{\text{Part. FI}} \\
&- \underbrace{\frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)^2}\mathbb{E}\left((h_1^2)^{-1}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})\Big|z_i,x_i\right)K\Big(\frac{x_i-c}{h}\Big)R_i}_{\text{Part. FII}} \\
&+ \underbrace{\frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i}_{\text{Part. FIII}} \\
&= \frac{T_i f_Z(z_i)}{f_{X,Z}(c,z_i)}K\Big(\frac{x_i-c}{h}\Big)R_i + O_p(h_1^2+h_2^2).
\end{aligned}
$$

Finally, we calculate $\mathbb{E}(\phi_{i,j}^{rd})$ using Lemma 9,

$$
\begin{aligned}
\mathbb{E}(\text{ Part.FIII}) &= h\int \frac{f_Z(z_i)}{f_{X,Z}(c,z_i)}M(c,z_i)dz_i \\
&= \frac{h}{2}\int(m_1(c^+,z_i)-\alpha_1)f_Z(z_i)dz_i - h^2\kappa_1\beta_1 - h^2\kappa_1\alpha_1\int\frac{\partial f_{X,Z}(c,z_i)}{\partial x_i}\frac{f_Z(z_i)}{f_{X,Z}(c,z_i)}dz_i \\
&+ h^2\kappa_1\int\left(\frac{\partial m_1(c^+,z_i)}{\partial x_i}f_Z(z_i) + m_1(c^+,z_i)\frac{\partial f_{X,Z}(c,z_i)}{\partial x_i}\frac{f_Z(z_i)}{f_{X,Z}(c,z_i)}\right)dz_i + O_p(h^3) \\
&= -h^2\kappa_1\alpha_1\int\frac{\partial f_{X,Z}(c,z_i)}{\partial x_i}\frac{f_Z(z_i)}{f_{X,Z}(c,z_i)}dz_i + O_p(h^3)
\end{aligned}
$$

where the O_p terms are valid uniformly over i and the last equality follows as

$$
\alpha_1 = \int m_1(c^+,z)f_Z(z)dz,
$$

and

$$
\beta_1 = \int\frac{\partial m_1(c^+,z_i)}{\partial x_i}f_Z(z_i)dz_i + \int m_1(c^+,z_i)\frac{\partial f_{X,Z}(c,z_i)}{\partial x_i}\frac{f_Z(z_i)}{f_{X,Z}(c,z_i)}dz_i.
$$

From Lemma 8, some tedious calculation implies

$$
\mathbb{E}(\text{Part.FI}) = \mathbb{E}(\text{Part.FIII}) + O_p(hh_2^2),
$$

and

$$
\mathbb{E}(\text{Part.FII}) = \mathbb{E}(\text{Part.FIII}) + O_p(hh_1^2).
$$

Thus when $h_1 \asymp h$

$$
\mathbb{E}(\phi_{i,j}) = O_p(h^3).
$$

This completes the proof. □

Lemma 12. When either $h \asymp h_1 \asymp h_2$ or $h \asymp \sqrt{h_1} \asymp h_2$, the condition in Theorem 12.3 in Van der Vaart (2000) holds, such that $\mathbb{E}(\phi_{ij}^2) < \infty$

Proof. $\mathbb{E}(\phi_{ij}^2)$ is a linear combination of the following quantities:

$$\mathbb{E}\left(\frac{1}{f_{X,Z(1)}(c^+,z_i)^2}(h_2)^{-2}K^2(\frac{z_i-z_j}{h_2})K^2\left(\frac{x_i-c}{h}\right)R_i^2\right) \tag{3.10}$$

$$\mathbb{E}\left(\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^4}(h_1)^{-4}K_1^2(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K^2\left(\frac{x_i-c}{h}\right)R_i^2\right) \tag{3.11}$$

$$\mathbb{E}\left(\frac{f_Z^2(z_i)}{f_{X,Z(1)}^2(c^+,z_i)}K^2\left(\frac{x_i-c}{h}\right)R_i^2\right) \tag{3.12}$$

$$\mathbb{E}\left(\frac{f_Z(z_i)}{f_{X,Z(1)}^3(c^+,z_i)}(h_2)^{-1}(h_1)^{-2}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K(\frac{z_i-z_j}{h_2})K^2\left(\frac{x_i-c}{h}\right)R_i^2\right) \tag{3.13}$$

$$\mathbb{E}\left(\frac{f_Z(z_i)}{f_{X,Z(1)}^2(c^+,z_i)}(h_2)^{-1}K(\frac{z_i-z_j}{h_2})K^2\left(\frac{x_i-c}{h}\right)R_i^2\right) \tag{3.14}$$

$$\mathbb{E}\left(\frac{f_Z^2(z_i)}{f_{X,Z(1)}(c^+,z_i)^3}(h_1^2)^{-1}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K^2\left(\frac{x_i-c}{h}\right)R_i^2\right) \tag{3.15}$$

- From Lemma 13, equation 3.10 can be written as

$$\begin{aligned}&\mathbb{E}\left(\frac{1}{f_{X,Z(1)}(c^+,z_i)^2}(h_2)^{-2}K^2(\frac{z_i-z_j}{h_2})K^2\left(\frac{x_i-c}{h}\right)R_i^2\Big|z_j\right)\\ &=\int_{z_i}\frac{1}{f_{X,Z(1)}(c^+,z_i)^2}(h_2)^{-1}K^2(\frac{z_i-z_j}{h_2})J(x_i,z_i)dz_i \qquad (z_j\perp\!\!\!\perp(x_i,z_i))\\ &=\kappa_{20}\int_{z_i}\frac{1}{f_{X,Z(1)}(c^+,z_i)}(h_2)^{-1}K^2(\frac{z_i-z_j}{h_2})(\sigma^2+d_1(c^+,z_i)^2)dz_i+O_p(h)\\ &=\frac{\kappa_{20}^2}{f_{X,Z(1)}(c^+,z_j)}(\sigma^2+d_1(c^+,z_j)^2)+O_p(h)\end{aligned}$$

  Thus

$$\begin{aligned}&\mathbb{E}\left(\frac{1}{f_{X,Z(1)}(c^+,z_i)^2}(h_2)^{-2}K^2(\frac{z_i-z_j}{h_2})K^2\left(\frac{x_i-c}{h}\right)R_i^2\right)\\ &=\kappa_{20}^2\left(\sigma^2\int\frac{f_Z(z_j)}{f_{X,Z(1)}(c^+,z_j)}dz_j+\int d_1(c^+,z_j)^2\frac{f_Z(z_j)}{f_{X,Z(1)}(c^+,z_j)}dz_j\right)+O_p(h)<\infty\end{aligned}$$

- When $h \asymp h_1 \asymp h_2$, by law of iterated expectation and Lemma 13, equation 3.11 can be written as

$$\begin{aligned}&\mathbb{E}\left(\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^4}(h_1)^{-4}K_1^2(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K^2\left(\frac{x_i-c}{h}\right)R_i^2\Big|x_j,z_j\right)\\ &=\int_{z_i}\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^4}(h_1)^{-3}K_1^2(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})J(c,z_i)dz_i\\ &=\kappa_{20}\int_{z_i}\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^3}(h_1)^{-3}K_1^2(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})(\sigma^2+d_1(c^+,z_i)^2)dz_i+O_p(h)\\ &=\kappa_{20}^2\frac{f_Z(z_j)^2}{f_{X,Z(1)}(c^+,z_j)^3}(\sigma^2+d_1(c^+,z_j)^2)(h_1)^{-1}K^2(\frac{c-x_j}{h_1})+O_p(h)\end{aligned}$$

Thus

$$
\begin{aligned}
&\mathbb{E}\left(\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^4}(h_1)^{-4}K_1^2(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K^2\left(\frac{x_i-c}{h}\right)R_i^2\right)\\
&=\kappa_{20}^2\int\int\frac{f_Z(z_j)^2}{f_{X,Z(1)}(c^+,z_j)^3}(\sigma^2+d_1(c^+,z_j)^2)(h_1)^{-1}K^2(\frac{c-x_j}{h_1})f_{X,Z(1)}(x_j,z_j)dz_jdx_j+O_p(h)\\
&=\kappa_{20}^3\int\frac{f_Z(z_j)^2}{f_{X,Z(1)}(c^+,z_j)^2}(\sigma^2+d_1(c^+,z_j)^2)dz_j+O_p(h)\\
&=\kappa_{20}^3\left(\sigma^2\int\frac{f_Z(z_j)^2}{f_{X,Z(1)}(c^+,z_j)^2}dz_j+\int d_1(c^+,z_j)^2\frac{f_Z(z_j)^2}{f_{X,Z(1)}(c^+,z_j)^2}dz_j\right)+O_p(h)<\infty
\end{aligned}
$$

When $h \asymp \sqrt{h_1} \asymp h_2$

$$
\begin{aligned}
&\mathbb{E}\left(\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^4}(h_1)^{-4}K_1^2(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K^2\left(\frac{x_i-c}{h}\right)R_i^2\middle|x_j,z_j\right)\\
&=\int_{z_i}\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^4}(h_1)^{-2}K_1^2(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})J(c,z_i)dz_i\\
&=\kappa_{20}\int_{z_i}\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^3}(h_1)^{-2}K_1^2(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})(\sigma^2+d_1(c^+,z_i)^2)dz_i+O_p(h)\\
&=\kappa_{20}^2\frac{f_Z(z_j)^2}{f_{X,Z(1)}(c^+,z_j)^3}(\sigma^2+d_1(c^+,z_j)^2)K^2(\frac{c-x_j}{h_1})+O_p(h)
\end{aligned}
$$

Thus

$$
\begin{aligned}
&\mathbb{E}\left(\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^4}(h_1)^{-4}K_1^2(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K^2\left(\frac{x_i-c}{h}\right)R_i^2\right)\\
&=\kappa_{20}^2\int\int\frac{f_Z(z_j)^2}{f_{X,Z(1)}(c^+,z_j)^3}(\sigma^2+d_1(c^+,z_j)^2)K^2(\frac{c-x_j}{h_1})f_{X,Z(1)}(x_j,z_j)+O_p(h)\\
&=O_p(h)<\infty
\end{aligned}
$$

- From Lemma 13, equation 3.12 can be written as

$$
\begin{aligned}
&\mathbb{E}\left(\frac{f_Z(z_i)^2}{f_{X,Z(1)^2}(c^+,z_i)}K^2\left(\frac{x_i-c}{h}\right)R_i^2\right)\\
&=h\int_{z_i}\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c,z_i)^2}J(c,z_i)dz_i\\
&=h\kappa_{20}\int_{z_i}\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c,z_i)}(\sigma^2+d_1(c^+,z_j)^2)dz_i+O_p(h^2)\\
&=O_p(h)<\infty
\end{aligned}
$$

- From Lemma 13, and using the fact that the kernel is bounded by a constant $\mathcal{K}$, equation 3.13 can be written as

$$
\begin{aligned}
&\mathbb{E}\left(\frac{f_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)^3}(h_2)^{-1}(h_1)^{-2}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K(\frac{z_i-z_j}{h_2})K^2\Big(\frac{x_i-c}{h}\Big)R_i^2\Big|x_j,z_j\right)\\
&=\int_{z_i}\frac{f_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)^3}(h_1)^{-2}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K(\frac{z_i-z_j}{h_2})J(c,z_j)dz_i\\
&=\kappa_{20}K(\frac{c-x_j}{h_1})\int_{z_i}\frac{f_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)^2}(h_1)^{-2}K(\frac{z_i-z_j}{h_1})K(\frac{z_i-z_j}{h_2})(\sigma^2+d_1(c^+,z_i)^2)dz_i+O_p(h)\\
&\leq\mathcal{K}\kappa_{20}K(\frac{c-x_j}{h_1})(h_1)^{-1}\frac{f_Z(z_j)}{f_{X,Z(1)}(c^+,z_j)^2}(\sigma^2+d_1(c^+,z_j)^2)+O_p(h)
\end{aligned}
$$

Thus

$$
\begin{aligned}
&\mathbb{E}\left(\frac{f_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)^3}(h_2)^{-1}(h_1)^{-2}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K(\frac{z_i-z_j}{h_2})K^2\Big(\frac{x_i-c}{h}\Big)R_i^2\right)\\
&\leq\mathcal{K}\kappa_{20}\int\int K(\frac{c-x_j}{h_1})(h_1)^{-1}\frac{f_Z(z_j)}{f_{X,Z(1)}(c^+,z_j)^2}(\sigma^2+d_1(c^+,z_j)^2)f_{X,Z(1)}(x_j,z_j)dz_jdx_j+O_p(h)\\
&=\mathcal{K}\kappa_{20}\int\frac{f_Z(z_j)}{f_{X,Z(1)}(c^+,z_j)}(\sigma^2+d_1(c^+,z_j)^2)dz_j+O_p(h)\\
&=\mathcal{K}\kappa_{20}\left(\sigma^2\int\frac{f_Z(z_j)}{f_{X,Z(1)}(c^+,z_j)}dz_j+\int d_1(c^+,z_j)^2\frac{f_Z(z_j)}{f_{X,Z(1)}(c^+,z_j)}dz_j\right)+O_p(h)<\infty
\end{aligned}
$$

- From Lemma 13, equation 3.14 can be written as

$$
\begin{aligned}
&\mathbb{E}\left(\frac{f_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)^2}(h_2)^{-1}K(\frac{z_i-z_j}{h_2})K^2\Big(\frac{x_i-c}{h}\Big)R_i^2\Big|x_j,z_j\right)\\
&=\int\frac{f_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)^2}K(\frac{z_i-z_j}{h_2})J(c,z_i)dz_i\\
&=\kappa_{20}\int\frac{f_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)}K(\frac{z_i-z_j}{h_2})(\sigma^2+d_1(c^+,z_i)^2)dz_i+O_p(h)\\
&=O_p(h)
\end{aligned}
$$

Thus

$$
\mathbb{E}\left(\frac{f_Z(z_i)}{f_{X,Z(1)}(c^+,z_i)^2}(h_2)^{-1}K(\frac{z_i-z_j}{h_2})K^2\Big(\frac{x_i-c}{h}\Big)R_i^2\right)<\infty
$$

- From Lemma 13, equation 3.15 can be written as

$$
\begin{aligned}
&\mathbb{E}\left(\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^3}(h_1^2)^{-1}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K^2\Big(\frac{x_i-c}{h}\Big)R_i^2\Big|x_j,z_j\right)\\
&=h\int\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^3}(h_1^2)^{-1}K_1(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})J(c,z_i)dz_i\\
&=h\kappa_{20}K(\frac{c-x_j}{h_1})\int\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^2}(h_1^2)^{-1}K(\frac{z_i-z_j}{h_1})(\sigma^2+d_1(c^+,z_i)^2)dz_i+O_p(h)\\
&=hh_1^{-1}\kappa_{20}K(\frac{c-x_j}{h_1})\frac{f_Z(z_j)^2}{f_{X,Z(1)}(c^+,z_j)^2}(\sigma^2+d_1(c^+,z_j)^2)dz_i+O_p(h)
\end{aligned}
$$

Thus when $h \asymp h_1 \asymp h_2$

$$\mathbb{E}\left(\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^2}K_1^2(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K^2\left(\frac{x_i-c}{h}\right)R_i^2\right)$$
$$= \kappa_{20}^2\int\int\frac{f_Z(z_j)^2}{f_{X,Z(1)}(c^+,z_j)^3}(\sigma^2+d_1(c^+,z_j)^2)K(\frac{c-x_j}{h_1})f_{X,Z(1)}(x_j,z_j)dz_jdx_j+O_p(h)$$
$$= O_p(h)<\infty$$

and when $h \asymp \sqrt{h_1} \asymp h_2$

$$\mathbb{E}\left(\frac{f_Z(z_i)^2}{f_{X,Z(1)}(c^+,z_i)^2}K_1^2(\frac{c-x_j}{h_1},\frac{z_i-z_j}{h_1})K^2\left(\frac{x_i-c}{h}\right)R_i^2\right)=O_p(\sqrt{h})<\infty$$

□

Lemma 13. Under the same condition as in Theorem 1,

$$J(c,z_i):=\frac{1}{h}\int K^2\left(\frac{x_i-c}{h}\right)\mathbb{E}(R_i^2|x_i,z_i)f_{X,Z(1)}(x_i,z_i)dx_i$$
$$=(\sigma^2+d_1(c^+,z_i)^2)\cdot f_{X,Z(1)}(c^+,z_i)\kappa_{20}+O_p(h),,$$

where O_p terms are valid uniformly over i.

Proof. Recall $R_i=\epsilon_i+m_1(x_i,z_i)-\alpha_1-(x_i-c)\beta_1=\epsilon_i+d_1(x_i,z_i)-(x_i-c)\beta_1$, thus

$$\mathbb{E}(R_i^2|x_i,z_i)=d_1(x_i,z_i)^2+\mathbb{E}(\epsilon_i^2|x_i,z_i)+(x_i-c)^2\beta_1^2$$
$$+2d_1(x_i,z_i)\mathbb{E}(\epsilon_i|x_i,z_i)+2(x_i-c)\beta_1\mathbb{E}(\epsilon_i|x_i,z_i)+2d_1(x_i,z_i)(x_i-c)\beta_1$$
$$=\sigma^2+d_1(x_i,z_i)^2+(x_i-c)^2\beta_1^2+2d_1(x_i,z_i)\cdot(x_i-c)\beta_1$$

Following the standard Taylor expansion, we can show that

$$J(c,z_i)=\frac{1}{h}\int K^2\left(\frac{x_i-c}{h}\right)\mathbb{E}(R_i^2|x_i,z_i)f_{X,Z(1)}(x_i,z_i)dx_i$$
$$=\sigma^2\int_u K^2(u)f_{X,Z(1)}(c+uh,z_i)du+\int_u K^2(u)d_1(c+uh,z_i)^2f_{X,Z(1)}(c+uh,z_i)du$$
$$+\beta_1^2h^2\int_u K^2(u)u^2f_{X,Z(1)}(c+uh,z_i)du+2h\beta_1\cdot\int_u K^2(u)ud_1(c+uh,z_i)f_{X,Z(1)}(c+uh,z_i)du$$
$$=(\sigma^2+d_1(c^+,z_i)^2)\cdot f_{X,Z(1)}(c^+,z_i)\int_u K^2(u)du+O_p(h),$$

where O_p terms are valid uniformly over i as the (mixed) third derivatives of $f_{X,Z(1)}(x_i,z_i)$ are all bounded.
□

References

Fan, J. (1993). Local linear regression smoothers and their minimax efficiencies. *Annals of Statistics* **21** 196–216.

Frölich, M. and Huber, M. (2018). Including covariates in the regression discontinuity design. *Journal of Business & Economic Statistics* **o** 0–0.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics* **142** 615–635.

Van der Vaart, A. W. (2000). *Asymptotic statistics*, vol. 3. Cambridge university press.