Supplement to "Inference on Individual Treatment Effects in Nonseparable Triangular Models"

Jun Ma* Vadim Marmer† Zhengfei Yu‡

Contents

S1 Proofs of Lemmas 2, 3, 4 and 5

Proof of Lemma 2. Let $\mathit{\Pi}_{dzx}(y) \coloneqq \Pr[Y \leq y, D = d, Z = z, X = x]$ and $\widehat{\mathit{\Pi}}_{dzx}(y) \coloneqq n^{-1}\sum_{i=1}^{n} \mathcal{D}_{dzx}(W_i, y)$, where $\mathcal{D}_{dzx}(W_i, y) \coloneqq \mathbb{1}(Y_i \leq y, D_i = d, Z_i = z, X_i = x)$. Let $I_{dx} \coloneqq \mathscr{S}_{g(d,x,\epsilon)|X=x} = \left[\underline{y}_{dx}, \overline{y}_{dx}\right]$. It follows from Kosorok (2007, Lemmas 9.7(iv) and 9.8) that the class $\mathfrak{D} \coloneqq \{\mathcal{D}_{dzx}(\cdot, y) : y \in I_{dx}\}$ is VC-subgraph with VC index being at most 2. Then it follows from Giné and Nickl (2016, Theorem 3.6.9) that $\mathfrak{D}$ is VC-type with respect to the constant envelope $F_{\mathfrak{D}} = 1$. It follows from Talagrand's inequality (Chernozhukov et al., 2016, Lemma 6.3, with $\mathcal{F} = \mathfrak{D}$, $\sigma = b = F_{\mathfrak{D}} = 1$ and $t = \log(n)$) that $\left\|\mathbb{G}_n^W\right\|_{\mathfrak{D}} = O_p^{\star}\left(\sqrt{\log(n)}\right)$. Note that $\left\|\widehat{\mathit{\Pi}}_{dzx} - \mathit{\Pi}_{dzx}\right\|_{I_{dx}} = n^{-1/2}\left\|\mathbb{G}_n^W\right\|_{\mathfrak{D}}$ and therefore, $\left\|\widehat{\mathit{\Pi}}_{dzx} - \mathit{\Pi}_{dzx}\right\|_{I_{dx}} = O_p^{\star}\left(\sqrt{\log(n)/n}\right)$. It is shown in the proof of Theorem 1 of FVX that

$$\left(\frac{\widehat{\mathit{\Pi}}_{d0x}\left(\widehat{\phi}_{dx}(y)\right)}{\widehat{p}_{0x}} - \frac{\widehat{\mathit{\Pi}}_{d1x}\left(\widehat{\phi}_{dx}(y)\right)}{\widehat{p}_{1x}}\right) + \left(\frac{\widehat{\mathit{\Pi}}_{d'0x}(y)}{\widehat{p}_{0x}} - \frac{\widehat{\mathit{\Pi}}_{d'1x}(y)}{\widehat{p}_{1x}}\right) = \xi_n, \tag{S1}$$

with an error term that satisfies $\xi_n = O_p^{\star}\left(n^{-1}\right)$. Note that $\phi_{dx}(y)$ satisfies

$$\left(\frac{\mathit{\Pi}_{d0x}(\phi_{dx}(y))}{p_{0x}} - \frac{\mathit{\Pi}_{d1x}(\phi_{dx}(y))}{p_{1x}}\right) + \left(\frac{\mathit{\Pi}_{d'0x}(y)}{p_{0x}} - \frac{\mathit{\Pi}_{d'1x}(y)}{p_{1x}}\right) = 0. \tag{S2}$$

September 26, 2022.

*School of Economics, Renmin University of China, P.R. China
†Vancouver School of Economics, University of British Columbia, Canada
‡Faculty of Humanities and Social Sciences, University of Tsukuba, Japan

Lemma 2.1 and Theorem 3.1 of Abadie (2003) imply that

$$F_{dx|C_x}(y) = \frac{\Pr[Y \le y, D = d|Z = 1, X = x] - \Pr[Y \le y, D = d|Z = 0, X = x]}{\Pr[D = d|Z = 1, X = x] - \Pr[D = d|Z = 0, X = x]}.$$

Then it is clear that $\zeta_{dx}(y) = \varPi'_{d1x}(y)/p_{1x} - \varPi'_{d0x}(y)/p_{0x}$. By Assumption 1(c,h), $\underline{\zeta}_{dx} \coloneqq \inf_{y \in I_{dx}} |\zeta_{dx}(y)| > 0$. It follows from Hoeffding's inequality that $\widehat{p}_{zx} - p_{zx} = O_p^\star\left(\sqrt{\log(n)/n}\right)$ and it is easy to check that this also implies $\widehat{p_{zx}^{-1}} - p_{zx}^{-1} = O_p^\star\left(\sqrt{\log(n)/n}\right)$. Note that by construction, $\widehat{\phi}_{dx}(y) \in I_{dx}$. By (S1) and (S2),

$$\begin{aligned}
&\left(\frac{\varPi_{d0x}\left(\widehat{\phi}_{dx}(y)\right)}{p_{0x}} - \frac{\varPi_{d1x}\left(\widehat{\phi}_{dx}(y)\right)}{p_{1x}}\right) - \left(\frac{\varPi_{d0x}(\phi_{dx}(y))}{p_{0x}} - \frac{\varPi_{d1x}(\phi_{dx}(y))}{p_{1x}}\right)\\
&\qquad = \xi_n - \left\{\left(\frac{\widehat{\varPi}_{d'0x}(y)}{\widehat{p}_{0x}} - \frac{\widehat{\varPi}_{d'1x}(y)}{\widehat{p}_{1x}}\right) - \left(\frac{\varPi_{d'0x}(y)}{p_{0x}} - \frac{\varPi_{d'1x}(y)}{p_{1x}}\right)\right\}\\
&\qquad - \left\{\left(\frac{\varPi_{d0x}\left(\widehat{\phi}_{dx}(y)\right)}{\widehat{p}_{0x}} - \frac{\varPi_{d1x}\left(\widehat{\phi}_{dx}(y)\right)}{\widehat{p}_{1x}}\right) - \left(\frac{\varPi_{d0x}\left(\widehat{\phi}_{dx}(y)\right)}{p_{0x}} - \frac{\varPi_{d1x}\left(\widehat{\phi}_{dx}(y)\right)}{p_{1x}}\right)\right\}\\
&\qquad - \left\{\left(\frac{\widehat{\varPi}_{d0x}\left(\widehat{\phi}_{dx}(y)\right)}{\widehat{p}_{0x}} - \frac{\widehat{\varPi}_{d1x}\left(\widehat{\phi}_{dx}(y)\right)}{\widehat{p}_{1x}}\right) - \left(\frac{\varPi_{d0x}\left(\widehat{\phi}_{dx}(y)\right)}{\widehat{p}_{0x}} - \frac{\varPi_{d1x}\left(\widehat{\phi}_{dx}(y)\right)}{\widehat{p}_{1x}}\right)\right\}.
\end{aligned}$$

Then it follows from this result, $\left\|\widehat{\varPi}_{dzx} - \varPi_{dzx}\right\|_{I_{dx}} = O_p^\star\left(\sqrt{\log(n)/n}\right)$, $\widehat{p_{zx}^{-1}} - p_{zx}^{-1} = O_p^\star\left(\sqrt{\log(n)/n}\right)$ and also $\underline{\zeta}_{dx} > 0$ that $\left\|\widehat{\phi}_{dx} - \phi_{dx}\right\|_{I_{d'x}} = O_p^\star\left(\sqrt{\log(n)/n}\right)$, i.e., for some constant C_1, C_2,

$$\Pr\left[\left\|\widehat{\phi}_{dx} - \phi_{dx}\right\|_{I_{d'x}} \le C_1\sqrt{\frac{\log(n)}{n}}\right] > 1 - C_2 n^{-1}. \tag{S3}$$

Decompose

$$\widehat{\varPi}_{dzx}\left(\widehat{\phi}_{dx}(y)\right) - \varPi_{dzx}(\phi_{dx}(y)) = \left\{\widehat{\varPi}_{dzx}\left(\widehat{\phi}_{dx}(y)\right) - \varPi_{dzx}\left(\widehat{\phi}_{dx}(y)\right)\right\} + \left\{\varPi_{dzx}\left(\widehat{\phi}_{dx}(y)\right) - \varPi_{dzx}(\phi_{dx}(y))\right\}.$$

And by this result, (S3) and $\left\|\widehat{\varPi}_{dzx} - \varPi_{dzx}\right\|_{I_{dx}} = O_p^\star\left(\sqrt{\log(n)/n}\right)$, we also have $\left\|\widehat{\varPi}_{dzx} \circ \widehat{\phi}_{dx} - \varPi_{dzx} \circ \phi_{dx}\right\|_{I_{d'x}} = O_p^\star\left(\sqrt{\log(n)/n}\right)$. By using this result, (S1), (S2), $\left\|\widehat{\varPi}_{dzx} - \varPi_{dzx}\right\|_{I_{dx}} = O_p^\star\left(\sqrt{\log(n)/n}\right)$ and the equality

$$\frac{a}{b} = \frac{a}{c} - \frac{a(b-c)}{c^2} + \frac{a(b-c)^2}{bc^2}, \tag{S4}$$

we have

$$\begin{aligned}
O_p^\star\left(\frac{\log(n)}{n}\right) &= \left(\frac{\widehat{\varPi}_{d0x}\left(\widehat{\phi}_{dx}(y)\right)}{p_{0x}} - \frac{\varPi_{d0x}(\phi_{dx}(y))}{p_{0x}}\right) - \left(\frac{\widehat{\varPi}_{d1x}\left(\widehat{\phi}_{dx}(y)\right)}{p_{1x}} - \frac{\varPi_{d1x}(\phi_{dx}(y))}{p_{1x}}\right)\\
&+ \left(\frac{\widehat{\varPi}_{d'0x}(y)}{p_{0x}} - \frac{\varPi_{d'0x}(y)}{p_{0x}}\right) - \left(\frac{\widehat{\varPi}_{d'1x}(y)}{p_{1x}} - \frac{\varPi_{d'1x}(y)}{p_{1x}}\right) - \frac{\varPi_{d0x}(\phi_{dx}(y))}{p_{0x}^2}(\widehat{p}_{0x} - p_{0x}) + \frac{\varPi_{d1x}(\phi_{dx}(y))}{p_{1x}^2}(\widehat{p}_{1x} - p_{1x})\\
&- \frac{\varPi_{d'0x}(y)}{p_{0x}^2}(\widehat{p}_{0x} - p_{0x}) + \frac{\varPi_{d'1x}(y)}{p_{1x}^2}(\widehat{p}_{1x} - p_{1x}).
\end{aligned} \tag{S5}$$

We will later show that

$$\widehat{\mathit{\Pi}}_{dzx}\left(\widehat{\phi}_{dx}\left(y\right)\right)-\mathit{\Pi}_{dzx}\left(\widehat{\phi}_{dx}\left(y\right)\right)=\widehat{\mathit{\Pi}}_{dzx}\left(\phi_{dx}\left(y\right)\right)-\mathit{\Pi}_{dzx}\left(\phi_{dx}\left(y\right)\right)+O_{p}^{\star}\left(\left(\frac{\log\left(n\right)}{n}\right)^{3/4}\right),\qquad\text{(S6)}$$

uniformly in $y\in I_{d'x}$. By a second-order Taylor expansion,

$$\left(\frac{\mathit{\Pi}_{d1x}\left(\widehat{\phi}_{dx}\left(y\right)\right)}{p_{1x}}-\frac{\mathit{\Pi}_{d0x}\left(\widehat{\phi}_{dx}\left(y\right)\right)}{p_{0x}}\right)-\left(\frac{\mathit{\Pi}_{d1x}\left(\phi_{dx}\left(y\right)\right)}{p_{1x}}-\frac{\mathit{\Pi}_{d0x}\left(\phi_{dx}\left(y\right)\right)}{p_{0x}}\right)=$$
$$\zeta_{dx}\left(\phi_{dx}\left(y\right)\right)\left(\widehat{\phi}_{dx}\left(y\right)-\phi_{dx}\left(y\right)\right)+O_{p}^{\star}\left(\frac{\log\left(n\right)}{n}\right),\qquad\text{(S7)}$$

uniformly in $y\in I_{d'x}$. Note that if $X=x$,

$$\begin{aligned}\mathbb{1}\left(Y\leq\phi_{dx}\left(y\right),D=d\right)+\mathbb{1}\left(Y\leq y,D=d'\right)&=\mathbb{1}\left(g\left(d,x,\epsilon\right)\leq\phi_{dx}\left(y\right),D=d\right)+\mathbb{1}\left(g\left(d',x,\epsilon\right)\leq y,D=d'\right)\\&=\mathbb{1}\left(g\left(d',x,\epsilon\right)\leq y\right).\qquad\text{(S8)}\end{aligned}$$

Since Z is conditionally independent of ϵ given X, $\forall z\in\{0,1\}$,

$$\begin{aligned}R_{d'x}\left(y\right)&=\mathrm{E}\left[\mathbb{1}\left(Y\leq\phi_{dx}\left(y\right),D=d\right)+\mathbb{1}\left(Y\leq y,D=d'\right)\mid X=x,Z=z\right]\\&=\frac{\mathit{\Pi}_{dzx}\left(\phi_{dx}\left(y\right)\right)}{p_{zx}}+\frac{\mathit{\Pi}_{d'zx}\left(y\right)}{p_{zx}}\\&=F_{g(d',x,\epsilon)|X}\left(y\mid x\right),\qquad\text{(S9)}\end{aligned}$$

where $F_{g(d',x,\epsilon)|X}\left(y\mid x\right)\coloneqq\Pr\left[g\left(d',x,\epsilon\right)\leq y\mid X=x\right]$. Combining (S5), (S6) and (S7) and then using (S9), we have

$$\zeta_{dx}\left(\phi_{dx}\left(y\right)\right)\left(\widehat{\phi}_{dx}\left(y\right)-\phi_{dx}\left(y\right)\right)=$$
$$\frac{\widehat{\mathit{\Pi}}_{d0x}\left(\phi_{dx}\left(y\right)\right)}{p_{0x}}-\frac{\widehat{\mathit{\Pi}}_{d1x}\left(\phi_{dx}\left(y\right)\right)}{p_{1x}}+\frac{\widehat{\mathit{\Pi}}_{d'0x}\left(y\right)}{p_{0x}}-\frac{\widehat{\mathit{\Pi}}_{d'1x}\left(y\right)}{p_{1x}}-R_{d'x}\left(y\right)\frac{\widehat{p}_{0x}}{p_{0x}}+R_{d'x}\left(y\right)\frac{\widehat{p}_{1x}}{p_{1x}}+O_{p}^{\star}\left(\left(\frac{\log\left(n\right)}{n}\right)^{3/4}\right)=$$
$$\frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{1}\left(Y_{i}\leq\phi_{dx}\left(y\right),D_{i}=d\right)+\mathbb{1}\left(Y_{i}\leq y,D_{i}=d'\right)-R_{d'x}\left(y\right)\right)\pi_{x}\left(Z_{i},X_{i}\right)+O_{p}^{\star}\left(\left(\frac{\log\left(n\right)}{n}\right)^{3/4}\right).$$

The assertion follows from this result. It remains to show (S6).

Denote $\widehat{\mathit{\Lambda}}_{dzx}\left(y,y'\right)\coloneqq\widehat{\mathit{\Pi}}_{dzx}\left(y\right)-\widehat{\mathit{\Pi}}_{dzx}\left(y'\right)$ and $\mathit{\Lambda}_{dzx}\left(y,y'\right)\coloneqq\mathit{\Pi}_{dzx}\left(y\right)-\mathit{\Pi}_{dzx}\left(y'\right)$. In view of (S3), denote $\overline{\xi}\coloneqq C_{1}\sqrt{\log\left(n\right)/n}$. For $\xi>0$, denote $\mathcal{P}_{dzx}^{+}\left(W_{i},y,\xi\right)\coloneqq\mathbb{1}\left(\phi_{dx}\left(y\right)<Y_{i}\leq\phi_{dx}\left(y\right)+\xi,D_{i}=d,Z_{i}=z,X_{i}=x\right)$ and $\mathcal{P}_{dzx}^{-}\left(W_{i},y,\xi\right)\coloneqq\mathbb{1}\left(\phi_{dx}\left(y\right)-\xi<Y_{i}\leq\phi_{dx}\left(y\right),D_{i}=d,Z_{i}=z,X_{i}=x\right)$. By Kosorok (2007, Lemmas 9.7(iv) and 9.8), the function class $\mathfrak{P}^{+}\coloneqq\left\{\mathcal{P}_{dzx}^{+}\left(\cdot,y,\xi\right):y\in I_{d'x},\xi\in\left(0,\overline{\xi}\right]\right\}$ is VC-subgraph with VC index being at most 3, $\forall n$, and by Giné and Nickl (2016, Theorem 3.6.9), $\mathfrak{P}^{+}$ is uniformly VC-type with respect to the constant envelope $F_{\mathfrak{P}^{+}}=1$. $\mathfrak{P}^{-}$ is defined similarly. Then,

$$\sup_{y\in I_{d'x}}\left|\left\{\widehat{\mathit{\Pi}}_{dzx}\left(\widehat{\phi}_{dx}\left(y\right)\right)-\widehat{\mathit{\Pi}}_{dzx}\left(\phi_{dx}\left(y\right)\right)\right\}-\left\{\mathit{\Pi}_{dzx}\left(\widehat{\phi}_{dx}\left(y\right)\right)-\mathit{\Pi}_{dzx}\left(\phi_{dx}\left(y\right)\right)\right\}\right|\leq$$
$$\sup_{(y,\xi)\in I_{d'x}\times\left[-\overline{\xi},\overline{\xi}\right]}\left|\widehat{\mathit{\Lambda}}_{dzx}\left(\phi_{dx}\left(y\right)+\xi,\phi_{dx}\left(y\right)\right)-\mathit{\Lambda}_{dzx}\left(\phi_{dx}\left(y\right)+\xi,\phi_{dx}\left(y\right)\right)\right|\leq$$
$$\left\|\mathbb{P}_{n}^{W}-\mathbb{P}^{W}\right\|_{\mathfrak{P}^{+}}+\left\|\mathbb{P}_{n}^{W}-\mathbb{P}^{W}\right\|_{\mathfrak{P}^{-}},\qquad\text{(S10)}$$

where the first inequality holds with probability at least $1 - C_2 n^{-1}$, in view of (S3). $f_{Y|DZX}$ is bounded under Assumption 1. By calculation, we have

$$\begin{aligned} \mathrm{E}\left[\mathcal{P}_{dzx}^{+}(W, y, \xi)^2\right] &= \mathrm{E}\left[\left(\mathbb{1}\left(Y \leq \phi_{dx}(y) + \xi\right) - \mathbb{1}\left(Y \leq \phi_{dx}(y)\right)^2 \mathbb{1}\left(D = d, Z = z, X = x\right)\right)\right] \\ &= \left(\int_{\phi_{dx}(y)}^{\phi_{dx}(y)+\xi} f_{Y|DZX}\left(y' \mid d, z, x\right) \mathrm{d}y'\right) \Pr\left[D = d, Z = z, X = x\right] \end{aligned}$$

and

$$\sigma_{\mathfrak{P}^+}^2 \coloneqq \sup_{f \in \mathfrak{P}^+} \mathbb{P}^W f^2 = \sup_{(y,\xi) \in I_{d'x} \times (0, \overline{\xi}]} \mathrm{E}\left[\mathcal{P}_{d0x}^{+}(W, y, \xi)^2\right] = O\left(\sqrt{\frac{\log(n)}{n}}\right).$$

Then we apply Talagrand's inequality (the version given by Chernozhukov et al., 2016, Lemma 6.3) with $\mathcal{F} = \mathfrak{P}^+$, $b = F_{\mathfrak{P}^+} = 1$, $\sigma = \sigma_{\mathfrak{P}^+} \vee b\sqrt{V_{\mathfrak{P}^+} \log(n)/n}$ and $t = \log(n)$. It is straightforward to check that $\sigma_{\mathfrak{P}^+}^2 \leq \sigma^2 \leq b^2$, $n\sigma^2/b^2 \geq \log(n)$ and $n\sigma^2/b^2 \geq V_{\mathfrak{P}^+} \log\left(A_{\mathfrak{P}^+} b/\sigma\right)$, when n is large enough so that $V_{\mathfrak{P}^+} \log(n)/n \leq 1$ and $n/\log(n) \geq A_{\mathfrak{P}^+}^2$. Therefore, the conditions of Chernozhukov et al. (2016, Lemma 6.3) are satisfied when n is sufficiently large and by Talagrand's inequality, we have $\left\|\mathbb{G}_n^W\right\|_{\mathfrak{P}^+} = O_p^{\star}\left(\log(n)^{3/4}/n^{1/4}\right)$ and $\left\|\mathbb{P}_n^W - \mathbb{P}^W\right\|_{\mathfrak{P}^+} = n^{-1/2}\left\|\mathbb{G}_n^W\right\|_{\mathfrak{P}^+} = O_p^{\star}\left(\left(\log(n)/n\right)^{3/4}\right)$. A similar result holds for $\left\|\mathbb{G}_n^W\right\|_{\mathfrak{P}^-}$ and $\left\|\mathbb{P}_n^W - \mathbb{P}^W\right\|_{\mathfrak{P}^-}$. (S6) follows from these results and (S10). ■

Proof of Lemma 3. By mean value expansion,

$$\begin{aligned} \widehat{f}_{\Delta X}(v, x; b) - \widetilde{f}_{\Delta X}(v, x; b) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{b^2} K'\left(\frac{\Delta_i - v}{b}\right)\left(\widehat{\Delta}_i - \Delta_i\right)\mathbb{1}\left(X_i = x\right) \\ + \frac{1}{2}\frac{1}{n}\sum_{i=1}^{n} \frac{1}{b^3} K''\left(\frac{\dot{\Delta}_i - v}{b}\right)\left(\widehat{\Delta}_i - \Delta_i\right)^2 \mathbb{1}\left(X_i = x\right), \end{aligned} \tag{S11}$$

where $\dot{\Delta}_i$ denotes the mean value that lies between $\widehat{\Delta}_i$ and Δ_i so that $\left|\dot{\Delta}_i - \Delta_i\right| \leq \left|\widehat{\Delta}_i - \Delta_i\right|$. It is easy to see from the proof of Lemma 2 that (S3) also holds for $\widehat{\phi}_{dx}^{(-i)}(y)$ uniformly in $i = 1, ..., n$. Therefore,

$$\left|\widehat{\Delta}_i - \Delta_i\right| = \mathbb{1}\left(D_i = 1\right)\left|\widehat{\phi}_{0X_i}^{(-i)}(Y_i) - \phi_{0X_i}(Y_i)\right| + \mathbb{1}\left(D_i = 0\right)\left|\widehat{\phi}_{1X_i}^{(-i)}(Y_i) - \phi_{1X_i}(Y_i)\right|$$

and by (S3),

$$\overline{\Delta} \coloneqq \max_{i=1,\dots,n}\left|\widehat{\Delta}_i - \Delta_i\right| \mathbb{1}\left(X_i = x\right) = O_p^{\star}\left(\sqrt{\frac{\log(n)}{n}}\right). \tag{S12}$$

In view of $\Pr\left[\overline{\Delta} > C_1\sqrt{\log(n)/n}\right] \leq C_2 n^{-1}$, we have

$$1 - C_2 n^{-1} \leq \Pr\left[\overline{\Delta} \leq C_1\sqrt{\frac{\log(n)}{n}}\right] \leq \Pr\left[\overline{\Delta} \leq \underline{h}\right],$$

when n is sufficiently large, since $\sqrt{\log(n)/n} = o(h)$ under the assumption that $\log(n)/\left(nh_n^3\right) \downarrow 0$. By the triangle inequality, $|\Delta_i - v| \leq \left|\Delta_i - \dot{\Delta}_i\right| + \left|\dot{\Delta}_i - v\right| \leq \overline{\Delta} + \left|\dot{\Delta}_i - v\right|$, if $X_i = x$. Therefore, since $\left|K''\left(\left(\dot{\Delta}_i - v\right)/b\right)\right| \leq \|K''\|_{\infty}\mathbb{1}\left(\left|\dot{\Delta}_i - v\right| \leq b\right)$, for some constant $C_3 > 0$,

$$1 - C_3 n^{-1} \leq \Pr\left[\overline{\Delta} \leq \underline{h}\right] \leq$$

$$\Pr\left[\left|K''\left(\frac{\dot{\Delta}_i - v}{b}\right)\right| \mathbb{1}\left(X_i = x\right) \leq \left\|K''\right\|_\infty \mathbb{1}_i\left(v;b\right)\mathbb{1}\left(X_i = x\right), \forall\left(i,v,b\right)\in\{1,...,n\}\times I_x\times\mathbb{H}\right], \quad \text{(S13)}$$

where we denote $\mathbb{1}_i\left(v;b\right) := \mathbb{1}\left(\left|\Delta_i - v\right| \leq 2b\right)$. Denote $\mathbb{1}_{\Delta X}\left(v,x;b\right) := \left(nb\right)^{-1}\sum_{i=1}^{n}\mathbb{1}_i\left(v;b\right)\mathbb{1}\left(X_i = x\right)$. By this result and the triangle inequality,

$$\begin{aligned}&1 - C_3 n^{-1} \leq \\ &\Pr\left[\sup_{(v,b)\in I_x\times\mathbb{H}}\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b^3}K''\left(\frac{\dot{\Delta}_i - v}{b}\right)\left(\widehat{\Delta}_i - \Delta_i\right)^2\mathbb{1}\left(X_i = x\right)\right| \lesssim \left\{\sup_{(v,b)\in I_x\times\mathbb{H}} b^{-2}\mathbb{1}_{\Delta X}\left(v,x;b\right)\right\}\overline{\Delta}^2\right]. \quad \text{(S14)}\end{aligned}$$

Let $\mathcal{I}_x\left(U_i,v;b\right) := b^{-1}\mathbb{1}\left(\left|\Delta_x\left(\epsilon_i\right) - v\right| \leq 2b\right)\mathbb{1}\left(X_i = x\right)$. It follows from Kosorok (2007, Lemmas 9.7(iv), 9.8 and 9.9(vii,viii)) that $\mathfrak{I} := \{\mathcal{I}_x\left(\cdot,v;b\right) : \left(v,b\right)\in I_x\times\mathbb{H}\}$ is VC-subgraph with VC index being at most 3, $\forall h$, and has a constant envelope $F_{\mathfrak{I}} = \underline{h}^{-1}$. By Giné and Nickl (2016, Theorem 3.6.9), $\mathfrak{I}$ is uniformly VC-type with respect to the constant envelope $F_{\mathfrak{I}}$. It is easy to check that $\left\|\mathbb{P}^U\right\|_{\mathfrak{I}} = O\left(1\right)$ and $\sigma_{\mathfrak{I}}^2 := \sup_{f\in\mathfrak{I}}\mathbb{P}^U f^2 = O\left(h^{-1}\right)$ follow from change of variables. By Talagrand's inequality ($\mathcal{F} = \mathfrak{I}$, $b = F_{\mathfrak{I}}$, $\sigma = \sigma_{\mathfrak{I}} \vee b\sqrt{V_{\mathfrak{I}}\log\left(n\right)/n}$, $t = \log\left(n\right)$), $\sqrt{n}\left\|\mathbb{P}_n^U - \mathbb{P}^U\right\|_{\mathfrak{I}} = \left\|\mathbb{G}_n^U\right\|_{\mathfrak{I}} = O_p^\star\left(\sqrt{\log\left(n\right)/h}\right)$ and therefore,

$$\left\|\mathbb{1}_{\Delta X}\left(\cdot,x;\cdot\right)\right\|_{I_x\times\mathbb{H}} = \left\|\mathbb{P}_n^U\right\|_{\mathfrak{I}} \leq \left\|\mathbb{P}_n^U - \mathbb{P}^U\right\|_{\mathfrak{I}} + \left\|\mathbb{P}^U\right\|_{\mathfrak{I}} = O_p^\star\left(1\right). \quad \text{(S15)}$$

By this result, (S12) and (S14), we have

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b^3}K''\left(\frac{\dot{\Delta}_i - v}{b}\right)\left(\widehat{\Delta}_i - \Delta_i\right)^2\mathbb{1}\left(X_i = x\right) = O_p^\star\left(\frac{\log\left(n\right)}{nh^2}\right), \quad \text{(S16)}$$

uniformly in $\left(v,b\right)\in I_x\times\mathbb{H}$. By the definition of $\widehat{\Delta}_i$ and Δ_i,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b^2}K'\left(\frac{\Delta_i - v}{b}\right)\left(\widehat{\Delta}_i - \Delta_i\right)\mathbb{1}\left(X_i = x\right) &= -\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b^2}K'\left(\frac{\Delta_i - v}{b}\right)D_i\left\{\widehat{\phi}_{0X_i}^{(-i)}\left(Y_i\right) - \phi_{0X_i}\left(Y_i\right)\right\}\mathbb{1}\left(X_i = x\right)\\ &+ \frac{1}{n}\sum_{i=1}^{n}\frac{1}{b^2}K'\left(\frac{\Delta_i - v}{b}\right)\left(1 - D_i\right)\left\{\widehat{\phi}_{1X_i}^{(-i)}\left(Y_i\right) - \phi_{1X_i}\left(Y_i\right)\right\}\mathbb{1}\left(X_i = x\right). \quad \text{(S17)}\end{aligned}$$

Write

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b^2}K'\left(\frac{\Delta_i - v}{b}\right)\mathbb{1}\left(D_i = d', X_i = x\right)\left\{\widehat{\phi}_{dx}^{(-i)}\left(Y_i\right) - \phi_{dx}\left(Y_i\right)\right\} &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{b^2}K'\left(\frac{\Delta_i - v}{b}\right)\mathbb{1}\left(D_i = d', X_i = x\right)\\ &\times\left\{\frac{1}{n-1}\sum_{j\neq i}\mathcal{L}_{dx}\left(W_j,Y_i\right) + \widehat{\phi}_{dx}^{(-i)}\left(Y_i\right) - \phi_{dx}\left(Y_i\right) - \frac{1}{n-1}\sum_{j\neq i}\mathcal{L}_{dx}\left(W_j,Y_i\right)\right\}.\end{aligned}$$

It is easy to see from the proof of Lemma 2 that (38) also holds for $\widehat{\phi}_{dx}^{(-i)}\left(y\right)$ and the remainder term is uniform in $i = 1,...,n$. By Lemma 2, $\left|K'\left(\left(\Delta_i - v\right)/b\right)\right| \leq \left\|K'\right\|_\infty\mathbb{1}_i\left(v;b\right)$ and (S15),

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b^2}K'\left(\frac{\Delta_i - v}{b}\right)\mathbb{1}\left(D_i = d', X_i = x\right)\left\{\widehat{\phi}_{dx}^{(-i)}\left(Y_i\right) - \phi_{dx}\left(Y_i\right) - \frac{1}{n-1}\sum_{j\neq i}\mathcal{L}_{dx}\left(W_j,Y_i\right)\right\} = O_p^\star\left(\frac{\log\left(n\right)^{3/4}}{n^{3/4}h}\right),$$

uniformly in $\left(v,b\right)\in I_x\times\mathbb{H}$, and therefore

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b^2}K'\left(\frac{\mathit{\Delta}_i-v}{b}\right)\left(\widehat{\mathit{\Delta}}_i-\mathit{\Delta}_i\right)\mathbb{1}\left(X_i=x\right)=$$

$$\frac{1}{n_{(2)}}\sum_{(i,j)}\frac{1}{b^2}K'\left(\frac{\mathit{\Delta}_i-v}{b}\right)\left\{\left(1-D_i\right)\mathcal{L}_{1x}\left(W_j,Y_i\right)-D_i\mathcal{L}_{0x}\left(W_j,Y_i\right)\right\}\mathbb{1}\left(X_i=x\right)+O_p^{\star}\left(\frac{\log\left(n\right)^{1/2}}{n^{3/4}h}\right)=$$

$$\frac{1}{n_{(2)}}\sum_{(i,j)}\mathcal{G}_x\left(W_i,W_j,v;b\right)+O_p^{\star}\left(\frac{\log\left(n\right)^{3/4}}{n^{3/4}h}\right),\quad\text{(S18)}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. The conclusion follows from this result, (S11), (S16) and (S17). ∎

Proof of Lemma 4. By definition, we have

$$\mathcal{H}_x^{[1]}\left(U_i,v;b\right)=\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}\frac{1}{b^2}K'\left(\frac{\mathit{\Delta}_x\left(e\right)-v}{b}\right)\rho_x\left(e\right)\left\{\mathbb{1}\left(\epsilon_i\leq e\right)-F_{\epsilon|X}\left(e\mid x\right)\right\}\mathrm{d}e\right\}\pi_x\left(Z_i,X_i\right),$$

where

$$\rho_x\left(e\right):=\frac{f_{\epsilon DX}\left(e,0,x\right)}{\zeta_{1x}\left(g\left(1,x,e\right)\right)}-\frac{f_{\epsilon DX}\left(e,1,x\right)}{\zeta_{0x}\left(g\left(0,x,e\right)\right)}.$$

Since ϵ is independent of Z given X,

$$\mathrm{E}\left[\mathcal{H}_x^{[1]}\left(U,v;b\right)^2\right]=\mathrm{E}\left[\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}\frac{1}{b^2}K'\left(\frac{\mathit{\Delta}_x\left(e\right)-v}{b}\right)\rho_x\left(e\right)\left\{\mathbb{1}\left(\epsilon\leq e\right)-F_{\epsilon|X}\left(e\mid x\right)\right\}\mathrm{d}e\right\}^2\mid X=x\right]$$
$$\times\mathrm{E}\left[\left\{\frac{\mathbb{1}\left(Z=1\right)}{p_{1x}}-\frac{\mathbb{1}\left(Z=0\right)}{p_{0x}}\right\}^2\mid X=x\right]p_x.\quad\text{(S19)}$$

Note that

$$\mathrm{E}\left[\left\{\frac{\mathbb{1}\left(Z=1\right)}{p_{1x}}-\frac{\mathbb{1}\left(Z=0\right)}{p_{0x}}\right\}^2\mid X=x\right]p_x=p_{1x}^{-1}+p_{0x}^{-1}.\quad\text{(S20)}$$

Then,

$$\mathrm{E}\left[\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}\frac{1}{b^2}K'\left(\frac{\mathit{\Delta}_x\left(e\right)-v}{b}\right)\rho_x\left(e\right)\left\{\mathbb{1}\left(\epsilon\leq e\right)-F_{\epsilon|X}\left(e\mid x\right)\right\}\mathrm{d}e\right\}^2\mid X=x\right]=$$

$$\mathrm{E}\left[\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}\frac{1}{b^2}K'\left(\frac{\mathit{\Delta}_x\left(e\right)-v}{b}\right)\rho_x\left(e\right)\mathbb{1}\left(\epsilon\leq e\right)\mathrm{d}e\right\}^2\mid X=x\right]-\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}\frac{1}{b^2}K'\left(\frac{\mathit{\Delta}_x\left(e\right)-v}{b}\right)\rho_x\left(e\right)F_{\epsilon|X}\left(e\mid x\right)\mathrm{d}e\right\}^2,\quad\text{(S21)}$$

and

$$\mathrm{E}\left[\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}\frac{1}{b^2}K'\left(\frac{\mathit{\Delta}_x\left(e\right)-v}{b}\right)\rho_x\left(e\right)\mathbb{1}\left(\epsilon\leq e\right)\mathrm{d}e\right\}^2\mid X=x\right]=$$

$$b^{-4}\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}K'\left(\frac{\mathit{\Delta}_x\left(e\right)-v}{b}\right)K'\left(\frac{\mathit{\Delta}_x\left(e'\right)-v}{b}\right)F_{\epsilon|X}\left(e\wedge e'\mid x\right)\rho_x\left(e\right)\rho_x\left(e'\right)\mathrm{d}e\mathrm{d}e'=$$

$$b^{-4}\sum_{k=1}^{m}\sum_{j=1}^{m}\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}K'\left(\frac{\mathit{\Delta}_x\left(e\right)-v}{b}\right)K'\left(\frac{\mathit{\Delta}_x\left(e'\right)-v}{b}\right)F_{\epsilon|X}\left(e\wedge e'\mid x\right)\rho_x\left(e\right)\rho_x\left(e'\right)\mathrm{d}e\mathrm{d}e'.\quad\text{(S22)}$$

If $j > k$, since $\epsilon_{x,k} \leq \epsilon_{x,j-1}$,

$$\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} K'\left(\frac{\Delta_x(e) - v}{b}\right) K'\left(\frac{\Delta_x(e') - v}{b}\right) F_{\epsilon|X}(e \wedge e' \mid x)\, \rho_x(e)\, \rho_x(e')\, \mathrm{d}e\mathrm{d}e' =$$
$$\left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} K'\left(\frac{\Delta_x(e') - v}{b}\right) F_{\epsilon|X}(e' \mid x)\, \rho_x(e')\, \mathrm{d}e'\right\}\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} K'\left(\frac{\Delta_x(e) - v}{b}\right) \rho_x(e)\, \mathrm{d}e\right\}. \quad \text{(S23)}$$

Note that Δ_x is strictly monotonic on $[\epsilon_{x,j-1}, \epsilon_{x,j}]$. We assume without loss of generality that the restriction $\Delta_{x,j}$ is strictly increasing. Since I_x is an inner closed sub-interval of $\Delta_{x,j}((\epsilon_{x,j-1}, \epsilon_{x,j})) = (\Delta_x(\epsilon_{x,j-1}), \Delta_x(\epsilon_{x,j}))$, $v \in I_x$ is an interior point of $(\Delta_x(\epsilon_{x,j-1}), \Delta_x(\epsilon_{x,j}))$. Let $\psi_{x,j}(t) \coloneqq \rho_x\left(\Delta_{x,j}^{-1}(t)\right)\left(\Delta_{x,j}^{-1}\right)'(t)$. By change of variables and mean value expansion,

$$\begin{aligned}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^2} K'\left(\frac{\Delta_x(e) - v}{b}\right) \rho_x(e)\, \mathrm{d}e &= b^{-1}\int_{\frac{\Delta_x(\epsilon_{x,j-1})-v}{b}}^{\frac{\Delta_x(\epsilon_{x,j})-v}{b}} K'(u)\, \rho_x\left(\Delta_{x,j}^{-1}(bu+v)\right)\left(\Delta_{x,j}^{-1}\right)'(bu+v)\, \mathrm{d}u \\ &= b^{-1}\int_{\frac{\Delta_x(\epsilon_{x,j-1})-v}{b}}^{\frac{\Delta_x(\epsilon_{x,j})-v}{b}} K'(u)\left\{\psi_{x,j}(v) + \psi_{x,j}'(\dot{v})\, bu\right\} \mathrm{d}u, \end{aligned} \quad \text{(S24)}$$

where the mean value $\dot{v}$ depends on u and satisfies $|\dot{v} - v| \leq b\,|u|$. Note that $\int K'(u)\, \mathrm{d}u = 0$, K' is supported on $[-1, 1]$ and therefore, $\int_{(\Delta_x(\epsilon_{x,j-1})-v)/b}^{(\Delta_x(\epsilon_{x,j})-v)/b} K'(u)\, \mathrm{d}u = 0$, $\forall (v, b) \in I_x \times \mathbb{H}$, when $\overline{h}$ is sufficiently small. Therefore,

$$\left|\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^2} K'\left(\frac{\Delta_x(e) - v}{b}\right) \rho_x(e)\, \mathrm{d}e\right| \lesssim \sup_{t \in [\Delta_x(\epsilon_{x,j-1}), \Delta_x(\epsilon_{x,j})]} \left|\psi_{x,j}'(t)\right|, \quad \text{(S25)}$$

uniformly in $(v, b) \in I_x \times \mathbb{H}$, when $\overline{h}$ is sufficiently small. Denote $\chi_{x,j}(t) \coloneqq F_{\epsilon|X}\left(\Delta_{x,j}^{-1}(t) \mid x\right)\psi_{x,j}(t)$. Similarly,

$$\left|\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2} K'\left(\frac{\Delta_x(e') - v}{b}\right) F_{\epsilon|X}(e' \mid x)\, \rho_x(e')\, \mathrm{d}e'\right| \lesssim \sup_{t \in [\Delta_x(\epsilon_{x,k-1}), \Delta_x(\epsilon_{x,k})]} \left|\chi_{x,k}'(t)\right|, \quad \text{(S26)}$$

uniformly in $(v, b) \in I_x \times \mathbb{H}$, when $\overline{h}$ is sufficiently small. Then, it follows that

$$\mathrm{E}\left[\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x} \frac{1}{b^2} K'\left(\frac{\Delta_x(e) - v}{b}\right) \rho_x(e)\, \mathbb{1}(\epsilon \leq e)\, \mathrm{d}e\right\}^2 \mid X = x\right] =$$
$$b^{-4}\sum_{j=1}^{m} \int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} K'\left(\frac{\Delta_x(e) - v}{b}\right) K'\left(\frac{\Delta_x(e') - v}{b}\right) F_{\epsilon|X}(e \wedge e' \mid x)\, \rho_x(e)\, \rho_x(e')\, \mathrm{d}e\mathrm{d}e' + O(1), \quad \text{(S27)}$$

uniformly in $(v, b) \in I_x \times \mathbb{H}$. By change of variables,

$$b^{-4}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} K'\left(\frac{\Delta_x(e) - v}{b}\right) K'\left(\frac{\Delta_x(e') - v}{b}\right) F_{\epsilon|X}(e \wedge e' \mid x)\, \rho_x(e)\, \rho_x(e')\, \mathrm{d}e\mathrm{d}e' =$$
$$b^{-2}\int_{\frac{\Delta_x(\epsilon_{x,j-1})-v}{b}}^{\frac{\Delta_x(\epsilon_{x,j})-v}{b}} \int_{\frac{\Delta_x(\epsilon_{x,j-1})-v}{b}}^{\frac{\Delta_x(\epsilon_{x,j})-v}{b}} K'(u)\, K'(w)\, F_{\epsilon|X}\left(\Delta_{x,j}^{-1}(bu+v) \wedge \Delta_{x,j}^{-1}(bw+v) \mid x\right) \psi_{x,j}(bu+v)\, \psi_{x,j}(bw+v)\, \mathrm{d}u\mathrm{d}w$$
$$= 2b^{-2}\int K'(w)\, \psi_{x,j}(bw+v)\left\{\int_{-\infty}^{w} K'(u)\, \chi_{x,j}(bu+v)\, \mathrm{d}u\right\} \mathrm{d}w, \quad \text{(S28)}$$

where the second equality follows from symmetry and holds when $\overline{h}$ is sufficiently small. By integration by parts,

$$\int_{-\infty}^{w} K'(u)\,\chi_{x,j}(bu+v)\,\mathrm{d}u = K(w)\,\chi_{x,j}(bw+v) - b\int_{-\infty}^{w} K(u)\,\chi'_{x,j}(bu+v)\,\mathrm{d}u. \tag{S29}$$

Then,

$$\begin{aligned} b^{-1}\int K'(w)\,K(w)\,\{&\psi_{x,j}(bw+v)\,\chi_{x,j}(bw+v) - \psi_{x,j}(v)\,\chi_{x,j}(v) \\ &- \left(\psi'_{x,j}(v)\,\chi_{x,j}(v) + \psi_{x,j}(v)\,\chi'_{x,j}(v)\right)(bw)\}\,\mathrm{d}w = \int K'(w)\,K(w)\,w\left\{\left(\psi'_{x,j}(\dot{v})\,\chi_{x,j}(\dot{v}) + \psi_{x,j}(\dot{v})\,\chi'_{x,j}(\dot{v})\right)\right. \\ &\left. - \left(\psi'_{x,j}(v)\,\chi_{x,j}(v) + \psi_{x,j}(v)\,\chi'_{x,j}(v)\right)\right\}\mathrm{d}w = o(1), \end{aligned} \tag{S30}$$

where the mean value $\dot{v}$ depends on w and satisfies $|\dot{v}-v| \le b\,|w|$ and the second equality holds uniformly in $(v,b) \in I_x \times \mathbb{H}$. Similarly,

$$\begin{aligned} &\int K'(w)\,\psi_{x,j}(bw+v)\left(\int_{-\infty}^{w} K(u)\,\chi'_{x,j}(bu+v)\,\mathrm{d}u\right)\mathrm{d}w - \psi_{x,j}(v)\,\chi'_{x,j}(v)\int K'(w)\left(\int_{-\infty}^{w} K(u)\,\mathrm{d}u\right)\mathrm{d}w \\ &= \int\int_{-\infty}^{w} K'(w)\,K(u)\left\{\psi_{x,j}(bw+v)\,\chi'_{x,j}(bu+v) - \psi_{x,j}(v)\,\chi'_{x,j}(v)\right\}\mathrm{d}u\mathrm{d}w = o(1), \end{aligned} \tag{S31}$$

the second equality holds uniformly in $(v,b) \in I_x \times \mathbb{H}$. By integration by parts, $\int K'(w)\left(\int_{-\infty}^{w} K(u)\,\mathrm{d}u\right)\mathrm{d}w = -\int K(u)^2\,\mathrm{d}u$ and $\int K'(w)\,K(w)\,w\mathrm{d}w = \left(-\int K(u)^2\,\mathrm{d}u\right)/2$. Now it follows from these equalities, (S28), (S29), (S30), (S31) and $\int K'(w)\,K(w)\,\mathrm{d}w = 0$ that

$$\begin{aligned} b^{-3}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} K'\left(\frac{\mathit{\Delta}_x(e)-v}{b}\right)K'\left(\frac{\mathit{\Delta}_x(e')-v}{b}\right)F_{\epsilon|X}(e\wedge e' \mid x)\,\rho_x(e)\,\rho_x(e')\,\mathrm{d}e\mathrm{d}e' = \\ 2\left(\psi'_{x,j}(v)\,\chi_{x,j}(v) + \psi_{x,j}(v)\,\chi'_{x,j}(v)\right)\int K'(w)\,K(w)\,w\mathrm{d}w \\ -2\psi_{x,j}(v)\,\chi'_{x,j}(v)\int K'(w)\left(\int_{-\infty}^{w} K(u)\,\mathrm{d}u\right)\mathrm{d}w + o(1) = \left(\psi_{x,j}(v)\,\chi'_{x,j}(v) - \psi'_{x,j}(v)\,\chi_{x,j}(v)\right)\int K(u)^2\,\mathrm{d}u + o(1), \end{aligned} \tag{S32}$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$. Note that

$$\begin{aligned} &\psi_{x,j}(v)\,\chi'_{x,j}(v) - \psi'_{x,j}(v)\,\chi_{x,j}(v) \\ &= \left\{\frac{f_{\epsilon DX}\left(\mathit{\Delta}_{x,j}^{-1}(v),0,x\right)}{\zeta_{1x}\left(g\left(1,x,\mathit{\Delta}_{x,j}^{-1}(v)\right)\right)} - \frac{f_{\epsilon DX}\left(\mathit{\Delta}_{x,j}^{-1}(v),1,x\right)}{\zeta_{0x}\left(g\left(0,x,\mathit{\Delta}_{x,j}^{-1}(v)\right)\right)}\right\}^2 f_{\epsilon|X}\left(\mathit{\Delta}_{x,j}^{-1}(v)\mid x\right)\left(\left(\mathit{\Delta}_{x,j}^{-1}\right)'(v)\right)^3. \end{aligned}$$

Now it follows from this equality, (S19), (S20), (S21), (S22), (S23), (S27) and (S32) that

$$\mathrm{E}\left[\mathcal{H}_x^{[1]}(U,v;b)^2\right] = b^{-1}\mathscr{V}^{\ddagger}_{\mathit{\Delta}X}(v,x) + o\left(h^{-1}\right),$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$. Denote $\mathfrak{H} := \{\mathcal{H}_x(\cdot,v;b) : (v,b) \in I_x \times \mathbb{H}\}$ and $\mathfrak{H}^{[1]} := \left\{\mathcal{H}_x^{[1]}(\cdot,v;b) : (v,b) \in I_x \times \mathbb{H}\right\}$. Therefore, we have $\sigma^2_{\mathfrak{H}^{[1]}} := \sup_{f\in\mathfrak{H}^{[1]}}\mathbb{P}^U f^2 = O\left(h^{-1}\right)$. By LIE,

$$\mathrm{E}\left[\mathcal{H}_x(U_1,U_2,v;b)^2\right] =$$

$$\mathrm{E}\left[\frac{1}{b^4}K'\left(\frac{\Delta_x(\epsilon)-v}{b}\right)^2\left\{\frac{\mathbb{1}(D=0,X=x)}{\zeta_{1x}(g(1,x,\epsilon))^2}+\frac{\mathbb{1}(D=1,X=x)}{\zeta_{1x}(g(0,x,\epsilon))^2}\right\}F_{\epsilon|X}(\epsilon\mid x)\left(1-F_{\epsilon|X}(\epsilon\mid x)\right)\right]\left(p_{1x}^{-1}+p_{0x}^{-1}\right)$$
$$=O\left(h^{-3}\right), \quad \text{(S33)}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Therefore, $\sigma_{\mathfrak{H}}^2 \coloneqq \sup_{f\in\mathfrak{H}}\mathrm{E}\left[f(U_1,U_2)^2\right]=O\left(h^{-3}\right)$.

Since $\|K''\|_\infty<\infty$, K' is of bounded variation. There exists a decomposition $K'=K_1-K_2$, where K_1 and K_2 are non-decreasing and bounded. It follows from Kosorok (2007, Lemma 9.6) that the function class $\{(\Delta_x(\cdot)-v)/b:(v,b)\in I_x\times\mathbb{H}\}$ is VC-subgraph with VC index being at most 4. Then, by Kosorok (2007, Lemma 9.9(viii)), $\mathfrak{C}_k\coloneqq\{K_k((\Delta_x(\cdot)-v)/b):(v,b)\in I_x\times\mathbb{H}\}$ is VC-subgraph with VC index being at most 4. By Giné and Nickl (2016, Theorem 3.6.9) and Chernozhukov et al. (2014a, Lemma B.2), $\mathfrak{C}\coloneqq\{K'((\Delta_x(\cdot)-v)/b):(v,b)\in I_x\times\mathbb{H}\}$ is uniformly VC-type with respect to a constant envelope $\|K_1\|_\infty+\|K_2\|_\infty$, since $\mathfrak{C}$ can be written as (a sub-class of) the pointwise difference of $\mathfrak{C}_1$ and $\mathfrak{C}_2$. By Kosorok (2007, Lemma 9.6), $\left\{b^{-2}\mathcal{C}_x:b\in\mathbb{H}\right\}$ is VC-subgraph with VC index being at most 3 and by Giné and Nickl (2016, Theorem 3.6.9), it is uniformly VC-type with respect to a constant envelope that is a multiple of $\underline{h}^{-2}$. By Chernozhukov et al. (2014a, Lemma B.2), $\mathfrak{H}$ is uniformly VC-type with respect to a constant envelope $F_{\mathfrak{H}}=O\left(h^{-2}\right)$.

For a centered VC-type class $\mathfrak{F}$ of functions defined on $\mathscr{S}_U$ with respect to an envelope $F_{\mathfrak{F}}$ ($\mathbb{P}^Uf=0$, $\forall f\in\mathfrak{F}$), by standard calculations (see, e.g., the proof of Chernozhukov et al., 2014b, Corollary 5.1) and Chernozhukov et al. (2014b, Lemma 2.1), there exists a zero-mean Gaussian process $\left\{G^U(f):f\in\mathfrak{F}\right\}$ that is a tight random element in $\ell^\infty(\mathfrak{F})$ and has the same covariance structure as the empirical process $\left\{\mathbb{G}_n^Uf:f\in\mathfrak{F}\right\}$ (i.e., $\mathrm{E}\left[G^U(f_1)G^U(f_2)\right]=\mathrm{Cov}\left[f_1(U),f_2(U)\right]$, $\forall f_1,f_2\in\mathfrak{F}$). The tightness of $\left\{G^U(f):f\in\mathfrak{F}\right\}$ is equivalent to the condition that $\mathfrak{F}$ (endowed with the intrinsic pseudo metric $\mathfrak{F}\times\mathfrak{F}\ni(f_1,f_2)\mapsto\sqrt{\mathrm{E}\left[\left(G^U(f_1)-G^U(f_2)\right)^2\right]}=\|f_1-f_2\|_{\mathbb{P}^U,2}$) is totally bounded and almost surely the sample paths $f\mapsto G^U(f)$ are uniformly continuous with respect to the intrinsic pseudo metric (therefore, $\Pr\left[\left\|G^U\right\|_{\mathfrak{F}}<\infty\right]=1$). And $\left\{G^U(f):f\in\mathfrak{F}\right\}$ is also separable as a stochastic process. See Kosorok (2007, Lemmas 7.2 and 7.4). Since $F_{\mathfrak{F}}$ is also an envelope of $\mathfrak{F}_\pm\coloneqq\mathfrak{F}\cup(-\mathfrak{F})$ and the covering number of $\mathfrak{F}_\pm$ is at most twice that of $\mathfrak{F}$, $\mathfrak{F}_\pm$ is VC-type with respect to $F_{\mathfrak{F}}$ and therefore, there exists a zero-mean Gaussian process $\left\{G^U(f):f\in\mathfrak{F}_\pm\right\}$ that is a tight random element in $\ell^\infty(\mathfrak{F}_\pm)$ and has the same covariance structure as that of $\left\{\mathbb{G}_n^Uf:f\in\mathfrak{F}_\pm\right\}$. Moreover, by Giné and Nickl (2016, Theorem 3.7.28), almost surely the sample paths $\mathfrak{F}_\pm\ni f\mapsto G^U(f)$ are prelinear and therefore, almost surely, $\forall f\in\mathfrak{F}$, $G^U(f)+G^U(-f)=0$, and $\sup_{f\in\mathfrak{F}_\pm}G^U(f)=\left\|G^U\right\|_{\mathfrak{F}}$. In our first application, $\mathfrak{F}=\mathfrak{H}^{[1]}$, which is also uniformly VC-type (see Lemma A.3 of CK) with respect to the constant envelope $F_{\mathfrak{H}^{[1]}}=F_{\mathfrak{H}}$.

By the coupling theorem (CK Proposition 2.1 with $\mathcal{H}=\mathfrak{H}_\pm$, $\overline{\sigma}_{\mathfrak{g}}=\sigma_{\mathfrak{H}^{[1]}}$, $\sigma_{\mathfrak{h}}=\sigma_{\mathfrak{H}}$, $b_{\mathfrak{g}}=b_{\mathfrak{h}}=F_{\mathfrak{H}}$, $\chi_n=0$, $\gamma=\sqrt{\log(n)/(nh^3)}$ and $q=\infty$), one can construct a coupling $\left(\left\|\mathbb{U}_n^{(2)}\right\|_{\mathfrak{H}},Z_{\mathfrak{H}_\pm}\right)$ that satisfies the following conditions: $Z_{\mathfrak{H}_\pm}=_d\sup_{f\in\mathfrak{H}_\pm^{[1]}}G^U(f)=\left\|G^U\right\|_{\mathfrak{H}^{[1]}}$, where $\left\{G^U(f):f\in\mathfrak{H}_\pm^{[1]}\right\}$ has zero mean and the same covariance structure as that of the (Hájek) empirical process $\left\{\mathbb{G}_nf:f\in\mathfrak{H}_\pm^{[1]}\right\}$, and $\left(\left\|\mathbb{U}_n^{(2)}\right\|_{\mathfrak{H}},Z_{\mathfrak{H}_\pm}\right)$ satisfies $\left\|\mathbb{U}_n^{(2)}\right\|_{\mathfrak{H}}-Z_{\mathfrak{H}_\pm}=O_p^\star\left(\sqrt{\log(n)/h},\sqrt{\log(n)/(nh^3)}\right)$. By Dudley's entropy integral bound (Giné and Nickl, 2016, Theorem 2.3.7), Lemma A.2 of CK, (36) and standard calculations (see, e.g., calculations in the proof of Chernozhukov et al., 2014b, Corollary 5.1),

$$\begin{aligned}
\mathrm{E}\left[\left\|G^U\right\|_{\mathfrak{H}^{[1]}}\right] &\lesssim \int_0^{\sigma_{\mathfrak{H}^{[1]}}\vee n^{-1/2}\left\|F_{\mathfrak{H}^{[1]}}\right\|_{\mathbb{P}^U,2}}\sqrt{1+\log\left(N\left(\varepsilon,\mathfrak{H}^{[1]},\|\cdot\|_{\mathbb{P}^U,2}\right)\right)}\mathrm{d}\varepsilon\\
&\leq 4\left\|F_{\mathfrak{H}^{[1]}}\right\|_{\mathbb{P}^U,2}\int_0^{\frac{\sigma_{\mathfrak{H}^{[1]}}\vee n^{-1/2}\left\|F_{\mathfrak{H}^{[1]}}\right\|_{\mathbb{P}^U,2}}{4\left\|F_{\mathfrak{H}^{[1]}}\right\|_{\mathbb{P}^U,2}}}\sqrt{1+\log\left(\sup_{Q\in\mathcal{Q}_{\mathscr{S}_U}^{\mathsf{fd}}}N\left(\varepsilon\left\|F_{\mathfrak{H}^{[1]}}\right\|_{Q,2},\mathfrak{H}^{[1]},\|\cdot\|_{Q,2}\right)\right)}\mathrm{d}\varepsilon
\end{aligned}$$

$$\lesssim \left(\sigma_{\mathfrak{H}^{[1]}} \vee n^{-1/2} \left\|F_{\mathfrak{H}^{[1]}}\right\|_{\mathbb{P}^U,2}\right) \sqrt{V_{\mathfrak{H}^{[1]}} \log\left(4A_{\mathfrak{H}^{[1]}} n^{1/2}\right)}, \tag{S34}$$

when n is sufficiently large. By the Borell-Sudakov-Tsirelson inequality (Giné and Nickl, 2016, Theorem 2.5.8), $\Pr\left[\left\|G^U\right\|_{\mathfrak{H}^{[1]}} \geq \mathrm{E}\left[\left\|G^U\right\|_{\mathfrak{H}^{[1]}}\right] + \sqrt{2 \cdot \log(n)}\sigma_{\mathfrak{H}^{[1]}}\right] \leq n^{-1}$. Therefore, since $Z_{\mathfrak{H}_\pm} =_d \left\|G^U\right\|_{\mathfrak{H}^{[1]}}$, we have $Z_{\mathfrak{H}_\pm} = O_p^\star\left(\sqrt{\log(n)/h}\right)$ and $\left\|\mathbb{U}_n^{(2)}\right\|_{\mathfrak{H}} = O_p^\star\left(\sqrt{\log(n)/h}, \sqrt{\log(n)/(nh^3)}\right)$. The assertion follows from these results. ∎

Proof of Lemma 5. Denote $\mathcal{H}_x^{\Delta[1]}(u,v;b,h) := \mathrm{E}\left[\mathcal{H}_x^\Delta(U,u,v;b,h)\right]$, $\mathfrak{H}^\Delta := \left\{\mathcal{H}_x^\Delta(\cdot,v;b,h) : (v,b) \in I_x \times \mathbb{H}\right\}$ and $\mathfrak{H}^{\Delta[1]} := \left\{\mathcal{H}_x^{\Delta[1]}(\cdot,v;b,h) : (v,b) \in I_x \times \mathbb{H}\right\}$. Then, $\left\|\mathbb{U}_n^{(2)}\right\|_{\mathfrak{H}^\Delta} = \sup_{(v,b)\in I_x\times\mathbb{H}} \left|\sqrt{n}\left(n_{(2)}^{-1}\sum_{(i,j)} \mathcal{H}_x^\Delta(U_i,U_j,v;b,h)\right)\right|$. By the same arguments for showing that $\mathfrak{H}$ is uniformly VC-type, $\left\{\sqrt{b}\cdot\mathcal{H}_x(\cdot,v;b) : (v,b) \in I_x \times \mathbb{H}\right\}$ uniformly VC-type with respect to a constant envelope that is $O\left(h^{-3/2}\right)$. By Chernozhukov et al. (2014a, Lemma B.2), $\mathfrak{H}^\Delta$ is uniformly VC-type with respect to a constant envelope that is $O\left(h^{-3/2}\right)$. We now use different arguments to show that $\mathfrak{H}^\Delta$ is uniformly VC-type with respect to a tighter envelope. Let $L(u;b,h) := (h/b)^{1/2} K((h/b)u) - K(u)$ and denote $L'(u;b,h) := \partial L(u;b,h)/\partial u$. Then it is clear that $\mathcal{H}_x^\Delta(U_i,U_j,v;b,h) = h^{-3/2} L'((\Delta_x(\epsilon_i) - v)/h;b,h)\,\mathcal{C}_x(U_i,U_j)$. Write $L'(u;b,h) = L_\dagger(u;b,h) + L_\ddagger(u;b,h)$, where $L_\dagger(u;b,h) := \left((h/b)^{3/2} - 1\right) K'((h/b)u)$ and $L_\ddagger(u;b,h) := K'((h/b)u) - K'(u)$. Then,

$$\begin{aligned}\mathcal{H}_x^\Delta(U_i,U_j,v;b,h) &= h^{-3/2} L_\dagger\left(\frac{\Delta_x(\epsilon_i) - v}{h};b,h\right)\mathcal{C}_x(U_i,U_j) + h^{-3/2} L_\ddagger\left(\frac{\Delta_x(\epsilon_i) - v}{h};b,h\right)\mathcal{C}_x(U_i,U_j)\\ &=: \mathcal{H}_x^\dagger(U_i,U_j,v;b,h) + \mathcal{H}_x^\ddagger(U_i,U_j,v;b,h).\end{aligned}$$

It follows from the fact that $\left\{\sqrt{b}\cdot\mathcal{H}_x(\cdot,v;b) : (v,b) \in I_x \times \mathbb{H}\right\}$ uniformly VC-type with respect to a constant envelope that is $O\left(h^{-3/2}\right)$, Chernozhukov et al. (2014a, Lemma B.2) and $\sup_{b\in\mathbb{H}} (b/h)^{3/2}\left|(h/b)^{3/2} - 1\right| = O(\varepsilon_n)$ that $\left\{\mathcal{H}_x^\dagger(\cdot,v;b,h) : (v,b) \in I_x \times \mathbb{H}\right\}$ is uniformly VC-type with respect to a constant envelope that is $O\left(\varepsilon_n/h^{3/2}\right)$. Denote $\mathfrak{V} := \left\{y \mapsto L_\ddagger((y-v)/h;b,h) : b \in \mathbb{H}, v \in I_x\right\}$ and $\mathfrak{V}(b) := \left\{y \mapsto L_\ddagger((y-v)/h;b,h) : v \in I_x\right\}$ so that $\mathfrak{V} = \bigcup_{b\in\mathbb{H}} \mathfrak{V}(b)$. We have $\left\|L_\ddagger(\cdot;b_1,h) - L_\ddagger(\cdot;b_2,h)\right\|_\infty \leq C_{1,n}|b_1 - b_2|$, where $C_{1,n} := \|K''\|_\infty\left(\overline{h}/\underline{h}^2\right)$. Clearly, $C_{1,n}\iota(\mathbb{H}) = O(\varepsilon_n)$. Let $\mathfrak{U} := \left\{u \mapsto L_\ddagger(u;b,h) : b \in \mathbb{H}\right\}$. Then, $N\left(C_{1,n}\varepsilon, \mathfrak{U}, \|\cdot\|_\infty\right) \leq N(\varepsilon, \mathbb{H}, |\cdot|) \leq 1 + \iota(\mathbb{H})/\varepsilon$ and therefore, $N\left(C_{1,n}\iota(\mathbb{H})\varepsilon, \mathfrak{U}, \|\cdot\|_\infty\right) \leq 1 + \varepsilon^{-1}$. By simple calculation, we have $\left\|L_\ddagger'(\cdot;\cdot,h)\right\|_{\mathbb{R}\times\mathbb{H}} \leq C_{2,n}\varepsilon_n$, where $C_{2,n} := \|K''\|_\infty/(1-\varepsilon_n) + C_{\mathsf{Lip}}(1+\varepsilon_n)/(1-\varepsilon_n)$ and C_{Lip} is the Lipschitz constant for K''. Therefore, the total variation (see Giné and Nickl, 2016, Page 220 for the definition) of $L_\ddagger(\cdot;b,h)$ is bounded by $2(1+\varepsilon_n)C_{2,n}\varepsilon_n$, uniformly in $b \in \mathbb{H}$. By Giné and Nickl (2016, Proposition 3.6.12), for each $b \in \mathbb{H}$,

$$\sup_{Q\in\mathcal{Q}_{\mathbb{R}}^{\mathsf{fd}}} N\left(2(1+\varepsilon_n)C_{2,n}\varepsilon_n\varepsilon, \mathfrak{V}(b), \|\cdot\|_{Q,2}\right) \leq \left(\frac{A_{\mathfrak{V}}}{\varepsilon}\right)^{V_{\mathfrak{V}}},$$

where $(A_{\mathfrak{V}}, V_{\mathfrak{V}})$ are independent of b, h and n. Now fix some finitely discrete probability measure Q on $\mathbb{R}$. Construct a minimal $C_{1,n}\iota(\mathbb{H})\varepsilon$-net (with respect to $\|\cdot\|_\infty$) $\left\{u \mapsto L_\ddagger(u;b,h) : b \in \mathbb{H}^\circ\right\}$ for $\mathfrak{U}$, where $\mathbb{H}^\circ \subseteq \mathbb{H}$ and $\#\mathbb{H}^\circ \leq 1 + \varepsilon^{-1}$. For each $b \in \mathbb{H}^\circ$, construct a minimal $2(1+\varepsilon_n)C_{2,n}\varepsilon_n\varepsilon$-net (with respect to $\|\cdot\|_{Q,2}$) $\left\{y \mapsto L_\ddagger((y-v)/h;b,h) : v \in I_x^\circ(b)\right\}$ for $\mathfrak{V}(b)$, where $I_x^\circ(b) \subseteq I_x$ and $\#I_x^\circ(b) \leq (A_{\mathfrak{V}}/\varepsilon)^{V_{\mathfrak{V}}}$. For any (v_1,b_1) and (v_2,b_2), by the triangle inequality,

$$\begin{aligned}\sqrt{\int\left(L_\ddagger\left(\frac{y-v_1}{h};b_1,h\right) - L_\ddagger\left(\frac{y-v_2}{h};b_2,h\right)\right)^2 Q(\mathrm{d}y)} &\leq \left\|L_\ddagger(\cdot;b_1,h) - L_\ddagger(\cdot;b_2,h)\right\|_\infty\\ &\quad + \sqrt{\int\left(L_\ddagger\left(\frac{y-v_1}{h};b_2,h\right) - L_\ddagger\left(\frac{y-v_2}{h};b_2,h\right)\right)^2 Q(\mathrm{d}y)}.\end{aligned}$$

Then, $\bigcup_{b\in\mathbb{H}^\circ}\{y\mapsto L_\ddagger((y-v)/h;b,h):v\in I_x^\circ(b)\}$ is a $(C_{1,n}\iota(\mathbb{H})+2(1+\varepsilon_n)C_{2,n}\varepsilon_n)\,\varepsilon$-net (with respect to $\|\cdot\|_{Q,2}$) for $\mathfrak{V}$ and has a cardinality bounded by $(A_{\mathfrak{V}}/\varepsilon)^{V_{\mathfrak{V}}}(1+\varepsilon^{-1})$. Note that $F_{\mathfrak{V}}\coloneqq C_{1,n}\iota(\mathbb{H})+2(1+\varepsilon_n)C_{2,n}\varepsilon_n=O(\varepsilon_n)$ is also a constant envelope for $\mathfrak{V}$. Therefore, $\mathfrak{V}$ is uniformly VC-type with respect to the constant envelope $F_{\mathfrak{V}}=O(\varepsilon_n)$. Then it is easy to see that $\left\{\mathcal{H}_x^\ddagger(\cdot,v;b,h):(v,b)\in I_x\times\mathbb{H}\right\}$ is uniformly VC-type with respect to a constant envelope that is $O\left(\varepsilon_n/h^{3/2}\right)$. By Chernozhukov et al. (2014a, Lemma B.2), $\mathfrak{H}^\Delta$ is uniformly VC-type with respect to a constant envelope $F_{\mathfrak{H}^\Delta}=O\left(\varepsilon_n/h^{3/2}\right)$.

By LIE, we have

$$\begin{aligned}&\mathrm{E}\left[\mathcal{H}_x^{\Delta[1]}(U,v;b,h)^2\right]\\&\quad=\mathrm{E}\left[\frac{1}{h^3}\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)\rho_x(e)\left\{\mathbb{1}(\epsilon\le e)-F_{\epsilon|X}(e\mid x)\right\}\mathrm{d}e\right\}^2\mid X=x\right]\left(p_{1x}^{-1}+p_{0x}^{-1}\right).\end{aligned}$$

Then,

$$\begin{aligned}&\mathrm{E}\left[\frac{1}{h^3}\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)\rho_x(e)\left\{\mathbb{1}(\epsilon\le e)-F_{\epsilon|X}(e\mid x)\right\}\mathrm{d}e\right\}^2\mid X=x\right]=\\&\qquad\mathrm{E}\left[\frac{1}{h^3}\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)\rho_x(e)\mathbb{1}(\epsilon\le e)\,\mathrm{d}e\right\}^2\mid X=x\right]\\&\qquad\qquad-\frac{1}{h^3}\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)\rho_x(e)F_{\epsilon|X}(e\mid x)\,\mathrm{d}e\right\}^2.\end{aligned}$$

By change of variables and mean value expansion,

$$\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)\rho_x(e)\,\mathrm{d}e\;=\;h\int_{\frac{\Delta_x(\epsilon_{x,j-1})-v}{h}}^{\frac{\Delta_x(\epsilon_{x,j})-v}{h}}L'(u;b,h)\left\{\psi_{x,j}(v)+\psi'_{x,j}(\dot{v})(hu)\right\}\mathrm{d}u,$$

where $\dot{v}$ denotes the mean value that lies between v and $v+hu$. It is easy to check by simple calculation that $\forall b\in\mathbb{H}$,

$$|L'(u;b,h)|\le\left(\|K''\|_\infty\left(\frac{1+\varepsilon_n}{(1-\varepsilon_n)^{5/2}}\right)+\frac{3\|K'\|_\infty}{(1-\varepsilon_n)^3}\right)\varepsilon_n\mathbb{1}(|u|\le1+\varepsilon_n).\tag{S35}$$

Then by these results and the fact $\int L'(u;b,h)\,\mathrm{d}u=0$, we have

$$\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)\rho_x(e)\,\mathrm{d}e=O\left(\varepsilon_nh^2\right),$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Similarly,

$$\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)F_{\epsilon|X}(e'\mid x)\rho_x(e')\,\mathrm{d}e'=O\left(\varepsilon_nh^2\right),$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Then by these results, we have

$$h^{-3}\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)L'\left(\frac{\Delta_x(e')-v}{h};b,h\right)F_{\epsilon|X}(e\wedge e'\mid x)\rho_x(e)\rho_x(e')\,\mathrm{d}e\mathrm{d}e'$$

$$= h^{-3}\left(\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} L'\left(\frac{\Delta_x(e)-v}{h};b,h\right) F_{\epsilon|X}(e\mid x)\,\rho_x(e)\,\mathrm{d}e\right)\left(\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} L'\left(\frac{\Delta_x(e')-v}{h};b,h\right)\rho_x(e')\,\mathrm{d}e'\right) = O\left(\varepsilon_n^2 h\right), \tag{S36}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$, if $j<k$. The same result holds if $j>k$. By change of variables,

$$h^{-3}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)L'\left(\frac{\Delta_x(e')-v}{h};b,h\right)F_{\epsilon|X}(e\wedge e'\mid x)\,\rho_x(e)\,\rho_x(e')\,\mathrm{d}e\mathrm{d}e' =$$
$$2h^{-1}\int L'(w;b,h)\,\psi_{x,j}(hw+v)\left\{\int_{-\infty}^{w} L'(u;b,h)\,\chi_{x,j}(hu+v)\,\mathrm{d}u\right\}\mathrm{d}w.$$

By integration by parts,

$$\int_{-\infty}^{w} L'(u;b,h)\,\chi_{x,j}(hu+v)\,\mathrm{d}u = L(w;b,h)\,\chi_{x,j}(hw+v) - b\int_{-\infty}^{w} L(u;b,h)\,\chi'_{x,j}(hu+v)\,\mathrm{d}u.$$

By these results, we have

$$h^{-3}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)L'\left(\frac{\Delta_x(e')-v}{h};b,h\right)F_{\epsilon|X}(e\wedge e'\mid x)\,\rho_x(e)\,\rho_x(e')\,\mathrm{d}e\mathrm{d}e' =$$
$$2h^{-1}\int L'(w;b,h)\,\psi_{x,j}(hw+v)\,L(w;b,h)\,\chi_{x,j}(hw+v)\,\mathrm{d}w$$
$$-\,2\int L'(w;b,h)\,\psi_{x,j}(hw+v)\int_{-\infty}^{w} L(u;b,h)\,\chi'_{x,j}(hu+v)\,\mathrm{d}u\mathrm{d}w. \tag{S37}$$

It is easy to check by simple calculation that $\forall b\in\mathbb{H}$,

$$|L(u;b,h)| \le \left(\|K'\|_\infty\left(\frac{1+\varepsilon_n}{1-\varepsilon_n}\right)+\frac{\|K\|_\infty}{1-\varepsilon_n}\right)\varepsilon_n\mathbb{1}\left(|u|\le 1+\varepsilon_n\right). \tag{S38}$$

By (S35), (S38) and mean value expansion,

$$h^{-1}\int L'(w;b,h)\,L(w;b,h)\left\{\psi_{x,j}(hw+v)\,\chi_{x,j}(hw+v)-\psi_{x,j}(v)\,\chi_{x,j}(v)\right.$$
$$\left.-\left(\psi'_{x,j}(v)\,\chi_{x,j}(v)+\psi_{x,j}(v)\,\chi'_{x,j}(v)\right)(hw)\right\}\mathrm{d}w = \int L'(w;b,h)\,L(w;b,h)\,w\left\{\left(\psi'_{x,j}(\dot v)\,\chi_{x,j}(\dot v)+\psi_{x,j}(\dot v)\,\chi'_{x,j}(\dot v)\right)\right.$$
$$\left.-\left(\psi'_{x,j}(v)\,\chi_{x,j}(v)+\psi_{x,j}(v)\,\chi'_{x,j}(v)\right)\right\}\mathrm{d}w = o\left(\varepsilon_n^2\right),$$

where $\dot v$ denotes the mean value that lies between v and $v+hw$ and the second equality holds uniformly in $(v,b)\in I_x\times\mathbb{H}$. Similarly,

$$\int L'(w;b,h)\,\psi_{x,j}(hw+v)\left(\int_{-\infty}^{w} L(u;b,h)\,\chi'_{x,j}(hu+v)\,\mathrm{d}u\right)\mathrm{d}w$$
$$-\,\psi_{x,j}(v)\,\chi'_{x,j}(v)\int L'(w;b,h)\left(\int_{-\infty}^{w} L(u;b,h)\,\mathrm{d}u\right)\mathrm{d}w$$
$$=\int\int_{-\infty}^{w} L'(w;b,h)\,L(u;b,h)\left\{\psi_{x,j}(hw+v)\,\chi'_{x,j}(hu+v)-\psi_{x,j}(v)\,\chi'_{x,j}(v)\right\}\mathrm{d}u\mathrm{d}w = O\left(\varepsilon_n^2 h\right),$$

where the second equality holds uniformly in $(v,b)\in I_x\times\mathbb{H}$. By integration by parts,

$$\int L'(w;b,h)\left(\int_{-\infty}^{w} L(u;b,h)\,\mathrm{d}u\right)\mathrm{d}w \quad = \quad -\int L(u;b,h)^2\,\mathrm{d}u$$

$$\begin{aligned}\int L'(w;b,h)\,L(w;b,h)\,w\mathrm{d}w &= -\frac{1}{2}\int L(u;b,h)^2\,\mathrm{d}u\\ \int L'(w;b,h)\,L(w;b,h)\,\mathrm{d}w &= 0.\end{aligned}\tag{S39}$$

By using these calculations, (S37) and (S38), we have

$$h^{-3}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)L'\left(\frac{\Delta_x(e')-v}{h};b,h\right)F_{\epsilon|X}(e\wedge e'\mid x)\,\rho_x(e)\,\rho_x(e')\,\mathrm{d}e\mathrm{d}e' = O\left(\varepsilon_n^2\right).$$

Then it follows from this result and (S36) that

$$\sigma^2_{\mathfrak{H}^{\Delta[1]}} \coloneqq \sup_{f\in\mathfrak{H}^{\Delta[1]}}\mathbb{P}^U f^2 = \sup_{(v,b)\in I_x\times\mathbb{H}} \mathrm{E}\left[\mathcal{H}_x^{\Delta[1]}(U,v;b,h)^2\right] = O\left(\varepsilon_n^2\right).$$

By LIE,

$$\begin{aligned}&\mathrm{E}\left[\mathcal{H}_x^{\Delta}(U_1,U_2,v;b)^2\right] =\\ &\mathrm{E}\left[\frac{1}{h^4}L'\left(\frac{\Delta_x(\epsilon)-v}{h};b,h\right)^2\left\{\frac{\mathbb{1}(D=0,X=x)}{\zeta_{1x}(g(1,x,\epsilon))^2}+\frac{\mathbb{1}(D=1,X=x)}{\zeta_{1x}(g(0,x,\epsilon))^2}\right\}F_{\epsilon|X}(\epsilon\mid x)\left(1-F_{\epsilon|X}(\epsilon\mid x)\right)\right]\left(p_{1x}^{-1}+p_{0x}^{-1}\right).\end{aligned}$$

By change of variables and (S35), we have

$$\sigma^2_{\mathfrak{H}^{\Delta}} \coloneqq \sup_{f\in\mathfrak{H}^{\Delta}}\mathrm{E}\left[f(U_1,U_2)^2\right] = \sup_{(v,b)\in I_x\times\mathbb{H}}\mathrm{E}\left[\mathcal{H}_x^{\Delta}(U_1,U_2,v;b,h)^2\right] = O\left(\varepsilon_n^2/h^3\right).$$

By the coupling theorem (CK Proposition 2.1 with $\mathcal{H}=\mathfrak{H}^{\Delta}_{\pm}$, $\overline{\sigma}_{\mathfrak{g}}=\sigma_{\mathfrak{H}^{\Delta[1]}}$, $\sigma_{\mathfrak{h}}=\sigma_{\mathfrak{H}^{\Delta}}$, $b_{\mathfrak{g}}=b_{\mathfrak{h}}=F_{\mathfrak{H}^{\Delta}}$, $\chi_n=0$, $\gamma=\sqrt{\log(n)/(nh^3)}$ and $q=\infty$), there exists a random variable $Z_{\mathfrak{H}^{\Delta}_{\pm}}=_d\left\|G^U\right\|_{\mathfrak{H}^{\Delta[1]}}$ such that $\left\|\mathbb{U}_n^{(2)}\right\|_{\mathfrak{H}^{\Delta}}-Z_{\mathfrak{H}^{\Delta}_{\pm}}=O_p^{\star}\left(\varepsilon_n\sqrt{\log(n)},\sqrt{\log(n)/(nh^3)}\right)$. By (S34) with $\mathfrak{H}^{[1]}$ replaced by $\mathfrak{H}^{\Delta[1]}$, $\mathrm{E}\left[\left\|G^U\right\|_{\mathfrak{H}^{\Delta[1]}}\right]=O\left(\varepsilon_n\sqrt{\log(n)}\right)$. By the Borell-Sudakov-Tsirelson inequality, $\Pr\left[\left\|G^U\right\|_{\mathfrak{H}^{\Delta[1]}}\geq\mathrm{E}\left[\left\|G^U\right\|_{\mathfrak{H}^{\Delta[1]}}\right]+\sqrt{2\cdot\log(n)}\sigma_{\mathfrak{H}^{\Delta[1]}}\right]\leq n^{-1}$. Therefore, $\left\|\mathbb{U}_n^{(2)}\right\|_{\mathfrak{H}^{\Delta}}=O_p^{\star}\left(\varepsilon_n\sqrt{\log(n)},\sqrt{\log(n)/(nh^3)}\right)$. The first assertion follows from this result.

Let $\mathcal{E}_x^{\Delta}(U_i,v;b,h)\coloneqq\sqrt{b}\cdot\mathcal{E}_x(U_i,v;b)-\sqrt{h}\cdot\mathcal{E}_x(U_i,v;h)$ and $\mathfrak{E}^{\Delta}\coloneqq\{\mathcal{E}_x^{\Delta}(\cdot,v;b,h):(v,b)\in I_x\times\mathbb{H}\}$. Then we have

$$\sup_{(v,b)\in I_x\times\mathbb{H}}\left|\sqrt{nb}\left(\widetilde{f}_{\Delta X}(v,x;b)-m_{\Delta X}(v,x;b)\right)-\sqrt{nh}\left(\widetilde{f}_{\Delta X}(v,x;h)-m_{\Delta X}(v,x;h)\right)\right|=\left\|\mathbb{G}_n^U\right\|_{\mathfrak{E}^{\Delta}}$$

and we write $\mathcal{E}_x^{\Delta}(U_i,v;b,h)=h^{-1/2}L\left((\Delta_x(\epsilon_i)-v)/h;b,h\right)\mathbb{1}(X_i=x)$. It follows from similar arguments that $\mathfrak{E}^{\Delta}$ is uniformly VC-type with respect to a constant envelope $F_{\mathfrak{E}^{\Delta}}=O\left(\varepsilon_n/h^{1/2}\right)$. By LIE and (S38), $\sigma^2_{\mathfrak{E}^{\Delta}}\coloneqq\sup_{f\in\mathfrak{E}^{\Delta}}\mathbb{P}^U f^2=O\left(\varepsilon_n^2\right)$. By Talagrand's inequality (Chernozhukov et al., 2016, Lemma 6.3 with $\mathcal{F}=\mathfrak{E}^{\Delta}$, $b=F_{\mathfrak{E}^{\Delta}}$, $\sigma=\sigma_{\mathfrak{E}^{\Delta}}\vee b\sqrt{V_{\mathfrak{E}^{\Delta}}\log(n)/n}$, $t=\log(n)$), $\left\|\mathbb{G}_n^U\right\|_{\mathfrak{E}^{\Delta}}=O_p^{\star}\left(\varepsilon_n\sqrt{\log(n)}\right)$. The second assertion follows from this result. ■

S2 Proofs of Lemmas 6, 7, 8, 9 and Theorem B1

Proof of Lemma 6. Denote $\Pi_{dx}(y)\coloneqq\Pi_{d1x}(y)+\Pi_{d0x}(y)$ and $\widehat{\Pi}_{dx}(y)\coloneqq\widehat{\Pi}_{d1x}(y)+\widehat{\Pi}_{d0x}(y)$. Then, $\widehat{R}_{d'x}(y)=\left(\widehat{\Pi}_{dx}\left(\widehat{\phi}_{dx}(y)\right)+\widehat{\Pi}_{d'x}(y)\right)/\widehat{p}_x$ and $R_{d'x}(y)=\left(\Pi_{dx}(\phi_{dx}(y))+\Pi_{d'x}(y)\right)/p_x$. Part (a) follows from the following facts: $\left\|\widehat{\Pi}_{dx}\circ\widehat{\phi}_{dx}-\Pi_{dx}\circ\phi_{dx}\right\|_{I_{d'x}}=O_p^\star\left(\sqrt{\log(n)/n}\right)$, $\left\|\widehat{\Pi}_{d'x}-\Pi_{d'x}\right\|_{I_{d'x}}=O_p^\star\left(\sqrt{\log(n)/n}\right)$ and $\widehat{p}_x-p_x=O_p^\star\left(\sqrt{\log(n)/n}\right)$. Denote $p_{z|x}\coloneqq\Pr[Z=z\mid X=x]$. Part (b) follows from similar arguments used to show $\left\|\widetilde{f}_{\Delta X}(\cdot,x;\cdot)-m_{\Delta X}(\cdot,x;\cdot)\right\|_{I_x\times\mathbb{H}}=O_p^\star\left(\sqrt{\log(n)/(nh)}\right)$ in the proof of Theorem A1, $\widehat{p}_{z|x}-p_{z|x}=O_p^\star\left(\sqrt{\log(n)/n}\right)$, Abadie (2003, Theorem 3.1) and standard arguments for the bias of kernel density estimators. ■

Proof of Lemma 7. Denote

$$\begin{aligned}\widehat{r}_{\Delta X}(v,x;b)&\coloneqq\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b}K\left(\frac{\widehat{\Delta}_i-v}{b}\right)^2\mathbb{1}(X_i=x)\\ \widetilde{r}_{\Delta X}(v,x;b)&\coloneqq\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b}K\left(\frac{\Delta_i-v}{b}\right)^2\mathbb{1}(X_i=x)\end{aligned}$$

and $r_{\Delta X}(v,x;b)\coloneqq\mathrm{E}[\widetilde{r}_{\Delta X}(v,x;b)]$. Then, $\widehat{V}_1(v,x;b)=\widehat{r}_{\Delta X}(v,x;b)-b\cdot\widehat{f}_{\Delta X}(v,x;b)^2$ and $V_1(v,x;b)=r_{\Delta X}(v,x;b)-b\cdot m_{\Delta X}(v,x;b)^2$. It follows from arguments used in previous proofs (see the proofs of Lemmas 3 and 4 and Theorem A1) that

$$\left\|\widehat{r}_{\Delta X}(\cdot,x;\cdot)-\widetilde{r}_{\Delta X}(\cdot,x;\cdot)\right\|_{I_x\times\mathbb{H}}=O_p^\star\left(\sqrt{\frac{\log(n)}{nh}},\sqrt{\frac{\log(n)}{nh^3}}\right)$$

and $\widetilde{r}_{\Delta X}(v,x;b)-r_{\Delta X}(v,x;b)=O_p^\star\left(\sqrt{\log(n)/(nh)}\right)$, uniformly in $(v,b)\in I_x\times\mathbb{H}$. And also, by using the fact

$$\left\|\widehat{f}_{\Delta X}(\cdot,x;\cdot)-m_{\Delta X}(\cdot,x;\cdot)\right\|_{I_x\times\mathbb{H}}=O_p^\star\left(\sqrt{\frac{\log(n)}{nh}},\sqrt{\frac{\log(n)}{nh^3}}\right)$$

shown in the proofs of Lemmas 3 and 4 and Theorem A1 that

$$\begin{aligned}\widehat{f}_{\Delta X}(v,x;b)^2-m_{\Delta X}(v,x;b)^2&=\left(\widehat{f}_{\Delta X}(v,x;b)-m_{\Delta X}(v,x;b)\right)\left(\widehat{f}_{\Delta X}(v,x;b)+m_{\Delta X}(v,x;b)\right)\\&=O_p^\star\left(\sqrt{\frac{\log(n)}{nh}},\sqrt{\frac{\log(n)}{nh^3}}\right),\end{aligned}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Therefore,

$$\left\|\widehat{V}_1(\cdot,x;\cdot)-V_1(\cdot,x;\cdot)\right\|_{I_x\times\mathbb{H}}=O_p^\star\left(\sqrt{\frac{\log(n)}{nh}},\sqrt{\frac{\log(n)}{nh^3}}\right).\tag{S40}$$

Denote

$$\begin{aligned}\widetilde{V}_2(v,x;b,b_\zeta)&\coloneqq\frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\Delta}_j-v}{b}\right)\widehat{q}_x(W_j,W_i;b_\zeta)K'\left(\frac{\widehat{\Delta}_k-v}{b}\right)\widehat{q}_x(W_k,W_i;b_\zeta)\mathbb{1}(X_i=x)\\ \dot{V}_2(v,x;b)&\coloneqq\frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\Delta}_j-v}{b}\right)q_x(W_j,W_i)K'\left(\frac{\widehat{\Delta}_k-v}{b}\right)q_x(W_k,W_i)\mathbb{1}(X_i=x)\\ \ddot{V}_2(v,x;b)&\coloneqq\frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\Delta_j-v}{b}\right)q_x(W_j,W_i)K'\left(\frac{\Delta_k-v}{b}\right)q_x(W_k,W_i)\mathbb{1}(X_i=x).\end{aligned}$$

Therefore, $\widehat{V}_2(v,x;b,b_\zeta) = \widetilde{V}_2(v,x;b,b_\zeta)\widehat{p}_x^{-1}\left(\widehat{p_{1x}^{-1}} + \widehat{p_{0x}^{-1}}\right)$. Then,

$$\begin{aligned}
&\dot{V}_2(v,x;b) - \ddot{V}_2(v,x;b) = \\
&\quad \frac{2}{n_{(3)}} \sum_{(i,j,k)} \frac{1}{b^3} \left\{ K'\left(\frac{\widehat{\Delta}_j - v}{b}\right) - K'\left(\frac{\Delta_j - v}{b}\right) \right\} q_x(W_j, W_i) K'\left(\frac{\Delta_k - v}{b}\right) q_x(W_k, W_i) \mathbb{1}(X_i = x) \\
&\qquad + \frac{1}{n_{(3)}} \sum_{(i,j,k)} \frac{1}{b^3} \left\{ K'\left(\frac{\widehat{\Delta}_j - v}{b}\right) - K'\left(\frac{\Delta_j - v}{b}\right) \right\} q_x(W_j, W_i) \\
&\qquad \times \left\{ K'\left(\frac{\widehat{\Delta}_k - v}{b}\right) - K'\left(\frac{\Delta_k - v}{b}\right) \right\} q_x(W_k, W_i) \mathbb{1}(X_i = x) = \\
&\frac{2}{n_{(3)}} \sum_{(i,j,k)} \frac{1}{b^3} \left\{ K'\left(\frac{\widehat{\Delta}_j - v}{b}\right) - K'\left(\frac{\Delta_j - v}{b}\right) \right\} q_x(W_j, W_i) K'\left(\frac{\Delta_k - v}{b}\right) q_x(W_k, W_i) \mathbb{1}(X_i = x) + O_p^\star\left(\frac{\log(n)}{nh^3}\right),
\end{aligned} \tag{S41}$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$, where the second equality follows from the triangle inequality, mean value expansion, (S12), (S13) and (S15). By mean value expansion, the first term can be written as

$$\begin{aligned}
&\frac{2}{n_{(3)}} \sum_{(i,j,k)} \frac{1}{b^4} K''\left(\frac{\Delta_j - v}{b}\right) \left(\widehat{\Delta}_j - \Delta_j\right) q_x(W_j, W_i) K'\left(\frac{\Delta_k - v}{b}\right) q_x(W_k, W_i) \mathbb{1}(X_i = x) \\
&+ \frac{1}{n_{(3)}} \sum_{(i,j,k)} \frac{1}{b^4} \left(K''\left(\frac{\dot{\Delta}_j - v}{b}\right) - K''\left(\frac{\Delta_j - v}{b}\right) \right) \left(\widehat{\Delta}_j - \Delta_j\right) q_x(W_j, W_i) K'\left(\frac{\Delta_k - v}{b}\right) q_x(W_k, W_i) \mathbb{1}(X_i = x) \\
&=: 2 \cdot T_1(v;b) + T_2(v;b),
\end{aligned} \tag{S42}$$

where $\dot{\Delta}_j$ is the mean value that lies between $\widehat{\Delta}_j$ and Δ_j. By using

$$\Pr\left[\left|K''\left(\frac{\dot{\Delta}_i - v}{b}\right)\right| \mathbb{1}(X_i = x)(1 - \mathbb{1}_i(v;b)) = 0, \forall (i,v,b) \in \{1,...,n\} \times I_x \times \mathbb{H}\right] = 1 - O\left(n^{-1}\right), \tag{S43}$$

which follows from (S13), we have

$$\begin{aligned}
T_2(v;b) = \frac{1}{n_{(3)}} \sum_{(i,j,k)} \frac{1}{b^4} &\left(K''\left(\frac{\dot{\Delta}_j - v}{b}\right) - K''\left(\frac{\Delta_j - v}{b}\right) \right) \\
&\times \mathbb{1}_j(v;b) \left(\widehat{\Delta}_j - \Delta_j\right) q_x(W_j, W_i) K'\left(\frac{\Delta_k - v}{b}\right) q_x(W_k, W_i) \mathbb{1}(X_i = x),
\end{aligned}$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$, with probability at least $1 - O\left(n^{-1}\right)$. Then, by the triangle inequality, (S15) and (S12),

$$\|T_2\|_{I_x \times \mathbb{H}} \leq \underline{h}^{-3}\left(n^3/n_{(3)}\right)\left(\underline{\zeta}_{1x}^{-1} + \underline{\zeta}_{0x}^{-1}\right)^2 C_{\mathsf{Lip}} \|\mathbb{1}_{\Delta X}(\cdot, x; \cdot)\|_{I_x \times \mathbb{H}}^2 \overline{\Delta}^2 = O_p^\star\left(\frac{\log(n)}{nh^3}\right). \tag{S44}$$

By Lemma 2 and (S15),

$$T_1(v;b) = \frac{1}{n_{(4)}} \sum_{(i,j,k,m)} \frac{1}{b^4} K''\left(\frac{\Delta_j - v}{b}\right) q_x(W_j, W_m) \pi_x(Z_m, X_m) q_x(W_j, W_i)$$

$$\times K'\left(\frac{\Delta_k - v}{b}\right) q_x\left(W_k, W_i\right) \mathbb{1}\left(X_i = x\right) + O_p^{\star}\left(\left(\frac{\log(n)}{n}\right)^{3/4} h^{-2}\right),$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$ and therefore, by (39),

$$\begin{aligned}
T_1(v;b) &= \frac{1}{n_{(4)}} \sum_{(i,j,k,m)} \frac{1}{b^4} K''\left(\frac{\Delta_x(\epsilon_j) - v}{b}\right) \varpi_x(U_j) \left\{\mathbb{1}(\epsilon_m \le \epsilon_j) - F_{\epsilon|X}(\epsilon_j \mid x)\right\} \pi_x(Z_m, X_m) \\
&\quad \times \varpi_x(U_j) \left\{\mathbb{1}(\epsilon_i \le \epsilon_j) - F_{\epsilon|X}(\epsilon_j \mid x)\right\} K'\left(\frac{\Delta_x(\epsilon_k) - v}{b}\right) \varpi_x(U_k) \\
&\quad \times \left\{\mathbb{1}(\epsilon_i \le \epsilon_k) - F_{\epsilon|X}(\epsilon_k \mid x)\right\} \mathbb{1}(X_i = x) + O_p^{\star}\left(\left(\frac{\log(n)}{n}\right)^{3/4} h^{-2}\right) \\
&=: \frac{1}{n_{(4)}} \sum_{(i,j,k,m)} \mathcal{K}_x(U_i, U_j, U_k, U_m, v; b) + O_p^{\star}\left(\left(\frac{\log(n)}{n}\right)^{3/4} h^{-2}\right),
\end{aligned} \tag{S45}$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$, where the leading term is written as a U-statistic. Let

$$\mathring{\mathcal{K}}_x(U_i, U_j, U_k, U_m, v; b) \coloneqq \frac{1}{4!} \sum_{(i',j',k',K') \in \operatorname{perm}\{i,j,k,m\}} \mathcal{K}_x(U_{i'}, U_{j'}, U_{k'}, U_{K'}, v; b) \tag{S46}$$

denote the symmetrization of the kernel and then,

$$T_1(v;b) = \frac{1}{n^{(4)}} \sum_{(i,j,k,m)} \mathring{\mathcal{K}}_x(U_i, U_j, U_k, U_m, v; b) + O_p^{\star}\left(\left(\frac{\log(n)}{n}\right)^{3/4} h^{-2}\right), \tag{S47}$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$. Denote

$$\begin{aligned}
\mathring{\mathcal{K}}_x^{\langle 1 \rangle}(u, v; b) &\coloneqq \mathrm{E}\left[\mathring{\mathcal{K}}_x(u, U_1, U_2, U_3, v; b)\right] \\
\mathring{\mathcal{K}}_x^{\langle 2 \rangle}(u_1, u_2, v; b) &\coloneqq \mathrm{E}\left[\mathring{\mathcal{K}}_x(u_1, u_2, U_1, U_2, v; b)\right].
\end{aligned}$$

It is easy to see that $\mathrm{E}\left[\mathring{\mathcal{K}}_x(U_1, U_2, U_3, U_4, v; b)\right] = 0$. Denote $\mathfrak{K} \coloneqq \left\{\mathring{\mathcal{K}}_x(\cdot, v; b) : (v,b) \in I_x \times \mathbb{H}\right\}$, $\mathfrak{K}^{\langle 1 \rangle} \coloneqq \left\{\mathring{\mathcal{K}}_x^{\langle 1 \rangle}(\cdot, v; b) : (v,b) \in I_x \times \mathbb{H}\right\}$ and $\mathfrak{K}^{\langle 2 \rangle} \coloneqq \left\{\mathring{\mathcal{K}}_x^{\langle 2 \rangle}(\cdot, v; b) : (v,b) \in I_x \times \mathbb{H}\right\}$. By similar arguments used in the proof of Lemma 4, $\mathfrak{K}$ is uniformly VC-type with respect to a constant envelope $F_{\mathfrak{K}} = O\left(h^{-4}\right)$. By Lemma A.3 of CK, $\mathfrak{K}^{\langle 1 \rangle}$ and $\mathfrak{K}^{\langle 2 \rangle}$ are also uniformly VC-type with respect to constant envelopes $F_{\mathfrak{K}^{\langle 1 \rangle}} = F_{\mathfrak{K}^{\langle 2 \rangle}} = F_{\mathfrak{K}}$. Since ϵ is independent of Z conditionally on X, $\mathrm{E}\left[\mathring{\mathcal{K}}_x^{\langle 1 \rangle}(U, v; b)^2\right] = \mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{K}_x(U_1, U_2, U_3, U_4, v; b) \mid U_4\right]\right)^2\right]$. Therefore,

$$\begin{aligned}
\mathrm{E}\left[\mathring{\mathcal{K}}_x^{\langle 1 \rangle}(U, v; b)^2\right] &= \mathrm{E}\Bigg[\frac{1}{b^4}\Bigg\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x} \int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x} \frac{1}{b^2} K''\left(\frac{\Delta_x(e) - v}{b}\right) \tilde{\rho}_x(e) \left\{\mathbb{1}(\epsilon \le e) - F_{\epsilon|X}(e \mid x)\right\} \\
&\quad \times \left\{F_{\epsilon|X}(e \wedge e' \mid x) - F_{\epsilon|X}(e \mid x) F_{\epsilon|X}(e' \mid x)\right\} K'\left(\frac{\Delta_x(e') - v}{b}\right) \rho_x(e') \, \mathrm{d}e' \mathrm{d}e\Bigg\}^2 \mid X = x\Bigg] \left(p_{1x}^{-1} + p_{0x}^{-1}\right),
\end{aligned} \tag{S48}$$

where

$$\tilde{\rho}_x(e) \coloneqq \frac{f_{\epsilon DX}(e, 0, x)}{\zeta_{1x}\left(g(1, x, e)\right)^2} + \frac{f_{\epsilon DX}(e, 1, x)}{\zeta_{0x}\left(g(0, x, e)\right)^2}.$$

By the c_r inequality,

$$\begin{aligned}&\mathrm{E}\left[\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}\frac{1}{b^2}K''\left(\frac{\Delta_x(e)-v}{b}\right)\tilde{\rho}_x(e)\left\{\mathbb{1}(\epsilon\le e)-F_{\epsilon|X}(e\mid x)\right\}\right.\right.\\&\qquad\left.\left.\times\left\{F_{\epsilon|X}(e\wedge e'\mid x)-F_{\epsilon|X}(e\mid x)F_{\epsilon|X}(e'\mid x)\right\}\frac{1}{b^2}K'\left(\frac{\Delta_x(e')-v}{b}\right)\rho_x(e')\,\mathrm{d}e'\mathrm{d}e\right\}^2\mid X=x\right]=\\&\qquad\mathrm{E}\left[\left\{\sum_{k=1}^m\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\frac{1}{b^2}K''\left(\frac{\Delta_x(e)-v}{b}\right)\tilde{\rho}_x(e)\left\{\mathbb{1}(\epsilon\le e)-F_{\epsilon|X}(e\mid x)\right\}\vartheta_x(e,v;b)\,\mathrm{d}e\right\}^2\mid X=x\right]\lesssim\\&\qquad\qquad\sum_{k=1}^m\sum_{j=1}^m\mathrm{E}\left[\left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\frac{1}{b^2}K''\left(\frac{\Delta_x(e)-v}{b}\right)\tilde{\rho}_x(e)\,\mathbb{1}(\epsilon\le e)\,\vartheta_{x,j}(e,v;b)\,\mathrm{d}e\right\}^2\mid X=x\right],\end{aligned}\tag{S49}$$

where $\vartheta_x(e,v;b):=\sum_{j=1}^m\vartheta_{x,j}(e,v;b)$ and

$$\vartheta_{x,j}(e,v;b):=\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\frac{1}{b^2}K'\left(\frac{\Delta_x(e')-v}{b}\right)\left\{F_{\epsilon|X}(e\wedge e'\mid x)-F_{\epsilon|X}(e\mid x)F_{\epsilon|X}(e'\mid x)\right\}\rho_x(e')\,\mathrm{d}e'.$$

If $k<j$,

$$\begin{aligned}&\mathrm{E}\left[\left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\frac{1}{b^2}K''\left(\frac{\Delta_x(e)-v}{b}\right)\tilde{\rho}_x(e)\,\mathbb{1}(\epsilon\le e)\,\vartheta_{x,j}(e,v;b)\,\mathrm{d}e\right\}^2\mid X=x\right]=\\&\qquad\mathrm{E}\left[\left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\frac{1}{b^2}K''\left(\frac{\Delta_x(e)-v}{b}\right)\tilde{\rho}_x(e)\,\mathbb{1}(\epsilon\le e)\,F_{\epsilon|X}(e\mid x)\,\mathrm{d}e\right\}^2\mid X=x\right]\\&\qquad\qquad\times\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\left\{1-F_{\epsilon|X}(e'\mid x)\right\}\frac{1}{b^2}K'\left(\frac{\Delta_x(e')-v}{b}\right)\rho_x(e')\,\mathrm{d}e'\right\}^2.\end{aligned}\tag{S50}$$

By change of variables and integration by parts,

$$\begin{aligned}&\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\left\{1-F_{\epsilon|X}(e'\mid x)\right\}\frac{1}{b^2}K'\left(\frac{\Delta_x(e')-v}{b}\right)\rho_x(e')\,\mathrm{d}e'=\\&\int_{\frac{\Delta_x(\epsilon_{x,j-1})-v}{b}}^{\frac{\Delta_x(\epsilon_{x,j})-v}{b}}b^{-1}\left\{\psi_{x,j}(bu+v)-\chi_{x,j}(bu+v)\right\}K'(u)\,\mathrm{d}u=-\int_{\frac{\Delta_x(\epsilon_{x,j-1})-v}{b}}^{\frac{\Delta_x(\epsilon_{x,j})-v}{b}}K(u)\left\{\psi'_{x,j}(bu+v)-\chi'_{x,j}(bu+v)\right\}\mathrm{d}u\\&\qquad=-\left(\psi'_{x,j}(v)-\chi'_{x,j}(v)\right)+o(1),\end{aligned}\tag{S51}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Let $\tilde{\psi}_{x,j}(t):=\tilde{\rho}_x\left(\Delta_{x,j}^{-1}(t)\right)\left(\Delta_{x,j}^{-1}\right)'(t)$ and $\tilde{\chi}_{x,j}(t):=\tilde{\psi}_{x,j}(t)F_{\epsilon|X}\left(\Delta_{x,j}^{-1}(t)\mid x\right)$. By similar arguments used to show (S32),

$$\begin{aligned}&\mathrm{E}\left[\left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\frac{1}{b^2}K''\left(\frac{\Delta_x(e)-v}{b}\right)\tilde{\rho}_x(e)\,\mathbb{1}(\epsilon\le e)\,F_{\epsilon|X}(e\mid x)\,\mathrm{d}e\right\}^2\mid X=x\right]=\\&b^{-4}\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}K''\left(\frac{\Delta_x(e)-v}{b}\right)K''\left(\frac{\Delta_x(e')-v}{b}\right)F_{\epsilon|X}(e\wedge e'\mid x)F_{\epsilon|X}(e\mid x)\tilde{\rho}_x(e)F_{\epsilon|X}(e'\mid x)\tilde{\rho}_x(e')\,\mathrm{d}e\mathrm{d}e'\\&=2b^{-2}\int_{\frac{\Delta_x(\epsilon_{x,k-1})-v}{b}}^{\frac{\Delta_x(\epsilon_{x,k})-v}{b}}\int_{\frac{\Delta_x(\epsilon_{x,k-1})-v}{b}}^{\frac{\Delta_x(\epsilon_{x,k})-v}{b}}\mathbb{1}(u\le w)K''(u)K''(w)F_{\epsilon|X}\left(\Delta_{x,k}^{-1}(bu+v)\mid x\right)\tilde{\chi}_{x,k}(bu+v)\tilde{\chi}_{x,k}(bw+v)\,\mathrm{d}u\mathrm{d}w\\&\qquad=2b^{-1}\left\{b^{-1}\left(\int\int\mathbb{1}(u\le w)K''(u)K''(w)\,\mathrm{d}u\mathrm{d}w\right)F_{\epsilon|X}\left(\Delta_{x,k}^{-1}(v)\mid x\right)\tilde{\chi}_{x,k}(v)^2+O(1)\right\}=O\left(h^{-1}\right),\end{aligned}\tag{S52}$$

uniformly in $(v, b) \in I_x \times \mathbb{H}$, where the last equality follows from

$$\int \int \mathbb{1}(u \leq w) K''(u) K''(w) \mathrm{d}u \mathrm{d}w = \int K''(w) K'(w) \mathrm{d}w = 0. \tag{S53}$$

It follows from (S52), (S50) and (S51) that if $k < j$,

$$\mathrm{E}\left[\left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2} K''\left(\frac{\Delta_x(e) - v}{b}\right) \tilde{\rho}_x(e) \mathbb{1}(\epsilon \leq e) \vartheta_{x,j}(e, v; b) \mathrm{d}e\right\}^2 \mid X = x\right] = O\left(h^{-1}\right), \tag{S54}$$

uniformly in $(v, b) \in I_x \times \mathbb{H}$. If $k > j$, by similar arguments, we have

$$\begin{aligned}
&\mathrm{E}\left[\left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2} K''\left(\frac{\Delta_x(e) - v}{b}\right) \tilde{\rho}_x(e) \mathbb{1}(\epsilon \leq e) \vartheta_{x,j}(e, v; b) \mathrm{d}e\right\}^2 \mid X = x\right] = \\
&\qquad \mathrm{E}\left[\left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2} K''\left(\frac{\Delta_x(e) - v}{b}\right) \tilde{\rho}_x(e) \mathbb{1}(\epsilon \leq e) \left(1 - F_{\epsilon|X}(e \mid x)\right) \mathrm{d}e\right\}^2 \mid X = x\right] \\
&\qquad\qquad \times \left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} F_{\epsilon|X}(e' \mid x) \frac{1}{b^2} K'\left(\frac{\Delta_x(e') - v}{b}\right) \rho_x(e') \mathrm{d}e'\right\}^2,
\end{aligned}$$

$$\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} F_{\epsilon|X}(e' \mid x) \frac{1}{b^2} K'\left(\frac{\Delta_x(e') - v}{b}\right) \rho_x(e') \mathrm{d}e' = O(1)$$

and

$$\mathrm{E}\left[\left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2} K''\left(\frac{\Delta_x(e) - v}{b}\right) \tilde{\rho}_x(e) \mathbb{1}(\epsilon \leq e) \left(1 - F_{\epsilon|X}(e \mid x)\right) \mathrm{d}e\right\}^2 \mid X = x\right] = O\left(h^{-1}\right).$$

Therefore, (S54) also holds if $k < j$. If $k = j$, by change of variables, integration by parts and using the equality

$$\begin{aligned}
&F_{\epsilon|X}(e \mid x) \left\{-\frac{1}{h} K\left(\frac{\Delta_x(e) - v}{h}\right) \left(\psi_{x,j}(\Delta_{x,j}(e)) - \chi_{x,j}(\Delta_{x,j}(e))\right)\right\} \\
&\qquad + \left(1 - F_{\epsilon|X}(e \mid x)\right) \frac{1}{h} K\left(\frac{\Delta_x(e) - v}{h}\right) \chi_{x,j}(\Delta_{x,j}(e)) = 0,
\end{aligned}$$

for $e \in (\epsilon_{x,j-1}, \epsilon_{x,j})$,

$$\begin{aligned}
\vartheta_{x,j}(e, v; b) &= F_{\epsilon|X}(e \mid x) \left\{\int_e^{\epsilon_{x,j}} \left(1 - F_{\epsilon|X}(e' \mid x)\right) \frac{1}{b^2} K'\left(\frac{\Delta_x(e') - v}{b}\right) \rho_x(e') \mathrm{d}e'\right\} \\
&\quad + \left(1 - F_{\epsilon|X}(e \mid x)\right) \left\{\int_{\epsilon_{x,j-1}}^{e} F_{\epsilon|X}(e' \mid x) \frac{1}{b^2} K'\left(\frac{\Delta_x(e') - v}{b}\right) \rho_x(e') \mathrm{d}e'\right\} \\
&= -F_{\epsilon|X}(e \mid x) \int_{\frac{\Delta_x(e) - v}{b}}^{\frac{\Delta_x(\epsilon_{x,j}) - v}{b}} K(u) \left\{\psi'_{x,j}(bu + v) - \chi'_{x,j}(bu + v)\right\} \mathrm{d}u \\
&\quad - \left(1 - F_{\epsilon|X}(e \mid x)\right) \int_{\frac{\Delta_x(\epsilon_{x,j-1}) - v}{b}}^{\frac{\Delta_x(e) - v}{b}} K(u) \chi'_{x,j}(bu + v) \mathrm{d}u.
\end{aligned} \tag{S55}$$

Let $\tilde{K}(u) := \int_{-\infty}^{u} K(w) \mathrm{d}w$ and $\bar{K}(u) := K''(u) \tilde{K}(u)$. Then, $\vartheta_{x,j}(e, v; b) = \bar{\vartheta}_{x,j}(e, v; b) + O(h)$, uniformly in

$(e, v, b) \in (\epsilon_{x,j-1}, \epsilon_{x,j}) \times I_x \times \mathbb{H}$, where

$$\begin{aligned}
\bar{\vartheta}_{x,j}(e, v; b) &:= -F_{\epsilon|X}(e \mid x)\left(1 - \tilde{K}\left(\frac{\Delta_x(e) - v}{b}\right)\right)\left(\psi'_{x,j}(v) - \chi'_{x,j}(v)\right) \\
&\quad - \left(1 - F_{\epsilon|X}(e \mid x)\right)\tilde{K}\left(\frac{\Delta_x(e) - v}{b}\right)\chi'_{x,j}(v) \\
&= -F_{\epsilon|X}(e \mid x)\left(\psi'_{x,j}(v) - \chi'_{x,j}(v)\right) + \left(F_{\epsilon|X}(e \mid x)\psi'_{x,j}(v) - \chi'_{x,j}(v)\right)\tilde{K}\left(\frac{\Delta_x(e) - v}{b}\right). \quad \text{(S56)}
\end{aligned}$$

Then,

$$\begin{aligned}
&\mathrm{E}\left[\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^2}K''\left(\frac{\Delta_x(e) - v}{b}\right)\tilde{\rho}_x(e)\,\mathbb{1}(\epsilon \leq e)\,\vartheta_{x,j}(e, v; b)\,\mathrm{d}e\right\}^2 \mid X = x\right] = \\
&\qquad \mathrm{E}\left[\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^2}K''\left(\frac{\Delta_x(e) - v}{b}\right)\tilde{\rho}_x(e)\,\mathbb{1}(\epsilon \leq e)\,\bar{\vartheta}_{x,j}(e, v; b)\,\mathrm{d}e\right\}^2 \mid X = x\right] + O\left(h^{-1}\right). \quad \text{(S57)}
\end{aligned}$$

Then we have

$$\begin{aligned}
&\mathrm{E}\left[\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^2}K''\left(\frac{\Delta_x(e) - v}{b}\right)\tilde{\rho}_x(e)\,\mathbb{1}(\epsilon \leq e)\,\bar{\vartheta}_{x,j}(e, v; b)\,\mathrm{d}e\right\}^2 \mid X = x\right] = \\
&\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^4}K''\left(\frac{\Delta_x(e) - v}{b}\right)\tilde{\rho}_x(e)\,\bar{\vartheta}_{x,j}(e, v; b)\,F_{\epsilon|X}(e \wedge e' \mid x)\,K''\left(\frac{\Delta_x(e') - v}{b}\right)\tilde{\rho}_x(e')\,\bar{\vartheta}_{x,j}(e', v; b)\,\mathrm{d}e\mathrm{d}e' \\
&= \int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^4}K''\left(\frac{\Delta_x(e) - v}{b}\right)\tilde{\rho}_x(e)\,F_{\epsilon|X}(e \mid x)\,F_{\epsilon|X}(e \wedge e' \mid x)\,K''\left(\frac{\Delta_x(e') - v}{b}\right)\tilde{\rho}_x(e')\,F_{\epsilon|X}(e' \mid x)\,\mathrm{d}e\mathrm{d}e' \\
&\quad \times \left(\psi'_{x,j}(v) - \chi'_{x,j}(v)\right)^2 - 2\left(\psi'_{x,j}(v) - \chi'_{x,j}(v)\right)\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^4}K''\left(\frac{\Delta_x(e) - v}{b}\right)\tilde{\rho}_x(e)\,F_{\epsilon|X}(e \mid x) \\
&\quad \times F_{\epsilon|X}(e \wedge e' \mid x)\,\bar{K}\left(\frac{\Delta_x(e') - v}{b}\right)\tilde{\rho}_x(e')\left(F_{\epsilon|X}(e' \mid x)\psi'_{x,j}(v) - \chi'_{x,j}(v)\right)\mathrm{d}e\mathrm{d}e' \\
&\quad + \int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^4}\bar{K}\left(\frac{\Delta_x(e) - v}{b}\right)\tilde{\rho}_x(e)\left(F_{\epsilon|X}(e \mid x)\psi'_{x,j}(v) - \chi'_{x,j}(v)\right)F_{\epsilon|X}(e \wedge e' \mid x) \\
&\quad \times \bar{K}\left(\frac{\Delta_x(e') - v}{b}\right)\tilde{\rho}_x(e')\left(F_{\epsilon|X}(e' \mid x)\psi'_{x,j}(v) - \chi'_{x,j}(v)\right)\mathrm{d}e\mathrm{d}e' =: I_1(v, b) + I_2(v, b) + I_3(v, b). \quad \text{(S58)}
\end{aligned}$$

It is shown by (S52) that $I_1(v, b) = O\left(h^{-1}\right)$, uniformly in $(v, b) \in I_x \times \mathbb{H}$. Then,

$$\begin{aligned}
I_2(v, b) &= -2b^{-2}\left(\psi'_{x,j}(v) - \chi'_{x,j}(v)\right)\int_{\frac{\Delta_x(\epsilon_{x,j-1}) - v}{b}}^{\frac{\Delta_x(\epsilon_{x,j}) - v}{b}}\int_{\frac{\Delta_x(\epsilon_{x,j-1}) - v}{b}}^{\frac{\Delta_x(\epsilon_{x,j}) - v}{b}} K''(u)\,\tilde{\chi}_{x,j}(bu + v) \\
&\quad \times F_{\epsilon|X}\left(\Delta_{x,j}^{-1}(bu + v) \wedge \Delta_{x,j}^{-1}(bw + v) \mid x\right)\bar{K}(w)\left(\tilde{\chi}_{x,j}(bw + v)\,\psi'_{x,j}(v) - \tilde{\psi}_{x,j}(bw + v)\,\chi'_{x,j}(v)\right)\mathrm{d}u\mathrm{d}w \\
&= -2b^{-2}\left(\psi'_{x,j}(v) - \chi'_{x,j}(v)\right)\left\{\int_{\frac{\Delta_x(\epsilon_{x,j-1}) - v}{b}}^{\frac{\Delta_x(\epsilon_{x,j}) - v}{b}}\int_{\frac{\Delta_x(\epsilon_{x,j-1}) - v}{b}}^{\frac{\Delta_x(\epsilon_{x,j}) - v}{b}} \mathbb{1}(u \leq w)\,K''(u)\,\tilde{\chi}_{x,j}(bu + v)\,F_{\epsilon|X}\left(\Delta_{x,j}^{-1}(bu + v) \mid x\right)\bar{K}(w)\right. \\
&\quad \times \left(\tilde{\chi}_{x,j}(bw + v)\,\psi'_{x,j}(v) - \tilde{\psi}_{x,j}(bw + v)\,\chi'_{x,j}(v)\right)\mathrm{d}u\mathrm{d}w + \int_{\frac{\Delta_x(\epsilon_{x,j-1}) - v}{b}}^{\frac{\Delta_x(\epsilon_{x,j}) - v}{b}}\int_{\frac{\Delta_x(\epsilon_{x,j-1}) - v}{b}}^{\frac{\Delta_x(\epsilon_{x,j}) - v}{b}} \mathbb{1}(u > w)\,K''(u)\,\tilde{\chi}_{x,j}(bu + v) \\
&\quad \left. \times F_{\epsilon|X}\left(\Delta_{x,j}^{-1}(bw + v) \mid x\right)\bar{K}(w)\left(\tilde{\chi}_{x,j}(bw + v)\,\psi'_{x,j}(v) - \tilde{\psi}_{x,j}(bw + v)\,\chi'_{x,j}(v)\right)\mathrm{d}u\mathrm{d}w\right\} = -2b^{-2}\left(\psi'_{x,j}(v) - \chi'_{x,j}(v)\right)
\end{aligned}$$

$$\times \left\{ \left(\int \int K''(u) \bar{K}(w) \,\mathrm{d}u\mathrm{d}w \right) \left(\tilde{\chi}_{x,j}(v) F_{\epsilon|X}\left(\Delta_{x,j}^{-1}(v) \mid x \right) \left(\tilde{\chi}_{x,j}(v) \psi'_{x,j}(v) - \tilde{\psi}_{x,j}(v) \chi'_{x,j}(v) \right) \right) + O(h) \right\},$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$. Then, it follows from this result and $\int \int K''(u) \bar{K}(w) \,\mathrm{d}u\mathrm{d}w = 0$ that $I_2(v,b) = O\left(h^{-1}\right)$. Similarly,

$$\begin{aligned} I_3(v,b) = &\int_{\frac{\Delta_x(\epsilon_{x,j-1})-v}{b}}^{\frac{\Delta_x(\epsilon_{x,j})-v}{b}} \int_{\frac{\Delta_x(\epsilon_{x,j-1})-v}{b}}^{\frac{\Delta_x(\epsilon_{x,j})-v}{b}} b^{-2} \bar{K}(u) \left(\tilde{\chi}_{x,j}(bu+v) \psi'_{x,j}(v) - \tilde{\psi}_{x,j}(bu+v) \chi'_{x,j}(v) \right) \\ &\times F_{\epsilon|X}\left(\Delta_{x,j}^{-1}(bu+v) \wedge \Delta_{x,j}^{-1}(bw+v) \mid x \right) \bar{K}(w) \left(\tilde{\chi}_{x,j}(bw+v) \psi'_{x,j}(v) - \tilde{\psi}_{x,j}(bw+v) \chi'_{x,j}(v) \right) \mathrm{d}u\mathrm{d}w \\ = b^{-2} &\left\{ \left(\int \int \bar{K}(u) \bar{K}(w) \,\mathrm{d}u\mathrm{d}w \right) \left(\tilde{\chi}_{x,j}(v) \psi'_{x,j}(v) - \tilde{\psi}_{x,j}(v) \chi'_{x,j}(v) \right)^2 F_{\epsilon|X}\left(\Delta_{x,j}^{-1}(v) \mid x \right) + O(h) \right\} = O\left(h^{-1}\right), \end{aligned}$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$, where the last equality follows from

$$\int \bar{K}(u) \,\mathrm{d}u = \int K''(u) \tilde{K}(u) \,\mathrm{d}u = -\int K'(u) K(u) \,\mathrm{d}u = 0. \tag{S59}$$

Then it follows from these results, (S57) and (S58) that

$$\mathrm{E}\left[\left\{ \int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^2} K''\left(\frac{\Delta_x(e) - v}{b} \right) \tilde{\rho}_x(e) \mathbb{1}(\epsilon \leq e) \vartheta_{x,j}(e,v;b) \,\mathrm{d}e \right\}^2 \mid X = x \right] = O\left(h^{-1}\right),$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$. It follows from this result, (S54), (S48) and (S49) that $\sigma^2_{\mathfrak{K}^{\langle 1 \rangle}} := \sup_{f \in \mathfrak{K}^{\langle 1 \rangle}} \mathbb{P}^U f^2 = O\left(h^{-1}\right)$. It follows from the c_r inequality, change of variables and tedious but straightforward calculations,

$$\begin{aligned} \mathrm{E}\left[\mathring{\mathcal{K}}_x^{\langle 2 \rangle}(U_1,U_2,v;b)^2 \right] \quad &\lesssim \quad \mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{K}_x(U_1,U_2,U_3,U_4,v;b) \mid U_1,U_4 \right] \right)^2 \right] + \mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{K}_x(U_1,U_2,U_3,U_4,v;b) \mid U_2,U_4 \right] \right)^2 \right] \\ &\quad + \mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{K}_x(U_1,U_2,U_3,U_4,v;b) \mid U_3,U_4 \right] \right)^2 \right] \\ &= \quad O\left(h^{-5}\right), \end{aligned}$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$. Therefore, $\sigma^2_{\mathfrak{K}^{\langle 2 \rangle}} := \sup_{f \in \mathfrak{K}^{\langle 2 \rangle}} \mathrm{E}\left[f(U_1,U_2)^2 \right] = O\left(h^{-5}\right)$. By the coupling theorem (Proposition 2.1 of CK with $\mathcal{H} = \mathfrak{K}_\pm$, $\overline{\sigma}_{\mathfrak{g}} = \sigma_{\mathfrak{K}^{\langle 1 \rangle}}$, $\sigma_{\mathfrak{h}} = \sigma_{\mathfrak{K}^{\langle 2 \rangle}}$, $b_{\mathfrak{g}} = b_{\mathfrak{h}} = F_{\mathfrak{K}}$ and $q = \infty$), when n is sufficiently large so that $V_{\mathfrak{K}_\pm} \cdot \log\left(A_{\mathfrak{K}_\pm} \vee n\right) \leq n^{1/3}$, $\forall \gamma_1 \in (0,1)$, there exists a random variable $Z_{\mathfrak{K}_\pm,\gamma_1} =_d \left\| G^U \right\|_{\mathfrak{K}^{\langle 1 \rangle}}$ such that

$$\Pr\left[\left| \left\| \mathbb{U}_n^{(4)} \right\|_{\mathfrak{K}} - Z_{\mathfrak{K}_\pm,\gamma_1} \right| > C_1 \left(\frac{\log(n)^{2/3}}{n^{1/6} h^{5/3} \gamma_1^{1/3}} + \frac{\log(n)}{n^{1/2} h^4 \gamma_1} \right) \right] \leq C_2\left(\gamma_1 + n^{-1}\right).$$

By the Borell-Sudakov-Tsirelson inequality, $\Pr\left[\left\| G^U \right\|_{\mathfrak{K}^{\langle 1 \rangle}} > \mathrm{E}\left[\left\| G^U \right\|_{\mathfrak{K}^{\langle 1 \rangle}} \right] + \sqrt{2 \cdot \log(n)} \sigma_{\mathfrak{K}^{\langle 1 \rangle}} \right] \leq n^{-1}$. By Dudley's entropy integral bound, $\mathrm{E}\left[\left\| G^U \right\|_{\mathfrak{K}^{\langle 1 \rangle}} \right] \lesssim \left(\sigma_{\mathfrak{K}^{\langle 1 \rangle}} \vee n^{-1/2} \left\| F_{\mathfrak{K}^{\langle 1 \rangle}} \right\|_{\mathbb{P}^U,2} \right) \sqrt{\log(n)}$. Therefore,

$$\Pr\left[\left\| \mathbb{U}_n^{(4)} \right\|_{\mathfrak{K}} > C_1 \left(\frac{\log(n)^{2/3}}{n^{1/6} h^{5/3} \gamma_1^{1/3}} + \frac{\log(n)}{n^{1/2} h^4 \gamma_1} + \sqrt{\frac{\log(n)}{h}} \right) \right] \leq C_2\left(\gamma_1 + n^{-1}\right). \tag{S60}$$

It follows from this result, (S41), (S42), (S45) and (S44) that when n is sufficiently large, $\forall \gamma_1 \in (0,1)$,

$$\Pr\left[\left\| \dot{V}_2(\cdot,x;\cdot) - \ddot{V}_2(\cdot,x;\cdot) \right\|_{I_x \times \mathbb{H}} > C_1 \left(\frac{\log(n)^{2/3}}{(nh)^{2/3} h \gamma_1^{1/3}} + \frac{\log(n)}{n h^4 \gamma_1} \right) \right] \leq C_2\left(\gamma_1 + n^{-1}\right). \tag{S61}$$

Next,

$$\widetilde{V}_2(v,x;b,b_\zeta)-\dot{V}_2(v,x;b)=$$
$$\frac{2}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\Delta}_j-v}{b}\right)\varOmega_x(W_j,W_i;b_\zeta)K'\left(\frac{\widehat{\Delta}_k-v}{b}\right)q_x(W_k,W_i)\mathbb{1}(X_i=x)$$
$$+\frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\Delta}_j-v}{b}\right)\varOmega_x(W_j,W_i;b_\zeta)K'\left(\frac{\widehat{\Delta}_k-v}{b}\right)\varOmega_x(W_k,W_i;b_\zeta)\mathbb{1}(X_i=x), \tag{S62}$$

where $\varOmega_x(W_j,W_i;b_\zeta):=\widehat{q}_x(W_j,W_i;b_\zeta)-q_x(W_j,W_i)$. Let $\varOmega_{dx}(W_j,W_i;b_\zeta):=\widehat{q}_{dx}(W_j,W_i;b_\zeta)-q_{dx}(W_j,W_i)$. Denote

$$\varphi_{dx}(W_j,W_i):=\mathbb{1}(Y_i\leq\phi_{dx}(Y_j),D_i=d)+\mathbb{1}(Y_i\leq Y_j,D_i=d')-R_{d'x}(Y_j)$$

and let $\widehat{\varphi}_{dx}(W_j,W_i)$ be defined by the same formula with $(\phi_{dx},R_{d'x})$ replaced by $\left(\widehat{\phi}_{dx},\widehat{R}_{d'x}\right)$. Then, by (S4),

$$\varOmega_{dx}(W_j,W_i;b_\zeta)=\mathbb{1}(D_j=d',X_j=x)\left\{\frac{\widehat{\varphi}_{dx}(W_j,W_i)}{\zeta_{dx}(\phi_{dx}(Y_j))}-\frac{\varphi_{dx}(W_j,W_i)}{\zeta_{dx}(\phi_{dx}(Y_j))}\right.$$
$$\left.-\frac{\widehat{\varphi}_{dx}(W_j,W_i)\left(\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(Y_j);b_\zeta\right)-\zeta_{dx}(\phi_{dx}(Y_j))\right)}{\zeta_{dx}(\phi_{dx}(Y_j))^2}+\frac{\widehat{\varphi}_{dx}(W_j,W_i)\left(\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(Y_j);b_\zeta\right)-\zeta_{dx}(\phi_{dx}(Y_j))\right)^2}{\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(Y_j);b_\zeta\right)\zeta_{dx}(\phi_{dx}(Y_j))^2}\right\}. \tag{S63}$$

Then, by this result,

$$\frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\Delta}_j-v}{b}\right)\varOmega_{dx}(W_j,W_i;b_\zeta)K'\left(\frac{\widehat{\Delta}_k-v}{b}\right)q_x(W_k,W_i)\mathbb{1}(X_i=x)=$$
$$T_3(v;b)+T_4(v;b)+T_5(v;b,b_\zeta)+T_6(v;b,b_\zeta),$$

where

$$\begin{aligned}
T_3(v;b) &:= \frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\Delta}_j-v}{b}\right)\mathbb{1}(D_j=d',X_j=x)\mathbb{1}(D_i=d)\\
&\quad\times\frac{\left\{\mathbb{1}\left(Y_i\leq\widehat{\phi}_{dx}(Y_j)\right)-\mathbb{1}(Y_i\leq\phi_{dx}(Y_j))\right\}}{\zeta_{dx}(\phi_{dx}(Y_j))}K'\left(\frac{\widehat{\Delta}_k-v}{b}\right)q_x(W_k,W_i)\mathbb{1}(X_i=x)\\
T_4(v;b) &:= \frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\Delta}_j-v}{b}\right)\mathbb{1}(D_j=d',X_j=x)\\
&\quad\times\frac{\widehat{R}_{d'x}(Y_j)-R_{d'x}(Y_j)}{\zeta_{dx}(\phi_{dx}(Y_j))}K'\left(\frac{\widehat{\Delta}_k-v}{b}\right)q_x(W_k,W_i)\mathbb{1}(X_i=x)\\
T_5(v;b,b_\zeta) &:= \frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\Delta}_j-v}{b}\right)\mathbb{1}(D_j=d',X_j=x)\\
&\quad\times\frac{\widehat{\varphi}_{dx}(W_j,W_i)\left\{\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(Y_j);b_\zeta\right)-\zeta_{dx}(\phi_{dx}(Y_j))\right\}}{\zeta_{dx}(\phi_{dx}(Y_j))^2}K'\left(\frac{\widehat{\Delta}_k-v}{b}\right)q_x(W_k,W_i)\mathbb{1}(X_i=x)\\
T_6(v;b,b_\zeta) &:= \frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\Delta}_j-v}{h}\right)\mathbb{1}(D_j=d',X_j=x)
\end{aligned}$$

$$\times \frac{\widehat{\varphi}_{dx}\left(W_j, W_i\right)\left\{\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}\left(Y_j\right); b_\zeta\right) - \zeta_{dx}\left(\phi_{dx}\left(Y_j\right)\right)\right\}^2}{\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}\left(Y_j\right); b_\zeta\right)\zeta_{dx}\left(\phi_{dx}\left(Y_j\right)\right)^2} K'\left(\frac{\widehat{\Delta}_k - v}{b}\right) q_x\left(W_k, W_i\right) \mathbb{1}\left(X_i = x\right).$$

In view of (S12), by using similar arguments for proving (S13),

$$1 - O\left(n^{-1}\right) \leq \Pr\left[\left|K'\left(\frac{\widehat{\Delta}_i - v}{b}\right)\right| \mathbb{1}\left(X_i = x\right) \leq \left\|K'\right\|_\infty \mathbb{1}_i\left(v; b\right) \mathbb{1}\left(X_i = x\right), \forall\left(i, v, b\right) \in \{1, ..., n\} \times I_x \times \mathbb{H}\right] \tag{S64}$$

and therefore,

$$\left|T_3\left(v; b\right)\right| \lesssim b^{-1}\left(n^3 / n_{(3)}\right) \underline{\zeta}_{dx}^{-1}\left(\underline{\zeta}_{1x}^{-1} + \underline{\zeta}_{0x}^{-1}\right) \mathbb{1}_{\Delta X}\left(v, x; b\right)^2 \overline{\phi},$$

where

$$\overline{\phi} \coloneqq \max_{j=1,\ldots,n} \mathbb{1}\left(D_j = d', X_j = x\right) \frac{1}{n} \sum_{i=1}^n \left|\mathbb{1}\left(Y_i \leq \widehat{\phi}_{dx}\left(Y_j\right)\right) - \mathbb{1}\left(Y_i \leq \phi_{dx}\left(Y_j\right)\right)\right| \mathbb{1}\left(D_i = d, X_i = x\right).$$

By using

$$\begin{aligned}\left|\mathbb{1}\left(Y_i \leq \widehat{\phi}_{dx}\left(y\right)\right) - \mathbb{1}\left(Y_i \leq \phi_{dx}\left(y\right)\right)\right| &= \mathbb{1}\left(Y_i \leq \widehat{\phi}_{dx}\left(y\right)\right) \mathbb{1}\left(Y_i > \phi_{dx}\left(y\right)\right) + \mathbb{1}\left(Y_i > \widehat{\phi}_{dx}\left(y\right)\right) \mathbb{1}\left(Y_i \leq \phi\left(y\right)\right) \\ &\leq \mathbb{1}\left(\phi_{dx}\left(y\right) < Y_i \leq \phi_{dx}\left(y\right) + \left\|\widehat{\phi}_{dx} - \phi_{dx}\right\|_{I_{d'x}}\right) + \mathbb{1}\left(\phi_{dx}\left(y\right) \geq Y_i > \phi_{dx}\left(y\right) - \left\|\widehat{\phi}_{dx} - \phi_{dx}\right\|_{I_{d'x}}\right),\end{aligned}$$

and letting

$$\begin{aligned}\tilde{\mathcal{P}}_{dx}^+\left(W_i, y, \xi\right) &\coloneqq \mathbb{1}\left(\phi_{dx}\left(y\right) < Y_i \leq \phi_{dx}\left(y\right) + \xi\right) \mathbb{1}\left(D_i = d, X_i = x\right) \\ \tilde{\mathfrak{P}}^+ &\coloneqq \left\{\tilde{\mathcal{P}}_{dx}^+\left(\cdot, y, \xi\right) : \left(y, \xi\right) \in I_{d'x} \times \left(0, \overline{\xi}\right]\right\} \\ \tilde{\mathcal{P}}_{dx}^-\left(W_i, y, \xi\right) &\coloneqq \mathbb{1}\left(\phi_{dx}\left(y\right) - \xi < Y_i \leq \phi_{dx}\left(y\right)\right) \mathbb{1}\left(D_i = d, X_i = x\right) \\ \tilde{\mathfrak{P}}^- &\coloneqq \left\{\tilde{\mathcal{P}}_{dx}^-\left(\cdot, y, \xi\right) : \left(y, \xi\right) \in I_{d'x} \times \left(0, \overline{\xi}\right]\right\},\end{aligned}$$

where $\overline{\xi} \coloneqq C_1 \sqrt{\log\left(n\right) / n}$ in view of (S3), we have

$$\begin{aligned}&\sup_{y \in I_{d'x}} \left|\frac{1}{n} \sum_{i=1}^n \left|\mathbb{1}\left(Y_i \leq \widehat{\phi}_{dx}\left(y\right)\right) - \mathbb{1}\left(Y_i \leq \phi_{dx}\left(y\right)\right)\right| \mathbb{1}\left(D_i = d, X_i = x\right)\right| \leq \\ &\qquad \sup_{y \in I_{d'x}} \left|\frac{1}{n} \sum_{i=1}^n \mathbb{1}\left(\phi_{dx}\left(y\right) < Y_i \leq \phi_{dx}\left(y\right) + \left\|\widehat{\phi}_{dx} - \phi_{dx}\right\|_{I_{d'x}}\right) \mathbb{1}\left(D_i = d, X_i = x\right)\right| \\ &\quad + \sup_{y \in I_{d'x}} \left|\frac{1}{n} \sum_{i=1}^n \mathbb{1}\left(\phi_{dx}\left(y\right) \geq Y_i > \phi_{dx}\left(y\right) - \left\|\widehat{\phi}_{dx} - \phi_{dx}\right\|_{I_{d'x}}\right) \mathbb{1}\left(D_i = d, X_i = x\right)\right| \leq \left\|\mathbb{P}_n^W\right\|_{\tilde{\mathfrak{P}}^+} + \left\|\mathbb{P}_n^W\right\|_{\tilde{\mathfrak{P}}^-},\end{aligned}$$

where the second inequality holds with probability at least $1 - C_2 n^{-1}$. By similar arguments used in the proof of Lemma 2, we have $\left\|\mathbb{G}_n^W\right\|_{\tilde{\mathfrak{P}}^+} = O_p^\star\left(\log\left(n\right)^{3/4} / n^{1/4}\right)$ and $\left\|\mathbb{G}_n^W\right\|_{\tilde{\mathfrak{P}}^+} = O_p^\star\left(\log\left(n\right)^{3/4} / n^{1/4}\right)$,

$$\left\|\mathbb{P}^W\right\|_{\tilde{\mathfrak{P}}^+} = \sup_{\left(y, \xi\right) \in I_{d'x} \times \left(0, \overline{\xi}\right]} \mathrm{E}\left[\mathbb{1}\left(\phi_{dx}\left(y\right) < Y \leq \phi_{dx}\left(y\right) + \xi\right) \mathbb{1}\left(D = d, X = x\right)\right] = O\left(\sqrt{\frac{\log\left(n\right)}{n}}\right)$$

and $\left\|\mathbb{P}^W\right\|_{\tilde{\mathfrak{P}}^-} = O\left(\sqrt{\log\left(n\right) / n}\right)$. By these results, $\left\|\mathbb{P}_n^W\right\|_{\tilde{\mathfrak{P}}^+} \leq n^{-1/2} \left\|\mathbb{G}_n^W\right\|_{\tilde{\mathfrak{P}}^+} + \left\|\mathbb{P}^W\right\|_{\tilde{\mathfrak{P}}^+} = O_p^\star\left(\sqrt{\log\left(n\right) / n}\right)$

and $\left\|\mathbb{P}_n^W\right\|_{\tilde{\mathfrak{P}}^-} = O_p^\star\left(\sqrt{\log(n)/n}\right)$. Now it follows that

$$\overline{\phi} \leq \sup_{y \in I_{d'x}} \left|\frac{1}{n}\sum_{i=1}^{n}\left|\mathbb{1}\left(Y_i \leq \widehat{\phi}_{dx}(y)\right) - \mathbb{1}\left(Y_i \leq \phi_{dx}(y)\right)\right| \mathbb{1}\left(D_i = d, X_i = x\right)\right| = O_p^\star\left(\sqrt{\frac{\log(n)}{n}}\right) \tag{S65}$$

and (S15), $\|T_3\|_{I_x \times \mathbb{H}} = O_p^\star\left(\sqrt{\log(n)/(nh^2)}\right)$. Similarly, when n is sufficiently large, with probability at least $1 - C_2 n^{-1}$,

$$|T_4(v;b)| \lesssim b^{-1}\left(n^3/n_{(3)}\right)\underline{\zeta}_{dx}^{-1}\left(\underline{\zeta}_{1x}^{-1} + \underline{\zeta}_{0x}^{-1}\right)\mathbb{1}_{\Delta X}(v,x;b)^2\,\overline{R} = O_p^\star\left(\sqrt{\frac{\log(n)}{nh^2}}\right)$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$, where

$$\overline{R} \coloneqq \max_{j=1,\ldots,n}\mathbb{1}\left(D_j = d', X_j = x\right)\left|\widehat{R}_{d'x}(Y_j) - R_{d'x}(Y_j)\right| = O_p^\star\left(\sqrt{\frac{\log(n)}{n}}\right)$$

and the equality follows from (S15) and Lemma 6.

For $T_5(v;b,b_\zeta)$, write

$$\begin{aligned}
T_5(v;b,b_\zeta) &= \frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\Delta}_j - v}{b}\right)\mathbb{1}\left(D_j = d', X_j = x\right)\frac{\varphi_{dx}(W_j,W_i)}{\zeta_{dx}(\phi_{dx}(Y_j))^2}\left\{\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(Y_j);b_\zeta\right) - \zeta_{dx}(\phi_{dx}(Y_j))\right\} \\
&\quad \times K'\left(\frac{\Delta_k - v}{b}\right)q_x(W_k,W_i)\,\mathbb{1}(X_i = x) + \frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\Delta}_j - v}{b}\right)\mathbb{1}\left(D_j = d', X_j = x\right)\frac{\varphi_{dx}(W_j,W_i)}{\zeta_{dx}(\phi_{dx}(Y_j))^2} \\
&\quad \times \left\{\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(Y_j);b_\zeta\right) - \zeta_{dx}(\phi_{dx}(Y_j))\right\}\left\{K'\left(\frac{\widehat{\Delta}_k - v}{b}\right) - K'\left(\frac{\Delta_k - v}{b}\right)\right\}q_x(W_k,W_i)\,\mathbb{1}(X_i = x) \\
&+ \frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\Delta}_j - v}{b}\right)\mathbb{1}\left(D_j = d', X_j = x\right)\frac{\widehat{\varphi}_{dx}(W_j,W_i) - \varphi_{dx}(W_j,W_i)}{\zeta_{dx}(\phi_{dx}(Y_j))^2}\left\{\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(Y_j);b_\zeta\right) - \zeta_{dx}(\phi_{dx}(Y_j))\right\} \\
&\quad \times K'\left(\frac{\widehat{\Delta}_k - v}{b}\right)q_x(W_k,W_i)\,\mathbb{1}(X_i = x) \eqqcolon T_{5.1}(v;b,b_\zeta) + T_{5.2}(v;b,b_\zeta) + T_{5.3}(v;b,b_\zeta).
\end{aligned}$$

By (S3),

$$\Pr\left[\left|\sup_{y \in \dot{I}_{d'x}}\widehat{\phi}_{dx}(y) - \sup_{y \in \dot{I}_{d'x}}\phi_{dx}(y)\right| \leq C_1\sqrt{\frac{\log(n)}{n}}\right] > 1 - C_2 n^{-1}$$

and

$$\Pr\left[\left|\inf_{y \in \dot{I}_{d'x}}\widehat{\phi}_{dx}(y) - \inf_{y \in \dot{I}_{d'x}}\phi_{dx}(y)\right| \leq C_1\sqrt{\frac{\log(n)}{n}}\right] > 1 - C_2 n^{-1},$$

where $\dot{I}_{d'x}$ is any closed sub-interval of $I_{d'x}$. By these results and Lemma 6, $\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(y);b_\zeta\right) - \zeta_{dx}\left(\widehat{\phi}_{dx}(y)\right) = O_p^\star\left(\sqrt{\log(n)/(nh_\zeta)} + h_\zeta^2\right)$, uniformly in $(y,b_\zeta) \in \dot{I}_{d'x} \times \mathbb{H}_\zeta$. By this result and $\zeta_{dx}\left(\widehat{\phi}_{dx}(y)\right) - \zeta_{dx}(\phi_{dx}(y)) = O_p^\star\left(\sqrt{\log(n)/n}\right)$, which follows from (S3) and mean value expansion,

$$\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(y);b_\zeta\right) - \zeta_{dx}(\phi_{dx}(y)) = O_p^\star\left(\sqrt{\frac{\log(n)}{nh_\zeta}} + h_\zeta^2\right), \tag{S66}$$

uniformly in $(y,b_\zeta)\in\dot{I}_{d'x}\times\mathbb{H}_\zeta$. (S66) implies that

$$\Pr\left[\inf_{(y,b_\zeta)\in\dot{I}_{d'x}\times\mathbb{H}_\zeta}\left|\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(y);b_\zeta\right)\right|>\frac{1}{2}\underline{\zeta}_{dx}\right]>1-C_2n^{-1}. \tag{S67}$$

Denote

$$\overline{\zeta}\coloneqq\max_{j=1,\ldots,n}\mathbb{1}\left(D_j=d',X_j=x\right)\mathbb{1}_j\left(v;\overline{h}\right)\left|\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(Y_j);b_\zeta\right)-\zeta_{dx}\left(\phi_{dx}(Y_j)\right)\right|. \tag{S68}$$

Then, $\overline{\zeta}=O_p^\star\left(\sqrt{\log(n)/(nh_\zeta)}+h_\zeta^2\right)$, uniformly in $(v,b_\zeta)\in I_x\times\mathbb{H}_\zeta$. Then,

$$|T_{5.3}(v;b,b_\zeta)|\lesssim b^{-1}\left(n^3/n_{(3)}\right)\underline{\zeta}_{dx}^{-1}\left(\underline{\zeta}_{1x}^{-1}+\underline{\zeta}_{0x}^{-1}\right)\mathbb{1}_{\Delta X}(v,x;b)^2\left(\overline{\phi}+\overline{R}\right)\overline{\zeta}=O_p^\star\left(\sqrt{\frac{\log(n)}{nh^2}}\left(\sqrt{\frac{\log(n)}{nh_\zeta}}+h_\zeta^2\right)\right), \tag{S69}$$

uniformly in $(v,b,b_\zeta)\in I_x\times\mathbb{H}\times\mathbb{H}_\zeta$. Then,

$$\begin{aligned}|T_{5.2}(v;b,b_\zeta)|&\lesssim\frac{1}{n}\sum_{j=1}^n\underline{\zeta}_{dx}^{-2}b^{-3}\mathbb{1}_j(v;b)\,\mathbb{1}\left(D_j=d',X_j=x\right)\left|\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(Y_j);b_\zeta\right)-\zeta_{dx}\left(\phi_{dx}(Y_j)\right)\right|\\&\left|\frac{1}{(n-1)(n-2)}\sum_{i\neq j}\sum_{k\neq j,k\neq j}\varphi_{dx}(W_j,W_i)\left\{K'\left(\frac{\widehat{\Delta}_k-v}{b}\right)-K'\left(\frac{\Delta_k-v}{b}\right)\right\}q_x(W_k,W_i)\,\mathbb{1}(X_i=x)\right|\lesssim\\&b^{-2}\underline{\zeta}_{dx}^{-2}\left(\underline{\zeta}_{1x}^{-1}+\underline{\zeta}_{0x}^{-1}\right)\mathbb{1}_{\Delta X}(v,x;b)^2\,\overline{\zeta}\cdot\overline{\Delta}=O_p^\star\left(\sqrt{\frac{\log(n)}{nh^4}}\left(\sqrt{\frac{\log(n)}{nh_\zeta}}+h_\zeta^2\right)\right),\end{aligned} \tag{S70}$$

uniformly in $(v,b,b_\zeta)\in I_x\times\mathbb{H}\times\mathbb{H}_\zeta$. Denote

$$\mathcal{Z}_x(U_i,U_k,v,e;b)\coloneqq\frac{1}{b^2}K'\left(\frac{\Delta_x(\epsilon_k)-v}{b}\right)\varpi_x(U_k)\left(\mathbb{1}(\epsilon_i\le\epsilon_k)-F_{\epsilon|X}(\epsilon_k\mid x)\right)\left(\mathbb{1}(\epsilon_i\le e)-F_{\epsilon|X}(e\mid x)\right)\mathbb{1}(X_i=x).$$

Then,

$$\begin{aligned}|T_{5.1}(v;b,b_\zeta)|&\lesssim\frac{1}{n_{(3)}}\sum_{j=1}^n\underline{\zeta}_{dx}^{-2}b^{-3}\mathbb{1}_j(v;b)\,\mathbb{1}\left(D_j=d',X_j=x\right)\left|\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}(Y_j);b_\zeta\right)-\zeta_{dx}\left(\phi_{dx}(Y_j)\right)\right|\\&\times\left|\sum_{i\neq j}\sum_{k\neq j,k\neq j}\left(\mathbb{1}(\epsilon_i\le\epsilon_j)-F_{\epsilon|X}(\epsilon_j\mid x)\right)K'\left(\frac{\Delta_x(\epsilon_k)-v}{b}\right)\varpi_x(U_k)\left(\mathbb{1}(\epsilon_i\le\epsilon_k)-F_{\epsilon|X}(\epsilon_k\mid x)\right)\mathbb{1}(X_i=x)\right|\\&\lesssim\underline{\zeta}_{dx}^{-2}\mathbb{1}_{\Delta X}(v,x;b)\,\overline{\zeta}\sup_{e\in\left[\underline{\epsilon}_x,\overline{\epsilon}_x\right]}\left|\frac{1}{n_{(2)}}\sum_{(i,k)}\mathcal{Z}_x(U_i,U_k,v,e;b)\right|.\end{aligned}$$

Let $\mathring{\mathcal{Z}}_x(U_i,U_k,v,e;b)\coloneqq\left(\mathcal{Z}_x(U_i,U_k,v,e;b)+\mathcal{Z}_x(U_k,U_i,v,e;b)\right)/2$ denote the symmetrization of $\mathcal{Z}_x(U_i,U_k,v,e;b)$ so that $n_{(2)}^{-1}\sum_{(i,k)}\mathcal{Z}_x(U_i,U_k,v,e;b)=n_{(2)}^{-1}\sum_{(i,k)}\mathring{\mathcal{Z}}_x(U_i,U_k,v,e;b)$. By Chernozhukov et al. (2014a, Lemma B.2), $\mathfrak{Z}\coloneqq\left\{\mathring{\mathcal{Z}}_x(\cdot,v,e;b):(v,e,b)\in I_x\times\left[\underline{\epsilon}_x,\overline{\epsilon}_x\right]\times\mathbb{H}\right\}$ is uniformly VC-type with respect to the constant envelope $F_{\mathfrak{Z}}=O\left(h^{-2}\right)$. Let $\mathring{\mathcal{Z}}_x^{\langle 1\rangle}(u,v,e;b)\coloneqq\mathrm{E}\left[\mathring{\mathcal{Z}}_x(u,U,v,e;b)\right]$ and $\mathfrak{Z}^{\langle 1\rangle}\coloneqq\left\{\mathring{\mathcal{Z}}_x^{\langle 1\rangle}(\cdot,v,e;b):(v,e,b)\in I_x\times\left[\underline{\epsilon}_x,\overline{\epsilon}_x\right]\times\mathbb{H}\right\}$. Note that

$$\sup_{(v,e,b)\in I_x\times\left[\underline{\epsilon}_x,\overline{\epsilon}_x\right]\times\mathbb{H}}\left|\frac{1}{n_{(2)}}\sum_{(i,k)}\mathcal{Z}_x(U_i,U_k,v,e;b)-\vartheta_x(e,v;b)\,p_x\right|=n^{-1/2}\left\|\mathbb{U}_n^{(2)}\right\|_{\mathfrak{Z}},$$

since by calculation, $\mathrm{E}\left[\mathcal{Z}_x(U_1,U_2,v,e;b)\right]=\vartheta_x(e,v;b)\,p_x$. It is shown by (S55) that $\vartheta_x(e,v;b)=O(1)$, uniformly in $(v, e, b) \in I_x \times [\underline{\epsilon}_x, \overline{\epsilon}_x] \times \mathbb{H}$. By the c_r inequality,

$$\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{Z}_x\left(U_1, U_2, v, e; b\right) \mid U_1\right]\right)^2\right] =$$

$$\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x} \left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x} \frac{1}{b^2} K'\left(\frac{\Delta_x\left(e'\right) - v}{b}\right) \rho_x\left(e'\right)\left(\mathbb{1}\left(e'' \leq e'\right) - F_{\epsilon|X}\left(e' \mid x\right)\right) \mathrm{d}e'\right\}^2 \left(\mathbb{1}\left(e'' \leq e\right) - F_{\epsilon|X}\left(e \mid x\right)\right)^2 f_{\epsilon X}\left(e'', x\right) \mathrm{d}e''$$

$$\lesssim \sum_{j=1}^{m}\sum_{k=1}^{m} \int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2} K'\left(\frac{\Delta_x\left(e'\right) - v}{b}\right) \rho_x\left(e'\right)\left(\mathbb{1}\left(e'' \leq e'\right) - F_{\epsilon|X}\left(e' \mid x\right)\right) \mathrm{d}e'\right\}^2 f_{\epsilon X}\left(e'', x\right) \mathrm{d}e'',$$

where $f_{\epsilon X}(e, x) \coloneqq f_{\epsilon|X}(e \mid x) p_x$. If $k < j$,

$$\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2} K'\left(\frac{\Delta_x\left(e'\right) - v}{b}\right) \rho_x\left(e'\right)\left(\mathbb{1}\left(e'' \leq e'\right) - F_{\epsilon|X}\left(e' \mid x\right)\right) \mathrm{d}e'\right\}^2 f_{\epsilon X}\left(e'', x\right) \mathrm{d}e'' =$$

$$\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2} K'\left(\frac{\Delta_x\left(e'\right) - v}{b}\right) \rho_x\left(e'\right) F_{\epsilon|X}\left(e' \mid x\right) \mathrm{d}e'\right\}^2 f_{\epsilon X}\left(e'', x\right) \mathrm{d}e'' = O(1)$$

and if $k > j$,

$$\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2} K'\left(\frac{\Delta_x\left(e'\right) - v}{b}\right) \rho_x\left(e'\right)\left(\mathbb{1}\left(e'' \leq e'\right) - F_{\epsilon|X}\left(e' \mid x\right)\right) \mathrm{d}e'\right\}^2 f_{\epsilon X}\left(e'', x\right) \mathrm{d}e'' =$$

$$\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2} K'\left(\frac{\Delta_x\left(e'\right) - v}{b}\right) \rho_x\left(e'\right)\left(1 - F_{\epsilon|X}\left(e' \mid x\right)\right) \mathrm{d}e'\right\}^2 f_{\epsilon X}\left(e'', x\right) \mathrm{d}e'' = O(1),$$

uniformly in $(v, e, b) \in I_x \times [\underline{\epsilon}_x, \overline{\epsilon}_x] \times \mathbb{H}$. If $k = j$, since $e'' \in (\epsilon_{x,j-1}, \epsilon_{x,j})$, by integration by parts and change of variables,

$$\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b} K'\left(\frac{\Delta_x\left(e'\right) - v}{b}\right) \rho_x\left(e'\right) \mathbb{1}\left(e'' \leq e'\right) \mathrm{d}e' = -K\left(\frac{\Delta_x\left(e''\right) - v}{b}\right) \psi_{x,j}\left(\Delta_x\left(e''\right)\right) - b\int_{\frac{\Delta_x\left(e''\right) - v}{b}}^{\frac{\Delta_x\left(\epsilon_{x,j}\right) - v}{b}} K(u)\, \psi'_{x,j}(bu + v)\, \mathrm{d}u \quad \text{(S71)}$$

and

$$\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^2} K'\left(\frac{\Delta_x\left(e'\right) - v}{b}\right) \rho_x\left(e'\right) F_{\epsilon|X}\left(e' \mid x\right) \mathrm{d}e' = -\chi'_{x,j}(v) + o(1), \quad \text{(S72)}$$

uniformly in $(v, b) \in I_x \times \mathbb{H}$. Therefore, $\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{Z}_x\left(U_1, U_2, v, e; b\right) \mid U_1\right]\right)^2\right] = O\left(h^{-1}\right)$, uniformly in $(v, e, b) \in I_x \times [\underline{\epsilon}_x, \overline{\epsilon}_x] \times \mathbb{H}$. By change of variables,

$$\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{Z}_x\left(U_1, U_2, v, e; b\right) \mid U_2\right]\right)^2\right] =$$

$$\sum_{j=1}^{m} \int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^4} K'\left(\frac{\Delta_x\left(e'\right) - v}{b}\right)^2 \tilde{\rho}_x\left(e'\right)\left\{F_{\epsilon|X}\left(e \wedge e' \mid x\right) - F_{\epsilon|X}(e \mid x) F_{\epsilon|X}\left(e' \mid x\right)\right\}^2 f_{\epsilon X}\left(e', x\right) \mathrm{d}e' = O\left(h^{-3}\right),$$

and $\mathrm{E}\left[\mathcal{Z}_x\left(U_1, U_2, v, e; b\right)^2\right] = O\left(h^{-3}\right)$, uniformly in $(v, e, b) \in I_x \times [\underline{\epsilon}_x, \overline{\epsilon}_x] \times \mathbb{H}$. Let $\sigma^2_{\mathfrak{Z}^{\langle 1\rangle}} \coloneqq \sup_{f \in \mathfrak{Z}^{\langle 1\rangle}} \mathbb{P}^U f^2$ and $\sigma^2_{\mathfrak{Z}^{\langle 2\rangle}} \coloneqq \sup_{f \in \mathfrak{Z}^{\langle 2\rangle}} \mathrm{E}\left[f\left(U_1, U_2\right)^2\right]$. We have shown that $\sigma^2_{\mathfrak{Z}^{\langle 1\rangle}} = O\left(h^{-3}\right)$ and $\sigma^2_{\mathfrak{Z}^{\langle 2\rangle}} = O\left(h^{-3}\right)$. By the coupling theorem (CK Proposition 2.1 with $\mathcal{H} = \mathfrak{Z}_{\pm}$, $\overline{\sigma}_{\mathfrak{g}} = \sigma_{\mathfrak{Z}^{\langle 1\rangle}}$, $\sigma_{\mathfrak{h}} = \sigma_{\mathfrak{Z}}$, $b_{\mathfrak{g}} = b_{\mathfrak{h}} = F_{\mathfrak{Z}}$, $\chi_n = 0$, $\gamma = \sqrt{\log(n)/(nh)}$

and $q = \infty$), Dudley's entropy integral bound and the Borell-Sudakov-Tsirelson inequality, $\left\|\mathbb{U}_n^{(2)}\right\|_{\mathfrak{Z}} = O_p^\star\left(\sqrt{\log(n)/h^3}, \sqrt{\log(n)/(nh)}\right)$. Therefore, $T_{5.1}(v;b,b_\zeta) = O_p^\star\left(\sqrt{\log(n)/(nh_\zeta)} + h_\zeta^2, \sqrt{\log(n)/(nh)}\right)$, uniformly in $(v,e,b) \in I_x \times [\underline{\epsilon}_x, \overline{\epsilon}_x] \times \mathbb{H}$. Now it follows that $\|T_5\|_{I_x\times\mathbb{H}\times\mathbb{H}_\zeta} = O_p^\star\left(\sqrt{\log(n)/(nh_\zeta)} + h_\zeta^2, \sqrt{\log(n)/(nh)}\right)$. By (S67) and similar arguments, $\|T_6\|_{I_x\times\mathbb{H}\times\mathbb{H}_\zeta} = O_p^\star\left(\left(\sqrt{\log(n)/(nh_\zeta)} + h_\zeta^2\right)^2, \sqrt{\log(n)/(nh)}\right)$. Then it follows that

$$\frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\varDelta}_j - v}{b}\right)\varOmega_{dx}(W_j,W_i;b_\zeta)K'\left(\frac{\widehat{\varDelta}_k - v}{b}\right)q_x(W_k,W_i)\mathbb{1}(X_i = x)$$
$$= O_p^\star\left(\sqrt{\frac{\log(n)}{nh^2}} + \sqrt{\frac{\log(n)}{nh_\zeta}} + h_\zeta^2, \sqrt{\frac{\log(n)}{nh}}\right).$$

By tedious algebra, (S63), Lemma 6 and (S65),

$$\frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\varDelta}_j - v}{b}\right)\varOmega(W_j,W_i;b_\zeta)K'\left(\frac{\widehat{\varDelta}_k - v}{b}\right)\varOmega(W_k,W_i;b_\zeta)\mathbb{1}(X_i = x) =$$
$$\frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\widehat{\varDelta}_j - v}{b}\right)\mathbb{1}(D_j = d', X_j = x)\frac{\mathbb{1}\left(Y_i \le \widehat{\phi}_{dx}(Y_j)\right) - \mathbb{1}(Y_i \le \phi_{dx}(Y_j))}{\zeta_{dx}(\phi_{dx}(Y_j))}$$
$$\times K'\left(\frac{\widehat{\varDelta}_k - v}{b}\right)\mathbb{1}(D_k = d', X_k = x)\frac{\mathbb{1}\left(Y_i \le \widehat{\phi}_{dx}(Y_k)\right) - \mathbb{1}(Y_i \le \phi_{dx}(Y_k))}{\zeta_{dx}(\phi_{dx}(Y_k))}\mathbb{1}(D_i = d, X_i = x)$$
$$+ O_p^\star\left(\sqrt{\frac{\log(n)}{nh^2}}\left(\sqrt{\frac{\log(n)}{nh_\zeta}} + h_\zeta^2\right)\right) =: T_7(v;b) + O_p^\star\left(\sqrt{\frac{\log(n)}{nh^2}}\left(\sqrt{\frac{\log(n)}{nh_\zeta}} + h_\zeta^2\right)\right),$$

uniformly in $(v,b,b_\zeta) \in I_x \times \mathbb{H} \times \mathbb{H}_\zeta$. Then, $|T_7(v;b)| \lesssim b^{-1}\underline{\zeta}_{dx}^{-2}\mathbb{1}_{\varDelta X}(v,x;b)^2\,\overline{\phi} = O_p^\star\left(\sqrt{\log(n)/(nh^2)}\right)$, uniformly in $(v,b,b_\zeta) \in I_x \times \mathbb{H} \times \mathbb{H}_\zeta$. Then, by these results and (S62),

$$\widetilde{V}_2(v,x;b,b_\zeta) - \dot{V}_2(v,x;b) = O_p^\star\left(\sqrt{\frac{\log(n)}{nh^2}} + \sqrt{\frac{\log(n)}{nh_\zeta}} + h_\zeta^2, \sqrt{\frac{\log(n)}{nh}}\right), \tag{S73}$$

uniformly in $(v,b,b_\zeta) \in I_x \times \mathbb{H} \times \mathbb{H}_\zeta$. By (39), we can write

$$\begin{aligned}\ddot{V}_2(v,x;b) &= \frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^3}K'\left(\frac{\varDelta_x(\epsilon_j) - v}{b}\right)\varpi_x(U_j)\left\{\mathbb{1}(\epsilon_i \le \epsilon_j) - F_{\epsilon|X}(\epsilon_j \mid x)\right\}\\ &\quad\times K'\left(\frac{\varDelta_x(\epsilon_k) - v}{b}\right)\varpi_x(U_k)\left\{\mathbb{1}(\epsilon_i \le \epsilon_k) - F_{\epsilon|X}(\epsilon_k \mid x)\right\}\mathbb{1}(X_i = x)\\ &=: \frac{1}{n_{(3)}}\sum_{(i,j,k)}\mathcal{J}_x(U_i,U_j,U_k,v;b).\end{aligned}$$

Let

$$\bar{V}_2(v,x;b) := \mathrm{E}\left[\frac{1}{b^3}K'\left(\frac{\varDelta_3 - v}{b}\right)q_x(W_3,W_1)K'\left(\frac{\varDelta_2 - v}{b}\right)q_x(W_2,W_1)\mathbb{1}(X_1 = x)\right] = \mathrm{E}\left[\mathcal{J}_x(U_1,U_2,U_3,v;b)\right]$$

and therefore, $V_2(v,x;b) = \bar{V}_2(v,x;b)\,p_x^{-1}\left(p_{1x}^{-1} + p_{0x}^{-1}\right)$. Write $\ddot{V}_2(v,x;b) = n_{(3)}^{-1}\sum_{(i,j,k)}\mathring{\mathcal{J}}_x(U_i,U_j,U_k,v;b)$, where $\mathring{\mathcal{J}}_x$ denotes the symmetrization of $\mathcal{J}_x$ (see (S46)). Denote $\mathring{\mathcal{J}}_x^{\langle 1\rangle}(u,v;b) := \mathrm{E}\left[\mathring{\mathcal{J}}_x(u,U_1,U_2,v;b)\right]$ and

$\mathring{\mathcal{J}}_x^{\langle 2\rangle}(u_1,u_2,v;b) \coloneqq \mathrm{E}\left[\mathring{\mathcal{J}}_x(u_1,u_2,U_1,v;b)\right]$. Let $\mathfrak{J} \coloneqq \left\{\mathring{\mathcal{J}}_x(\cdot,v;b) : v \in I_x\right\}$, $\mathfrak{J}^{\langle 1\rangle} \coloneqq \left\{\mathring{\mathcal{J}}_x^{\langle 1\rangle}(\cdot,v;b) : (v,b) \in I_x \times \mathbb{H}\right\}$ and $\mathfrak{J}^{\langle 2\rangle} \coloneqq \left\{\mathring{\mathcal{J}}_x^{\langle 2\rangle}(\cdot,v;b) : (v,b) \in I_x \times \mathbb{H}\right\}$. Then, we have $\left\|\ddot{V}_2(\cdot,x;\cdot) - \bar{V}_2(\cdot,x;\cdot)\right\|_{I_x\times\mathbb{H}} = n^{-1/2}\left\|\mathbb{U}_n^{(3)}\right\|_{\mathfrak{J}}$. By similar arguments used in the proof of Lemma 4, $\mathfrak{J}$ is uniformly VC-type with respect to a constant envelope $F_{\mathfrak{J}} = O\left(h^{-3}\right)$. By Lemma A.3 of CK, $\mathfrak{J}^{\langle 1\rangle}$ and $\mathfrak{J}^{\langle 2\rangle}$ are also uniformly VC-type with respect to constant envelopes $F_{\mathfrak{J}^{\langle 1\rangle}} = F_{\mathfrak{J}^{\langle 2\rangle}} = F_{\mathfrak{J}}$. Then we have

$$\begin{aligned}
&\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x(U_1,U_2,U_3,v;b) \mid U_1\right]\right)^2\right] = \\
&\qquad \int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x} \frac{1}{b^2}\left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x} \frac{1}{b}K'\left(\frac{\Delta_x(e)-v}{b}\right)\rho_x(e)\left\{\mathbb{1}(e'\le e) - F_{\epsilon|X}(e\mid x)\right\}\mathrm{d}e\right\}^4 f_{\epsilon X}(e',x)\,\mathrm{d}e' \lesssim \\
&\sum_{k=1}^m\sum_{j=1}^m \int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2}\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b}K'\left(\frac{\Delta_x(e)-v}{b}\right)\rho_x(e)\left\{\mathbb{1}(e'\le e) - F_{\epsilon|X}(e\mid x)\right\}\mathrm{d}e\right\}^4 f_{\epsilon X}(e',x)\,\mathrm{d}e'. \quad \text{(S74)}
\end{aligned}$$

If $k<j$, by (S51),

$$\begin{aligned}
&\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2}\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b}K'\left(\frac{\Delta_x(e)-v}{b}\right)\rho_x(e)\left\{\mathbb{1}(e'\le e) - F_{\epsilon|X}(e\mid x)\right\}\mathrm{d}e\right\}^4 f_{\epsilon X}(e',x)\,\mathrm{d}e' = \\
&\qquad = \int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2}\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b}K'\left(\frac{\Delta_x(e)-v}{b}\right)\rho_x(e)\left\{1 - F_{\epsilon|X}(e\mid x)\right\}\mathrm{d}e\right\}^4 f_{\epsilon X}(e',x)\,\mathrm{d}e' = O\left(h^2\right),
\end{aligned}$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$. Similarly, if $k>j$,

$$\begin{aligned}
&\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2}\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b}K'\left(\frac{\Delta_x(e)-v}{b}\right)\rho_x(e)\left\{\mathbb{1}(e'\le e) - F_{\epsilon|X}(e\mid x)\right\}\mathrm{d}e\right\}^4 f_{\epsilon X}(e',x)\,\mathrm{d}e' = \\
&\qquad \int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2}\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b}K'\left(\frac{\Delta_x(e)-v}{b}\right)\rho_x(e)F_{\epsilon|X}(e\mid x)\,\mathrm{d}e\right\}^4 f_{\epsilon X}(e',x)\,\mathrm{d}e' = O\left(h^2\right).
\end{aligned}$$

If $k=j$, by (S71) and (S72),

$$\begin{aligned}
&\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^2}\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b}K'\left(\frac{\Delta_x(e)-v}{b}\right)\rho_x(e)\left\{\mathbb{1}(e'\le e) - F_{\epsilon|X}(e\mid x)\right\}\mathrm{d}e\right\}^4 f_{\epsilon X}(e',x)\,\mathrm{d}e' \lesssim \\
&\qquad \int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} \frac{1}{b^2}K\left(\frac{\Delta_x(e')-v}{b}\right)^4 \psi_{x,j}\left(\Delta_x(e')\right)^4 f_{\epsilon X}(e',x)\,\mathrm{d}e' + O\left(h^2\right) = O\left(h^{-1}\right),
\end{aligned}$$

uniformly in $(v,b) \in I_x \times \mathbb{H}$. It follows from these calculations and (S74) that $\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x(U_1,U_2,U_3,v;b) \mid U_1\right]\right)^2\right] = O\left(h^{-1}\right)$, uniformly in $(v,b) \in I_x \times \mathbb{H}$. We have

$$\begin{aligned}
&\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x(U_1,U_2,U_3,v;b) \mid U_3\right]\right)^2\right] = \left\{\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x} \frac{1}{b^2}K'\left(\frac{\Delta_x(e)-v}{b}\right)^2 \tilde{\rho}_x(e)^2\right. \\
&\quad \left.\times\left(\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x} \frac{1}{b^2}K'\left(\frac{\Delta_x(e')-v}{b}\right)\rho_x(e')\left\{F_{\epsilon|X}(e\wedge e'\mid x) - F_{\epsilon|X}(e\mid x)F_{\epsilon|X}(e'\mid x)\right\}\mathrm{d}e'\right)^2 \mathrm{d}e\right\} p_x^2 \lesssim \\
&\qquad \sum_{k=1}^m\sum_{j=1}^m\left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} \frac{1}{b^2}K'\left(\frac{\Delta_x(e)-v}{b}\right)^2 \tilde{\rho}_x(e)^2\,\vartheta_j(e,v;b)^2\,\mathrm{d}e\right\} p_x^2 \quad \text{(S75)}
\end{aligned}$$

and $\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x\left(U_1,U_2,U_3,v;b\right)\mid U_2\right]\right)^2\right]=\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x\left(U_1,U_2,U_3,v;b\right)\mid U_3\right]\right)^2\right]$. If $k<j$, by change of variables, integration by parts and (S51),

$$\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\frac{1}{b^2}K'\left(\frac{\Delta_x\left(e\right)-v}{b}\right)^2\tilde{\rho}_x\left(e\right)^2\vartheta_{x,j}\left(e,v;b\right)^2\mathrm{d}e=\left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\frac{1}{b^2}K'\left(\frac{\Delta_x\left(e\right)-v}{b}\right)^2\tilde{\rho}_x\left(e\right)^2F_{\epsilon|X}\left(e\mid x\right)^2\mathrm{d}e\right\}$$
$$\times\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\left\{1-F_{\epsilon|X}\left(e'\mid x\right)\right\}\frac{1}{b^2}K'\left(\frac{\Delta_x\left(e'\right)-v}{b}\right)\rho_x\left(e'\right)\mathrm{d}e'\right\}^2=O\left(h^{-1}\right),$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Similarly, if $k>j$, by change of variables, integration by parts and (S72),

$$\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\frac{1}{b^2}K'\left(\frac{\Delta_x\left(e\right)-v}{b}\right)^2\tilde{\rho}_x\left(e\right)^2\vartheta_{x,j}\left(e,v;b\right)^2\mathrm{d}e=\left\{\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\frac{1}{b^2}K'\left(\frac{\Delta_x\left(e\right)-v}{b}\right)^2\tilde{\rho}_x\left(e\right)^2\left(1-F_{\epsilon|X}\left(e\mid x\right)\right)^2\mathrm{d}e\right\}$$
$$\times\left\{\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}F_{\epsilon|X}\left(e'\mid x\right)\frac{1}{b^2}K'\left(\frac{\Delta_x\left(e'\right)-v}{b}\right)\rho_x\left(e'\right)\mathrm{d}e'\right\}^2=O\left(h^{-1}\right),$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. If $k=j$, by (S55), $\vartheta_{x,j}\left(e,v;b\right)=\bar{\vartheta}_{x,j}\left(e,v;b\right)+O\left(h\right)$, (S56) and change of variables,

$$\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\frac{1}{b^2}K'\left(\frac{\Delta_x\left(e\right)-v}{b}\right)^2\tilde{\rho}_x\left(e\right)^2\vartheta_{x,j}\left(e,v;b\right)^2\mathrm{d}e=O\left(h^{-1}\right),$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. It follows from these calculations and (S75) that $\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x\left(U_1,U_2,U_3,v;b\right)\mid U_2\right]\right)^2\right]=\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x\left(U_1,U_2,U_3,v;b\right)\mid U_3\right]\right)^2\right]=O\left(h^{-1}\right)$. Then, by the c_r inequality,

$$\begin{aligned}\mathrm{E}\left[\mathring{\mathcal{J}}_x^{\langle 1\rangle}\left(U,v;b\right)^2\right]&\lesssim\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x\left(U_1,U_2,U_3,v;b\right)\mid U_1\right]\right)^2\right]+\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x\left(U_1,U_2,U_3,v;b\right)\mid U_2\right]\right)^2\right]\\&\quad+\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x\left(U_1,U_2,U_3,v;b\right)\mid U_3\right]\right)^2\right]\\&=O\left(h^{-1}\right),\end{aligned}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Therefore, $\sigma^2_{\mathfrak{J}^{\langle 1\rangle}}\coloneqq\sup_{f\in\mathfrak{J}^{\langle 1\rangle}}\mathbb{P}^Uf^2=O\left(h^{-1}\right)$. It is easy to check that

$$\begin{aligned}\mathrm{E}\left[\mathring{\mathcal{J}}_x^{\langle 2\rangle}\left(U_1,U_2,v;b\right)^2\right]&\lesssim\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x\left(U_1,U_2,U_3,v;b\right)\mid U_2,U_3\right]\right)^2\right]+\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x\left(U_1,U_2,U_3,v;b\right)\mid U_1,U_3\right]\right)^2\right]\\&\quad+\mathrm{E}\left[\left(\mathrm{E}\left[\mathcal{J}_x\left(U_1,U_2,U_3,v;b\right)\mid U_1,U_2\right]\right)^2\right]\\&=O\left(h^{-4}\right),\end{aligned}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Therefore, $\sigma^2_{\mathfrak{J}^{\langle 2\rangle}}\coloneqq\sup_{f\in\mathfrak{J}^{\langle 2\rangle}}\mathrm{E}\left[f\left(U_1,U_2\right)^2\right]=O\left(h^{-4}\right)$. By the coupling theorem (Proposition 2.1 of CK with $\mathcal{H}=\mathfrak{J}_\pm$, $\overline{\sigma}_{\mathfrak{g}}=\sigma_{\mathfrak{J}^{\langle 1\rangle}}$, $\sigma_{\mathfrak{h}}=\sigma_{\mathfrak{J}^{\langle 2\rangle}}$, $b_{\mathfrak{g}}=b_{\mathfrak{h}}=F_{\mathfrak{J}}$ and $q=\infty$), when n is sufficiently large so that $V_{\mathfrak{J}_\pm}\cdot\log\left(A_{\mathfrak{J}_\pm}\vee n\right)\le n^{1/3}$, $\forall\gamma_2\in(0,1)$, there exists a random variable $Z_{\mathfrak{J}_\pm,\gamma_2}=_d\left\|G^U\right\|_{\mathfrak{J}^{\langle 1\rangle}}$ such that

$$\Pr\left[\left|\left\|\mathbb{U}_n^{(3)}\right\|_{\mathfrak{J}}-Z_{\mathfrak{J}_\pm,\gamma_2}\right|>C_1\left(\frac{\log\left(n\right)^{2/3}}{n^{1/6}h^{4/3}\gamma_2^{1/3}}+\frac{\log\left(n\right)}{n^{1/2}h^3\gamma_2}\right)\right]\le C_2\left(\gamma_2+n^{-1}\right).$$

By the Borell-Sudakov-Tsirelson inequality, $\Pr\left[\left\|G^U\right\|_{\mathfrak{J}^{\langle 1\rangle}}>\mathrm{E}\left[\left\|G^U\right\|_{\mathfrak{J}^{\langle 1\rangle}}\right]+\sqrt{2\cdot\log\left(n\right)}\sigma_{\mathfrak{J}^{\langle 1\rangle}}\right]\le n^{-1}$. By Dudley's entropy integral bound, $\mathrm{E}\left[\left\|G^{U}\right\|_{\mathfrak{J}^{\langle 1\rangle}}\right] \lesssim \left(\sigma_{\mathfrak{J}^{\langle 1\rangle}} \vee n^{-1/2}\left\|F_{\mathfrak{J}^{\langle 1\rangle}}\right\|_{\mathbb{P}^{U},2}\right)\sqrt{\log(n)}$. It now follows that

$$\Pr\left[\left\|\mathbb{U}_{n}^{(3)}\right\|_{\mathfrak{J}} > C_{1}\left(\frac{\log(n)^{2/3}}{n^{1/6}h^{4/3}\gamma_{2}^{1/3}} + \frac{\log(n)}{n^{1/2}h^{3}\gamma_{2}} + \sqrt{\frac{\log(n)}{h}}\right)\right] \leq C_{2}\left(\gamma_{2} + n^{-1}\right) \tag{S76}$$

and therefore, we have

$$\Pr\left[\left\|\ddot{V}_{2}(\cdot, x; \cdot) - \bar{V}_{2}(\cdot, x; \cdot)\right\|_{I_{x}\times\mathbb{H}} > C_{1}\left(\frac{\log(n)^{2/3}}{n^{2/3}h^{4/3}\gamma_{2}^{1/3}} + \frac{\log(n)}{nh^{3}\gamma_{2}} + \sqrt{\frac{\log(n)}{nh}}\right)\right] \leq C_{2}\left(\gamma_{2} + n^{-1}\right). \tag{S77}$$

Since $\widehat{V}_{2}(v, x; b, b_{\zeta}) = \widetilde{V}_{2}(v, x; b, b_{\zeta})\,\widehat{p}_{x}^{-1}\left(\widehat{p}_{1x}^{-1} + \widehat{p}_{0x}^{-1}\right)$, $V_{2}(v, x; b) = \bar{V}_{2}(v, x; b)\,p_{x}^{-1}\left(p_{1x}^{-1} + p_{0x}^{-1}\right)$ and

$$\begin{aligned}\widetilde{V}_{2}(v, x; b, b_{\zeta}) - \bar{V}_{2}(v, x; b) = {}& \left(\widetilde{V}_{2}(v, x; b, b_{\zeta}) - \dot{V}_{2}(v, x; b)\right) + \left(\dot{V}_{2}(v, x; b) - \ddot{V}_{2}(v, x; b)\right)\\ &+ \left(\ddot{V}_{2}(v, x; b) - \bar{V}_{2}(v, x; b)\right),\end{aligned}$$

it follows from (S73) and taking $\gamma_{1} = \gamma_{2} = \gamma$ in (S61) and (S77), $\widehat{p}_{x} - p_{x} = O_{p}^{\star}\left(n^{-1/2}\right)$ and $\widehat{p}_{zx} - p_{zx} = O_{p}^{\star}\left(n^{-1/2}\right)$ that $\forall\gamma \in (0, 1)$,

$$\Pr\left[\sup_{(v,b,b_{\zeta})\in I_{x}\times\mathbb{H}\times\mathbb{H}_{\zeta}}\left|\widehat{V}_{2}(v, x; b, b_{\zeta}) - V_{2}(v, x; b)\right| > C_{1}\kappa_{1}^{V}(\gamma)\right] \leq C_{2}\kappa_{2}^{V}(\gamma),$$

when n is sufficiently large. The first assertion follows from this result, (S40), and

$$\begin{aligned}\widehat{V}(v \mid x; b, b_{\zeta}) - V(v \mid x; b) = {}& \frac{\widehat{V}_{1}(v, x; b) + \widehat{V}_{2}(v, x; b, b_{\zeta})}{p_{x}^{2}}\left(\frac{p_{x}^{2}}{\widehat{p}_{x}^{2}} - 1\right)\\ &+ \frac{1}{p_{x}^{2}}\left\{\left(\widehat{V}_{1}(v, x; b) - V_{1}(v, x; b)\right) + \left(\widehat{V}_{2}(v, x; b, b_{\zeta}) - V_{2}(v, x; b)\right)\right\}.\end{aligned}$$

For the second part, note that $V_{1}(v, x; b) = r_{\Delta X}(v, x; b) - b\cdot m_{\Delta X}(v, x; b)^{2}$. Let

$$\tilde{L}(u; b, h) := \left(\frac{h}{b}\right)K\left(\frac{h}{b}u\right)^{2} - K(u)^{2} = \left(\left(\frac{h}{b}\right)^{1/2}K\left(\frac{h}{b}u\right) + K(u)\right)L(u; b, h).$$

It is easy to check that

$$\int\tilde{L}(u; b, h)\,\mathrm{d}u = \int u\tilde{L}(u; b, h)\,\mathrm{d}u = 0 \tag{S78}$$

follows from change of variables. By (S38), $\forall b \in \mathbb{H}$,

$$\left|\tilde{L}(u; b, h)\right| \leq \left(\frac{2\left\|K\right\|_{\infty}}{\sqrt{1-\varepsilon_{n}}}\right)\left(\left\|K'\right\|_{\infty}\left(\frac{1+\varepsilon_{n}}{1-\varepsilon_{n}}\right) + \frac{\left\|K\right\|_{\infty}}{1-\varepsilon_{n}}\right)\varepsilon_{n}\mathbb{1}\left(|u| \leq 1+\varepsilon_{n}\right).$$

Then, by this result, change of variables, Taylor expansion and (S78),

$$r_{\Delta X}(v, x; b) - r_{\Delta X}(v, x; h) = \mathrm{E}\left[\frac{1}{h}\tilde{L}\left(\frac{\Delta - v}{h}; b, h\right)\mathbb{1}(X = x)\right] = O\left(h^{2}\varepsilon_{n}\right),$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Similarly, by change of variables and (S38),

$$\sqrt{b}\cdot m_{\Delta X}(v,x;b)-\sqrt{h}\cdot m_{\Delta X}(v,x;h)=\mathrm{E}\left[\frac{1}{\sqrt{h}}L\left(\frac{\Delta-v}{h};b,h\right)\mathbb{1}(X=x)\right]=O\left(h^{1/2}\varepsilon_n\right), \tag{S79}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Therefore, $b\cdot m_{\Delta X}(v,x;b)^2-h\cdot m_{\Delta X}(v,x;h)^2=O(h\varepsilon_n)$, uniformly in $(v,b)\in I_x\times\mathbb{H}$. Then it follows that

$$V_1(v,x;b)-V_1(v,x;h)=O(h\varepsilon_n), \tag{S80}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Then note that $V_2(v,x;b)=\bar{V}_2(v,x;b)\,p_x^{-1}\left(p_{1x}^{-1}+p_{0x}^{-1}\right)$ and

$$\begin{aligned}\bar{V}_2(v,x;b)=b^{-3}\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}K'\left(\frac{\Delta_x(e)-v}{b}\right)K'\left(\frac{\Delta_x(e')-v}{b}\right)F_{\epsilon X}(e\wedge e',x)\,\rho_x(e)\,\rho_x(e')\,\mathrm{d}e\mathrm{d}e'\\-b^{-3}\left\{\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}K'\left(\frac{\Delta_x(e)-v}{b}\right)F_{\epsilon|X}(e\mid x)\,\rho_x(e)\,\mathrm{d}e\right\}^2p_x,\end{aligned} \tag{S81}$$

where $F_{\epsilon X}(e,x)\coloneqq F_{\epsilon|X}(e\mid x)\,p_x$. By change of variables and integration by parts ,

$$\begin{aligned}\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}b^{-3/2}K'\left(\frac{\Delta_x(e)-v}{b}\right)\rho_x(e)\,F_{\epsilon|X}(e\mid x)\,\mathrm{d}e-\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}h^{-3/2}K'\left(\frac{\Delta_x(e)-v}{h}\right)\rho_x(e)\,F_{\epsilon|X}(e\mid x)\,\mathrm{d}e\\=h^{-3/2}\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)\rho_x(e)\,F_{\epsilon|X}(e\mid x)\,\mathrm{d}e=O\left(h^{1/2}\varepsilon_n\right),\end{aligned} \tag{S82}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$, where the second equality follows from $\int L'(u;b,h)\,\mathrm{d}u=0$ and (S35). And similarly,

$$\begin{aligned}\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}b^{-3/2}K'\left(\frac{\Delta_x(e)-v}{b}\right)\rho_x(e)\,\mathrm{d}e-\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}h^{-3/2}K'\left(\frac{\Delta_x(e)-v}{h}\right)\rho_x(e)\,\mathrm{d}e\\=h^{-3/2}\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)\rho_x(e)\,\mathrm{d}e=O\left(h^{1/2}\varepsilon_n\right),\end{aligned} \tag{S83}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Therefore, by this result and (S26),

$$b^{-3}\left\{\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}K'\left(\frac{\Delta_x(e)-v}{b}\right)\rho_x(e)\,F_{\epsilon|X}(e\mid x)\,\mathrm{d}e\right\}^2-h^{-3}\left\{\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}K'\left(\frac{\Delta_x(e)-v}{h}\right)\rho_x(e)\,F_{\epsilon|X}(e\mid x)\,\mathrm{d}e\right\}^2=O(h\varepsilon_n). \tag{S84}$$

Then, it is easy to see that

$$\begin{aligned}&b^{-3}\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}\int_{\underline{\epsilon}_x}^{\bar{\epsilon}_x}K'\left(\frac{\Delta_x(e)-v}{b}\right)K'\left(\frac{\Delta_x(e')-v}{b}\right)F_{\epsilon X}(e\wedge e',x)\,\rho_x(e)\,\rho_x(e')\,\mathrm{d}e\mathrm{d}e'=\\&b^{-3}\sum_{k=1}^{m}\sum_{j=1}^{m}\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}K'\left(\frac{\Delta_x(e)-v}{b}\right)K'\left(\frac{\Delta_x(e')-v}{b}\right)F_{\epsilon X}(e\wedge e',x)\,\rho_x(e)\,\rho_x(e')\,\mathrm{d}e\mathrm{d}e'=\\&\sum_{j=1}^{m}\bar{V}_{2,j}(v,x;b)\,p_x+2b^{-3}\sum_{k<j}\left(\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}}K'\left(\frac{\Delta_x(e)-v}{b}\right)F_{\epsilon X}(e,x)\,\rho_x(e)\,\mathrm{d}e\right)\left(\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}K'\left(\frac{\Delta_x(e')-v}{b}\right)\rho_x(e')\,\mathrm{d}e'\right),\end{aligned} \tag{S85}$$

where

$$\bar{V}_{2,j}(v,x;b) := b^{-3}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}}\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} K'\left(\frac{\Delta_x(e)-v}{b}\right)K'\left(\frac{\Delta_x(e')-v}{b}\right)F_{\epsilon|X}(e\wedge e' \mid x)\,\rho_x(e)\,\rho_x(e')\,\mathrm{d}e\mathrm{d}e'.$$

By(S25), (S26), (S82) and (S83),

$$\begin{aligned}
&b^{-3}\sum_{k<j}\left(\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} K'\left(\frac{\Delta_x(e)-v}{b}\right)F_{\epsilon X}(e,x)\,\rho_x(e)\,\mathrm{d}e\right)\left(\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} K'\left(\frac{\Delta_x(e')-v}{b}\right)\rho_x(e')\,\mathrm{d}e'\right)-\\
&h^{-3}\sum_{k<j}\left(\int_{\epsilon_{x,k-1}}^{\epsilon_{x,k}} K'\left(\frac{\Delta_x(e)-v}{h}\right)F_{\epsilon X}(e,x)\,\rho_x(e)\,\mathrm{d}e\right)\left(\int_{\epsilon_{x,j-1}}^{\epsilon_{x,j}} K'\left(\frac{\Delta_x(e')-v}{h}\right)\rho_x(e')\,\mathrm{d}e'\right)=O(h\varepsilon_n),
\end{aligned}\tag{S86}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. By change of variables and integration by parts,

$$\begin{aligned}
\bar{V}_{2,j}(v,x;b) = 2h^{-1}&\int\left(\frac{h}{b}\right)^2 K'\left(\left(\frac{h}{b}\right)w\right)K\left(\left(\frac{h}{b}\right)w\right)\psi_{x,j}(hw+v)\,\chi_{x,j}(hw+v)\,\mathrm{d}w\\
&-2\int\int_{-\infty}^{w}\left(\frac{h}{b}\right)^2 K'\left(\left(\frac{h}{b}\right)w\right)K\left(\left(\frac{h}{b}\right)u\right)\psi_{x,j}(hw+v)\,\chi'_{x,j}(hu+v)\,\mathrm{d}u\mathrm{d}w.
\end{aligned}$$

Note that by change of variables and integration by parts,

$$\begin{aligned}
\int\left(\frac{h}{b}\right)^2 K'\left(\left(\frac{h}{b}\right)w\right)K\left(\left(\frac{h}{b}\right)w\right)\mathrm{d}w &= \int K'(w)\,K(w)\,\mathrm{d}w = 0\\
\int\left(\frac{h}{b}\right)^2 K'\left(\left(\frac{h}{b}\right)w\right)K\left(\left(\frac{h}{b}\right)w\right)w\mathrm{d}w &= \int K'(w)\,K(w)\,w\mathrm{d}w = -\frac{1}{2}\int K(u)^2\,\mathrm{d}u\\
\int\left(\frac{h}{b}\right)^2 K'\left(\left(\frac{h}{b}\right)w\right)K\left(\left(\frac{h}{b}\right)w\right)w^2\mathrm{d}w &= \int K'(w)\,K(w)\,w^2\mathrm{d}w = 0.
\end{aligned}\tag{S87}$$

By tedious but straightforward calculations, $\forall b\in\mathbb{H}$,

$$\begin{aligned}
&\left|\left(\frac{h}{b}\right)^2 K'\left(\left(\frac{h}{b}\right)w\right)K\left(\left(\frac{h}{b}\right)u\right)-K'(w)\,K(u)\right|\\
&\leq\left(\frac{2\|K\|_\infty\|K'\|_\infty}{(1-\varepsilon_n)^2}+\|K\|_\infty\|K''\|_\infty\left(\frac{1+\varepsilon_n}{1-\varepsilon_n}\right)+\|K'\|_\infty^2\left(\frac{1+\varepsilon_n}{1-\varepsilon_n}\right)\right)\varepsilon_n\mathbb{1}(|w|\leq 1+\varepsilon_n)\,\mathbb{1}(|u|\leq 1+\varepsilon_n).
\end{aligned}\tag{S88}$$

By (S87), (S88) and Taylor expansion,

$$\begin{aligned}
h^{-1}&\int\left(\frac{h}{b}\right)^2 K'\left(\left(\frac{h}{b}\right)w\right)K\left(\left(\frac{h}{b}\right)w\right)\psi_{x,j}(hw+v)\,\chi_{x,j}(hw+v)\,\mathrm{d}w\\
&-h^{-1}\int K'(w)\,K(w)\,\psi_{x,j}(hw+v)\,\chi_{x,j}(hw+v)\,\mathrm{d}w = O\left(h^2\varepsilon_n\right),
\end{aligned}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. By change of variables and integration by parts,

$$\int\int_{-\infty}^{w}\left(\frac{h}{b}\right)^2 K'\left(\left(\frac{h}{b}\right)w\right)K\left(\left(\frac{h}{b}\right)u\right)\mathrm{d}u\mathrm{d}w = \int\int_{-\infty}^{w} K'(w)\,K(u)\,\mathrm{d}u\mathrm{d}w = -\int K(u)^2\,\mathrm{d}u.$$

By this result and (S88),

$$\int\int_{-\infty}^{w}\left(\frac{h}{b}\right)^{2}K'\left(\left(\frac{h}{b}\right)w\right)K\left(\left(\frac{h}{b}\right)u\right)\psi_{x,j}\left(hw+v\right)\chi'_{x,j}\left(hu+v\right)\mathrm{d}u\mathrm{d}w$$
$$-\int\int_{-\infty}^{w}K'\left(w\right)K\left(u\right)\psi_{x,j}\left(hw+v\right)\chi'_{x,j}\left(hu+v\right)\mathrm{d}u\mathrm{d}w=O\left(h\varepsilon_n\right),$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Then, $\bar{V}_{2,j}(v,x;b)-\bar{V}_{2,j}(v,x;h)=O(h\varepsilon_n)$, uniformly in $(v,b)\in I_x\times\mathbb{H}$. Then it follows from this result, (S81), (S84), (S85) and (S86) that $V_2(v,x;b)-V_2(v,x;h)=O(h\varepsilon_n)$, uniformly in $(v,b)\in I_x\times\mathbb{H}$. The second assertion follows from this result and (S80). ■

Proof of Theorem B1. If follows from Lemma 7(a), $\Pr\left[\widehat{h}\in\mathbb{H}\right]>1-\delta_n$, $\Pr\left[\widehat{h}_\zeta\in\mathbb{H}_\zeta\right]>1-\delta_n^\zeta$ and

$$\begin{aligned}\Pr&\left[\left\|\widehat{V}\left(\cdot\mid x;\widehat{h},\widehat{h}_\zeta\right)-V\left(\cdot\mid x;\widehat{h}\right)\right\|_{I_x}>C_1\kappa_1^V(\gamma)\right]\\&\le\Pr\left[\left\|\widehat{V}\left(\cdot\mid x;\widehat{h},\widehat{h}_\zeta\right)-V\left(\cdot\mid x;\widehat{h}\right)\right\|_{I_x}>C_1\kappa_1^V(\gamma),\left(\widehat{h},\widehat{h}_\zeta\right)\in\mathbb{H}\times\mathbb{H}_\zeta\right]+\delta_n+\delta_n^\zeta\\&\le\Pr\left[\sup_{(v,b,b_\zeta)\in I_x\times\mathbb{H}\times\mathbb{H}_\zeta}\left|\widehat{V}\left(v\mid x;b,b_\zeta\right)-V\left(v\mid x;b\right)\right|>C_1\kappa_1^V(\gamma)\right]+\delta_n+\delta_n^\zeta\end{aligned}\qquad\text{(S89)}$$

that

$$\Pr\left[\left\|\widehat{V}\left(\cdot\mid x;\widehat{h},\widehat{h}_\zeta\right)-V\left(\cdot\mid x;\widehat{h}\right)\right\|_{I_x}>C_1\kappa_1^V(\gamma)\right]\le\kappa_2^V(\gamma)+\delta_n+\delta_n^\zeta.$$

The conclusion of the theorem follows from this result and the fact that with probability $1-C_2\delta_n$,

$$\left|V\left(v\mid x;\widehat{h}\right)-V\left(v\mid x;h\right)\right|\le\sup_{(v,b)\in I_x\times\mathbb{H}}\left|V\left(v\mid x;b\right)-V\left(v\mid x;h\right)\right|=O\left(\varepsilon_n h\right),$$

where the equality follows from Lemma 7(b). ■

Proof of Lemma 8. By definition, we have $\Pr\left[Y_n>C_2\alpha_n\right]\le C_3\beta_n$ and $\Pr\left[\Pr_{|W_1^n}\left[|X_n|>C_1Y_n\right]>C_4\gamma_n\right]\le C_5\delta_n$. Part (a) follows from

$$1-C_3\beta_n-C_5\delta_n\le\Pr\left[Y_n\le C_2\alpha_n,\Pr_{|W_1^n}\left[|X_n|>C_1Y_n\right]\le C_4\gamma_n\right]\le\Pr\left[\Pr_{|W_1^n}\left[|X_n|>C_1C_2\alpha_n\right]\le C_4\gamma_n\right].$$

Part (b) follows from

$$\Pr\left[\Pr_{|W_1^n}\left[|Y_n|>C_2\alpha_n\right]>\varepsilon_n\right]=\Pr\left[\mathbb{1}\left(|Y_n|>C_2\alpha_n\right)>\varepsilon_n\right]=\Pr\left[|Y_n|>C_2\alpha_n\right]\le C_3\beta_n.$$

■

Proof of Lemma 9. Denote $S_{\mathsf{jmb}}(v,x;b):=p_xS_{\mathsf{jmb}}(v\mid x;b)$, $\widehat{S}_{\mathsf{jmb}}(v,x;b,b_\zeta):=\widehat{p}_x\widehat{S}_{\mathsf{jmb}}(v\mid x;b,b_\zeta)$ and $\widehat{\mu}_{\mathcal{U}_x}(v;b):=\sqrt{b}\cdot\widehat{f}_{\Delta X}(v,x;b)$. Then, we write

$$\begin{aligned}&\widehat{S}_{\mathsf{jmb}}\left(v,x;\widehat{h},\widehat{h}_\zeta\right)-S_{\mathsf{jmb}}(v,x;h)=\\&\quad\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\left\{\widehat{\mathcal{U}}_x^{[1]}\left(W_i,v;\widehat{h},\widehat{h}_\zeta\right)-\widetilde{\mathcal{U}}_x^{[1]}\left(W_i,v;\widehat{h}\right)\right\}+\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\left\{\widetilde{\mathcal{U}}_x^{[1]}\left(W_i,v;\widehat{h}\right)-\widetilde{\mathcal{U}}_x^{[1]}\left(W_i,v;h\right)\right\}\\&\quad-\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\right)\left(\widehat{\mu}_{\mathcal{U}_x}\left(v;\widehat{h}\right)-\widetilde{\mu}_{\mathcal{U}_x}(v;h)\right)-\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\left\{\widetilde{\mathcal{U}}_x^{[2]}\left(W_i,v;h\right)-\widetilde{\mu}_{\mathcal{U}_x}(v;h)\right\}.\end{aligned}\qquad\text{(S90)}$$

By Taylor expansion,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\left\{\widehat{\mathcal{U}}_x^{[1]}\left(W_i,v;\widehat{h},\widehat{h}_\zeta\right)-\widetilde{\mathcal{U}}_x^{[1]}\left(W_i,v;\widehat{h}\right)\right\}=$$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\widehat{h}^{-3/2}K'\left(\frac{\Delta_i-v}{\widehat{h}}\right)\left(\widehat{\Delta}_i-\Delta_i\right)\mathbb{1}\left(X_i=x\right)+\frac{1}{2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\widehat{h}^{-5/2}K''\left(\frac{\dot{\Delta}_i-v}{\widehat{h}}\right)\left(\widehat{\Delta}_i-\Delta_i\right)^2\mathbb{1}\left(X_i=x\right)$$
$$+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\left\{\frac{1}{n-1}\sum_{j\neq i}\widehat{h}^{-3/2}K'\left(\frac{\widehat{\Delta}_j-v}{\widehat{h}}\right)\left(\widehat{q}_x\left(W_j,W_i;\widehat{h}_\zeta\right)\widehat{\pi}_x\left(Z_i,X_i\right)-q_x\left(W_j,W_i\right)\pi_x\left(Z_i,X_i\right)\right)\right\}$$
$$+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\left\{\frac{1}{n-1}\sum_{j\neq i}\widehat{h}^{-3/2}\left(K'\left(\frac{\widehat{\Delta}_j-v}{\widehat{h}}\right)-K'\left(\frac{\Delta_j-v}{\widehat{h}}\right)\right)q_x\left(W_j,W_i\right)\pi_x\left(Z_i,X_i\right)\right\}$$
$$=:T_1^\sharp(v)+T_2^\sharp(v)+T_3^\sharp(v)+T_4^\sharp(v),\quad\text{(S91)}$$

where $\dot{\Delta}_i$ denotes the mean value that lies between Δ_i and $\widehat{\Delta}_i$. By the concentration inequality for the maximum of normal random variables (e.g., Giné and Nickl, 2016, (2.3)) and $\mathrm{E}\left[\max_{1\leq i\leq n}|\nu_i|\right]\lesssim\sqrt{\log(n)}$ (Giné and Nickl, 2016, Lemma 2.3.4), we have $\Pr_{|W_1^n}\left[\max_{1\leq i\leq n}|\nu_i|>C_1\sqrt{\log(n)}\right]\leq 2n^{-1}$. By (S15) and $\Pr\left[\widehat{h}\in\mathbb{H}\right]>1-\delta_n$, $\left\|\mathbb{1}_{\Delta X}\left(\cdot,x;\widehat{h}\right)\right\|_{I_x}=O_p^\star\left(1,\delta_n+n^{-1}\right)$. Then, by these results, Lemma 8, (S13), $\Pr\left[\widehat{h}\in\mathbb{H}\right]>1-\delta_n$ and (S12),

$$\left\|T_2^\sharp\right\|_{I_x}\leq\|K''\|_\infty\sqrt{n}\cdot\widehat{h}^{-3/2}\left(\max_{1\leq i\leq n}|\nu_i|\right)\left\|\mathbb{1}_{\Delta X}\left(\cdot,x;\widehat{h}\right)\right\|_{I_x}\overline{\Delta}^2=O_p^\sharp\left(\sqrt{\frac{\log(n)^3}{nh^3}},n^{-1},\delta_n+n^{-1}\right),$$

where the inequality holds on the event given by $\overline{\Delta}\leq\underline{h}$ with probability $1-O\left(n^{-1}\right)$. Then by

$$\Pr\left[\Pr_{|W_1^n}\left[\left\|T_2^\sharp\right\|_{I_x}>C_1\sqrt{\frac{\log(n)^3}{nh^3}}\right]>C_2n^{-1}\right]\leq\Pr\left[\Pr_{|W_1^n}\left[\left\|T_2^\sharp\right\|_{I_x}>C_1\sqrt{\frac{\log(n)^3}{nh^3}}\right]>C_2n^{-1},\,\overline{\Delta}\leq\underline{h}\right]$$
$$+O\left(n^{-1}\right)\leq\Pr\left[\Pr_{|W_1^n}\left[\|K''\|_\infty\sqrt{n}\cdot\widehat{h}^{-3/2}\left(\max_{1\leq i\leq n}|\nu_i|\right)\left\|\mathbb{1}_{\Delta X}\left(\cdot,x;\widehat{h}\right)\right\|_{I_x}\overline{\Delta}^2>C_1\sqrt{\frac{\log(n)^3}{nh^3}}\right]>C_2n^{-1}\right]+O\left(n^{-1}\right),$$

we have $\left\|T_2^\sharp\right\|_{I_x}=O_p^\sharp\left(\sqrt{\log(n)^3/(nh^3)},n^{-1},\delta_n+n^{-1}\right)$.

Let $s_i(v):=\widehat{h}^{-3/2}K'\left((\Delta_i-v)/\widehat{h}\right)\left(\widehat{\Delta}_i-\Delta_i\right)\mathbb{1}\left(X_i=x\right)$ and $S:=\left\{(s_1(v),...,s_n(v))\in\mathbb{R}^n:v\in I_x\right\}$. Then $\left\|T_1^\sharp\right\|_{I_x}=\sup_{(s_1,...,s_n)\in S\cup\{0\}}\left|n^{-1/2}\sum_{i=1}^n\nu_is_i\right|$, where $\left\{n^{-1/2}\sum_{i=1}^n\nu_is_i:(s_1,...,s_n)\in S\cup\{0\}\right\}$ is a centered Gaussian process, conditionally on the data. Let $\|\cdot\|_{n,2}$ be the implicit norm on S induced by the Gaussian process: $\|(s_1,...,s_n)\|_{n,2}:=\sqrt{\mathrm{E}_{|W_1^n}\left[\left(n^{-1/2}\sum_{i=1}^n\nu_is_i\right)^2\right]}=\sqrt{n^{-1}\sum_{i=1}^n s_i^2}$. It is easy to see

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(s_i(v)-s_i(v')\right)^2}\leq\|K''\|_\infty\widehat{h}^{-5/2}\overline{\Delta}\left|v-v'\right|.$$

Therefore,

$$N\left(\varepsilon\cdot\|K''\|_\infty\widehat{h}^{-5/2}\overline{\Delta},S,\|\cdot\|_{n,2}\right)\leq N\left(\varepsilon,I_x,|\cdot|\right)\leq 1+\frac{\iota(I_x)}{\varepsilon},$$

where $\iota(I_x)$ denotes the length of I_x. Let $\sigma_n^2 := \sup_{(s_1,...,s_n)\in S} \|(s_1,...,s_n)\|_{n,2}^2 = \sup_{v\in I_x} \mathrm{E}_{|W_1^n} \left[T_1^\sharp(v)^2\right]$ and then,

$$\sigma_n \leq \overline{\Delta}\sqrt{\sup_{v\in I_x}\frac{1}{n}\sum_{i=1}^n \frac{1}{\widehat{h}^3}K'\left(\frac{\Delta_i - v}{\widehat{h}}\right)^2 \mathbb{1}(X_i = x)} \lesssim \left(\frac{\overline{\Delta}}{\widehat{h}}\right)\sqrt{\left\|\mathbb{1}_{\Delta X}\left(\cdot, x; \widehat{h}\right)\right\|_{I_x}} = O_p^\star\left(\sqrt{\frac{\log(n)}{nh^2}}, \delta_n + n^{-1}\right), \tag{S92}$$

where the second inequality follows from $\left|K'\left(\left(\Delta_i - v\right)/\widehat{h}\right)\right| \leq \|K'\|_\infty \mathbb{1}_i\left(v;\widehat{h}\right)$ and the last equality follows from (S12) and $\Pr\left[\widehat{h} \geq \underline{h}\right] > 1-\delta_n$. By Dudley's metric entropy bound (Giné and Nickl, 2016, Theorem 2.3.7), the above inequality and calculations in the proof of Chernozhukov et al. (2014b, Corollary 5.1),

$$\begin{aligned}
\mathrm{E}_{|W_1^n}\left[\left\|T_1^\sharp\right\|_{I_x}\right] &\lesssim \int_0^{\sigma_n \vee n^{-1/2}\|K''\|_\infty \widehat{h}^{-5/2}\overline{\Delta}} \sqrt{1+\log\left(N\left(\varepsilon, S, \|\cdot\|_{n,2}\right)\right)}\mathrm{d}\varepsilon \\
&\leq \left(\|K''\|_\infty \widehat{h}^{-5/2}\overline{\Delta}\right)\int_0^{\frac{\sigma_n \vee n^{-1/2}\|K''\|_\infty \widehat{h}^{-5/2}\overline{\Delta}}{\|K''\|_\infty \widehat{h}^{-5/2}\overline{\Delta}}} \sqrt{1+\log\left(N\left(\varepsilon, I_x, |\cdot|\right)\right)}\mathrm{d}\varepsilon \\
&\lesssim \left(\sigma_n \vee n^{-1/2}\|K''\|_\infty \widehat{h}^{-5/2}\overline{\Delta}\right)\sqrt{\log\left(\iota(I_x)\, n^{1/2}\right)},
\end{aligned} \tag{S93}$$

when n is sufficiently large. Then, by the Borell-Sudakov-Tsirelson concentration inequality (Giné and Nickl, 2016, Theorem 2.5.8), $\Pr_{|W_1^n}\left[\left\|T_1^\sharp\right\|_{I_x} > \mathrm{E}_{|W_1^n}\left[\left\|T_1^\sharp\right\|_{I_x}\right] + \sigma_n\sqrt{2\cdot\log(n)}\right] \leq n^{-1}$. By Lemma 8, (S92) and (S93),

$$\left\|T_1^\sharp\right\|_{I_x} = O_p^\sharp\left(\frac{\log(n)}{\sqrt{nh^2}}, n^{-1}, \delta_n + n^{-1}\right). \tag{S94}$$

Write

$$T_3^\sharp(v) = \frac{1}{\sqrt{n_{(2)}}}\sum_{j=1}^n \widehat{h}^{-3/2}K'\left(\frac{\widehat{\Delta}_j - v}{\widehat{h}}\right)\frac{1}{\sqrt{n-1}}\sum_{i\neq j}\nu_i\left(\widehat{q}_x\left(W_j, W_i; \widehat{h}_\zeta\right)\widehat{\pi}_x(Z_i, X_i) - q_x(W_j, W_i)\,\pi_x(Z_i, X_i)\right).$$

Then, in view of (S64), with probability $1 - O\left(n^{-1} + \delta_n\right)$,

$$\left\|T_3^\sharp\right\|_{I_x} \lesssim \widehat{h}^{-1/2}\sqrt{\frac{n}{n-1}}\left\|\mathbb{1}_{\Delta X}\left(\cdot, x; \widehat{h}\right)\right\|_{I_x}\max_{1\leq j\leq n}|\Xi_j|, \tag{S95}$$

where

$$\Xi_j := \mathbb{1}_j\left(v;\widehat{h}\right)\frac{1}{\sqrt{n-1}}\sum_{i\neq j}\nu_i\left(\widehat{q}_x\left(W_j, W_i; \widehat{h}_\zeta\right)\widehat{\pi}_x(Z_i, X_i) - q_x(W_j, W_i)\,\pi_x(Z_i, X_i)\right).$$

Conditionally on the data (W_1^n), $(\Xi_1, ..., \Xi_n)$ are centered and jointly normal. Then, since $\Pr\left[\widehat{h} \in \mathbb{H}\right] > 1-\delta_n$, by calculating the conditional variance and the c_r inequality, with probability $1 - O(\delta_n)$,

$$\begin{aligned}
\mathrm{E}_{|W_1^n}\left[{\Xi_j}^2\right] \lesssim\ & \mathbb{1}_j\left(v;\overline{h}\right)\sum_{d\in\{0,1\}}\frac{1}{n-1}\sum_{i\neq j}\left(\widehat{q}_{dx}\left(W_j, W_i; \widehat{h}_\zeta\right) - q_{dx}(W_j, W_i)\right)^2 \pi_x(Z_i, X_i)^2 \\
&+ \mathbb{1}_j\left(v;\overline{h}\right)\sum_{d\in\{0,1\}}\frac{1}{n-1}\sum_{i\neq j}\widehat{q}_{dx}\left(W_j, W_i; \widehat{h}_\zeta\right)^2\left(\widehat{\pi}_x(Z_i, X_i) - \pi_x(Z_i, X_i)\right)^2 =: T_{1,j}^\Xi + T_{2,j}^\Xi.
\end{aligned}$$

By (S66) and $\Pr\left[\widehat{h}_\zeta \in \mathbb{H}_\zeta\right] > 1-\delta_n^\zeta$,

$$\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}\left(y\right);\widehat{h}_\zeta\right) - \zeta_{dx}\left(\phi_{dx}\left(y\right)\right) = O_p^\star\left(\sqrt{\frac{\log\left(n\right)}{nh_\zeta}} + h_\zeta^2, n^{-1} + \delta_n^\zeta\right), \tag{S96}$$

uniformly in $y \in \dot{I}_{d'x}$, where $\dot{I}_{d'x}$ is any closed sub-interval of $I_{d'x}$. (S96) implies that

$$\Pr\left[\inf_{y\in \dot{I}_{d'x}}\left|\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}\left(y\right);\widehat{h}_\zeta\right)\right| > \frac{1}{2}\underline{\zeta}_{dx}\right] > 1 - C_2\left(n^{-1}+\delta_n^\zeta\right). \tag{S97}$$

By (S63) and the c_r inequality,

$$\begin{aligned}
&\mathbb{1}_j\left(v;\overline{h}\right)\frac{1}{n-1}\sum_{i\neq j}\left(\widehat{q}_{dx}\left(W_j, W_i;\widehat{h}_\zeta\right) - q_{dx}\left(W_j,W_i\right)\right)^2\pi_x\left(Z_i,X_i\right)^2 \lesssim \\
&\qquad \mathbb{1}_j\left(v;\overline{h}\right)\frac{\mathbb{1}\left(D_j = d', X_j = x\right)}{\zeta_{dx}\left(\phi_{dx}\left(Y_j\right)\right)^2}\frac{1}{n-1}\sum_{i\neq j}\left(\widehat{\varphi}_{dx}\left(W_j,W_i\right)-\varphi_{dx}\left(W_j,W_i\right)\right)^2\pi_x\left(Z_i,X_i\right)^2 \\
&\quad + \mathbb{1}_j\left(v;\overline{h}\right)\frac{\widehat{\varphi}_{dx}\left(W_j,W_i\right)^2\mathbb{1}\left(D_j=d',X_j=x\right)}{\zeta_{dx}\left(\phi_{dx}\left(Y_j\right)\right)^4}\frac{1}{n-1}\sum_{i\neq j}\left\{\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}\left(Y_j\right);\widehat{h}_\zeta\right)-\zeta_{dx}\left(\phi_{dx}\left(Y_j\right)\right)\right\}^2\pi_x\left(Z_i,X_i\right)^2 \\
&\quad + \mathbb{1}_j\left(v;\overline{h}\right)\frac{\widehat{\varphi}_{dx}\left(W_j,W_i\right)^2\mathbb{1}\left(D_j=d',X_j=x\right)}{\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}\left(Y_j\right);\widehat{h}_\zeta\right)^2\zeta_{dx}\left(\phi_{dx}\left(Y_j\right)\right)^4}\frac{1}{n-1}\sum_{i\neq j}\left\{\widehat{\zeta}_{dx}\left(\widehat{\phi}_{dx}\left(Y_j\right);\widehat{h}_\zeta\right)-\zeta_{dx}\left(\phi_{dx}\left(Y_j\right)\right)\right\}^4\pi_x\left(Z_i,X_i\right)^2.
\end{aligned}$$

By this result, (S97) and (S96), and using the fact

$$\begin{aligned}
&\left(\mathbb{1}\left(Y_i \le \widehat{\phi}_{dx}\left(Y_j\right)\right) - \mathbb{1}\left(Y_i\le\phi_{dx}\left(Y_j\right)\right)\right)^2 = \\
&\left(\mathbb{1}\left(Y_i\le\widehat{\phi}_{dx}\left(Y_j\right), Y_i>\phi_{dx}\left(Y_j\right)\right) - \mathbb{1}\left(Y_i>\widehat{\phi}_{dx}\left(Y_j\right), Y_i\le\phi_{dx}\left(Y_j\right)\right)\right)^2 = \left|\mathbb{1}\left(Y_i\le\widehat{\phi}_{dx}\left(Y_j\right)\right)-\mathbb{1}\left(Y_i\le\phi_{dx}\left(Y_j\right)\right)\right|,
\end{aligned}$$

we have

$$\begin{aligned}
&\max_{j=1,\ldots,n}\mathbb{1}_j\left(v;\overline{h}\right)\frac{1}{n-1}\sum_{i\neq j}\left(\widehat{q}_{dx}\left(W_j,W_i;\widehat{h}_\zeta\right)-q_{dx}\left(W_j,W_i\right)\right)^2\pi_x\left(Z_i,X_i\right)^2 \lesssim \\
&\qquad \left(n/\left(n-1\right)\right)\left(p_{0x}^{-1}+p_{1x}^{-1}\right)\left\{\underline{\zeta}_{dx}^{-2}\left(\overline{\phi}+\overline{R}^2\right)+\underline{\zeta}_{dx}^{-4}\overline{\zeta}^2+2\underline{\zeta}_{dx}^{-6}\overline{\zeta}^4\right\} = O_p^\star\left(\sqrt{\frac{\log\left(n\right)}{n}}, n^{-1}+\delta_n\right),
\end{aligned}$$

where $\overline{\zeta}$ is given by (S68) and the inequality holds with probability $1-O\left(n^{-1}+\delta_n^\zeta\right)$ in view of (S97). It follows from this result that $\max_{1\le j\le n}\left|T_{1,j}^{\Xi}\right| = O_p^\star\left(\sqrt{\log\left(n\right)/n}, n^{-1}+\delta_n+\delta_n^\zeta\right)$. By (S97),

$$\begin{aligned}
&\max_{j=1,\ldots,n}\mathbb{1}_j\left(v;\overline{h}\right)\frac{1}{n-1}\sum_{i\neq j}\widehat{q}_{dx}\left(W_j,W_i;\widehat{h}_\zeta\right)^2\left(\widehat{\pi}_x\left(Z_i,X_i\right)-\pi_x\left(Z_i,X_i\right)\right)^2 \lesssim \\
&\qquad \left(n/\left(n-1\right)\right)\underline{\zeta}_{dx}^{-1}\left\{\left(\widehat{p_{0x}^{-1}}-p_{0x}^{-1}\right)^2+\left(\widehat{p_{1x}^{-1}}-p_{1x}^{-1}\right)^2\right\} = O_p^\star\left(\frac{\log\left(n\right)}{n}, n^{-1}+\delta_n\right),
\end{aligned}$$

where the inequality holds with probability $1-O\left(n^{-1}+\delta_n^\zeta\right)$. Therefore, $\max_{1\le j\le n}\left|T_{2,j}^{\Xi}\right| = O_p^\star\left(\log\left(n\right)/n, n^{-1}+\delta_n\right)$ and it follows that $\sigma_\Xi^2 = O_p^\star\left(\sqrt{\log\left(n\right)/n}, n^{-1}+\delta_n\right)$, where $\sigma_\Xi^2 \coloneqq \max_{1\le j\le n}\mathrm{E}_{|W_1^n}\left[\Xi_j^2\right]$. By Giné and Nickl (2016, Lemma 2.3.4), $\mathrm{E}_{|W_1^n}\left[\max_{1\le j\le n}\left|\Xi_j\right|\right] \lesssim \sigma_\Xi\sqrt{\log\left(n\right)}$. Then by using the concentration inequality for Gaussian maxima (Giné and Nickl, 2016, (2.3)), we have $\Pr_{|W_1^n}\left[\max_{1\le j\le n}\left|\Xi_j\right| > C\sqrt{\log\left(n\right)}\sigma_\Xi\right] \le 2n^{-1}$. It now follows from these results, (S95), $\left\|\mathbb{1}_{\Delta X}\left(\cdot,x;\widehat{h}\right)\right\|_{I_x}=O_p^\star\left(1,n^{-1}+\delta_n\right)$ and Lemma 8 that

$$\left\|T_3^\sharp\right\|_{I_x}=O_p^\sharp\left(\left(\frac{\log(n)^3}{nh^2}\right)^{1/4},n^{-1},n^{-1}+\delta_n\right). \tag{S98}$$

Clearly,

$$\left|T_4^\sharp(v)\right|\leq\left(\frac{1}{\sqrt{n_{(2)}}}\sum_{j=1}^n\widehat{h}^{-3/2}\left|K'\left(\frac{\widehat{\Delta}_j-v}{\widehat{h}}\right)-K'\left(\frac{\Delta_j-v}{\widehat{h}}\right)\right|\right)\left(\max_{1\leq j\leq n}\left|\frac{1}{\sqrt{n-1}}\sum_{i\neq j}\nu_iq_x\left(W_j,W_i\right)\pi_x\left(Z_i,X_i\right)\right|\right), \tag{S99}$$

where $(n-1)^{-1/2}\sum_{i\neq j}\nu_iq_x\left(W_j,W_i\right)\pi_x\left(Z_i,X_i\right)$, $j=1,...,n$, are centered and jointly normal, conditionally on the data. By calculating the conditional variances,

$$\max_{j=1,\ldots,n}\mathrm{E}_{|W_1^n}\left[\left\{\frac{1}{\sqrt{n-1}}\sum_{i\neq j}\nu_iq_x\left(W_j,W_i\right)\pi_x\left(Z_i,X_i\right)\right\}^2\right]=\max_{j=1,\ldots,n}\frac{1}{n-1}\sum_{i\neq j}q_x\left(W_j,W_i\right)^2\pi_x\left(Z_i,X_i\right)^2=O_p^\star(1).$$

By mean value expansion, (S13), $\Pr\left[\widehat{h}\geq\underline{h}\right]=1-O\left(\delta_n\right)$ and $\left\|\mathbb{1}_{\Delta X}\left(\cdot,x;\widehat{h}\right)\right\|_{I_x}=O_p^\star\left(1,n^{-1}+\delta_n\right)$,

$$\begin{aligned}\sup_{v\in I_x}\frac{1}{\sqrt{n_{(2)}}}\sum_{j=1}^n\widehat{h}^{-3/2}\left|K'\left(\frac{\widehat{\Delta}_j-v}{\widehat{h}}\right)-K'\left(\frac{\Delta_j-v}{\widehat{h}}\right)\right|&\lesssim\sqrt{\frac{n}{n-1}}\widehat{h}^{-3/2}\overline{\Delta}\left\|\mathbb{1}_{\Delta X}\left(\cdot,x;\widehat{h}\right)\right\|_{I_x}\\&=O_p^\star\left(\sqrt{\frac{\log(n)}{nh^3}},n^{-1}+\delta_n\right).\end{aligned}$$

Then if follows from the concentration inequality for Gaussian maxima (Giné and Nickl, 2016, (2.3)), Lemma 8 and (S99) that $\left\|T_4^\sharp\right\|_{I_x}=O_p^\sharp\left(\log(n)/\sqrt{nh^3},n^{-1},n^{-1}+\delta_n\right)$. Now we have shown that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\left\{\widehat{\mathcal{U}}_x^{[1]}\left(W_i,v;\widehat{h},\widehat{h}_\zeta\right)-\widetilde{\mathcal{U}}_x^{[1]}\left(W_i,v;\widehat{h}\right)\right\}=O_p^\sharp\left(\sqrt{\frac{\log(n)^3}{nh^3}}+\left(\frac{\log(n)^3}{nh^2}\right)^{1/4},n^{-1},n^{-1}+\delta_n+\delta_n^\zeta\right), \tag{S100}$$

uniformly in $v\in I_x$.

Since $n^{-1/2}\sum_{i=1}^n\nu_i\sim\mathrm{N}(0,1)$ is independent of the data,

$$\Pr_{|W_1^n}\left[\left|n^{-1/2}\sum_{i=1}^n\nu_i\right|>\sqrt{2\cdot\log(n)}\right]\leq 2n^{-1}. \tag{S101}$$

Decompose

$$\begin{aligned}\widehat{\mu}_{\mathcal{U}_x}\left(v;\widehat{h}\right)-\widetilde{\mu}_{\mathcal{U}_x}(v;h)=&\left(\sqrt{\widehat{h}}\cdot\widehat{f}_{\Delta X}\left(v,x;\widehat{h}\right)-\sqrt{\widehat{h}}\cdot\widetilde{f}_{\Delta X}\left(v,x;\widehat{h}\right)\right)+\left(\sqrt{\widehat{h}}\cdot\widetilde{f}_{\Delta X}\left(v,x;\widehat{h}\right)-\sqrt{h}\cdot\widetilde{f}_{\Delta X}(v,x;h)\right)\\&-\sqrt{h}\left(\frac{1}{n_{(2)}}\sum_{(j,k)}\mathcal{H}_x\left(U_j,U_k,v;h\right)\right).\end{aligned}$$

By Lemmas 3 and 5,

$$\left(\sqrt{\widehat{h}}\cdot\widehat{f}_{\Delta X}\left(v,x;\widehat{h}\right)-\sqrt{\widehat{h}}\cdot\widetilde{f}_{\Delta X}\left(v,x;\widehat{h}\right)\right)-\sqrt{h}\left(\frac{1}{n_{(2)}}\sum_{(j,k)}\mathcal{H}_x\left(U_j,U_k,v;h\right)\right)$$
$$=O_p^{\star}\left(\varepsilon_n\sqrt{\frac{\log(n)}{n}}+\frac{\log(n)}{nh^{3/2}}+\frac{\log(n)^{3/4}}{n^{3/4}h^{1/2}},\sqrt{\frac{\log(n)}{nh^3}}+\delta_n\right),$$

uniformly in $v\in I_x$. By Lemma 5 and (S79),

$$\sqrt{\widehat{h}}\cdot\widetilde{f}_{\Delta X}\left(v,x;\widehat{h}\right)-\sqrt{h}\cdot\widetilde{f}_{\Delta X}\left(v,x;h\right)=\sqrt{\widehat{h}}\left(\widetilde{f}_{\Delta X}\left(v,x;\widehat{h}\right)-m_{\Delta X}\left(v,x;\widehat{h}\right)\right)-\sqrt{h}\left(\widetilde{f}_{\Delta X}\left(v,x;h\right)-m_{\Delta X}\left(v,x;h\right)\right)$$
$$+\left(\sqrt{\widehat{h}}\cdot m_{\Delta X}\left(v,x;\widehat{h}\right)-\sqrt{h}\cdot m_{\Delta X}\left(v,x;h\right)\right)=O_p^{\star}\left(\sqrt{\log(n)}\varepsilon_n,n^{-1}+\delta_n\right),$$

uniformly in $v\in I_x$. It follows from these results and Lemma 8 that

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\right)\left(\widehat{\mu}_{\mathcal{U}_x}\left(v;\widehat{h}\right)-\widetilde{\mu}_{\mathcal{U}_x}\left(v;h\right)\right)=O_p^{\sharp}\left(\frac{\log(n)}{nh^{3/2}}+\frac{\log(n)^{3/4}}{n^{3/4}h^{1/2}}+\varepsilon_n\sqrt{\log(n)},n^{-1},\sqrt{\frac{\log(n)}{nh^3}}+\delta_n\right),\quad\text{(S102)}$$

uniformly in $v\in I_x$.

Decompose

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\left\{\widetilde{\mathcal{U}}_x^{[2]}\left(W_i,v;h\right)-\widetilde{\mu}_{\mathcal{U}_x}\left(v;h\right)\right\}=$$
$$-\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\right)\sqrt{h}\left(\frac{1}{n_{(2)}}\sum_{(j,k)}\mathcal{H}_x\left(U_j,U_k,v;h\right)\right)+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\left\{\frac{1}{n-1}\sum_{j\neq i}h^{-3/2}K'\left(\frac{\Delta_i-v}{h}\right)q_x\left(W_i,W_j\right)\pi_x\left(Z_j,X_j\right)\right\}$$
$$-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\left\{\sqrt{h}\cdot\widetilde{f}_{\Delta X}\left(v,x;h\right)-\frac{1}{n-1}\sum_{j\neq i}h^{-1/2}K\left(\frac{\Delta_j-v}{h}\right)\mathbb{1}\left(X_j=x\right)\right\}$$
$$=:T_5^{\sharp}(v)+T_6^{\sharp}(v)+T_7^{\sharp}(v).$$

It follows from (S101), Lemmas 4 and 8 that $\left\|T_5^{\sharp}\right\|_{I_x}=O_p^{\sharp}\left(\log(n)/\sqrt{n},n^{-1},\sqrt{\log(n)/(nh^3)}\right)$. Write

$$T_6^{\sharp}(v)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_ih^{-3/2}K'\left(\frac{\Delta_i-v}{h}\right)\mathbb{1}\left(X_i=x\right)\left\{\frac{1}{n-1}\sum_{j\neq i}\left(\left(1-D_i\right)\mathcal{L}_{1x}\left(W_j,Y_i\right)-D_i\mathcal{L}_{0x}\left(W_j,Y_i\right)\right)\right\}.$$

Let $\mathcal{N}_{dx}\left(U_i,y\right):=\mathcal{L}_{dx}\left(g\left(D_i,X_i,\epsilon_i\right),D_i,Z_i,X_i,y\right)$. Then, by (S8), (S9), Kosorok (2007, Lemmas 9.6), Kosorok (2007, Lemmas 9.6, 9.8 and 9.9(vi,vii)) and Chernozhukov et al. (2014b, Lemma A.6), $\mathfrak{N}:=\left\{\mathcal{N}_{dx}\left(\cdot,y\right):y\in I_{d'x}\right\}$ is uniformly VC-type with respect to a constant envelope. By Talagrand's inequality, $\left\|\mathbb{G}_n^U\right\|_{\mathfrak{N}}=O_p^{\star}\left(\sqrt{\log(n)}\right)$ and therefore, $n^{-1}\sum_{j=1}^{n}\mathcal{L}_{dx}\left(W_j,y\right)=O_p^{\star}\left(\sqrt{\log(n)/n}\right)$, uniformly in $y\in I_{d'x}$. And, therefore, by this result, (S15) and

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}K'\left(\frac{\Delta_i-v}{h}\right)^2\mathbb{1}\left(X_i=x\right)\lesssim\left\|\mathbb{1}_{\Delta X}\left(\cdot,x;h\right)\right\|_{I_x}=O_p^{\star}(1),$$

uniformly in $v\in I_x$, we have

$$\mathrm{E}_{|W_1^n}\left[T_6^\sharp(v)^2\right]=\frac{1}{n}\sum_{i=1}^n h^{-3}K'\left(\frac{\mathit{\Delta}_i-v}{h}\right)^2\mathbb{1}(X_i=x)\left\{\frac{1}{n-1}\sum_{j\neq i}\left((1-D_i)\mathcal{L}_{1x}(W_j,Y_i)-D_i\mathcal{L}_{0x}(W_j,Y_i)\right)\right\}^2$$
$$=O_p^\star\left(\frac{\log(n)}{nh^2}\right),$$

uniformly in $v\in I_x$. Then it follows from arguments used to show (S94) that $\left\|T_6^\sharp\right\|_{I_x}=O_p^\sharp\left(\log(n)/\sqrt{nh^2}\right)$. Write

$$T_7^\sharp(v)=\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\left(\frac{1}{n-1}\widetilde{f}_{\mathit{\Delta}X}(v,x;h)\right)-\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\left(\frac{1}{n-1}h^{-1/2}K\left(\frac{\mathit{\Delta}_i-v}{h}\right)\mathbb{1}(X_i=x)\right).$$

By (S101), the fact $\left\|\widetilde{f}_{\mathit{\Delta}X}(\cdot,x;h)\right\|_{I_x}=O_p^\star(1)$ and Lemma 8, $\left(n^{-1/2}\sum_{i=1}^n\nu_i\right)\widetilde{f}_{\mathit{\Delta}X}(v,x;h)=O_p^\sharp\left(\sqrt{\log(n)}\right)$, uniformly in $v\in I_x$. It follows from arguments used to show (S94) that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i h^{-1/2}K\left(\frac{\mathit{\Delta}_i-v}{h}\right)\mathbb{1}(X_i=x)=O_p^\sharp\left(\sqrt{\log(n)}\right),\tag{S103}$$

uniformly in $v\in I_x$. Therefore, $\left\|T_7^\sharp\right\|_{I_x}=O_p^\sharp\left(\sqrt{\log(n)}/n\right)$ and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\left\{\widetilde{\mathcal{U}}_x^{[2]}(W_i,v;h)-\widetilde{\mu}_{\mathcal{U}_x}(v;h)\right\}=O_p^\sharp\left(\frac{\log(n)}{\sqrt{nh^2}},n^{-1},\sqrt{\frac{\log(n)}{nh^3}}\right),\tag{S104}$$

uniformly in $v\in I_x$.

By (S104), we can write

$$\begin{aligned}S_{\mathsf{jmb}}(v,x;h)&=\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\left\{\widetilde{\mathcal{U}}_x^{[1]}(W_i,v;h)-\widetilde{\mu}_{\mathcal{U}_x}(v;h)\right\}+O_p^\sharp\left(\frac{\log(n)}{\sqrt{nh^2}},n^{-1},\sqrt{\frac{\log(n)}{nh^3}}\right)\\&=T_5^\sharp(v)+T_8^\sharp(v)+T_9^\sharp(v)+O_p^\sharp\left(\frac{\log(n)}{\sqrt{nh^2}},n^{-1},\sqrt{\frac{\log(n)}{nh^3}}\right),\end{aligned}\tag{S105}$$

where

$$\begin{aligned}T_8^\sharp(v)&:=\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\left\{h^{-1/2}K\left(\frac{\mathit{\Delta}_i-v}{h}\right)\mathbb{1}(X_i=x)-\frac{1}{n}\sum_{j=1}^n h^{-1/2}K\left(\frac{\mathit{\Delta}_j-v}{h}\right)\mathbb{1}(X_j=x)\right\}\\T_9^\sharp(v)&:=\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\left\{\frac{1}{n-1}\sum_{j\neq i}h^{-3/2}K'\left(\frac{\mathit{\Delta}_j-v}{h}\right)q_x(W_j,W_i)\pi_x(Z_i,X_i)\right\}.\end{aligned}$$

We have

$$T_8^\sharp(v)=\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i h^{-1/2}K\left(\frac{\mathit{\Delta}_i-v}{h}\right)\mathbb{1}(X_i=x)-\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\right)\left(\sqrt{h}\cdot\widetilde{f}_{\mathit{\Delta}X}(v,x;h)\right)=O_p^\sharp\left(\sqrt{\log(n)}\right),$$

where the second equality follows from $\left(n^{-1/2}\sum_{i=1}^n\nu_i\right)\widetilde{f}_{\mathit{\Delta}X}(v,x;h)=O_p^\sharp\left(\sqrt{\log(n)}\right)$ and (S103). Write

$$T_9^\sharp(v)=\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\mathcal{H}_x^\sharp(U_i,v;h)+\frac{1}{\sqrt{n}}\sum_{i=1}^n\nu_i\left\{\frac{1}{n-1}h^{-3/2}K'\left(\frac{\mathit{\Delta}_i-v}{h}\right)q_x(W_i,W_i)\pi_x(Z_i,X_i)\right\},$$

where $\mathcal{H}_x^\sharp(U_i, v; h) \coloneqq n^{-1} \sum_{j=1}^n \sqrt{h} \cdot \mathcal{H}_x(U_j, U_i, v; h)$. By arguments used to show (S94),

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \nu_i h^{-3/2} K'\left(\frac{\Delta_i - v}{h}\right) q_x(W_i, W_i) \pi_x(Z_i, X_i) = O_p^\sharp\left(\frac{\log(n)}{h}\right).$$

By CK Lemma 5.4, the (data-dependent) function class $\mathfrak{H}^\sharp \coloneqq \left\{\mathcal{H}_x^\sharp(\cdot, v; h) : v \in I_x\right\}$ is uniformly VC-type (conditionally on the data) with respect to a constant envelope $F_{\mathfrak{H}^\sharp} = F_{\mathfrak{H}} = O\left(h^{-3/2}\right)$ so that we have $N\left(\varepsilon \|F_{\mathfrak{H}^\sharp}\|_{\mathbb{P}_n^U,2}, \mathfrak{H}^\sharp, \|\cdot\|_{\mathbb{P}_n^U,2}\right) \leq \left(4\sqrt{A_{\mathfrak{H}}}/\varepsilon\right)^{2V_{\mathfrak{H}}}$, $\forall \varepsilon \in (0,1]$. Denote $G^\nu(f) \coloneqq n^{-1/2} \sum_{i=1}^n \nu_i f(U_i)$. Then, $\left\{G^\nu(f) : f \in \mathfrak{H}^\sharp\right\}$ is a centered Gaussian process, conditional on the data. The intrinsic pseudo metric for $\mathfrak{H}^\sharp$ induced by $\left\{G^\nu(f) : f \in \mathfrak{H}^\sharp\right\}$ is given by $\mathfrak{H}^\sharp \times \mathfrak{H}^\sharp \ni (f, g) \mapsto \mathrm{E}_{|W_1^n}\left[(G^\nu(f) - G^\nu(g))^2\right] = \|f - g\|_{\mathbb{P}_n^U,2}$. Clearly, all sample paths of $\left\{G^\nu(f) : f \in \mathfrak{H}^\sharp\right\}$ are continuous with respect to $(f, g) \mapsto \|f - g\|_{\mathbb{P}_n^U,2}$. Let $\widehat{\sigma}_{\mathfrak{H}^\sharp}^2 \coloneqq \sup_{f \in \mathfrak{H}^\sharp} \mathbb{P}_n^U f^2 = \sup_{v \in I_x} n^{-1} \sum_{i=1}^n \mathcal{H}_x^\sharp(U_i, v; h)^2$. Then we have

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^n \mathcal{H}_x^\sharp(U_i, v; h)^2 &= \frac{1}{n^3}\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n \frac{1}{h^3} K'\left(\frac{\Delta_x(\epsilon_j) - v}{h}\right) \mathcal{C}_x(U_j, U_i) K'\left(\frac{\Delta_x(\epsilon_k) - v}{h}\right) \mathcal{C}_x(U_k, U_i) \\
&=: \frac{1}{n^3}\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n \tilde{\mathcal{J}}_x(U_i, U_j, U_k, v; h) = \frac{1}{n_{(3)}} \sum_{(i,j,k)} \tilde{\mathcal{J}}_x(U_i, U_j, U_k, v; h) + \frac{O(n^{-1})}{3n^2 - 2n}\left\{2\sum_{(i,k)} \tilde{\mathcal{J}}_x(U_i, U_i, U_k, v; h)\right. \\
&\quad \left. + \sum_{(i,k)} \tilde{\mathcal{J}}_x(U_k, U_i, U_i, v; h) + \sum_{i=1}^n \tilde{\mathcal{J}}_x(U_i, U_i, U_i, v; h)\right\} = \frac{1}{n_{(3)}} \sum_{(i,j,k)} \tilde{\mathcal{J}}_x(U_i, U_j, U_k, v; h) + O_p^\star\left(\left(nh^3\right)^{-1}\right),
\end{aligned}$$

uniformly in $v \in I_x$, where the third equality follows from V-statistic decomposition (Serfling, 2009, 5.7.3) and the fourth equality follows from the fact that $n_{(2)}^{-1} \sum_{(i,k)} \tilde{\mathcal{J}}_x(U_i, U_i, U_k, v; h)$, $n_{(2)}^{-1} \sum_{(i,k)} \tilde{\mathcal{J}}_x(U_k, U_i, U_i, v; h)$ and $n^{-1} \sum_{i=1}^n \tilde{\mathcal{J}}_x(U_i, U_i, U_i, v; h)$ are all bounded by a constant that is $O\left(h^{-3}\right)$. By using similar arguments that are used to show (S77),

$$\frac{1}{n_{(3)}} \sum_{(i,j,k)} \tilde{\mathcal{J}}_x(U_i, U_j, U_k, v; h) = V_2(v, x; h) + O_p^\star\left(\sqrt{\frac{\log(n)}{nh^3}}, \sqrt{\frac{\log(n)}{nh^3}}\right),$$

uniformly in $v \in I_x$. Therefore, since $\|V_2(\cdot, x; h)\|_{I_x} = O(1)$, we have $\widehat{\sigma}_{\mathfrak{H}^\sharp}^2 = O_p^\star\left(1, \sqrt{\log(n)/(nh^3)}\right)$. By Dudley's metric entropy bound,

$$\begin{aligned}
\mathrm{E}_{|W_1^n}\left[\|G^\nu\|_{\mathfrak{H}^\sharp}\right] &\lesssim \int_0^{\widehat{\sigma}_{\mathfrak{H}^\sharp} \vee n^{-1/2}\|F_{\mathfrak{H}^\sharp}\|_{\mathbb{P}_n^U,2}} \sqrt{1 + \log\left(N\left(\varepsilon, \mathfrak{H}^\sharp, \|\cdot\|_{\mathbb{P}_n^U,2}\right)\right)} \mathrm{d}\varepsilon \\
&= \|F_{\mathfrak{H}^\sharp}\|_{\mathbb{P}_n^U,2} \int_0^{\frac{\widehat{\sigma}_{\mathfrak{H}^\sharp} \vee n^{-1/2}\|F_{\mathfrak{H}^\sharp}\|_{\mathbb{P}_n^U,2}}{\|F_{\mathfrak{H}^\sharp}\|_{\mathbb{P}_n^U,2}}} \sqrt{1 + \log\left(N\left(\varepsilon, \mathfrak{H}^\sharp, \|\cdot\|_{\mathbb{P}_n^U,2}\right)\right)} \mathrm{d}\varepsilon \\
&\lesssim \left(\widehat{\sigma}_{\mathfrak{H}^\sharp} \vee n^{-1/2}\|F_{\mathfrak{H}^\sharp}\|_{\mathbb{P}_n^U,2}\right) \sqrt{V_{\mathfrak{H}} \log\left(16 A_{\mathfrak{H}}^{1/2} n^{1/2}\right)}.
\end{aligned}$$

By the Borell-Sudakov-Tsirelson concentration inequality, $\Pr_{|W_1^n}\left[\|G^\nu\|_{\mathfrak{H}^\sharp} > \mathrm{E}_{|W_1^n}\left[\|G^\nu\|_{\mathfrak{H}^\sharp}\right] + \widehat{\sigma}_{\mathfrak{H}^\sharp}\sqrt{2 \cdot \log(n)}\right] \leq n^{-1}$. Therefore, by Lemma 8,

$$\sup_{v \in I_x}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n \nu_i \mathcal{H}_x^\sharp(U_i, v; h)\right| = \|G^\nu\|_{\mathfrak{H}^\sharp} = O_p^\sharp\left(\sqrt{\log(n)}, n^{-1}, \sqrt{\frac{\log(n)}{nh^3}}\right).$$

We have shown that $\left\|T_5^{\sharp}\right\|_{I_x}=O_p^{\sharp}\left(\log(n)/\sqrt{n},n^{-1},\sqrt{\log(n)/(nh^3)}\right)$. Then it follows that $\|S_{\mathsf{jmb}}(\cdot,x;h)\|_{I_x}=O_p^{\sharp}\left(\sqrt{\log(n)},n^{-1},\sqrt{\log(n)/(nh^3)}\right)$.

Write

$$S^{\triangle}(v,x;b,h)\coloneqq\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\left\{\widetilde{\mathcal{U}}_x^{[1]}(W_i,v;b)-\widetilde{\mathcal{U}}_x^{[1]}(W_i,v;h)\right\}=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\left\{h^{-1/2}L\left(\frac{\Delta_i-v}{h};b,h\right)\mathbb{1}(X_i=x)\right\}$$
$$+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nu_i\left\{\frac{1}{n-1}\sum_{j\neq i}h^{-3/2}L'\left(\frac{\Delta_j-v}{h};b,h\right)q_x(W_j,W_i)\pi_x(Z_i,X_i)\right\}\eqqcolon T_1^{\triangle}(v;b,h)+T_2^{\triangle}(v;b,h).$$

Then we have $T_1^{\triangle}(v;b,h)=n^{-1/2}\sum_{i=1}^{n}\nu_i\mathcal{E}_x^{\triangle}(U_i,v;b,h)$ and

$$T_2^{\triangle}(v;b,h)=\left(\frac{n}{n-1}\right)\left(n^{-1/2}\sum_{i=1}^{n}\nu_i\mathcal{H}_x^{\triangle\sharp}(U_i,v;b,h)\right)+\frac{1}{\sqrt{n}(n-1)}\sum_{i=1}^{n}\nu_i\mathcal{A}_x^{\triangle}(U_i,v;b,h),$$

where $\mathcal{H}_x^{\triangle\sharp}(U_i,v;b,h)\coloneqq n^{-1}\sum_{j=1}^{n}\mathcal{H}_x^{\triangle}(U_j,U_i,v;b,h)$ and

$$\mathcal{A}_x^{\triangle}(U_i,v;b,h)\coloneqq h^{-3/2}L'\left(\frac{\Delta_x(\epsilon_i)-v}{h};b,h\right)\varpi_x(U_i)\left(1-F_{\epsilon|X}(\epsilon_i\mid x)\right)\pi_x(Z_i,X_i).$$

Let $\widehat{\sigma}^2_{\mathfrak{E}^{\triangle}}\coloneqq\sup_{f\in\mathfrak{E}^{\triangle}}\mathbb{P}_n^Uf^2\leq\sigma^2_{\mathfrak{E}^{\triangle}}+\left\|\mathbb{P}_n^U-\mathbb{P}^U\right\|_{\mathfrak{E}^{\triangle\triangle}}$, where $\mathfrak{E}^{\triangle\triangle}\coloneqq\left\{\mathcal{E}_x^{\triangle}(\cdot,v;b,h)^2:(v,b)\in I_x\times\mathbb{H}\right\}$. It is shown in the proof of Lemma 5 that $\sigma^2_{\mathfrak{E}^{\triangle}}=O\left(\varepsilon_n^2\right)$. By Chernozhukov et al. (2014a, Lemma B.2) and the fact that $\mathfrak{E}^{\triangle}$ is uniformly VC-type with respect to an $O\left(\varepsilon_n/h^{1/2}\right)$ constant envelope, $\mathfrak{E}^{\triangle\triangle}$ is uniformly VC-type with respect to a constant envelope $F_{\mathfrak{E}^{\triangle\triangle}}=O\left(\varepsilon_n^2/h\right)$. Let $\sigma^2_{\mathfrak{E}^{\triangle\triangle}}\coloneqq\sup_{f\in\mathfrak{E}^{\triangle\triangle}}\mathbb{P}^Uf^2=O\left(\varepsilon_n^4/h\right)$, where the second equality follows from Taylor expansion, change of variables and (S38). By Talagrand's inequality (Chernozhukov et al., 2016, Lemma 6.3 with $\mathcal{F}=\mathfrak{E}^{\triangle\triangle}$, $b=F_{\mathfrak{E}^{\triangle\triangle}}$, $\sigma=\sigma_{\mathfrak{E}^{\triangle\triangle}}\vee b\sqrt{V_{\mathfrak{E}^{\triangle\triangle}}\log(n)/n}$, $t=\log(n)$), $\left\|\mathbb{G}_n^U\right\|_{\mathfrak{E}^{\triangle\triangle}}=O_p^{\star}\left(\varepsilon_n^2\sqrt{\log(n)/h}\right)$. Therefore, $\left\|\mathbb{P}_n^U-\mathbb{P}^U\right\|_{\mathfrak{E}^{\triangle\triangle}}=O_p^{\star}\left(\varepsilon_n^2\sqrt{\log(n)/(nh)}\right)$ and $\widehat{\sigma}^2_{\mathfrak{E}^{\triangle}}=O_p^{\star}\left(\varepsilon_n^2\right)$. By Dudley's metric entropy bound, $\mathrm{E}_{|W_1^n}\left[\|G^{\nu}\|_{\mathfrak{E}^{\triangle}}\right]=O_p^{\star}\left(\varepsilon_n\sqrt{\log(n)}\right)$. By the Borell-Sudakov-Tsirelson concentration inequality, , $\Pr_{|W_1^n}\left[\|G^{\nu}\|_{\mathfrak{E}^{\triangle}}>\mathrm{E}_{|W_1^n}\left[\|G^{\nu}\|_{\mathfrak{E}^{\triangle}}\right]+\widehat{\sigma}_{\mathfrak{E}^{\triangle}}\sqrt{2\cdot\log(n)}\right]\leq n^{-1}$. By Lemma 8, $\|T_1^{\triangle}(\cdot;\cdot,h)\|_{I_x\times\mathbb{H}}=\|G^{\nu}\|_{\mathfrak{E}^{\triangle}}=O_p^{\sharp}\left(\varepsilon_n\sqrt{\log(n)}\right)$. Let $\mathfrak{H}^{\triangle\sharp}\coloneqq\left\{\mathcal{H}_x^{\triangle\sharp}(\cdot,v;b,h):(v,b)\in I_x\times\mathbb{H}\right\}$. Then we have $\sup_{(v,b)\in I_x\times\mathbb{H}}\left|n^{-1/2}\sum_{i=1}^{n}\nu_i\mathcal{H}_x^{\triangle\sharp}(U_i,v;b,h)\right|=\|G^{\nu}\|_{\mathfrak{H}^{\triangle\sharp}}$. By CK Lemma 5.4 and the fact that $\mathfrak{H}^{\triangle}$ is uniformly VC-type with respect to an $O\left(\varepsilon_n/h^{3/2}\right)$ constant envelope, $\mathfrak{H}^{\triangle\sharp}$ is uniformly VC-type with respect to a constant envelope $F_{\mathfrak{H}^{\triangle\sharp}}=O\left(\varepsilon_n/h^{3/2}\right)$, conditionally on the data. Let $\widehat{\sigma}^2_{\mathfrak{H}^{\triangle\sharp}}\coloneqq\sup_{f\in\mathfrak{H}^{\triangle\sharp}}\mathbb{P}_n^Uf^2=\sup_{v\in I_x}n^{-1}\sum_{i=1}^{n}\mathcal{H}_x^{\triangle\sharp}(U_i,v;b,h)^2$, where

$$\frac{1}{n}\sum_{i=1}^{n}\mathcal{H}_x^{\triangle\sharp}(U_i,v;b,h)^2=\frac{1}{n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{1}{h^3}L'\left(\frac{\Delta_x(\epsilon_j)-v}{h};b,h\right)\mathcal{C}_x(U_j,U_i)L'\left(\frac{\Delta_x(\epsilon_k)-v}{h};b,h\right)\mathcal{C}_x(U_k,U_i)$$
$$\eqqcolon\frac{1}{n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\tilde{\mathcal{J}}_x^{\triangle}(U_i,U_j,U_k,v;b,h)=\frac{1}{n_{(3)}}\sum_{(i,j,k)}\tilde{\mathcal{J}}_x^{\triangle}(U_i,U_j,U_k,v;b,h)$$
$$+\frac{O(n^{-1})}{3n^2-2n}\left\{2\sum_{(i,k)}\tilde{\mathcal{J}}_x^{\triangle}(U_i,U_i,U_k,v;b,h)+\sum_{(i,k)}\tilde{\mathcal{J}}_x^{\triangle}(U_k,U_i,U_i,v;b,h)+\sum_{i=1}^{n}\tilde{\mathcal{J}}_x^{\triangle}(U_i,U_i,U_i,v;b,h)\right\}$$
$$=\frac{1}{n_{(3)}}\sum_{(i,j,k)}\tilde{\mathcal{J}}_x^{\triangle}(U_i,U_j,U_k,v;b,h)+O\left(\varepsilon_n^2/\left(nh^3\right)\right),$$

where the third equality follows from (S35). By Chernozhukov et al. (2014a, Lemma B.2) and the fact that $\mathfrak{H}^{\triangle}$

is uniformly VC-type with respect to an $O\left(\varepsilon_n/h^{3/2}\right)$ constant envelope, $\tilde{\mathfrak{J}}^{\Delta}:=\left\{\tilde{\mathcal{J}}_x^{\Delta}\left(\cdot,v;b,h\right):(v,b)\in I_x\times\mathbb{H}\right\}$ is uniformly VC-type with respect to a constant envelope $F_{\tilde{\mathfrak{J}}^{\Delta}}=O\left(\varepsilon_n^2/h^3\right)$. Therefore,

$$\widehat{\sigma}^2_{\mathfrak{H}^{\Delta\sharp}}\leq\sup_{f\in\tilde{\mathfrak{J}}^{\Delta}}\mathrm{E}\left[f\left(U_1,U_2,U_3\right)\right]+n^{-1/2}\left\|\mathbb{U}_n^{(3)}\right\|_{\tilde{\mathfrak{J}}^{\Delta}}+O\left(\varepsilon_n^2/\left(nh^3\right)\right).$$

It is shown in the proof of Lemma 5 that

$$\sup_{f\in\tilde{\mathfrak{J}}^{\Delta}}\mathrm{E}\left[f\left(U_1,U_2,U_3\right)\right]=\sup_{(v,b)\in I_x\times\mathbb{H}}\mathrm{E}\left[\tilde{\mathcal{J}}_x^{\Delta}\left(U_1,U_2,U_3,v;b,h\right)\right]=\sup_{(v,b)\in I_x\times\mathbb{H}}\mathrm{E}\left[\mathcal{H}_x^{\Delta[1]}\left(U,v;b,h\right)^2\right]=O\left(\varepsilon_n^2\right).$$

Then we use similar arguments for proving (S77), which involve

$$\mathrm{E}\left[\left(\mathrm{E}\left[\tilde{\mathcal{J}}_x^{\Delta}\left(U_1,U_2,U_3,v;b,h\right)\mid U_1\right]\right)^2\right]+\mathrm{E}\left[\left(\mathrm{E}\left[\tilde{\mathcal{J}}_x^{\Delta}\left(U_1,U_2,U_3,v;b,h\right)\mid U_3\right]\right)^2\right]=O\left(\varepsilon_n^4/h\right)\tag{S106}$$

and

$$\begin{aligned}\mathrm{E}\left[\left(\mathrm{E}\left[\tilde{\mathcal{J}}_x^{\Delta}\left(U_1,U_2,U_3,v;b,h\right)\mid U_2,U_3\right]\right)^2\right]+\mathrm{E}\left[\left(\mathrm{E}\left[\tilde{\mathcal{J}}_x^{\Delta}\left(U_1,U_2,U_3,v;b,h\right)\mid U_2,U_3\right]\right)^2\right]&\\+\mathrm{E}\left[\left(\mathrm{E}\left[\tilde{\mathcal{J}}_x^{\Delta}\left(U_1,U_2,U_3,v;b,h\right)\mid U_2,U_3\right]\right)^2\right]&=O\left(\varepsilon_n^4/h^4\right),\end{aligned}\tag{S107}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. (S106) and (S107) follow from calculation and (S35). By CK Proposition 2.1 with $\mathcal{H}=\tilde{\mathfrak{J}}^{\Delta}_{\pm}$, $\overline{\sigma}_{\mathfrak{g}}=\sigma_{\tilde{\mathfrak{J}}^{\Delta\langle 1\rangle}}$, $\sigma_{\mathfrak{h}}=\sigma_{\tilde{\mathfrak{J}}^{\Delta\langle 2\rangle}}$, $b_{\mathfrak{g}}=b_{\mathfrak{h}}=F_{\tilde{\mathfrak{J}}^{\Delta}}$, $\chi_n=0$, $q=\infty$ and $\gamma=\sqrt{\log\left(n\right)/\left(nh^3\right)}$, we have $n^{-1/2}\left\|\mathbb{U}_n^{(3)}\right\|_{\tilde{\mathfrak{J}}^{\Delta}}=O_p^{\star}\left(\varepsilon_n^2\sqrt{\log\left(n\right)/\left(nh^3\right)},\sqrt{\log\left(n\right)/\left(nh^3\right)}\right)$. Therefore, $\widehat{\sigma}^2_{\mathfrak{H}^{\Delta\sharp}}=O_p^{\star}\left(\varepsilon_n^2,\sqrt{\log\left(n\right)/\left(nh^3\right)}\right)$. By Dudley's metric entropy bound, $\mathrm{E}_{|W_1^n}\left[\left\|G^{\nu}\right\|_{\mathfrak{H}^{\Delta\sharp}}\right]=O_p^{\star}\left(\sqrt{\log\left(n\right)}\varepsilon_n,\sqrt{\log\left(n\right)/\left(nh^3\right)}\right)$. By the Borell-Sudakov-Tsirelson concentration inequality, $\Pr_{|W_1^n}\left[\left\|G^{\nu}\right\|_{\mathfrak{H}^{\Delta\sharp}}>\mathrm{E}_{|W_1^n}\left[\left\|G^{\nu}\right\|_{\mathfrak{H}^{\Delta\sharp}}\right]+\widehat{\sigma}_{\mathfrak{H}^{\Delta\sharp}}\sqrt{2\cdot\log\left(n\right)}\right]\leq n^{-1}$. Then it follows from Lemma 8 that $\left\|G^{\nu}\right\|_{\mathfrak{H}^{\Delta\sharp}}=O_p^{\sharp}\left(\sqrt{\log\left(n\right)}\varepsilon_n,n^{-1},\sqrt{\log\left(n\right)/\left(nh^3\right)}\right)$. Let $\mathfrak{A}^{\Delta}:=\left\{\mathcal{A}_x^{\Delta}\left(\cdot,v;b,h\right):(v,b)\in I_x\times\mathbb{H}\right\}$ and therefore, $\sup_{(v,b)\in I_x\times\mathbb{H}}\left|n^{-1/2}\sum_{i=1}^n\nu_i\mathcal{A}_x^{\Delta}\left(U_i,v;b,h\right)\right|=\left\|G^{\nu}\right\|_{\mathfrak{A}^{\Delta}}$. Let $\widehat{\sigma}^2_{\mathfrak{A}^{\Delta}}:=\sup_{f\in\mathfrak{A}^{\Delta}}\mathbb{P}_n^U f^2\leq\sup_{f\in\mathfrak{A}^{\Delta}}\mathbb{P}^U f^2+\left\|\mathbb{P}_n^U-\mathbb{P}^U\right\|_{\mathfrak{A}^{\Delta\Delta}}$, where $\mathfrak{A}^{\Delta\Delta}:=\left\{\mathcal{A}_x^{\Delta}\left(\cdot,v;b,h\right)^2:(v,b)\in I_x\times\mathbb{H}\right\}$. By using similar arguments for proving the fact that $\mathfrak{H}^{\Delta}$ is uniformly VC-type with respect to an $O\left(\varepsilon_n/h^{3/2}\right)$ constant envelope, $\mathfrak{A}^{\Delta}$ is also uniformly VC-type with respect to an $O\left(\varepsilon_n/h^{3/2}\right)$ constant envelope. By Chernozhukov et al. (2014a, Lemma B.2), $\mathfrak{A}^{\Delta\Delta}$ is uniformly VC-type with respect to a constant envelope $F_{\mathfrak{A}^{\Delta\Delta}}=O\left(\varepsilon_n^2/h^3\right)$. Then, by change of variables and (S35), $\sigma^2_{\mathfrak{A}^{\Delta\Delta}}:=\sup_{f\in\mathfrak{A}^{\Delta\Delta}}\mathbb{P}^U f^2=O\left(\varepsilon_n^4/h^5\right)$. By Talagrand's inequality (Chernozhukov et al., 2016, Lemma 6.3 with $\mathcal{F}=\mathfrak{A}^{\Delta\Delta}$, $b=F_{\mathfrak{A}^{\Delta\Delta}}$, $\sigma=\sigma_{\mathfrak{A}^{\Delta\Delta}}\vee b\sqrt{V_{\mathfrak{A}^{\Delta\Delta}}\log\left(n\right)/n}$, $t=\log\left(n\right)$), $\left\|\mathbb{G}_n\right\|_{\mathfrak{A}^{\Delta\Delta}}=O_p^{\star}\left(\varepsilon_n^2\sqrt{\log\left(n\right)/h^5}\right)$ and therefore, $\left\|\mathbb{P}_n^U-\mathbb{P}^U\right\|_{\mathfrak{A}^{\Delta\Delta}}=O_p^{\star}\left(\varepsilon_n^2\sqrt{\log\left(n\right)/\left(nh^5\right)}\right)$. By change of variables and (S38), $\sup_{f\in\mathfrak{A}^{\Delta}}\mathbb{P}^U f^2=O\left(\varepsilon_n^2/h^2\right)$. Therefore, $\widehat{\sigma}^2_{\mathfrak{A}^{\Delta}}=O_p^{\star}\left(\varepsilon_n^2/h^2\right)$ and by Dudley's metric entropy bound, $\mathrm{E}_{|W_1^n}\left[\left\|G^{\nu}\right\|_{\mathfrak{A}^{\Delta}}\right]=O_p^{\star}\left(\varepsilon_n\sqrt{\log\left(n\right)/h^2}\right)$ by Borell-Sudakov-Tsirelson concentration inequality, $\left\|G^{\nu}\right\|_{\mathfrak{A}^{\Delta}}=O_p^{\sharp}\left(\varepsilon_n\sqrt{\log\left(n\right)/h^2}\right)$. Therefore, $\left\|T_2^{\Delta}\left(\cdot;\cdot,h\right)\right\|_{I_x\times\mathbb{H}}\leq\left(n/\left(n-1\right)\right)\left\|G^{\nu}\right\|_{\mathfrak{H}^{\Delta\sharp}}+\left\|G^{\nu}\right\|_{\mathfrak{A}^{\Delta}}/\left(n-1\right)=O_p^{\sharp}\left(\varepsilon_n\sqrt{\log\left(n\right)},n^{-1},\sqrt{\log\left(n\right)/\left(nh^3\right)}\right)$. Therefore,

$$\left\|S^{\Delta}\left(\cdot,x;\cdot,h\right)\right\|_{I_x\times\mathbb{H}}=O_p^{\sharp}\left(\varepsilon_n\sqrt{\log\left(n\right)},n^{-1},\sqrt{\frac{\log\left(n\right)}{nh^3}}\right).\tag{S108}$$

By $\Pr\left[\widehat{h}\in\mathbb{H}\right]>1-\delta_n$ and monotonicity of conditional expectations,

$$\begin{aligned}
&\Pr\left[\Pr_{|W_1^n}\left[\left\|S^{\Delta}\left(\cdot,x;\widehat{h},h\right)\right\|_{I_x}>C_1\varepsilon_n\sqrt{\log(n)}\right]>C_2n^{-1}\right]\\
&\qquad\leq\Pr\left[\Pr_{|W_1^n}\left[\left\|S^{\Delta}\left(\cdot,x;\widehat{h},h\right)\right\|_{I_x}>C_1\varepsilon_n\sqrt{\log(n)}\right]>C_2n^{-1},\widehat{h}\in\mathbb{H}\right]+\delta_n\\
&\qquad\quad\leq\Pr\left[\Pr_{|W_1^n}\left[\left\|S^{\Delta}\left(\cdot,x;\cdot,h\right)\right\|_{I_x\times\mathbb{H}}>C_1\varepsilon_n\sqrt{\log(n)}\right]>C_2n^{-1}\right]+\delta_n\leq C_3\sqrt{\frac{\log(n)}{nh^3}}+\delta_n.
\end{aligned}$$

Then it follows that $\left\|S^{\Delta}\left(\cdot,x;\widehat{h},h\right)\right\|_{I_x}=O_p^{\sharp}\left(\varepsilon_n\sqrt{\log(n)},n^{-1},\sqrt{\log(n)/(nh^3)}+\delta_n\right)$. It then follows from this result, (S90), (S100), (S102) and (S104) that

$$\widehat{S}_{\mathsf{jmb}}\left(v,x;\widehat{h},\widehat{h}_{\zeta}\right)-S_{\mathsf{jmb}}(v,x;h)=O_p^{\sharp}\left(\sqrt{\frac{\log(n)^3}{nh^3}}+\left(\frac{\log(n)^3}{nh^2}\right)^{1/4}+\varepsilon_n\sqrt{\log(n)},n^{-1},\sqrt{\frac{\log(n)}{nh^3}}+\delta_n+\delta_n^{\zeta}\right).$$

It follows from this result,

$$\widehat{S}_{\mathsf{jmb}}\left(v\mid x;\widehat{h},\widehat{h}_{\zeta}\right)-S_{\mathsf{jmb}}(v\mid x;h)=\frac{\widehat{S}_{\mathsf{jmb}}\left(v,x;\widehat{h},\widehat{h}_{\zeta}\right)}{p_x}\left(\frac{p_x}{\widehat{p}_x}-1\right)+\frac{\widehat{S}_{\mathsf{jmb}}\left(v,x;\widehat{h},\widehat{h}_{\zeta}\right)-S_{\mathsf{jmb}}(v,x;h)}{p_x},$$

$p_x/\widehat{p}_x-1=O_p^{\star}\left(\sqrt{\log(n)/n}\right)$ and $\|S_{\mathsf{jmb}}(\cdot,x;h)\|_{I_x}=O_p^{\sharp}\left(\sqrt{\log(n)},n^{-1},\sqrt{\log(n)/(nh^3)}\right)$ that

$$\widehat{S}_{\mathsf{jmb}}\left(v\mid x;\widehat{h},\widehat{h}_{\zeta}\right)-S_{\mathsf{jmb}}(v\mid x;h)=O_p^{\sharp}\left(\sqrt{\frac{\log(n)^3}{nh^3}}+\left(\frac{\log(n)^3}{nh^2}\right)^{1/4}+\varepsilon_n\sqrt{\log(n)},n^{-1},\sqrt{\frac{\log(n)}{nh^3}}+\delta_n+\delta_n^{\zeta}\right).\tag{S109}$$

Write

$$\widehat{Z}_{\mathsf{jmb}}\left(v\mid x;\widehat{h},\widehat{h}_{\zeta}\right)-Z_{\mathsf{jmb}}(v\mid x;h)=\frac{\widehat{S}_{\mathsf{jmb}}\left(v\mid x;\widehat{h},\widehat{h}_{\zeta}\right)}{\sqrt{V(v\mid x;h)}}\left(\frac{\sqrt{V(v\mid x;h)}}{\sqrt{\widehat{V}\left(v\mid x;\widehat{h},\widehat{h}_{\zeta}\right)}}-1\right)+\frac{\widehat{S}_{\mathsf{jmb}}\left(v\mid x;\widehat{h},\widehat{h}_{\zeta}\right)-S_{\mathsf{jmb}}(v\mid x;h)}{\sqrt{V(v\mid x;h)}}.\tag{S110}$$

The conclusion follows from this result, (55), (S110), (S109), $\|S_{\mathsf{jmb}}(\cdot,x;h)\|_{I_x}=O_p^{\sharp}\left(\sqrt{\log(n)},n^{-1},\sqrt{\log(n)/(nh^3)}\right)$ and Lemma 8. ∎

S3 Bias correction

We maintain all the notations with K replaced by the bias-corrected version $M(\cdot;b,b_{\mathsf{b}})$. We also change the definitions of $\widehat{f}_{\Delta X}(v,x;b)$, $\widetilde{f}_{\Delta X}(v,x;b)$, $m_{\Delta X}(v,x;b)$, $\mathcal{H}_x(U_i,U_j,v;b)$ and $\mathcal{H}_x^{[1]}(U_i,v;b)$ by replacing $K(\cdot)$ with $M(\cdot;b,b_{\mathsf{b}})$ to get $\widehat{f}_{\Delta X}^{\mathsf{bc}}(v,x;b,b_{\mathsf{b}})$, $\widetilde{f}_{\Delta X}^{\mathsf{bc}}(v,x;b,b_{\mathsf{b}})$, $m_{\Delta X}^{\mathsf{bc}}(v,x;b,b_{\mathsf{b}})$, $\mathcal{H}_x^{\mathsf{bc}}(U_i,U_j,v;b,b_{\mathsf{b}})$ and $\mathcal{H}_x^{[1],\mathsf{bc}}(U_i,v;b,b_{\mathsf{b}})$. Similarly, we use the superscript "bc" to denote the bias-corrected versions constructed by replacing $K(\cdot)$ with its bias-correcting counterpart. The result presented in Section 4.3 is implied by the following theorem.

Theorem S1. *Suppose that Assumptions 1-5 hold with $P=2$, the third-order derivative functions in Assumption 2(a) are Lipschitz continuous, $h\propto n^{-\lambda}$ with $1/7<\lambda<1/4$ and $h_{\zeta}\propto n^{-\lambda_{\zeta}}$ with $1/8<\lambda_{\zeta}<1/2$. Assume that there exists some deterministic bandwidth h_{b} and positive sequences $\varepsilon_n^{\mathsf{b}},\delta_n^{\mathsf{b}}$ that decay to zeros such that $\Pr\left[\left|\widehat{h}_{\mathsf{b}}/h_{\mathsf{b}}-1\right|>\varepsilon_n^{\mathsf{b}}\right]\leq\delta_n^{\mathsf{b}}$. Assume that $h/h_{\mathsf{b}}\to\varsigma\in[0,\infty)$. Then,*

$$\Pr\left[f_{\Delta|X}(v\mid x)\in CB^{\mathsf{bc}}_{\mathsf{jmb}}\left(v\mid x;\widehat{h},\widehat{h}_{\varsigma},\widehat{h}_{\mathsf{b}}\right),\forall v\in I_x\right]=(1-\alpha)$$
$$+O\left(\left(\frac{\log(n)^5}{nh^3}\right)^{1/16}+\log(n)\,\kappa^V_{1,n}+\log(n)\left(\varepsilon_n+\varepsilon^{\mathsf{b}}_n\right)+\delta_n+\delta^{\mathsf{b}}_n+\delta^{\varsigma}_n+\sqrt{\log(n)}\sqrt{nh^5}h_{\mathsf{b}}\right). \quad \text{(S111)}$$

We show that Lemmas 3 and 4 still hold if $\widehat{f}_{\Delta X}(v,x;b)$, $\widetilde{f}_{\Delta X}(v,x;b)$ and $\mathcal{H}_x(U_i,U_j,v;b)$ are replaced by their bias-corrected versions. The proofs are based on modifications of the proofs of Lemmas 3 and 4. We present these results as lemmas and sketch how the proofs of 3 and 4 are modified. Denote $\mathbb{H}_{\mathsf{b}}\coloneqq\left[\underline{h}_{\mathsf{b}},\overline{h}_{\mathsf{b}}\right]$, where $\underline{h}_{\mathsf{b}}\coloneqq\left(1-\varepsilon^{\mathsf{b}}_n\right)h_{\mathsf{b}}$ and $\overline{h}_{\mathsf{b}}\coloneqq\left(1+\varepsilon^{\mathsf{b}}_n\right)h_{\mathsf{b}}$. $M'(u;b,b_{\mathsf{b}})$ and $M''(u;b,b_{\mathsf{b}})$ are defined by $M'(u;b,b_{\mathsf{b}})\coloneqq\partial M(u;b,b_{\mathsf{b}})/\partial u$ and $M''(u;b,b_{\mathsf{b}})\coloneqq\partial^2 M(u;b,b_{\mathsf{b}})/\partial u^2$.

Lemma S1. *Under the assumptions of Theorem S1,*

$$\widehat{f}^{\mathsf{bc}}_{\Delta X}(v,x;b,b_{\mathsf{b}})-\widetilde{f}^{\mathsf{bc}}_{\Delta X}(v,x;b,b_{\mathsf{b}})=\frac{1}{n_{(2)}}\sum_{(i,j)}\mathcal{H}^{\mathsf{bc}}_x(U_i,U_j,v;b,b_{\mathsf{b}})+O^{\star}_p\left(\frac{\log(n)}{nh^2}+\frac{\log(n)^{3/4}}{n^{3/4}h}\right),$$

uniformly in $(v,b,b_{\mathsf{b}})\in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}$.

Proof of Lemma S1. Denote $C^{(k)}_K\coloneqq\left\|K^{(k)}\right\|_\infty\vee\left\|K^{(k+2)}_{\mathsf{b}}\right\|_\infty\mu_{K,2}$, $\mathbb{1}^{(k)}_i(v;b,b_{\mathsf{b}})\coloneqq\mathbb{1}_i(v;b)+(b/b_{\mathsf{b}})^{3+k}\,\mathbb{1}_i(v;b_{\mathsf{b}})$ and

$$\mathbb{1}^{(k)}_{\Delta X}(v,x;b,b_{\mathsf{b}})\coloneqq\frac{1}{n}\sum_{i=1}^n\frac{1}{b}\mathbb{1}^{(k)}_i(v;b,b_{\mathsf{b}})\,\mathbb{1}(X_i=x)=\mathbb{1}_{\Delta X}(v,x;b)+\left(\frac{b}{b_{\mathsf{b}}}\right)^{2+k}\mathbb{1}_{\Delta X}(v,x;b_{\mathsf{b}}).$$

By using

$$\left|M''\left(\frac{\dot{\Delta}_i-v}{b};b,b_{\mathsf{b}}\right)\right|\leq C^{(2)}_K\left(\mathbb{1}\left(\left|\dot{\Delta}_i-v\right|\leq b\right)+\left(\frac{b}{b_{\mathsf{b}}}\right)^5\mathbb{1}\left(\left|\dot{\Delta}_i-v\right|\leq b_{\mathsf{b}}\right)\right),$$

we have

$$1-O\left(n^{-1}\right)=\Pr\left[\overline{\Delta}\leq\underline{h}\wedge\underline{h}_{\mathsf{b}}\right]\leq$$
$$\Pr\left[\left|M''\left(\frac{\dot{\Delta}_i-v}{b};b,b_{\mathsf{b}}\right)\right|\mathbb{1}(X_i=x)\leq C^{(2)}_K\mathbb{1}^{(2)}_i(v;b,b_{\mathsf{b}})\,\mathbb{1}(X_i=x),\forall(i,v,b,b_{\mathsf{b}})\in\{1,...,n\}\times I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}\right]. \quad \text{(S112)}$$

Then, by (S15), $\|\mathbb{1}_{\Delta X}(\cdot,x;\cdot)\|_{I_x\times\mathbb{H}_{\mathsf{b}}}=O^{\star}_p(1)$, which follows from similar arguments used to show (S15), and $h/h_{\mathsf{b}}\to\varsigma\in[0,\infty)$,

$$\left\|\mathbb{1}^{(k)}_{\Delta X}(\cdot,x;\cdot,\cdot)\right\|_{I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}}\leq\|\mathbb{1}_{\Delta X}(\cdot,x;\cdot)\|_{I_x\times\mathbb{H}}+\left(\frac{\overline{h}}{\underline{h}_{\mathsf{b}}}\right)^{2+k}\|\mathbb{1}_{\Delta X}(\cdot,x;\cdot)\|_{I_x\times\mathbb{H}_{\mathsf{b}}}=O^{\star}_p(1). \quad \text{(S113)}$$

Then, by this result and (S112), we have

$$\sup_{(v,b,b_{\mathsf{b}})\in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}}\left|\frac{1}{n}\sum_{i=1}^n\frac{1}{b^3}M''\left(\frac{\dot{\Delta}_i-v}{b};b,b_{\mathsf{b}}\right)\left(\widehat{\Delta}_i-\Delta_i\right)^2\mathbb{1}(X_i=x)\right|\lesssim\left\|\mathbb{1}^{(2)}_{\Delta X}(\cdot,x;\cdot,\cdot)\right\|_{I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}}b^{-2}\overline{\Delta}^2$$
$$=O^{\star}_p\left(\frac{\log(n)}{nh^2}\right), \quad \text{(S114)}$$

where the inequality holds with probability $1-O\left(n^{-1}\right)$. By

$$\left|M^{(k)}\left(u;b,b_{\mathsf{b}}\right)\right| \leq C_K^{(k)}\left(\mathbb{1}\left(|u|\leq 1\right)+\left(\frac{b}{b_{\mathsf{b}}}\right)^{3+k}\mathbb{1}\left(|u|\leq\frac{b_{\mathsf{b}}}{b}\right)\right), \tag{S115}$$

(S15) and $\left\|\mathbb{1}_{\Delta X}\left(\cdot,x;\cdot\right)\right\|_{I_x\times\mathbb{H}_{\mathsf{b}}}=O_p^{\star}\left(1\right)$,

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b^2}M'\left(\frac{\Delta_i-v}{b};b,b_{\mathsf{b}}\right)\left(\widehat{\Delta}_i-\Delta_i\right)\mathbb{1}\left(X_i=x\right)=\frac{1}{n_{(2)}}\sum_{(i,j)}\mathcal{H}_x^{\mathsf{bc}}\left(W_i,W_j,v;b,b_{\mathsf{b}}\right)+O_p^{\star}\left(\frac{\log\left(n\right)^{3/4}}{n^{3/4}h}\right). \tag{S116}$$

Then the assertion of the lemma follows from (S114) and (S116). ■

Lemma S2. *Under the assumptions of Theorem S1,*

$$\frac{1}{n_{(2)}}\sum_{(i,j)}\mathcal{H}_x^{\mathsf{bc}}\left(U_i,U_j,v;b,b_{\mathsf{b}}\right)=O_p^{\star}\left(\sqrt{\frac{\log\left(n\right)}{nh}},\sqrt{\frac{\log\left(n\right)}{nh^3}}\right),$$

uniformly in $(v,b,b_{\mathsf{b}})\in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}$.

Proof of Lemma S2. It is easy to check by simple calculations that $\int M'\left(u;b,b_{\mathsf{b}}\right)\mathrm{d}u=0$ and $\int M'\left(u;b,b_{\mathsf{b}}\right)M\left(u;b,b_{\mathsf{b}}\right)\mathrm{d}u=0$. By using these results, (S115), repeating the same arguments with $K(\cdot)$ replaced by its bias-corrected version $M\left(\cdot;b,b_{\mathsf{b}}\right)$ and $h/h_{\mathsf{b}}\to\varsigma\in[0,\infty)$, we have $\mathrm{E}\left[\mathcal{H}_x^{[1],\mathsf{bc}}\left(U,v;b,b_{\mathsf{b}}\right)^2\right]=O\left(h^{-1}\right)$ and $\mathrm{E}\left[\mathcal{H}_x^{\mathsf{bc}}\left(U_1,U_2,v;b,b_{\mathsf{b}}\right)^2\right]=O\left(h^{-3}\right)$, uniformly in $(v,b,b_{\mathsf{b}})\in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}$. Since

$$\begin{aligned}\mathcal{H}_x^{\mathsf{bc}}\left(U_1,U_2,v;b,b_{\mathsf{b}}\right) &:= \frac{1}{b^2}M'\left(\frac{\Delta_x\left(\epsilon_i\right)-v}{b};b,b_{\mathsf{b}}\right)\mathcal{C}_x\left(U_i,U_j\right)\\ &= \frac{1}{b^2}\left\{K'\left(\frac{\Delta_x\left(\epsilon_i\right)-v}{b}\right)-\left(\frac{b}{b_{\mathsf{b}}}\right)^4\mu_{K,2}K_{\mathsf{b}}^{(3)}\left(\frac{\Delta_x\left(\epsilon_i\right)-v}{b_{\mathsf{b}}}\right)\right\}\mathcal{C}_x\left(U_i,U_j\right)\end{aligned}$$

and it follows from the same arguments that $\left\{K_{\mathsf{b}}^{(3)}\left(\left(\Delta_x\left(\cdot\right)-v\right)/b_{\mathsf{b}}\right):\left(v,b_{\mathsf{b}}\right)\in I_x\times\mathbb{H}_{\mathsf{b}}\right\}$ is uniformly VC-type with respect to a constant envelope. By Chernozhukov et al. (2014a, Lemma B.2), $\left\{\mathcal{H}_x^{\mathsf{bc}}\left(\cdot,v;b,b_{\mathsf{b}}\right):\left(v,b,b_{\mathsf{b}}\right)\in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}\right\}$ is uniformly VC-type with respect to an $O\left(h^{-2}\right)$ constant envelope. Then the assertion follows from the same arguments. ■

The following result is an analogue of Lemma 5 under bias correction.

Lemma S3. *Under the assumptions of Theorem S1, (a)*

$$\sqrt{n}\left(\frac{1}{n_{(2)}}\sum_{(i,j)}\mathcal{H}_x^{\Delta,\mathsf{bc}}\left(U_i,U_j,v;b,b_{\mathsf{b}},h,h_{\mathsf{b}}\right)\right)=O_p^{\star}\left(\left(\varepsilon_n+\varepsilon_n^{\mathsf{b}}\right)\sqrt{\log\left(n\right)},\sqrt{\frac{\log\left(n\right)}{nh^3}}\right), \tag{S117}$$

uniformly in $(v,b,b_{\mathsf{b}})\in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}$. *(b)*

$$\sqrt{nb}\left(\widetilde{f}_{\Delta X}^{\mathsf{bc}}\left(v,x;b,b_{\mathsf{b}}\right)-m_{\Delta X}^{\mathsf{bc}}\left(v,x;b,b_{\mathsf{b}}\right)\right)-\sqrt{nh}\left(\widetilde{f}_{\Delta X}^{\mathsf{bc}}\left(v,x;h,h_{\mathsf{b}}\right)-m_{\Delta X}^{\mathsf{bc}}\left(v,x;h,h_{\mathsf{b}}\right)\right)=O_p^{\star}\left(\left(\varepsilon_n+\varepsilon_n^{\mathsf{b}}\right)\sqrt{\log\left(n\right)}\right), \tag{S118}$$

uniformly in $(v,b,b_{\mathsf{b}})\in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}$.

Proof of Lemma S2. Let $N(u;b,b_{\mathsf{b}},h,h_{\mathsf{b}}) \coloneqq (h/b)^{1/2}(b/b_{\mathsf{b}})^3 K''_{\mathsf{b}}((h/b_{\mathsf{b}})u) - (h/h_{\mathsf{b}})^3 K''_{\mathsf{b}}((h/h_{\mathsf{b}})u)$. Then,

$$\left(\frac{h}{b}\right)^{1/2} M\left(\left(\frac{h}{b}\right)u;b,b_{\mathsf{b}}\right) - M(u;h,h_{\mathsf{b}}) = L(u;b,h) - N(u;b,b_{\mathsf{b}},h,h_{\mathsf{b}})\mu_{K,2}. \tag{S119}$$

Let

$$\begin{aligned}
N'(u;b,b_{\mathsf{b}},h,h_{\mathsf{b}}) &\coloneqq \frac{\partial N(u;b,b_{\mathsf{b}},h,h_{\mathsf{b}})}{\partial u} \\
&= \left(\frac{h}{b}\right)^{3/2}\left(\frac{b}{b_{\mathsf{b}}}\right)^4 K^{(3)}_{\mathsf{b}}\left(\left(\frac{h}{b_{\mathsf{b}}}\right)u\right) - \left(\frac{h}{h_{\mathsf{b}}}\right)^4 K^{(3)}_{\mathsf{b}}\left(\left(\frac{h}{h_{\mathsf{b}}}\right)u\right) \\
&= \left\{\left(\frac{h}{b}\right)^{3/2}\left(\frac{b}{b_{\mathsf{b}}}\right)^4 - \left(\frac{h}{h_{\mathsf{b}}}\right)^4\right\} K^{(3)}_{\mathsf{b}}\left(\left(\frac{h}{b_{\mathsf{b}}}\right)u\right) + \left(\frac{h}{h_{\mathsf{b}}}\right)^4\left\{K^{(3)}_{\mathsf{b}}\left(\left(\frac{h}{b_{\mathsf{b}}}\right)u\right) - K^{(3)}_{\mathsf{b}}\left(\left(\frac{h}{h_{\mathsf{b}}}\right)u\right)\right\} \\
&=: N_{\dagger}(u;b,b_{\mathsf{b}},h,h_{\mathsf{b}}) + N_{\ddagger}(u;b,b_{\mathsf{b}},h,h_{\mathsf{b}})
\end{aligned}$$

And let

$$\begin{aligned}
\mathcal{H}^{\dagger,\mathsf{bc}}_x(U_i,U_j,v;b,b_{\mathsf{b}},h,h_{\mathsf{b}}) &\coloneqq h^{-3/2}N_{\dagger}\left(\frac{\Delta_x(\epsilon_i)-v}{h};b,b_{\mathsf{b}},h,h_{\mathsf{b}}\right)\mathcal{C}_x(U_i,U_j) \\
\mathcal{H}^{\ddagger,\mathsf{bc}}_x(U_i,U_j,v;b,b_{\mathsf{b}},h,h_{\mathsf{b}}) &\coloneqq h^{-3/2}N_{\ddagger}\left(\frac{\Delta_x(\epsilon_i)-v}{h};b,b_{\mathsf{b}},h,h_{\mathsf{b}}\right)\mathcal{C}_x(U_i,U_j).
\end{aligned}$$

Then

$$\begin{aligned}
\mathcal{H}^{\triangle,\mathsf{bc}}_x(U_i,U_j,v;b,b_{\mathsf{b}},h,h_{\mathsf{b}}) &\coloneqq \sqrt{b}\cdot\mathcal{H}^{\mathsf{bc}}_x(U_i,U_j,v;b,b_{\mathsf{b}}) - \sqrt{h}\cdot\mathcal{H}^{\mathsf{bc}}_x(U_i,U_j,v;h,h_{\mathsf{b}}) \\
&= h^{-3/2}\left\{L'\left(\frac{\Delta_x(\epsilon_i)-v}{h};b,h\right) - N'\left(\frac{\Delta_x(\epsilon_i)-v}{h};b,b_{\mathsf{b}},h,h_{\mathsf{b}}\right)\mu_{K,2}\right\}\mathcal{C}_x(U_i,U_j) \\
&=: \mathcal{H}^{\triangle}_x(U_i,U_j,v;b,h) - \left\{\mathcal{H}^{\dagger,\mathsf{bc}}_x(U_i,U_j,v;b,b_{\mathsf{b}},h,h_{\mathsf{b}}) + \mathcal{H}^{\ddagger,\mathsf{bc}}_x(U_i,U_j,v;b,b_{\mathsf{b}},h,h_{\mathsf{b}})\right\}\mu_{K,2}.
\end{aligned} \tag{S120}$$

It is easy to check that

$$C_{3,n} \coloneqq \sup_{(b,b_{\mathsf{b}})\in\mathbb{H}\times\mathbb{H}_{\mathsf{b}}}\left|\left(\frac{h}{b}\right)^{1/2}\left(\frac{b}{b_{\mathsf{b}}}\right)^3 - \left(\frac{h}{h_{\mathsf{b}}}\right)^3\right| = \left(\frac{h}{h_{\mathsf{b}}}\right)^3 \sup_{(b,b_{\mathsf{b}})\in\mathbb{H}\times\mathbb{H}_{\mathsf{b}}}\left|\frac{(b/h)^{5/2}}{(b_{\mathsf{b}}/h_{\mathsf{b}})^3} - 1\right| = O\left(\left(\frac{h}{h_{\mathsf{b}}}\right)^3\left(\varepsilon_n + \varepsilon^{\mathsf{b}}_n\right)\right). \tag{S121}$$

And similarly,

$$C_{4,n} \coloneqq \sup_{(b,b_{\mathsf{b}})\in\mathbb{H}\times\mathbb{H}_{\mathsf{b}}}\left|\left(\frac{h}{b}\right)^{3/2}\left(\frac{b}{b_{\mathsf{b}}}\right)^4 - \left(\frac{h}{h_{\mathsf{b}}}\right)^4\right| = O\left(\left(\frac{h}{h_{\mathsf{b}}}\right)^4\left(\varepsilon_n + \varepsilon^{\mathsf{b}}_n\right)\right). \tag{S122}$$

By similar arguments and using $h/h_{\mathsf{b}} \to \varsigma \in [0,\infty)$, $\left\{\mathcal{H}^{\dagger,\mathsf{bc}}_x(\cdot,v;b,b_{\mathsf{b}},h,h_{\mathsf{b}}) : (v,b,b_{\mathsf{b}}) \in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}\right\}$ is uniformly VC-type with respect to an $O\left(\left(\varepsilon_n+\varepsilon^{\mathsf{b}}_n\right)/h^{3/2}\right)$ constant envelope. Similarly, $\left\{\mathcal{H}^{\ddagger,\mathsf{bc}}_x(\cdot,v;b,b_{\mathsf{b}},h,h_{\mathsf{b}}) : (v,b,b_{\mathsf{b}}) \in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}\right\}$ is also uniformly VC-type with respect to an $O\left(\left(\varepsilon_n+\varepsilon^{\mathsf{b}}_n\right)/h^{3/2}\right)$ constant envelope. Therefore, by (S120) and Chernozhukov et al. (2014a, Lemma B.2), $\left\{\mathcal{H}^{\triangle,\mathsf{bc}}_x(\cdot,v;b,b_{\mathsf{b}},h,h_{\mathsf{b}}) : (v,b,b_{\mathsf{b}}) \in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}\right\}$ is uniformly VC-type with respect to an $O\left(\left(\varepsilon_n+\varepsilon^{\mathsf{b}}_n\right)/h^{3/2}\right)$ constant envelope. (S39) holds if $L(\cdot;b,h)$ is replaced by $L(\cdot;b,h) - N(\cdot;b,b_{\mathsf{b}},h,h_{\mathsf{b}})\mu_{K,2}$. By this result and

$$\begin{aligned}
|N(u;b,b_{\mathsf{b}},h,h_{\mathsf{b}})| &\le \left\{C_{3,n}\|K''_{\mathsf{b}}\|_\infty + \left(\frac{h}{h_{\mathsf{b}}}\right)^3\left\|K^{(3)}_{\mathsf{b}}\right\|_\infty\left(\frac{1+\varepsilon^{\mathsf{b}}_n}{1-\varepsilon^{\mathsf{b}}_n}\right)\varepsilon^{\mathsf{b}}_n\right\}\mathbb{1}\left(|u| \le \left(1+\varepsilon^{\mathsf{b}}_n\right)\frac{h_{\mathsf{b}}}{h}\right) \\
|N'(u;b,b_{\mathsf{b}},h,h_{\mathsf{b}})| &\le \left\{C_{4,n}\left\|K^{(3)}_{\mathsf{b}}\right\|_\infty + \left(\frac{h}{h_{\mathsf{b}}}\right)^4\left\|K^{(4)}_{\mathsf{b}}\right\|_\infty\left(\frac{1+\varepsilon^{\mathsf{b}}_n}{1-\varepsilon^{\mathsf{b}}_n}\right)\varepsilon^{\mathsf{b}}_n\right\}\mathbb{1}\left(|u| \le \left(1+\varepsilon^{\mathsf{b}}_n\right)\frac{h_{\mathsf{b}}}{h}\right),
\end{aligned} \tag{S123}$$

we have the first assertion.

It is easy to check that

$$\begin{aligned}
&\sqrt{nb}\left(\widetilde{f}^{\mathsf{bc}}_{\Delta X}(v,x;b,b_{\mathsf{b}})-m^{\mathsf{bc}}_{\Delta X}(v,x;b,b_{\mathsf{b}})\right)-\sqrt{nh}\left(\widetilde{f}^{\mathsf{bc}}_{\Delta X}(v,x;h,h_{\mathsf{b}})-m^{\mathsf{bc}}_{\Delta X}(v,x;h,h_{\mathsf{b}})\right)\\
&\quad=\left\{\sqrt{nb}\left(\widetilde{f}_{\Delta X}(v,x;b)-m_{\Delta X}(v,x;b)\right)-\sqrt{nh}\left(\widetilde{f}_{\Delta X}(v,x;h)-m_{\Delta X}(v,x;h)\right)\right\}\\
&\qquad+\left\{\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\left(N\left(\frac{\Delta_i-v}{h};b,b_{\mathsf{b}},h,h_{\mathsf{b}}\right)-\mathrm{E}\left[N\left(\frac{\Delta-v}{h};b,b_{\mathsf{b}},h,h_{\mathsf{b}}\right)\right]\right)\right\}\mu_{K,2}.
\end{aligned}$$

We have

$$\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\left(N\left(\frac{\Delta_i-v}{h};b,b_{\mathsf{b}},h,h_{\mathsf{b}}\right)\mathbb{1}(X_i=x)-\mathrm{E}\left[N\left(\frac{\Delta-v}{h};b,b_{\mathsf{b}},h,h_{\mathsf{b}}\right)\mathbb{1}(X=x)\right]\right)=O_p^{\star}\left(\left(\varepsilon_n+\varepsilon_n^{\mathsf{b}}\right)\sqrt{\log(n)}\right),$$

uniformly in $(v,b,b_{\mathsf{b}})\in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}$. ■

Let $V^{\mathsf{bc}}(v\mid x;b,b_{\mathsf{b}})$ be defined by the formula of $V(v\mid x;b)$ with $K(\cdot)$ replaced by the bias-correcting kernel $M(\cdot;b,b_{\mathsf{b}})$. Let $V^{\mathsf{bc}}(v,x;b,b_{\mathsf{b}})$, $V_1^{\mathsf{bc}}(v,x;b,b_{\mathsf{b}})$, $V_2^{\mathsf{bc}}(v,x;b,b_{\mathsf{b}})$, $\widehat{V}_1^{\mathsf{bc}}(v,x;b,b_{\mathsf{b}})$ and $\widehat{V}_2^{\mathsf{bc}}(v,x;b,b_{\zeta},b_{\mathsf{b}})$ be defined by the formulae of $V(v,x;b)$, $V_1(v,x;b)$, $V_2(v,x;b)$, $\widehat{V}_1(v,x;b)$ and $\widehat{V}_2(v,x;b,b_{\zeta})$ with $K(\cdot)$ replaced by $M(\cdot;b,b_{\mathsf{b}})$. Similarly, we replace all the notations defined in the proof of Lemma 7 with their bias-corrected versions, which simply replace $K(\cdot)$ with $M(\cdot;b,b_{\mathsf{b}})$. The following result is an analogue of Theorem B1 under bias correction.

Lemma S4. *Suppose the assumptions of Theorem S1 hold. For some constants $C_1,C_2>0$, when n is sufficiently large,*

$$\Pr\left[\left\|\widehat{V}^{\mathsf{bc}}\left(\cdot\mid x;\widehat{h},\widehat{h}_{\zeta},\widehat{h}_{\mathsf{b}}\right)-V^{\mathsf{bc}}(\cdot\mid x;h,h_{\mathsf{b}})\right\|_{I_x}>C_1\left(\kappa_1^V(\gamma)+\varepsilon_n+\varepsilon_n^{\mathsf{b}}\right)\right]\leq C_2\left(\kappa_2^V(\gamma)+\delta_n+\delta_n^{\zeta}+\delta_n^{\mathsf{b}}\right),\ \forall\gamma\in(0,1).$$

Proof of Lemma S4. We apply similar arguments used in the proof of Lemma 7. It can be shown that all intermediate results are still valid. E.g., by using

$$\begin{aligned}
\left\|\widehat{r}^{\mathsf{bc}}_{\Delta X}(\cdot,x;\cdot,\cdot)-r^{\mathsf{bc}}_{\Delta X}(\cdot,x;\cdot,\cdot)\right\|_{I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}}&=O_p^{\star}\left(\sqrt{\frac{\log(n)}{nh}},\sqrt{\frac{\log(n)}{nh^3}}\right)\\
\left\|\widehat{f}^{\mathsf{bc}}_{\Delta X}(\cdot,x;\cdot,\cdot)-m^{\mathsf{bc}}_{\Delta X}(\cdot,x;\cdot,\cdot)\right\|_{I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}}&=O_p^{\star}\left(\sqrt{\frac{\log(n)}{nh}},\sqrt{\frac{\log(n)}{nh^3}}\right),
\end{aligned}$$

we can show that (S40) is still valid:

$$\left\|\widehat{V}_1^{\mathsf{bc}}(\cdot,x;\cdot,\cdot)-V_1^{\mathsf{bc}}(\cdot,x;\cdot,\cdot)\right\|_{I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}}=O_p^{\star}\left(\sqrt{\frac{\log(n)}{nh}},\sqrt{\frac{\log(n)}{nh^3}}\right).$$

Similarly, we use a similar decomposition for $\dot{V}_2^{\mathsf{bc}}(v,x;b,b_{\mathsf{b}})-\ddot{V}_2^{\mathsf{bc}}(v,x;b,b_{\mathsf{b}})$, which is given by the right hand side of the first equality of (S41) with $K(\cdot)$ replaced by $M(\cdot;b,b_{\mathsf{b}})$. By (S12), (S112) and mean value expansion, we get the second equality of (S41). It is easy to see that (S43) also holds for $K_{\mathsf{b}}^{(4)}\left(\left(\dot{\Delta}_j-v\right)/b_{\mathsf{b}}\right)$. Then, by this result and (S43),

$$T_2^{\mathsf{bc}}(v;b,b_{\mathsf{b}})\quad:=\quad\frac{1}{n_{(3)}}\sum_{(i,j,k)}\frac{1}{b^4}\left(M''\left(\frac{\dot{\Delta}_j-v}{b};b,b_{\mathsf{b}}\right)-M''\left(\frac{\Delta_j-v}{b};b,b_{\mathsf{b}}\right)\right)$$

$$
\begin{aligned}
&\times \left(\widehat{\Delta}_j - \Delta_j\right) q_x\left(W_j, W_i\right) M'\left(\frac{\Delta_k - v}{b}; b, b_{\mathsf{b}}\right) q_x\left(W_k, W_i\right) \mathbb{1}\left(X_i = x\right) \\
= \;& \frac{1}{n_{(3)}} \sum_{(i,j,k)} \frac{1}{b^4} \left(K''\left(\frac{\dot{\Delta}_j - v}{b}\right) - K''\left(\frac{\Delta_j - v}{b}\right)\right) \mathbb{1}_j\left(v; b\right) \\
&\times \left(\widehat{\Delta}_j - \Delta_j\right) q_x\left(W_j, W_i\right) M'\left(\frac{\Delta_k - v}{b}; b, b_{\mathsf{b}}\right) q_x\left(W_k, W_i\right) \mathbb{1}\left(X_i = x\right) \\
&- \frac{\mu_{K,2}}{n_{(3)}} \sum_{(i,j,k)} \frac{1}{b^4} \left(\frac{b}{b_{\mathsf{b}}}\right)^5 \left(K_{\mathsf{b}}^{(4)}\left(\frac{\dot{\Delta}_j - v}{b_{\mathsf{b}}}\right) - K_{\mathsf{b}}^{(4)}\left(\frac{\Delta_j - v}{b_{\mathsf{b}}}\right)\right) \mathbb{1}_j\left(v; b_{\mathsf{b}}\right) \\
&\times \left(\widehat{\Delta}_j - \Delta_j\right) q_x\left(W_j, W_i\right) M'\left(\frac{\Delta_k - v}{b}; b, b_{\mathsf{b}}\right) q_x\left(W_k, W_i\right) \mathbb{1}\left(X_i = x\right),
\end{aligned}
$$

uniformly in $(v, b, b_{\mathsf{b}}) \in I_x \times \mathbb{H} \times \mathbb{H}_{\mathsf{b}}$, where the second equality holds with probability $1 - O\left(n^{-1}\right)$. Then by the triangle inequality, (S15), (S12), $\left\|\mathbb{1}_{\Delta X}\left(\cdot, x; \cdot\right)\right\|_{I_x \times \mathbb{H}_{\mathsf{b}}} = O_p^{\star}(1)$, $h/h_{\mathsf{b}} \to \varsigma \in [0, \infty)$, (S115) and (S113), we have $\left\|T_2^{\mathsf{bc}}\right\|_{I_x \times \mathbb{H} \times \mathbb{H}_{\mathsf{b}}} = O_p^{\star}\left(\log(n) / \left(nh^3\right)\right)$. By (S115), (S113) and Lemma 2, we have

$$
T_1^{\mathsf{bc}}\left(v; b, b_{\mathsf{b}}\right) = \frac{1}{n_{(4)}} \sum_{(i,j,k,m)} \mathcal{K}_x^{\mathsf{bc}}\left(U_i, U_j, U_k, U_m, v; b, b_{\mathsf{b}}\right) + O_p^{\star}\left(\left(\frac{\log(n)}{n}\right)^{3/4} h^{-2}\right),
$$

uniformly in $(v, b, b_{\mathsf{b}}) \in I_x \times \mathbb{H} \times \mathbb{H}_{\mathsf{b}}$. By Chernozhukov et al. (2014a, Lemma B.2) and $h/h_{\mathsf{b}} \to \varsigma \in [0, \infty)$, $\mathfrak{K}^{\mathsf{bc}}$ is uniformly VC-type with respect to an $O\left(h^{-4}\right)$ constant envelope. Note that (S53) and (S59) with $K(\cdot)$ replaced by $M(\cdot; b, b_{\mathsf{b}})$ hold. It also holds that $\int M''\left(u; b, b_{\mathsf{b}}\right) \mathrm{d}u = 0$. By calculations with $K(\cdot)$ replaced by $M(\cdot; b, b_{\mathsf{b}})$, (S115) and $h/h_{\mathsf{b}} \to \varsigma \in [0, \infty)$, we have $\sigma^2_{\mathfrak{K}^{\langle 1 \rangle, \mathsf{bc}}} = O\left(h^{-1}\right)$ and $\sigma^2_{\mathfrak{K}^{\langle 2 \rangle, \mathsf{bc}}} = O\left(h^{-5}\right)$. Therefore, by the same arguments, (S60) with $\mathfrak{K}$ replaced by $\mathfrak{K}^{\mathsf{bc}}$ holds. Therefore, (S61) with $\left\|\dot{V}_2\left(\cdot, x; \cdot\right) - \ddot{V}_2\left(\cdot, x; \cdot\right)\right\|_{I_x \times \mathbb{H}}$ replaced by $\left\|\dot{V}_2^{\mathsf{bc}}\left(\cdot, x; \cdot, \cdot\right) - \ddot{V}_2^{\mathsf{bc}}\left(\cdot, x; \cdot, \cdot\right)\right\|_{I_x \times \mathbb{H} \times \mathbb{H}_{\mathsf{b}}}$ holds. By

$$
\begin{aligned}
\Pr\left[\left|M'\left(\frac{\widehat{\Delta}_i - v}{b}; b, b_{\mathsf{b}}\right)\right| \mathbb{1}\left(X_i = x\right) \leq C_K^{(1)} \mathbb{1}_i^{(1)}\left(v; b, b_{\mathsf{b}}\right) \mathbb{1}\left(X_i = x\right), \forall \left(i, v, b, b_{\mathsf{b}}\right) \in \{1, \ldots, n\} \times I_x \times \mathbb{H} \times \mathbb{H}_{\mathsf{b}}\right] \\
= 1 - O\left(n^{-1}\right), \quad \text{(S124)}
\end{aligned}
$$

(S113) and (S65), $\left\|T_3^{\mathsf{bc}}\right\|_{I_x \times \mathbb{H} \times \mathbb{H}_{\mathsf{b}}} = O_p^{\star}\left(\sqrt{\log(n) / nh^2}\right)$. Similarly, we show that $\left\|T_4^{\mathsf{bc}}\right\|_{I_x \times \mathbb{H} \times \mathbb{H}_{\mathsf{b}}} = O_p^{\star}\left(\sqrt{\log(n) / nh^2}\right)$. By (S112), (S124), $\overline{\zeta} = O_p^{\star}\left(\sqrt{\log(n) / \left(nh_{\zeta}\right)} + h_{\zeta}^2\right)$, $\overline{R} = O_p^{\star}\left(\sqrt{\log(n) / n}\right)$, $\overline{\phi} = O_p^{\star}\left(\sqrt{\log(n) / n}\right)$ and (S113), the last equalities of (S69) and (S70) also hold for $\left\|T_{5.3}^{\mathsf{bc}}\right\|_{I_x \times \mathbb{H} \times \mathbb{H}_{\zeta} \times \mathbb{H}_{\mathsf{b}}}$ and $\left\|T_{5.2}^{\mathsf{bc}}\right\|_{I_x \times \mathbb{H} \times \mathbb{H}_{\zeta} \times \mathbb{H}_{\mathsf{b}}}$. By (S124) and

$$
\sup_{e \in \left[\underline{\epsilon}_x, \overline{\epsilon}_x\right]} \left|\frac{1}{n_{(2)}} \sum_{(i,k)} \mathcal{Z}_x^{\mathsf{bc}}\left(U_i, U_k, v, e; b, b_{\mathsf{b}}\right)\right| = O_p^{\star}(1),
$$

uniformly in $(v, b, b_{\mathsf{b}}) \in I_x \times \mathbb{H} \times \mathbb{H}_{\mathsf{b}}$, which follows from similar arguments and calculations, we have $\left\|T_{5.1}^{\mathsf{bc}}\right\|_{I_x \times \mathbb{H} \times \mathbb{H}_{\zeta} \times \mathbb{H}_{\mathsf{b}}} = O_p^{\star}\left(\sqrt{\log(n) / \left(nh_{\zeta}\right)} + h_{\zeta}^2, \sqrt{\log(n) / (nh)}\right)$. By $\overline{\zeta} = O_p^{\star}\left(\sqrt{\log(n) / \left(nh_{\zeta}\right)} + h_{\zeta}^2\right)$, (S67) and (S124), $\left\|T_6^{\mathsf{bc}}\right\|_{I_x \times \mathbb{H} \times \mathbb{H}_{\zeta} \times \mathbb{H}_{\mathsf{b}}} = O_p^{\star}\left(\left(\sqrt{\log(n) / \left(nh_{\zeta}\right)} + h_{\zeta}^2\right)^2, \sqrt{\log(n) / (nh)}\right)$. Then by similar arguments,

$$
\frac{1}{n_{(3)}} \sum_{(i,j,k)} \frac{1}{b^3} M'\left(\frac{\widehat{\Delta}_j - v}{b}; b, b_{\mathsf{b}}\right) \Omega\left(W_j, W_i; b_{\zeta}\right) M'\left(\frac{\widehat{\Delta}_k - v}{b}; b, b_{\mathsf{b}}\right) \Omega\left(W_k, W_i; b_{\zeta}\right) \mathbb{1}\left(X_i = x\right) = O_p^{\star}\left(\sqrt{\frac{\log(n)}{nh^2}}\right).
$$

Then we have

$$\widetilde{V}_2^{\mathsf{bc}}(v,x;b,b_\zeta,b_{\mathsf{b}})-\dot{V}_2^{\mathsf{bc}}(v,x;b,b_{\mathsf{b}})=O_p^\star\left(\sqrt{\frac{\log(n)}{nh^2}}+\sqrt{\frac{\log(n)}{nh_\zeta}}+h_\zeta^2,\sqrt{\frac{\log(n)}{nh}}\right),$$

uniformly in $(v,b,b_\zeta,b_{\mathsf{b}})\in I_x\times\mathbb{H}\times\mathbb{H}_\zeta\times\mathbb{H}_{\mathsf{b}}$. By similar arguments, (S77) with $\left\|\ddot{V}_2(\cdot,x;\cdot)-\bar{V}_2(\cdot,x;\cdot)\right\|_{I_x\times\mathbb{H}}$ replaced by $\left\|\ddot{V}_2^{\mathsf{bc}}(\cdot,x;\cdot,\cdot)-\bar{V}_2^{\mathsf{bc}}(\cdot,x;\cdot,\cdot)\right\|_{I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}}$ holds. And, therefore, $\forall\gamma\in(0,1)$,

$$\Pr\left[\sup_{(v,b,b_\zeta,b_{\mathsf{b}})\in I_x\times\mathbb{H}\times\mathbb{H}_\zeta\times\mathbb{H}_{\mathsf{b}}}\left|\widehat{V}_2(v,x;b,b_\zeta,b_{\mathsf{b}})-V_2(v,x;b,b_{\mathsf{b}})\right|>C_1\kappa_1^V(\gamma)\right]\le C_2\kappa_2^V(\gamma),$$

when n is sufficiently large. By change of variables and using (S119), (S123), (S38), (S115) and

$$\begin{aligned}&\left(\frac{h}{b}\right)M\left(\left(\frac{h}{b}\right)u;b,b_{\mathsf{b}}\right)^2-M(u;h,h_{\mathsf{b}})^2\\&\qquad=\left(\left(\frac{h}{b}\right)^{1/2}M\left(\left(\frac{h}{b}\right)u;b,b_{\mathsf{b}}\right)+M(u;h,h_{\mathsf{b}})\right)(L(u;b,h)-N(u;b,b_{\mathsf{b}},h,h_{\mathsf{b}})\mu_{K,2}),\end{aligned}$$

we have $r_{\Delta X}^{\mathsf{bc}}(v,x;b)-r_{\Delta X}^{\mathsf{bc}}(v,x;h)=O\left(\varepsilon_n+\varepsilon_n^{\mathsf{b}}\right)$. By (S119), (S38), (S123) and change of variables,

$$\begin{aligned}\sqrt{b}\cdot m_{\Delta X}^{\mathsf{bc}}(v,x;b,b_{\mathsf{b}})-\sqrt{h}\cdot m_{\Delta X}^{\mathsf{bc}}(v,x;h,h_{\mathsf{b}})&=O\left(\sqrt{h}\left(\varepsilon_n+\varepsilon_n^{\mathsf{b}}\right)\right)\\ b\cdot m_{\Delta X}^{\mathsf{bc}}(v,x;b,b_{\mathsf{b}})^2-h\cdot m_{\Delta X}^{\mathsf{bc}}(v,x;h,h_{\mathsf{b}})^2&=O\left(h\left(\varepsilon_n+\varepsilon_n^{\mathsf{b}}\right)\right),\end{aligned}\tag{S125}$$

uniformly in $(v,b,b_{\mathsf{b}})\in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}$. By change of variables, integration by parts, (S38) and (S123),

$$\begin{aligned}&\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}b^{-3/2}M'\left(\frac{\Delta_x(e)-v}{b};b,b_{\mathsf{b}}\right)\rho_x(e)F_{\epsilon|X}(e\mid x)\,\mathrm{d}e-\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}h^{-3/2}M'\left(\frac{\Delta_x(e)-v}{h};h,h_{\mathsf{b}}\right)\rho_x(e)F_{\epsilon|X}(e\mid x)\,\mathrm{d}e\\&=h^{-3/2}\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}\left(L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)-N'\left(\frac{\Delta_x(e)-v}{h};b,b_{\mathsf{b}},h,h_{\mathsf{b}}\right)\mu_{K,2}\right)\rho_x(e)F_{\epsilon|X}(e\mid x)\,\mathrm{d}e=O\left(h^{1/2}\left(\varepsilon_n+\varepsilon_n^{\mathsf{b}}\right)\right)\end{aligned}$$

and

$$\begin{aligned}&\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}b^{-3/2}M'\left(\frac{\Delta_x(e)-v}{b};b,b_{\mathsf{b}}\right)\rho_x(e)\,\mathrm{d}e-\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}h^{-3/2}M'\left(\frac{\Delta_x(e)-v}{h};h,h_{\mathsf{b}}\right)\rho_x(e)\,\mathrm{d}e\\&\qquad=h^{-3/2}\int_{\underline{\epsilon}_x}^{\overline{\epsilon}_x}\left(L'\left(\frac{\Delta_x(e)-v}{h};b,h\right)-N'\left(\frac{\Delta_x(e)-v}{h};b,b_{\mathsf{b}},h,h_{\mathsf{b}}\right)\mu_{K,2}\right)\rho_x(e)\,\mathrm{d}e=O\left(h^{1/2}\left(\varepsilon_n+\varepsilon_n^{\mathsf{b}}\right)\right),\end{aligned}$$

uniformly in $(v,b,b_{\mathsf{b}})\in I_x\times\mathbb{H}\times\mathbb{H}_{\mathsf{b}}$. Then, by tedious calculations, we have

$$\begin{aligned}&\left|\left(\frac{h}{b}\right)^2M'\left(\frac{h}{b}w;b,b_{\mathsf{b}}\right)M\left(\frac{h}{b}u;b,b_{\mathsf{b}}\right)-M'(w;h,h_{\mathsf{b}})M(u;h,h_{\mathsf{b}})\right|\\&\le C_{\varepsilon,n}\left(\mathbb{1}(|u|\le1+\varepsilon_n)+\left(\frac{h}{h_{\mathsf{b}}}\right)^3\mathbb{1}\left(|u|\le\left(1+\varepsilon_n^{\mathsf{b}}\right)\frac{h_{\mathsf{b}}}{h}\right)\right)\left(\mathbb{1}(|w|\le1+\varepsilon_n)+\left(\frac{h}{h_{\mathsf{b}}}\right)^3\mathbb{1}\left(|w|\le\left(1+\varepsilon_n^{\mathsf{b}}\right)\frac{h_{\mathsf{b}}}{h}\right)\right),\end{aligned}$$

for some $C_{\varepsilon,n} = O\left(\varepsilon_n + \varepsilon_n^{\mathsf{b}}\right)$. By using this result and

$$\int \left(\frac{h}{b}\right)^2 M'\left(\left(\frac{h}{b}\right) w; b, b_{\mathsf{b}}\right) M\left(\left(\frac{h}{b}\right) w; b, b_{\mathsf{b}}\right) \mathrm{d}w = \int M'\left(w; h, h_{\mathsf{b}}\right) M\left(w; h, h_{\mathsf{b}}\right) \mathrm{d}w = 0,$$

we have

$$\begin{aligned}
&\bar{V}_{2,j}^{\mathsf{bc}}\left(v, x; b, b_{\mathsf{b}}\right) - \bar{V}_{2,j}^{\mathsf{bc}}\left(v, x; h, h_{\mathsf{b}}\right) \\
&\quad = 2h^{-1} \int \left\{\left(\frac{h}{b}\right)^2 M'\left(\left(\frac{h}{b}\right) w; b, b_{\mathsf{b}}\right) M\left(\left(\frac{h}{b}\right) w; b, b_{\mathsf{b}}\right) - M'\left(w; h, h_{\mathsf{b}}\right) M\left(w; h, h_{\mathsf{b}}\right)\right\} \\
&\qquad\qquad \times \psi_{x,j}\left(hw + v\right) \chi_{x,j}\left(hw + v\right) \mathrm{d}w \\
&\quad - 2 \int \int_{-\infty}^{w} \left\{\left(\frac{h}{b}\right)^2 M'\left(\left(\frac{h}{b}\right) w; b, b_{\mathsf{b}}\right) M\left(\left(\frac{h}{b}\right) u; b, b_{\mathsf{b}}\right) - M'\left(w; h, h_{\mathsf{b}}\right) M\left(u; h, h_{\mathsf{b}}\right)\right\} \\
&\qquad\qquad \times \psi_{x,j}\left(hw + v\right) \chi'_{x,j}\left(hu + v\right) \mathrm{d}u \mathrm{d}w = O\left(\varepsilon_n + \varepsilon_n^{\mathsf{b}}\right).
\end{aligned}$$

Then it follows that $V_2^{\mathsf{bc}}\left(v, x; b, b_{\mathsf{b}}\right) - V_2^{\mathsf{bc}}\left(v, x; h, h_{\mathsf{b}}\right) = O\left(\varepsilon_n + \varepsilon_n^{\mathsf{b}}\right)$, uniformly in $\left(v, b, b_{\mathsf{b}}\right) \in I_x \times \mathbb{H} \times \mathbb{H}_{\mathsf{b}}$. Then the assertion follows from arguments in the proof of Theorem B1. ■

The following result is the analogue of Lemma 9 under bias correction.

Lemma S5. *Suppose the assumptions of Theorem S1 hold. Then,*

$$\begin{aligned}
&\widehat{Z}_{\mathsf{jmb}}^{\mathsf{bc}}\left(v \mid x; \widehat{h}, \widehat{h}_{\varsigma}, \widehat{h}_{\mathsf{b}}\right) - Z_{\mathsf{jmb}}^{\mathsf{bc}}\left(v \mid x; h, h_{\mathsf{b}}\right) \\
&\qquad = O_p^{\sharp}\left(\kappa_{1,n}^{V} \sqrt{\log(n)} + \left(\frac{\log(n)^3}{nh^2}\right)^{1/4} + \left(\varepsilon_n + \varepsilon_n^{\mathsf{b}}\right) \sqrt{\log(n)}, n^{-1}, \kappa_{2,n}^{V} + \delta_n + \delta_n^{\varsigma} + \delta_n^{\mathsf{b}}\right).
\end{aligned}$$

Proof of Lemma S5. We use an expansion similar to (S91), where we replace $K(\cdot)$ with $M\left(\cdot; \widehat{h}, \widehat{h}_{\mathsf{b}}\right)$. By $\Pr\left[\left(\widehat{h}, \widehat{h}_{\mathsf{b}}\right) \in \mathbb{H} \times \mathbb{H}_{\mathsf{b}}\right] > 1 - \left(\delta_n + \delta_n^{\mathsf{b}}\right)$ and (S113), $\left\|\mathbb{1}_{\Delta X}^{(k)}\left(\cdot, x; \widehat{h}, \widehat{h}_{\mathsf{b}}\right)\right\|_{I_x} = O_p^{\star}\left(1, n^{-1} + \delta_n + \delta_n^{\mathsf{b}}\right)$. Then by this result and (S112),

$$\left\|T_2^{\sharp,\mathsf{bc}}\right\|_{I_x} \leq C_K^{(2)} \sqrt{n} \cdot \widehat{h}^{-3/2} \left(\max_{1 \leq i \leq n} |\nu_i|\right) \left\|\mathbb{1}_{\Delta X}^{(2)}\left(\cdot, x; \widehat{h}, \widehat{h}_{\mathsf{b}}\right)\right\|_{I_x} \overline{\Delta}^2 = O_p^{\sharp}\left(\sqrt{\frac{\log(n)^3}{nh^3}}, n^{-1}, n^{-1} + \delta_n + \delta_n^{\mathsf{b}}\right).$$

By (S115),

$$\sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(s_i^{\mathsf{bc}}(v) - s_i^{\mathsf{bc}}(v')\right)^2} \leq C_K^{(2)} \left(1 + \left(\frac{\widehat{h}}{\widehat{h}_{\mathsf{b}}}\right)^5\right) \widehat{h}^{-5/2} \overline{\Delta} \left|v - v'\right|.$$

Then by this result, $\Pr\left[\left(\widehat{h}, \widehat{h}_{\mathsf{b}}\right) \in \mathbb{H} \times \mathbb{H}_{\mathsf{b}}\right] > 1 - \left(\delta_n + \delta_n^{\mathsf{b}}\right)$, $h/h_{\mathsf{b}} \to \varsigma \in [0, \infty)$ and repeating the same arguments in the proof of (S94), we have $\left\|T_1^{\sharp,\mathsf{bc}}\right\|_{I_x} = O_p^{\sharp}\left(\log(n) / \sqrt{nh^2}, n^{-1}, n^{-1} + \delta_n + \delta_n^{\mathsf{b}}\right)$. In view of (S124), with probability $1 - O\left(\delta_n + \delta_n^{\mathsf{b}} + n^{-1}\right)$,

$$\left\|T_3^{\sharp,\mathsf{bc}}\right\|_{I_x} \leq \widehat{h}^{-1/2} \sqrt{\frac{n}{n-1}} \left\|\mathbb{1}_{\Delta X}^{(1)}\left(\cdot, x; \widehat{h}, \widehat{h}_{\mathsf{b}}\right)\right\|_{I_x} \max_{1 \leq j \leq n} \left|\Xi_j^{\mathsf{bc}}\right|,$$

where

$$\Xi_j^{\mathsf{bc}} := \mathbb{1}_j\left(v; \widehat{h} \vee \widehat{h}_{\mathsf{b}}\right) \frac{1}{\sqrt{n-1}} \sum_{i \neq j} \nu_i \left(\widehat{q}_x\left(W_j, W_i; \widehat{h}_\zeta\right) \widehat{\pi}_x\left(Z_i, X_i\right) - q_x\left(W_j, W_i\right) \pi_x\left(Z_i, X_i\right)\right).$$

Then by $\Pr\left[\left(\widehat{h}, \widehat{h}_{\mathsf{b}}\right) \in \mathbb{H} \times \mathbb{H}_{\mathsf{b}}\right] > 1 - \left(\delta_n + \delta_n^{\mathsf{b}}\right)$, $\left\|\mathbb{1}_{\Delta X}^{(k)}\left(\cdot, x; \widehat{h}, \widehat{h}_{\mathsf{b}}\right)\right\|_{I_x} = O_p^{\star}\left(1, n^{-1} + \delta_n + \delta_n^{\mathsf{b}}\right)$, and repeating the arguments in the proof of (S98), we have $\left\|T_3^{\sharp,\mathsf{bc}}\right\|_{I_x} = O_p^{\sharp}\left(\left(\log(n)^3 / \left(nh^2\right)\right)^{1/4}, n^{-1}, n^{-1} + \delta_n + \delta_n^{\mathsf{b}}\right)$. It also follows from similar arguments and (S112) that $\left\|T_4^{\sharp,\mathsf{bc}}\right\|_{I_x} = O_p^{\sharp}\left(\log(n) / \sqrt{nh^3}, n^{-1}, n^{-1} + \delta_n + \delta_n^{\mathsf{b}}\right)$. Therefore, (S100) with δ_n replaced by $\delta_n + \delta_n^{\mathsf{b}}$ still holds for the bias-corrected version. It follows from (S101), (S125) and Lemmas S1 and S3 that the bias-corrected version of (S102) holds. The bias-corrected version of (S104) follows from repeating the arguments in the proof of (S104), Lemma S2, (S115) and

$$\frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} M'\left(\frac{\Delta_i - v}{h}; h, h_{\mathsf{b}}\right)^2 \mathbb{1}\left(X_i = x\right) \lesssim \left\|\mathbb{1}_{\Delta X}\left(\cdot, x; h\right)\right\|_{I_x} + \left(\frac{h}{h_{\mathsf{b}}}\right)^7 \left\|\mathbb{1}_{\Delta X}\left(\cdot, x; h_{\mathsf{b}}\right)\right\|_{I_x} = O_p^{\star}(1),$$

uniformly in $v \in I_x$. Then by repeating the arguments in the proof of $\left\|S_{\mathsf{jmb}}\left(\cdot, x; h\right)\right\|_{I_x} = O_p^{\sharp}\left(\sqrt{\log(n)}, n^{-1}, \sqrt{\log(n) / \left(nh^3\right)}\right)$, where $K(\cdot)$ is replaced by $M\left(\cdot; h, h_{\mathsf{b}}\right)$, and using (S115) and $h/h_{\mathsf{b}} \to \varsigma \in [0, \infty)$, we have $\left\|S_{\mathsf{jmb}}^{\mathsf{bc}}\left(\cdot, x; h, h_{\mathsf{b}}\right)\right\|_{I_x} = O_p^{\sharp}\left(\sqrt{\log(n)}, n^{-1}, \sqrt{\log(n) / \left(nh^3\right)}\right)$. Similarly, we can simply modify the proof of (S108) by replacing $K(\cdot)$ with $M\left(\cdot; b, b_{\mathsf{b}}\right)$ and $M\left(\cdot; h, h_{\mathsf{b}}\right)$. Then,

$$\begin{aligned} S^{\Delta,\mathsf{bc}}\left(v, x; b, b_{\mathsf{b}}, h, h_{\mathsf{b}}\right) &:= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \nu_i \left\{\widetilde{\mathcal{U}}_x^{[1],\mathsf{bc}}\left(W_i, v; b, b_{\mathsf{b}}\right) - \widetilde{\mathcal{U}}_x^{[1],\mathsf{bc}}\left(W_i, v; h, h_{\mathsf{b}}\right)\right\} \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \nu_i \left\{h^{-1/2}\left(L\left(\frac{\Delta_i - v}{h}; b, h\right) - N\left(\frac{\Delta_i - v}{h}; b, b_{\mathsf{b}}, h, h_{\mathsf{b}}\right) \mu_{K,2}\right) \mathbb{1}\left(X_i = x\right)\right\} \\ + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \nu_i &\left\{\frac{1}{n-1} \sum_{j \neq i} h^{-3/2}\left(L'\left(\frac{\Delta_j - v}{h}; b, h\right) - N'\left(\frac{\Delta_j - v}{h}; b, b_{\mathsf{b}}, h, h_{\mathsf{b}}\right) \mu_{K,2}\right) q_x\left(W_j, W_i\right) \pi_x\left(Z_i, X_i\right)\right\}. \end{aligned}$$

Then by a modification of the arguments used in the proof of (S108), where (S123) and the fact that the bias corrected versions of the function classes are uniformly VC-type with respect to constant envelopes of order $O\left(\left(\varepsilon_n + \varepsilon_n^{\mathsf{b}}\right) / h^{1/2}\right)$ or $O\left(\left(\varepsilon_n + \varepsilon_n^{\mathsf{b}}\right) / h^{3/2}\right)$ are used, we have

$$\left\|S^{\Delta,\mathsf{bc}}\left(\cdot, x; \cdot, \cdot, h, h_{\mathsf{b}}\right)\right\|_{I_x \times \mathbb{H} \times \mathbb{H}_{\mathsf{b}}} = O_p^{\sharp}\left(\left(\varepsilon_n + \varepsilon_n^{\mathsf{b}}\right) \sqrt{\log(n)}, n^{-1}, \sqrt{\frac{\log(n)}{nh^3}}\right).$$

Then, by $\Pr\left[\left(\widehat{h}, \widehat{h}_{\mathsf{b}}\right) \in \mathbb{H} \times \mathbb{H}_{\mathsf{b}}\right] > 1 - \delta_n - \delta_n^{\mathsf{b}}$,

$$\left\|S^{\Delta,\mathsf{bc}}\left(\cdot, x; \widehat{h}, \widehat{h}_{\mathsf{b}}, h, h_{\mathsf{b}}\right)\right\|_{I_x} = O_p^{\sharp}\left(\left(\varepsilon_n + \varepsilon_n^{\mathsf{b}}\right) \sqrt{\log(n)}, n^{-1}, \sqrt{\frac{\log(n)}{nh^3}} + \delta_n + \delta_n^{\mathsf{b}}\right).$$

Then the assertion follows from repeating the arguments in the proof of Lemma 9 and using Lemma S4. ■

Proof of Theorem S1. It follows from standard arguments for kernel density estimators that $m_{\Delta X}^{\mathsf{bc}}\left(v, x; b, b_{\mathsf{b}}\right) - f_{\Delta X}(v, x) = O\left(h^2 h_b\right)$, uniformly in $\left(v, b, b_{\mathsf{b}}\right) \in I_x \times \mathbb{H} \times \mathbb{H}_{\mathsf{b}}$. The assertion follows from using this result in place of (11), using Lemmas S1, S2, S3, S4 and Lemma S5 in place of Lemmas 3, 4, 5 and 9 and Theorem B1 and repeating the arguments in the proof of Theorem B2. ■

S4 Nonparametric bootstrap

We denote $\mathbb{P}_n^{W^*} f := n^{-1} \sum_{i=1}^n f(W_i^*)$ and $\mathbb{G}_n^{W^*} := \sqrt{n}\left(\mathbb{P}_n^{W^*} - \mathbb{P}_n^{W}\right)$. $\mathbb{P}_n^{U^*}$ and $\mathbb{G}_n^{U^*}$ are defined similarly. Let $\widehat{p}_x^* := n^{-1} \sum_{i=1}^n \mathbb{1}(X_i^* = x)$, $\widehat{f}_{\Delta X}^*(v, x; b) := \widehat{p}_x^* \widehat{f}_{\Delta|X}^*(v \mid x; b)$ and let $\widetilde{f}_{\Delta X}^*(v, x; b)$ be the nonparametric bootstrap analogue of $\widetilde{f}_{\Delta X}(v, x; b)$: $\widetilde{f}_{\Delta X}^*(v, x; b) := (nb)^{-1} \sum_{i=1}^n K((\Delta_i^* - v)/b) \mathbb{1}(X_i^* = x)$. The following result is a nonparametric bootstrap analogue of Lemma 3. We prove it by adapting the proofs of Lemmas 2 and 3 and replacing the intermediate results with their bootstrap analogues.

Lemma S6. *Suppose that the assumptions in the statement of Theorem C1 hold. Then,*

$$\widehat{f}_{\Delta X}^*(v, x; b) - \widetilde{f}_{\Delta X}^*(v, x; b) = \frac{1}{n_{(2)}} \sum_{(i,j)} \mathcal{G}_x\left(W_i^*, W_j^*, v; b\right) + O_p^\sharp\left(\frac{\log(n)}{nh^2} + \frac{\log(n)^{3/4}}{n^{3/4} h}\right),$$

where the remainder is uniform in $(v, b) \in I_x \times \mathbb{H}$.

Proof of Lemma S6. By Talagrand's inequality (Chernozhukov et al., 2016, Lemma 6.3 with $\mathcal{F} = \mathfrak{D}$, $\sigma = b = F_{\mathfrak{D}} = 1$, $t = \log(n)$), we have the deviation bound $\Pr_{|W_1^n}\left[\left\|\mathbb{G}_n^{W^*}\right\|_{\mathfrak{D}} > C\sqrt{\log(n)}\right] \leq n^{-1}$ and therefore, by Lemma 8, $\left\|\widehat{\Pi}_{dzx}^* - \widehat{\Pi}_{dzx}\right\|_{I_{dx}} = O_p^\sharp\left(\sqrt{\log(n)/n}\right)$. Then, by (S1), we have a nonparametric bootstrap analogue of (S1):

$$\left(\frac{\widehat{\Pi}_{d0x}^*\left(\widehat{\phi}_{dx}^*(y)\right)}{\widehat{p}_{0x}^*} - \frac{\widehat{\Pi}_{d1x}^*\left(\widehat{\phi}_{dx}^*(y)\right)}{\widehat{p}_{1x}^*}\right) + \left(\frac{\widehat{\Pi}_{d'0x}^*(y)}{\widehat{p}_{0x}^*} - \frac{\widehat{\Pi}_{d'1x}^*(y)}{\widehat{p}_{1x}^*}\right) = \varepsilon_n^*,$$

where $\varepsilon_n^* = O_p^\sharp(n^{-1})$, and by Lemma 8, we have $\left\|\widehat{\phi}_{dx}^* - \widehat{\phi}_{dx}\right\|_{I_{d'x}} = O_p^\sharp\left(\sqrt{\log(n)/n}\right)$ and also $\left\|\widehat{\phi}_{dx}^* - \phi_{dx}\right\|_{I_{d'x}} = O_p^\sharp\left(\sqrt{\log(n)/n}\right)$. Then we can easily show a bootstrap analogue of (S5). Then, similarly, we decompose $\widehat{\Pi}_{dzx}^*\left(\widehat{\phi}_{dx}^*(y)\right) - \widehat{\Pi}_{dzx}\left(\widehat{\phi}_{dx}(y)\right)$ into the sum of $\widehat{\Pi}_{dzx}^*\left(\widehat{\phi}_{dx}^*(y)\right) - \widehat{\Pi}_{dzx}\left(\widehat{\phi}_{dx}^*(y)\right)$ and $\widehat{\Pi}_{dzx}\left(\widehat{\phi}_{dx}^*(y)\right) - \widehat{\Pi}_{dzx}\left(\widehat{\phi}_{dx}(y)\right)$. Next, we show that

$$\begin{aligned}
\widehat{\Pi}_{dzx}^*\left(\widehat{\phi}_{dx}^*(y)\right) - \widehat{\Pi}_{dzx}\left(\widehat{\phi}_{dx}^*(y)\right) &= \widehat{\Pi}_{dzx}^*(\phi_{dx}(y)) - \widehat{\Pi}_{dzx}(\phi_{dx}(y)) + O_p^\sharp\left(\left(\frac{\log(n)}{n}\right)^{3/4}\right) \\
\widehat{\Pi}_{dzx}\left(\widehat{\phi}_{dx}^*(y)\right) - \widehat{\Pi}_{dzx}\left(\widehat{\phi}_{dx}(y)\right) &= \Pi_{dzx}\left(\widehat{\phi}_{dx}^*(y)\right) - \Pi_{dzx}\left(\widehat{\phi}_{dx}(y)\right) + O_p^\sharp\left(\left(\frac{\log(n)}{n}\right)^{3/4}\right), \qquad \text{(S126)}
\end{aligned}$$

uniformly in $y \in I_{d'x}$. Denote $\widehat{\Lambda}_{dzx}^*(y, y') := \widehat{\Pi}_{dzx}^*(y) - \widehat{\Pi}_{dzx}^*(y')$. Let $\widehat{\sigma}_{\mathfrak{P}^+}^2 := \sup_{f \in \mathfrak{P}^+} \mathbb{P}_n^W f^2$. Then, $\widehat{\sigma}_{\mathfrak{P}^+}^2 \leq \sigma_{\mathfrak{P}^+}^2 + \left\|\mathbb{P}_n^W - \mathbb{P}^W\right\|_{\mathfrak{P}^+} = O_p^\star\left(\sqrt{\log(n)/n}\right)$. By Talagrand's inequality (Chernozhukov et al., 2016, Lemma 6.3 with $\mathcal{F} = \mathfrak{P}^+$, $b = F_{\mathfrak{P}^+} = 1$, $\sigma = \widehat{\sigma}_{\mathfrak{P}^+} \vee b\sqrt{V_{\mathfrak{P}^+} \log(n)/n}$, $t = \log(n)$), $\Pr_{|W_1^n}\left[\left\|\mathbb{G}_n^{W^*}\right\|_{\mathfrak{P}^+} > C\left(\widehat{\sigma}_{\mathfrak{P}^+} \vee \sqrt{\log(n)/n}\right)\sqrt{\log(n)}\right] \leq n^{-1}$. Similarly, we define $\widehat{\sigma}_{\mathfrak{P}^-}^2$ and have a deviation bound $\Pr_{|W_1^n}\left[\left\|\mathbb{G}_n^{W^*}\right\|_{\mathfrak{P}^-} > C\left(\widehat{\sigma}_{\mathfrak{P}^-} \vee \sqrt{\log(n)/n}\right)\sqrt{\log(n)}\right] \leq n^{-1}$. With probability greater than $1 - C_3 n^{-1}$, $\Pr_{|W_1^n}\left[\left\|\widehat{\phi}_{dx}^* - \phi_{dx}\right\|_{I_{d'x}} > C_1\sqrt{\log(n)/n}\right] > C_2 n^{-1}$. Then the first result in (S126) follows from Lemma 8, $\widehat{\sigma}_{\mathfrak{P}^+} = O_p^\star\left((\log(n)/n)^{1/4}\right)$, $\widehat{\sigma}_{\mathfrak{P}^-} = O_p^\star\left((\log(n)/n)^{1/4}\right)$ and

$$\begin{aligned}
&\Pr_{|W_1^n}\left[\sup_{y \in I_{d'x}}\left|\widehat{\Lambda}_{dzx}^*\left(\widehat{\phi}_{dx}^*(y), \phi_{dx}(y)\right) - \widehat{\Lambda}_{dzx}\left(\widehat{\phi}_{dx}^*(y), \phi_{dx}(y)\right)\right| > C\left(\widehat{\sigma}_{\mathfrak{P}^+} \vee \sqrt{\frac{\log(n)}{n}} + \widehat{\sigma}_{\mathfrak{P}^-} \vee \sqrt{\frac{\log(n)}{n}}\right)\sqrt{\frac{\log(n)}{n}}\right] \\
&\leq \Pr_{|W_1^n}\left[\left\|\mathbb{G}_n^{W^*}\right\|_{\mathfrak{P}^+} > C\left(\widehat{\sigma}_{\mathfrak{P}^+} \vee \sqrt{\frac{\log(n)}{n}}\right)\sqrt{\log(n)}\right] + \Pr_{|W_1^n}\left[\left\|\mathbb{G}_n^{W^*}\right\|_{\mathfrak{P}^-} > C\left(\widehat{\sigma}_{\mathfrak{P}^-} \vee \sqrt{\frac{\log(n)}{n}}\right)\sqrt{\log(n)}\right]
\end{aligned}$$

$$+\operatorname{Pr}_{|W_1^n}\left[\left\|\widehat{\phi}_{dx}^* - \phi_{dx}\right\|_{I_{d'x}} > C_1\sqrt{\frac{\log(n)}{n}}\right] = O_p^\star\left(n^{-1}\right).$$

Similarly, the second result in (S126) follows from Lemma 8, $\sigma_{\mathfrak{P}^+} = O\left((\log(n)/n)^{1/4}\right)$, $\sigma_{\mathfrak{P}^-} = O\left((\log(n)/n)^{1/4}\right)$ and

$$\operatorname{Pr}_{|W_1^n}\left[\sup_{y\in I_{d'x}}\left|\widehat{\Lambda}_{dzx}\left(\widehat{\phi}_{dx}^*(y), \widehat{\phi}_{dx}(y)\right) - \Lambda_{dzx}\left(\widehat{\phi}_{dx}^*(y), \widehat{\phi}_{dx}(y)\right)\right| > C\left(\sigma_{\mathfrak{P}^+}\vee\sqrt{\frac{\log(n)}{n}} + \sigma_{\mathfrak{P}^-}\vee\sqrt{\frac{\log(n)}{n}}\right)\sqrt{\frac{\log(n)}{n}}\right]$$
$$\leq \operatorname{Pr}_{|W_1^n}\left[\left\|\mathbb{G}_n^W\right\|_{\mathfrak{P}^+} > C\left(\sigma_{\mathfrak{P}^+}\vee\sqrt{\frac{\log(n)}{n}}\right)\sqrt{\frac{\log(n)}{n}}\right] + \operatorname{Pr}_{|W_1^n}\left[\left\|\mathbb{G}_n^W\right\|_{\mathfrak{P}^-} > C\left(\sigma_{\mathfrak{P}^-}\vee\sqrt{\frac{\log(n)}{n}}\right)\sqrt{\frac{\log(n)}{n}}\right]$$
$$+\operatorname{Pr}_{|W_1^n}\left[\left\|\widehat{\phi}_{dx}^* - \widehat{\phi}_{dx}\right\|_{I_{d'x}} > C_1\sqrt{\frac{\log(n)}{n}}\right] = O_p^\star\left(n^{-1}\right).$$

Then, by using (S126), a bootstrap analogue of (S5) and tedious algebra, we have

$$\widehat{\phi}_{dx}^*(y) - \widehat{\phi}_{dx}(y) = \frac{1}{n}\sum_{i=1}^n \mathcal{L}_{dx}\left(W_i^*, y\right) - \frac{1}{n}\sum_{i=1}^n \mathcal{L}_{dx}\left(W_i, y\right) + O_p^\sharp\left(\left(\frac{\log(n)}{n}\right)^{3/4}\right),$$

uniformly in $y \in I_{d'x}$, and by Lemmas 2 and 8, a linear representation in the bootstrap world holds:

$$\widehat{\phi}_{dx}^*(y) - \phi_{dx}(y) = \frac{1}{n}\sum_{i=1}^n \mathcal{L}_{dx}\left(W_i^*, y\right) + O_p^\sharp\left(\left(\frac{\log(n)}{n}\right)^{3/4}\right), \tag{S127}$$

uniformly in $y \in I_{d'x}$. By Taylor expansion, we get the bootstrap analogue of (S11), where $\dot{\Delta}_i^*$ denotes the mean value. The bootstrap analogue of (S12) (i.e., $\overline{\Delta}^* = O_p^\sharp\left(\sqrt{\log(n)/n}\right)$, where $\overline{\Delta}^* \coloneqq \max_{1\leq i\leq n}\left|\widehat{\Delta}_i^* - \Delta_i^*\right|\mathbb{1}\left(X_i^* = x\right)$) follows from $\left\|\widehat{\phi}_{dx}^* - \phi_{dx}\right\|_{I_{d'x}} = O_p^\sharp\left(\sqrt{\log(n)/n}\right)$. Then since $\sqrt{\log(n)/n} = o(h)$, for some constants $C_2, C_3 > 0$, with probability $1 - C_3 n^{-1}$, the bootstrap analogue of (S13) holds:

$$1 - C_2 n^{-1} \leq \operatorname{Pr}_{|W_1^n}\left[\overline{\Delta}^* \leq \underline{h}\right]$$
$$\leq \operatorname{Pr}_{|W_1^n}\left[\left|K''\left(\frac{\dot{\Delta}_i^* - v}{b}\right)\right|\mathbb{1}\left(X_i^* = x\right) \leq \left\|K''\right\|_\infty \mathbb{1}_i^*(v;b)\,\mathbb{1}\left(X_i^* = x\right), \forall (i,v,b) \in \{1,...,n\}\times I_x \times \mathbb{H}\right], \tag{S128}$$

where $\mathbb{1}_i^*(v;b) \coloneqq \mathbb{1}\left(\left|\Delta_i^* - v\right| \leq 2b\right)$. Then the bootstrap analogue of (S14) holds with probability $1 - C_3 n^{-1}$. Let $\widehat{\sigma}_{\mathfrak{J}}^2 \coloneqq \sup_{f\in\mathfrak{J}}\mathbb{P}_n^U f^2$. Then we have $\widehat{\sigma}_{\mathfrak{J}}^2 \leq h^{-1}\left(\left\|\mathbb{P}^U\right\|_{\mathfrak{J}} + \left\|\mathbb{P}_n^U - \mathbb{P}^U\right\|_{\mathfrak{J}}\right) = O_p^\star\left(h^{-1}\right)$. By Talagrand's inequality (Chernozhukov et al., 2016, Lemma 6.3 with $\mathcal{F} = \mathfrak{J}$, $b = F_{\mathfrak{J}} = h^{-1}$, $\sigma = \widehat{\sigma}_{\mathfrak{J}} \vee b\sqrt{V_{\mathfrak{J}}\log(n)/n}$, $t = \log(n)$), we have

$$\operatorname{Pr}_{|W_1^n}\left[\left\|\mathbb{G}_n^{U^*}\right\|_{\mathfrak{J}} > C\left(\widehat{\sigma}_{\mathfrak{J}}\vee\sqrt{\frac{\log(n)}{nh^2}}\right)\sqrt{\log(n)}\right] \leq n^{-1}.$$

By Lemma 8, $\left\|\mathbb{G}_n^{U^*}\right\|_{\mathfrak{J}} = O_p^\sharp\left(\sqrt{\log(n)/h}\right)$. Then, by Lemma 8,

$$\left\|\mathbb{P}_n^{U^*}\right\|_{\mathfrak{J}} \leq \left\|\mathbb{P}_n^{U^*} - \mathbb{P}_n^U\right\|_{\mathfrak{J}} + \left\|\mathbb{P}_n^U - \mathbb{P}^U\right\|_{\mathfrak{J}} + \left\|\mathbb{P}^U\right\|_{\mathfrak{J}} = O_p^\sharp(1).$$

Then, by these results, the bootstrap analogue of (S16) holds and then we have

$$\widehat{f}^*_{\Delta X}(v,x;b)-\widetilde{f}^*_{\Delta X}(v,x;b)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b^2}K'\left(\frac{\Delta_i^*-v}{b}\right)\left(\widehat{\Delta}_i^*-\Delta_i^*\right)\mathbb{1}\left(X_i^*=x\right)+O_p^\sharp\left(\frac{\log(n)}{nh^2}\right),$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Then the assertion follows from this result, (S127) and $\left\|\mathbb{P}_n^{U^*}\right\|_{\mathfrak{J}}=O_p^\sharp(1)$. ■

Lemma S7. *Suppose that the assumptions in the statement of Theorem C1 hold. Then,*

$$\widehat{f}^*_{\Delta X}(v,x;b)-\widetilde{f}^*_{\Delta X}(v,x;b)=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathcal{H}_x\left(U_i,U_j,v;b\right)+\left\{\frac{1}{n}\sum_{i=1}^{n}\mathcal{H}_x^{[1]}\left(U_i^*,v;b\right)-\frac{1}{n}\sum_{i=1}^{n}\mathcal{H}_x^{[1]}\left(U_i,v;b\right)\right\}$$
$$+O_p^\sharp\left(\left(\frac{\log(n)}{n^3h^5}\right)^{1/4},\left(\frac{\log(n)}{nh^3}\right)^{1/4},\sqrt{\frac{\log(n)}{nh^3}}\right),$$

where the remainder is uniform in $(v,b)\in I_x\times\mathbb{H}$.

Proof of Lemma S7. It is easy to check that

$$\mathcal{G}_x\left(W_i^*,W_j^*,v;b\right)=\mathcal{G}_x\left(\left(g\left(D_i^*,X_i^*,\epsilon_i^*\right),D_i^*,Z_i^*,X_i^*\right),\left(g\left(D_j^*,X_j^*,\epsilon_j^*\right),D_j^*,Z_j^*,X_j^*\right),v;b\right)=\mathcal{H}_x\left(U_i^*,U_j^*;v\right).$$

Denote

$$\begin{aligned}
\overline{\mathcal{H}}_x^{[1]}(u,v;b) &:= \mathrm{E}_{|W_1^n}\left[\mathcal{H}_x\left(U^*,u,v;b\right)\right]=\frac{1}{n}\sum_{j=1}^{n}\mathcal{H}_x\left(U_j,u,v;b\right)\\
\overline{\mathcal{H}}_x^{[2]}(u,v;b) &:= \mathrm{E}_{|W_1^n}\left[\mathcal{H}_x\left(u,U^*,v;b\right)\right]=\frac{1}{n}\sum_{j=1}^{n}\mathcal{H}_x\left(u,U_j,v;b\right)\\
\overline{\mu}_{\mathcal{H}_x}(v;b) &:= \mathrm{E}_{|W_1^n}\left[\mathcal{H}_x\left(U_1^*,U_2^*,v;b\right)\right]=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathcal{H}_x\left(U_i,U_j,v;b\right).
\end{aligned}$$

Then, by Hoeffding decomposition,

$$\frac{1}{n_{(2)}}\sum_{(i,j)}\mathcal{H}_x\left(U_i^*,U_j^*,v;b\right)=\overline{\mu}_{\mathcal{H}_x}(v;b)+\left\{\frac{1}{n}\sum_{i=1}^{n}\overline{\mathcal{H}}_x^{[1]}\left(U_i^*,v;b\right)-\overline{\mu}_{\mathcal{H}_x}(v;b)\right\}+\left\{\frac{1}{n}\sum_{i=1}^{n}\overline{\mathcal{H}}_x^{[2]}\left(U_i^*,v;b\right)-\overline{\mu}_{\mathcal{H}_x}(v;b)\right\}$$
$$+\frac{1}{n_{(2)}}\sum_{(i,j)}\left\{\mathcal{H}_x\left(U_i^*,U_j^*,v;b\right)-\overline{\mathcal{H}}_x^{[1]}\left(U_j^*,v;b\right)-\overline{\mathcal{H}}_x^{[2]}\left(U_i^*,v;b\right)+\overline{\mu}_{\mathcal{H}_x}(v;b)\right\}. \quad \text{(S129)}$$

Denote

$$\begin{aligned}
T_1^*(v;b) &:= \frac{1}{n}\sum_{i=1}^{n}\left(\overline{\mathcal{H}}_x^{[1]}\left(U_i^*,v;b\right)-\mathcal{H}_x^{[1]}\left(U_i^*,v;b\right)\right)-\left(\overline{\mu}_{\mathcal{H}_x}(v;b)-\frac{1}{n}\sum_{i=1}^{n}\mathcal{H}_x^{[1]}\left(U_i,v;b\right)\right)\\
T_2^*(v;b) &:= \frac{1}{n}\sum_{i=1}^{n}\overline{\mathcal{H}}_x^{[2]}\left(U_i^*,v;b\right)-\overline{\mu}_{\mathcal{H}_x}(v;b)\\
T_3^*(v;b) &:= \frac{1}{n_{(2)}}\sum_{(i,j)}\left\{\mathcal{H}_x\left(U_i^*,U_j^*,v;b\right)-\overline{\mathcal{H}}_x^{[1]}\left(U_j^*,v;b\right)-\overline{\mathcal{H}}_x^{[2]}\left(U_i^*,v;b\right)+\overline{\mu}_{\mathcal{H}_x}(v;b)\right\}.
\end{aligned}$$

Then, by (S129),

$$\frac{1}{n_{(2)}}\sum_{(i,j)}\mathcal{H}_x\left(U_i^*,U_j^*,v;b\right)=\overline{\mu}_{\mathcal{H}_x}\left(v;b\right)+\frac{1}{n}\sum_{i=1}^{n}\mathcal{H}_x^{[1]}\left(U_i^*,v;b\right)-\frac{1}{n}\sum_{i=1}^{n}\mathcal{H}_x^{[1]}\left(U_i,v;b\right)+T_1^*\left(v;b\right)+T_2^*\left(v;b\right)+T_3^*\left(v;b\right).$$

Note that $\mathrm{E}_{|W_1^n}\left[\left(\overline{\mathcal{H}}_x^{[1]}\left(U^*,v;b\right)-\mathcal{H}_x^{[1]}\left(U^*,v;b\right)\right)^2\right]$ can be represented by a V-statistic:

$$\begin{aligned}\mathrm{E}_{|W_1^n}&\left[\left(\overline{\mathcal{H}}_x^{[1]}\left(U^*,v;b\right)-\mathcal{H}_x^{[1]}\left(U^*,v;b\right)\right)^2\right]=\\&\frac{1}{n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\left(\mathcal{H}_x\left(U_j,U_i,v;b\right)-\mathcal{H}_x^{[1]}\left(U_i,v;b\right)\right)\left(\mathcal{H}_x\left(U_k,U_i,v;b\right)-\mathcal{H}_x^{[1]}\left(U_i,v;b\right)\right)\\&=:\frac{1}{n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathcal{V}_x\left(U_i,U_j,U_k,v;b\right).\end{aligned}\tag{S130}$$

It is easy to check that by using the V-statistic decomposition (Serfling (2009, 5.7.3)),

$$\frac{1}{n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathcal{V}_x\left(U_i,U_j,U_k,v;b\right)-\frac{1}{n_{(3)}}\sum_{(i,j,k)}\mathcal{V}_x\left(U_i,U_j,U_k,v;b\right)=O\left(\left(nh^4\right)^{-1}\right),\tag{S131}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$, and the kernel of the U-statistic $n_{(3)}^{-1}\sum_{(i,j,k)}\mathcal{V}_x\left(U_i,U_j,U_k,v;b\right)$ is degenerate of order one (see, e.g., Definition 5.1 of CK). Since both of $\mathfrak{H}:=\{\mathcal{H}_x\left(\cdot,v;b\right):(v,b)\in I_x\times\mathbb{H}\}$ and $\mathfrak{H}^{[1]}:=\left\{\mathcal{H}_x^{[1]}\left(\cdot,v;b\right):(v,b)\in I_x\times\mathbb{H}\right\}$ are uniformly VC-type with respect to a constant envelope that is a multiple of $\underline{h}^{-2}$ (see the proof of Lemma 4), by Chernozhukov et al. (2014a, Lemma B.2), $\mathfrak{V}:=\{\mathcal{V}_x\left(\cdot,v;b\right):(v,b)\in I_x\times\mathbb{H}\}$ is uniformly VC-type with respect to a constant envelope $F_{\mathfrak{V}}=O\left(h^{-4}\right)$. By Hoeffding decomposition (see Equation (18) in CK) of the U-process $\left\{\mathbb{U}_n^{(3)}f:f\in\mathfrak{V}\right\}$ and using the maximal inequality given by Corollary 5.6 of CK ($\mathcal{F}=\mathfrak{V}$, $r=3$, $k=2,3$, $p=1$, $F=F_{\mathfrak{V}}$), $\mathrm{E}\left[\left\|\mathbb{U}_n^{(3)}\right\|_{\mathfrak{V}}\right]=O\left(n^{-1/2}h^{-4}\right)$. Denote $\mathfrak{R}:=\left\{\overline{\mathcal{H}}_x^{[1]}\left(\cdot,v;b\right)-\mathcal{H}_x^{[1]}\left(\cdot,v;b\right):(v,b)\in I_x\times\mathbb{H}\right\}$ and

$$\widehat{\sigma}_{\mathfrak{R}}^2:=\sup_{f\in\mathfrak{R}}\mathbb{P}_n^Uf^2=\sup_{(v,b)\in I_x\times\mathbb{H}}\mathrm{E}_{|W_1^n}\left[\left(\overline{\mathcal{H}}_x^{[1]}\left(U^*,v;b\right)-\mathcal{H}_x^{[1]}\left(U^*,v;b\right)\right)^2\right].$$

By (S130) and (S131), $\mathrm{E}\left[\widehat{\sigma}_{\mathfrak{R}}^2\right]=n^{-1/2}\mathrm{E}\left[\left\|\mathbb{U}_n^{(3)}\right\|_{\mathfrak{V}}\right]=O\left(\left(nh^4\right)^{-1}\right)$. By Markov's inequality,

$$\Pr\left[\widehat{\sigma}_{\mathfrak{R}}>\left(\frac{\mathrm{E}\left[\widehat{\sigma}_{\mathfrak{R}}^2\right]}{\log\left(n\right)h}\right)^{1/4}\right]\leq\sqrt{\log\left(n\right)h\cdot\mathrm{E}\left[\widehat{\sigma}_{\mathfrak{R}}^2\right]}$$

and, therefore, $\widehat{\sigma}_{\mathfrak{R}}=O_p^{\star}\left(\log\left(n\right)^{-1/4}\left(nh^5\right)^{-1/4},\sqrt{\log\left(n\right)/\left(nh^3\right)}\right)$. By CK Lemma 5.4, the (data-dependent) function class $\overline{\mathfrak{H}}^{[1]}:=\left\{\overline{\mathcal{H}}_x^{[1]}\left(\cdot,v;b\right):(v,b)\in I_x\times\mathbb{H}\right\}$ is uniformly VC-type (conditionally on the data) with respect to a constant envelope $F_{\overline{\mathfrak{H}}^{[1]}}=F_{\mathfrak{H}}=O\left(h^{-2}\right)$, i.e., (36) with $\mathfrak{F}=\overline{\mathfrak{H}}^{[1]}$ and $F_{\mathfrak{F}}=F_{\overline{\mathfrak{H}}^{[1]}}$ is satisfied with VC characteristics that are functions of $(A_{\mathfrak{H}},V_{\mathfrak{H}})$ and do not depend on the data. Then, by Chernozhukov et al. (2014a, Lemma B.2), $\mathfrak{R}$ is also uniformly VC-type (conditionally on the data) with respect to the constant envelope $F_{\mathfrak{R}}=2F_{\mathfrak{H}}=O\left(h^{-2}\right)$ and its VC characteristics depend only on $(A_{\mathfrak{H}},V_{\mathfrak{H}})$. Then, by Talagrand's inequality

(Chernozhukov et al., 2016, Lemma 6.3 with $\mathcal{F} = \mathfrak{R}$, $b = F_{\mathfrak{R}}$, $\sigma = \widehat{\sigma}_{\mathfrak{R}} \vee b\sqrt{V_{\mathfrak{R}} \log(n)/n}$, $t = \log(n)$), we have

$$\Pr_{|W_1^n}\left[\left\|\mathbb{G}_n^{U^*}\right\|_{\mathfrak{R}} > C\left(\widehat{\sigma}_{\mathfrak{R}} \vee \sqrt{\frac{\log(n)}{nh^4}}\right)\sqrt{\log(n)}\right] \le n^{-1}.$$

Then, by $\|T_1^*\|_{I_x\times\mathbb{H}} = n^{-1/2}\left\|\mathbb{G}_n^{U^*}\right\|_{\mathfrak{R}}$ and Lemma 8, $\|T_1^*\|_{I_x\times\mathbb{H}} = O_p^\sharp\left(\left(\log(n)/\left(n^3h^5\right)\right)^{1/4}, n^{-1}, \sqrt{\log(n)/\left(nh^3\right)}\right)$. Denote $\overline{H}_x(e) \coloneqq n^{-1}\sum_{i=1}^n\left\{\mathbb{1}(\epsilon_i \le e) - F_{\epsilon|X}(e \mid x)\right\}\pi_x(Z_i, X_i)$. By Talagrand's inequality, we can easily show that $\left\|\overline{H}_x\right\|_{[\underline{\epsilon}_x, \overline{\epsilon}_x]} = O_p^\star\left(\sqrt{\log(n)/n}\right)$. Then, $\overline{\mathcal{H}}_x^{[2]}(U_i^*, v; b) = b^{-2}K'((\Delta_x(\epsilon_i^*) - v)/b)\varpi_x(U_i^*)\overline{H}_x(\epsilon_i^*)$. It can be easily verified by using arguments in the proof of Lemma 4 that the (data-dependent) function class $\overline{\mathfrak{H}}^{[2]} \coloneqq \left\{\overline{\mathcal{H}}_x^{[2]}(\cdot, v; b) : (v, b) \in I_x \times \mathbb{H}\right\}$ is uniformly VC-type (conditionally on the data) with respect to a constant envelope $F_{\overline{\mathfrak{H}}^{[2]}}$ that is a multiple of $\underline{h}^{-2}\left\|\overline{H}_x\right\|_{[\underline{\epsilon}_x, \overline{\epsilon}_x]}$, with VC characteristics that do not depend on the data. Then, by CK Corollary 5.6 (with $\mathcal{F} = \overline{\mathfrak{H}}^{[2]}$, $r = k = 1$, $p = 1$ and $F = F_{\overline{\mathfrak{H}}^{[2]}}$), $\mathrm{E}_{|W_1^n}\left[\|\mathbb{G}_n^{U^*}\|_{\overline{\mathfrak{H}}^{[2]}}\right] = O_p^\star\left(\sqrt{\log(n)/\left(nh^4\right)}\right)$. Then, by Markov's inequality,

$$\Pr_{|W_1^n}\left[\left\|\mathbb{G}_n^{U^*}\right\|_{\overline{\mathfrak{H}}^{[2]}} > h^{-1/4}\sqrt{\mathrm{E}_{|W_1^n}\left[\|\mathbb{G}_n^{U^*}\|_{\overline{\mathfrak{H}}^{[2]}}\right]}\right] \le h^{1/4}\sqrt{\mathrm{E}_{|W_1^n}\left[\|\mathbb{G}_n^{U^*}\|_{\overline{\mathfrak{H}}^{[2]}}\right]} = O_p^\star\left(\left(\frac{\log(n)}{nh^3}\right)^{1/4}\right).$$

By Lemma 8 and $\|T_2^*\|_{I_x\times\mathbb{H}} = n^{-1/2}\left\|\mathbb{G}_n^{U^*}\right\|_{\overline{\mathfrak{H}}^{[2]}}$, $\|T_2^*\|_{I_x\times\mathbb{H}} = O_p^\sharp\left(\left(\log(n)/\left(n^3h^5\right)\right)^{1/4}, \left(\log(n)/\left(nh^3\right)\right)^{1/4}, n^{-1}\right)$. By CK Corollary 5.6 (with $\mathcal{F} = \mathfrak{H}$, $r = k = 2$, $p = 1$ and $F = F_{\mathfrak{H}}$), $\mathrm{E}_{|W_1^n}\left[\|T_3^*\|_{I_x\times\mathbb{H}}\right] = O_p^\star\left(\left(nh^2\right)^{-1}\right)$. Then, by Markov's inequality,

$$\Pr_{|W_1^n}\left[\|T_3^*\|_{I_x\times\mathbb{H}} > \frac{\sqrt{\mathrm{E}_{|W_1^n}\left[\|T_3^*\|_{I_x\times\mathbb{H}}\right]}}{(nh)^{1/4}}\right] \le (nh)^{1/4}\sqrt{\mathrm{E}_{|W_1^n}\left[\|T_3^*\|_{I_x\times\mathbb{H}}\right]} = O_p^\star\left(\left(nh^3\right)^{-1/4}\right).$$

Therefore, by Lemma 8, we have $\|T_3^*\|_{I_x\times\mathbb{H}} = O_p^\sharp\left(\left(n^3h^5\right)^{-1/4}, \left(nh^3\right)^{-1/4}, n^{-1}\right)$. ■

Lemma S8. *Suppose that the assumptions of Theorem C1 hold. Then, (a)*

$$S_{\mathsf{npb}}(v \mid x; b) = \bar{S}_{\mathsf{npb}}(v \mid x; b) + O_p^\sharp\left(\left(\frac{\log(n)}{nh^3}\right)^{1/4} + \sqrt{\log(n)}\,h, \left(\frac{\log(n)}{nh^3}\right)^{1/4}, \sqrt{\frac{\log(n)}{nh^3}}\right),$$

where

$$\bar{S}_{\mathsf{npb}}(v \mid x; b) \coloneqq \frac{1}{\sqrt{n}}\sum_{i=1}^n\left(\mathcal{M}_x^{[1]}(U_i^*, v; b) - \frac{1}{n}\sum_{j=1}^n\mathcal{M}_x^{[1]}(U_j, v; b)\right),$$

uniformly in $(v, b) \in I_x \times \mathbb{H}$. (b) $\bar{S}_{\mathsf{npb}}(v \mid x; b) - \bar{S}_{\mathsf{npb}}(v \mid x; h) = O_p^\sharp\left(\varepsilon_n\sqrt{\log(n)}\right)$, uniformly in $(v, b) \in I_x \times \mathbb{H}$.

Proof of Lemma S8. By V-statistic decomposition Serfling (2009, 5.7.3) and the fact that $\mathfrak{H}$ has an $O\left(h^{-2}\right)$ envelope,

$$\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\mathcal{H}_x(U_i, U_j, v; b) - \frac{1}{n_{(2)}}\sum_{(i,j)}\mathcal{H}_x(U_i, U_j, v; b) = O\left(\left(nh^2\right)^{-1}\right),$$

uniformly in $(v, b) \in I_x \times \mathbb{H}$. We decompose

$$\widehat{f}^*_{\Delta X}(v, x; b) - \widehat{f}_{\Delta X}(v, x; b) = \left(\widehat{f}^*_{\Delta X}(v, x; b) - \widetilde{f}^*_{\Delta X}(v, x; b)\right) + \left(\widetilde{f}^*_{\Delta X}(v, x; b) - \widetilde{f}_{\Delta X}(v, x; b)\right)$$

$$-\left(\widehat{f}_{\Delta X}(v,x;b)-\widetilde{f}_{\Delta X}(v,x;b)\right).$$

Then, by Lemmas 3, 8 and S7,

$$\widehat{f}^*_{\Delta X}(v,x;b)-\widehat{f}_{\Delta X}(v,x;b)=\left\{\frac{1}{n}\sum_{i=1}^{n}\mathcal{H}_x^{[1]}(U_i^*,v;b)-\frac{1}{n}\sum_{i=1}^{n}\mathcal{H}_x^{[1]}(U_i,v;b)\right\}+\left\{\widetilde{f}^*_{\Delta X}(v,x;b)-\widetilde{f}_{\Delta X}(v,x;b)\right\}$$
$$+O_p^{\sharp}\left(\left(\frac{\log(n)}{n^3h^5}\right)^{1/4},\left(\frac{\log(n)}{nh^3}\right)^{1/4},\sqrt{\frac{\log(n)}{nh^3}}\right),\quad\text{(S132)}$$

uniformly in $(v,b)\in I_x\times\mathbb{H}$. Then, let $\widehat{\sigma}^2_{\mathfrak{H}^{[1]}}:=\sup_{f\in\mathfrak{H}^{[1]}}\mathbb{P}_n^U f^2$ and $\ddot{\mathfrak{H}}:=\left\{\mathcal{H}_x^{[1]}(\cdot,v;b)^2:(v,b)\in I_x\times\mathbb{H}\right\}$. By Chernozhukov et al. (2014a, Lemma B.2), $\ddot{\mathfrak{H}}$ is uniformly VC-type with respect to a constant envelope $F_{\ddot{\mathfrak{H}}}=O(h^{-4})$. Then, $\widehat{\sigma}^2_{\mathfrak{H}^{[1]}}\leq\sigma^2_{\mathfrak{H}^{[1]}}+\left\|\mathbb{P}_n^U-\mathbb{P}^U\right\|_{\ddot{\mathfrak{H}}}$. It was shown in the proof of Lemma 4 that $\sigma^2_{\mathfrak{H}^{[1]}}=O(h^{-1})$. Let $\sigma^2_{\ddot{\mathfrak{H}}}:=\sup_{f\in\ddot{\mathfrak{H}}}\mathbb{P}^U f^2$. It is easy to check that by change of variables, $\sigma^2_{\ddot{\mathfrak{H}}}=O(h^{-4})$ and therefore, by Talagrand's inequality (Chernozhukov et al., 2016, Lemma 6.3 with $\mathcal{F}=\ddot{\mathfrak{H}}$, $b=F_{\ddot{\mathfrak{H}}}$, $\sigma=\sigma_{\ddot{\mathfrak{H}}}\vee b\sqrt{V_{\ddot{\mathfrak{H}}}\log(n)/n}$, $t=\log(n)$), $\left\|\mathbb{G}_n^U\right\|_{\ddot{\mathfrak{H}}}=O_p^{\star}\left(\sqrt{\log(n)/h^4}+\log(n)/(n^{1/2}h^4)\right)$ and therefore, $\left\|\mathbb{P}_n^U-\mathbb{P}^U\right\|_{\ddot{\mathfrak{H}}}=n^{-1/2}\left\|\mathbb{G}_n^U\right\|_{\ddot{\mathfrak{H}}}=O_p^{\star}\left(\sqrt{\log(n)/(nh^4)}+\log(n)/(nh^4)\right)$ and $\widehat{\sigma}^2_{\mathfrak{H}^{[1]}}=O_p^{\star}(h^{-1})$. By using Talagrand's inequality (Chernozhukov et al., 2016, Lemma 6.3 with $\mathcal{F}=\mathfrak{H}^{[1]}$, $b=F_{\mathfrak{H}^{[1]}}=O(h^{-2})$, $\sigma=\widehat{\sigma}_{\mathfrak{H}^{[1]}}\vee b\sqrt{V_{\mathfrak{H}^{[1]}}\log(n)/n}$ and $t=\log(n)$) and Lemma 8, $\left\|\mathbb{G}_n^{U^*}\right\|_{\mathfrak{H}^{[1]}}=O_p^{\sharp}\left(\sqrt{\log(n)/h}\right)$ and therefore, $n^{-1}\sum_{i=1}^n\mathcal{H}_x^{[1]}(U_i^*,v;b)-n^{-1}\sum_{i=1}^n\mathcal{H}_x^{[1]}(U_i,v;b)$ is $O_p^{\sharp}\left(\sqrt{\log(n)/(nh)}\right)$, uniformly in $(v,b)\in I_x\times\mathbb{H}$. Similarly, by Talagrand's inequality, $\left\|\widetilde{f}^*_{\Delta X}(\cdot,x;\cdot)-\widetilde{f}_{\Delta X}(\cdot,x;\cdot)\right\|_{I_x\times\mathbb{H}}=O_p^{\sharp}\left(\sqrt{\log(n)/(nh)}\right)$. By these results and (S132),

$$\left\|\widehat{f}^*_{\Delta X}(\cdot,x;\cdot)-\widehat{f}_{\Delta X}(\cdot,x;\cdot)\right\|_{I_x\times\mathbb{H}}=O_p^{\sharp}\left(\sqrt{\frac{\log(n)}{nh}},\left(\frac{\log(n)}{nh^3}\right)^{1/4},\sqrt{\frac{\log(n)}{nh^3}}\right).\quad\text{(S133)}$$

By Lemma 8 and Hoeffding's inequality, we have $\widehat{p}^*_x-\widehat{p}_x=O_p^{\sharp}\left(\sqrt{\log(n)/n}\right)$. By $\widehat{p}_x-p_x=O_p^{\star}\left(\sqrt{\log(n)/n}\right)$ and Lemma 8, we have $\widehat{p}^*_x-p_x=O_p^{\sharp}\left(\sqrt{\log(n)/n}\right)$ ($\exists C_1,C_2,C_3>0,\Pr\left[\Pr_{|W_1^n}\left[|\widehat{p}^*_x-p_x|>C_1\sqrt{\log(n)/n}\right]>C_2n^{-1}\right]\leq C_3n^{-1}$). By using

$$\Pr_{|W_1^n}\left[\left|\frac{p_x}{\widehat{p}^*_x}-1\right|>\frac{2C_1}{p_x}\sqrt{\frac{\log(n)}{n}}\right]\leq\Pr_{|W_1^n}\left[|\widehat{p}^*_x-p_x|>C_1\sqrt{\frac{\log(n)}{n}}\right]+\Pr_{|W_1^n}\left[\widehat{p}^*_x<\frac{p_x}{2}\right]$$
$$\leq 2\cdot\Pr_{|W_1^n}\left[|\widehat{p}^*_x-p_x|>C_1\sqrt{\frac{\log(n)}{n}}\right],$$

where the second inequality holds when n is sufficiently large, we have $p_x/\widehat{p}^*_x-1=O_p^{\sharp}\left(\sqrt{\log(n)/n}\right)$. By this result, $p_x/\widehat{p}_x-1=O_p^{\star}\left(\sqrt{\log(n)/n}\right)$, Lemma 8, (S132), (S133) and

$$S_{\mathsf{npb}}(v\mid x;b)=\frac{1}{p_x}\sqrt{nb}\left(\widehat{f}^*_{\Delta X}(v,x;b)-\widehat{f}_{\Delta X}(v,x;b)\right)+\left(\frac{1}{\widehat{p}_x}-\frac{1}{p_x}\right)\sqrt{nb}\left(\widehat{f}^*_{\Delta X}(v,x;b)-\widehat{f}_{\Delta X}(v,x;b)\right)$$
$$+\frac{\sqrt{nb}\widehat{f}^*_{\Delta X}(v,x;b)}{\widehat{p}_x}\left(\frac{\widehat{p}_x}{\widehat{p}^*_x}-1\right),$$

we have the first assertion and also

$$\|S_{\mathsf{npb}}(\cdot \mid x;\cdot)\|_{I_x\times\mathbb{H}} = O_p^{\sharp}\left(\sqrt{\log(n)}, \left(\frac{\log(n)}{nh^3}\right)^{1/4}, \sqrt{\frac{\log(n)}{nh^3}}\right). \tag{S134}$$

Note that

$$\begin{aligned}
&\bar{S}_{\mathsf{npb}}(v \mid x; b) - \bar{S}_{\mathsf{npb}}(v \mid x; h) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\mathcal{E}_x^{\triangle}(U_i^*, v; b, h) - \frac{1}{n}\sum_{j=1}^{n}\mathcal{E}_x^{\triangle}(U_j, v; b, h)\right) + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\mathcal{H}_x^{\triangle[1]}(U_i^*, v; b, h) - \frac{1}{n}\sum_{j=1}^{n}\mathcal{H}_x^{\triangle[1]}(U_j, v; b, h)\right).
\end{aligned}$$

It is shown in the proof of Lemma 9 that $\widehat{\sigma}^2_{\mathfrak{E}^{\triangle}} := \sup_{f\in\mathfrak{E}^{\triangle}}\mathbb{P}_n^U f^2 = O_p^{\star}(\varepsilon_n^2)$. By Talagrand's inequality (Chernozhukov et al., 2016, Lemma 6.3 with $\mathcal{F} = \mathfrak{E}^{\triangle}$, $b = F_{\mathfrak{E}^{\triangle}}$, $\sigma = \widehat{\sigma}_{\mathfrak{E}^{\triangle}} \vee b\sqrt{V_{\mathfrak{E}^{\triangle}}\log(n)/n}$, $t = \log(n)$), $\left\|\mathbb{G}_n^{U^*}\right\|_{\mathfrak{E}^{\triangle}} = O_p^{\sharp}\left(\varepsilon_n\sqrt{\log(n)}\right)$. Let $\widehat{\sigma}^2_{\mathfrak{H}^{\triangle[1]}} := \sup_{f\in\mathfrak{H}^{\triangle[1]}}\mathbb{P}_n^U f^2 \le \sigma^2_{\mathfrak{H}^{\triangle[1]}} + \left\|\mathbb{P}_n^U - \mathbb{P}^U\right\|_{\mathfrak{H}^{\triangle\triangle}}$, where by CK Lemma 5.4 and Chernozhukov et al. (2014a, Lemma B.2), $\mathfrak{H}^{\triangle\triangle} := \left\{\mathcal{H}_x^{\triangle[1]}(\cdot, v; b, h)^2 : (v, b) \in I_x \times \mathbb{H}\right\}$ is uniformly VC-type with respect to a constant envelope $F_{\mathfrak{H}^{\triangle\triangle}} = O\left(\varepsilon_n^2/h^3\right)$. Let $\sigma^2_{\mathfrak{H}^{\triangle\triangle}} := \sup_{f\in\mathfrak{H}^{\triangle\triangle}}\mathbb{P}^U f^2 = O\left(\varepsilon_n^4/h^2\right)$, where the second equality follows from change of variables and (S35). By Talagrand's inequality (Chernozhukov et al. (2016, Lemma 6.3) with $\mathcal{F} = \mathfrak{H}^{\triangle\triangle}$, $b = F_{\mathfrak{H}^{\triangle\triangle}}$, $\sigma = \sigma_{\mathfrak{H}^{\triangle\triangle}} \vee b\sqrt{V_{\mathfrak{H}^{\triangle\triangle}}\log(n)/n}$, $t = \log(n)$), $\left\|\mathbb{G}_n^U\right\|_{\mathfrak{H}^{\triangle\triangle}} = O_p^{\star}\left(\varepsilon_n^2\sqrt{\log(n)/h^2}\right)$. Then we have $\left\|\mathbb{P}_n^U - \mathbb{P}^U\right\|_{\mathfrak{H}^{\triangle\triangle}} = O_p^{\star}\left(\varepsilon_n^2\sqrt{\log(n)/(nh^2)}\right)$. It is shown in the proof of Lemma 5 that $\sigma^2_{\mathfrak{H}^{\triangle[1]}} = O\left(\varepsilon_n^2\right)$. Therefore, $\widehat{\sigma}^2_{\mathfrak{H}^{\triangle[1]}} = O_p^{\star}\left(\varepsilon_n^2\right)$ and it follows from Talagrand's inequality that $\left\|\mathbb{G}_n^{U^*}\right\|_{\mathfrak{H}^{\triangle[1]}} = O_p^{\sharp}\left(\varepsilon_n\sqrt{\log(n)}\right)$. We have the second assertion. ■

Proof of Theorem C1. By using $\Pr\left[\widehat{h} \in \mathbb{H}\right] > 1 - \delta_n$, Lemma S8 and monotonicity of conditional expectations, we have

$$\begin{aligned}
&\Pr\left[\Pr_{|W_1^n}\left[\left\|S_{\mathsf{npb}}\left(\cdot \mid x; \widehat{h}\right) - \bar{S}_{\mathsf{npb}}\left(\cdot \mid x; \widehat{h}\right)\right\|_{I_x} > C_1\left(\left(\frac{\log(n)}{nh^3}\right)^{1/4} + \sqrt{\log(n)\,h}\right)\right] > C_2\left(\frac{\log(n)}{nh^3}\right)^{1/4}\right] \\
&\le \Pr\left[\Pr_{|W_1^n}\left[\left\|S_{\mathsf{npb}}\left(\cdot \mid x; \widehat{h}\right) - \bar{S}_{\mathsf{npb}}\left(\cdot \mid x; \widehat{h}\right)\right\|_{I_x} > C_1\left(\left(\frac{\log(n)}{nh^3}\right)^{1/4} + \sqrt{\log(n)\,h}\right)\right] > C_2\left(\frac{\log(n)}{nh^3}\right)^{1/4}, \widehat{h} \in \mathbb{H}\right] \\
&+\delta_n \le \Pr\left[\Pr_{|W_1^n}\left[\left\|S_{\mathsf{npb}}(\cdot \mid x; \cdot) - \bar{S}_{\mathsf{npb}}(\cdot \mid x; \cdot)\right\|_{I_x\times\mathbb{H}} > C_1\left(\left(\frac{\log(n)}{nh^3}\right)^{1/4} + \sqrt{\log(n)\,h}\right)\right] > C_2\left(\frac{\log(n)}{nh^3}\right)^{1/4}\right] + \delta_n
\end{aligned}$$

and therefore,

$$S_{\mathsf{npb}}\left(v \mid x; \widehat{h}\right) = \bar{S}_{\mathsf{npb}}\left(v \mid x; \widehat{h}\right) + O_p^{\sharp}\left(\left(\frac{\log(n)}{nh^3}\right)^{1/4} + \sqrt{\log(n)\,h}, \left(\frac{\log(n)}{nh^3}\right)^{1/4}, \sqrt{\frac{\log(n)}{nh^3}} + \delta_n\right),$$

uniformly in $v \in I_x$. By similar arguments, $\bar{S}_{\mathsf{npb}}\left(v \mid x; \widehat{h}\right) = \bar{S}_{\mathsf{npb}}(v \mid x; h) + O_p^{\sharp}\left(\varepsilon_n\sqrt{\log(n)}, n^{-1}, n^{-1} + \delta_n\right)$, uniformly in $v \in I_x$. Therefore,

$$S_{\mathsf{npb}}\left(v \mid x; \widehat{h}\right) - \bar{S}_{\mathsf{npb}}(v \mid x; h) = O_p^{\sharp}\left(\left(\frac{\log(n)}{nh^3}\right)^{1/4} + \sqrt{\log(n)\,h} + \varepsilon_n\sqrt{\log(n)}, \left(\frac{\log(n)}{nh^3}\right)^{1/4}, \sqrt{\frac{\log(n)}{nh^3}} + \delta_n\right), \tag{S135}$$

uniformly in $v \in I_x$. By (S134) and similar arguments,

$$\left\| S_{\mathsf{npb}}\left(\cdot \mid x; \widehat{h}\right) \right\|_{I_x \times \mathbb{H}} = O_p^{\sharp}\left(\sqrt{\log(n)}, \left(\frac{\log(n)}{nh^3}\right)^{1/4}, \sqrt{\frac{\log(n)}{nh^3}} + \delta_n\right).$$

Write

$$Z_{\mathsf{npb}}\left(v \mid x; \widehat{h}, \widehat{h}_{\zeta}\right) - \frac{\bar{S}_{\mathsf{npb}}(v \mid x; h)}{\sqrt{V(v \mid x; h)}} = \frac{S_{\mathsf{npb}}\left(v \mid x; \widehat{h}\right)}{\sqrt{V(v \mid x; h)}}\left(\frac{\sqrt{V(v \mid x; h)}}{\sqrt{\widehat{V}\left(v \mid x; \widehat{h}, \widehat{h}_{\zeta}\right)}} - 1\right) + \frac{S_{\mathsf{npb}}\left(v \mid x; \widehat{h}\right) - \bar{S}_{\mathsf{npb}}(v \mid x; h)}{\sqrt{V(v \mid x; h)}}.$$

By these results, (55), Theorem B1, (S135) and $\left\| \bar{S}_{\mathsf{npb}}(\cdot \mid x; h) / \sqrt{V(\cdot \mid x; h)} \right\|_{I_x} = \left\| \mathbb{G}_n^{U^*} \right\|_{\bar{\mathfrak{M}}^{[1]}}$, we have

$$\begin{aligned} &\left\| Z_{\mathsf{npb}}\left(\cdot \mid x; \widehat{h}, \widehat{h}_{\zeta}\right) \right\|_{I_x} - \left\| \mathbb{G}_n^{U^*} \right\|_{\bar{\mathfrak{M}}^{[1]}} \\ &\quad = O_p^{\sharp}\left(\left(\frac{\log(n)}{nh^3}\right)^{1/4} + \sqrt{\log(n)\, h} + \varepsilon_n \sqrt{\log(n)} + \kappa_{1,n}^V \sqrt{\log(n)}, \left(\frac{\log(n)}{nh^3}\right)^{1/4}, \kappa_{2,n}^V + \delta_n + \delta_n^{\zeta}\right). \end{aligned} \tag{S136}$$

We apply Chernozhukov et al. (2016, Theorem 2.3) with $\mathcal{F} = \bar{\mathfrak{M}}_{\pm}^{[1]}$, $B = 0$, $\sigma = \overline{\sigma}_{\bar{\mathfrak{M}}^{[1]}}$, $b = F_{\bar{\mathfrak{M}}^{[1]}}$ and $q = \infty$. When n is sufficiently large, for any coupling error $\gamma \in (0, 1)$, there exists a random variable $Z^*_{\bar{\mathfrak{M}}_{\pm}^{[1]},\gamma}$ such that (1) $Z^*_{\bar{\mathfrak{M}}_{\pm}^{[1]},\gamma}$ is independent of the data; (2) $Z^*_{\bar{\mathfrak{M}}_{\pm}^{[1]},\gamma}$ has the same distribution as $\left\| G^U \right\|_{\bar{\mathfrak{M}}^{[1]}}$; (3) $Z^*_{\bar{\mathfrak{M}}_{\pm}^{[1]},\gamma}$ and $\left\| \mathbb{G}_n^{U^*} \right\|_{\bar{\mathfrak{M}}^{[1]}}$ satisfies the deviation bound:

$$\Pr\left[\left| \left\| \mathbb{G}_n^{U^*} \right\|_{\bar{\mathfrak{M}}^{[1]}} - Z^*_{\bar{\mathfrak{M}}_{\pm}^{[1]},\gamma} \right| > C_1 \kappa^*_{\bar{\mathfrak{M}}_{\pm}^{[1]}}(\gamma)\right] \leq C_2\left(\gamma + n^{-1}\right),$$

where $\kappa^*_{\bar{\mathfrak{M}}_{\pm}^{[1]}}(\gamma) := \log(n)^{2/3} / \left(\gamma^{1/3}\left(nh^3\right)^{1/6}\right) + \log(n)^{3/4} / \left(\gamma\left(nh^3\right)^{1/4}\right)$, and by Markov's inequality,

$$\Pr\left[\Pr_{|W_1^n}\left[\left| \left\| \mathbb{G}_n^{U^*} \right\|_{\bar{\mathfrak{M}}^{[1]}} - Z^*_{\bar{\mathfrak{M}}_{\pm}^{[1]},\gamma} \right| > C_1 \kappa^*_{\bar{\mathfrak{M}}_{\pm}^{[1]}}(\gamma)\right] > \sqrt{C_2\left(\gamma + n^{-1}\right)}\right] \leq \sqrt{C_2\left(\gamma + n^{-1}\right)}.$$

By this result and (S136), when n is sufficiently large, $\forall \gamma \in (0, 1)$,

$$\begin{aligned} \Pr\Bigg[\Pr_{|W_1^n}\Bigg[\left| \left\| Z_{\mathsf{npb}}\left(\cdot \mid x; \widehat{h}, \widehat{h}_{\zeta}\right) \right\|_{I_x} - Z^*_{\bar{\mathfrak{M}}_{\pm}^{[1]},\gamma} \right| > C_1\left(\kappa^*_{\bar{\mathfrak{M}}_{\pm}^{[1]}}(\gamma) + \sqrt{\log(n)\, h} + \kappa_{1,n}^V \sqrt{\log(n)} + \varepsilon_n \sqrt{\log(n)}\right)\Bigg] > \\ C_2\left(\sqrt{\gamma} + \left(\frac{\log(n)}{nh^3}\right)^{1/4}\right)\Bigg] \leq C_3\left(\sqrt{\gamma} + \kappa_{2,n}^V + \delta_n + \delta_n^{\zeta}\right). \end{aligned}$$

Then, since $Z^*_{\bar{\mathfrak{M}}_{\pm}^{[1]},\gamma}$ is independent of the data and $Z^*_{\bar{\mathfrak{M}}_{\pm}^{[1]},\gamma} =_d \left\| G^U \right\|_{\bar{\mathfrak{M}}^{[1]}}$, by the above deviation bound and Chernozhukov et al. (2016, Lemma 2.1), with probability greater than $1 - C_3\left(\sqrt{\gamma} + \kappa_{2,n}^V + \delta_n + \delta_n^{\zeta}\right)$,

$$\begin{aligned} &\sup_{t \in \mathbb{R}} \left| \Pr_{|W_1^n}\left[\left\| Z_{\mathsf{npb}}\left(\cdot \mid x; \widehat{h}, \widehat{h}_{\zeta}\right) \right\|_{I_x} \leq t\right] - \Pr\left[\left\| G^U \right\|_{\bar{\mathfrak{M}}^{[1]}} \leq t\right] \right| \leq \\ &\quad \sup_{t \in \mathbb{R}} \Pr\left[\left| \left\| G^U \right\|_{\bar{\mathfrak{M}}^{[1]}} - t \right| \leq C_1\left(\kappa^*_{\bar{\mathfrak{M}}_{\pm}^{[1]}}(\gamma) + \sqrt{\log(n)\, h} + \kappa_{1,n}^V \sqrt{\log(n)} + \varepsilon_n \sqrt{\log(n)}\right)\right] \\ &\quad\quad + C_2\left(\sqrt{\gamma} + \left(\frac{\log(n)}{nh^3}\right)^{1/4}\right). \end{aligned}$$

By (S139), (S140) and optimally choosing $\gamma = \log(n)^{5/6} / \left(nh^3\right)^{1/6}$ (which balances $\sqrt{\gamma}$ and $\log(n)^{5/4} / \left(\gamma n^{1/4} h^{3/4}\right)$), we have (S136). By repeating the arguments used to show (62), we have $\left|\Pr\left[\left\|Z\left(\cdot\mid x;\widehat{h},\widehat{h}_{\zeta}\right)\right\|_{I_x}\leq z_{1-\alpha}^{\mathsf{npb}}\right]-(1-\alpha)\right|\leq C_1\bar{\kappa}_{1,n}+C_2\bar{\kappa}_{2,n}^{*}+C_3\bar{\kappa}_{3,n}^{*}$. ■

Proof of Theorem C2. By similar arguments used in the proof of Lemma 4, the centered function class $\mathfrak{M}:=\{\mathcal{M}_x(\cdot,v;h):v\in I_x\}$ is uniformly VC-type with respect to a constant envelope $F_{\mathfrak{M}}=O\left(h^{-3/2}\right)$. By (53) and (49), we have

$$\left\|S\left(\cdot\mid x;\widehat{h}\right)\right\|_{I_x}-\left\|\mathbb{U}_n^{(2)}\right\|_{\mathfrak{M}}=O_p^{\star}\left(\varepsilon_n\sqrt{\log(n)}+\upsilon_n+\sqrt{nh^5},\sqrt{\frac{\log(n)}{nh^3}}+\delta_n\right).\tag{S137}$$

By Lemma A.3 of CK, $\mathfrak{M}^{[1]}:=\left\{\mathcal{M}_x^{[1]}(\cdot,v;h):v\in I_x\right\}$ is also uniformly VC-type with respect to a constant envelope $F_{\mathfrak{M}^{[1]}}=F_{\mathfrak{M}}$. Let $\overline{\sigma}_{\mathfrak{M}^{[1]}}^2:=\sup_{f\in\mathfrak{M}^{[1]}}\mathbb{P}^U f^2=\sup_{v\in I_x}\mathscr{V}(v\mid x)+o(1)$ and $\sigma_{\mathfrak{M}}^2:=\sup_{f\in\mathfrak{M}}\mathrm{E}\left[f(U_1,U_2)^2\right]$. By calculations in the proof of Lemma 4, $\overline{\sigma}_{\mathfrak{M}^{[1]}}^2=O(1)$ and $\sigma_{\mathfrak{M}}^2=O\left(h^{-2}\right)$. By CK Proposition 2.1 (with $\mathcal{H}=\mathfrak{M}_{\pm}$, $\overline{\sigma}_{\mathfrak{g}}=\overline{\sigma}_{\mathfrak{M}^{[1]}}$, $\sigma_{\mathfrak{h}}=\sigma_{\mathfrak{M}}$, $b_{\mathfrak{g}}=b_{\mathfrak{h}}=F_{\mathfrak{M}}$, $\chi_n=0$ and $q=\infty$), when n is sufficiently large, for each coupling error $\gamma\in(0,1)$, one can construct a random variable $Z_{\mathfrak{M}_{\pm},\gamma}$ that satisfies the following conditions: $Z_{\mathfrak{M}_{\pm},\gamma}=_d\sup_{f\in\mathfrak{M}_{\pm}^{[1]}}G^U(f)=\left\|G^U\right\|_{\mathfrak{M}^{[1]}}$, where $\left\{G^U(f):f\in\mathfrak{M}_{\pm}^{[1]}\right\}$ is a centered separable Gaussian process that has the same covariance structure as the Hájek process $\left\{\mathbb{G}_n^U f:f\in\mathfrak{M}_{\pm}^{[1]}\right\}$ ($\mathrm{E}\left[G^U(f)G^U(g)\right]=\mathrm{Cov}\left[f(U),g(U)\right]$, $\forall f,g\in\mathfrak{M}_{\pm}^{[1]}$), and the difference between $\sup_{f\in\mathfrak{M}_{\pm}}\mathbb{U}_n^{(2)}f=\left\|\mathbb{U}_n^{(2)}\right\|_{\mathfrak{M}}$ and $Z_{\mathfrak{M}_{\pm},\gamma}$ satisfies the deviation bound:

$$\Pr\left[\left|\left\|\mathbb{U}_n^{(2)}\right\|_{\mathfrak{M}}-Z_{\mathfrak{M}_{\pm},\gamma}\right|>C_1\left(\frac{\log(n)^{2/3}}{\gamma^{1/3}\left(nh^3\right)^{1/6}}+\frac{\log(n)}{\gamma\sqrt{nh^3}}\right)\right]\leq C_2\left(\gamma+n^{-1}\right).\tag{S138}$$

We denote $\underline{V}:=\inf_{f\in\mathfrak{M}^{[1]}}\mathrm{Var}\left[f(U)\right]=\inf_{v\in I_x}V(v\mid x;h)$. We show in the proof of Theorem B2 that since $\underline{V}\to\inf_{v\in I_x}\mathscr{V}(v\mid x)>0$ as $h\downarrow 0$, when h is sufficiently small, $\underline{V}>\inf_{v\in I_x}\mathscr{V}(v\mid x)/2>0$. Similarly, let $\overline{V}:=\sup_{f\in\mathfrak{M}^{[1]}}\mathrm{Var}\left[f(U)\right]=\sup_{v\in I_x}V(v\mid x;h)$. By (43), we have $\overline{V}\to\sup_{v\in I_x}\mathscr{V}(v\mid x)\in(0,\infty)$. By the Gaussian anti-concentration inequality (CK Lemma A.1),

$$\sup_{t\in\mathbb{R}}\Pr\left[\left|\left\|G^U\right\|_{\mathfrak{M}^{[1]}}-t\right|\leq\varepsilon\right]\leq C_{\sigma}\varepsilon\left(\mathrm{E}\left[\left\|G^U\right\|_{\mathfrak{M}^{[1]}}\right]+\sqrt{1\vee\log\left(\underline{V}^{1/2}/\varepsilon\right)}\right),\ \forall\varepsilon>0,\tag{S139}$$

where C_{σ} is a constant that depends on $\underline{V}^{1/2}$ and $\overline{V}^{1/2}$. Since $\underline{V}\to\inf_{v\in I_x}\mathscr{V}(v\mid x)$ and $\overline{V}\to\sup_{v\in I_x}\mathscr{V}(v\mid x)$ as $h\downarrow 0$, we have $C_{\sigma}=O(1)$. By Dudley's metric entropy bound (Giné and Nickl, 2016, Theorem 2.3.7),

$$\mathrm{E}\left[\left\|G^U\right\|_{\mathfrak{M}^{[1]}}\right]\lesssim\left(\overline{\sigma}_{\mathfrak{M}^{[1]}}\vee n^{-1/2}\left\|F_{\mathfrak{M}^{[1]}}\right\|_{\mathbb{P}^U,2}\right)\sqrt{\log(n)},\tag{S140}$$

when n is sufficiently large. Then, since $Z_{\mathfrak{M}_{\pm},\gamma}=_d\left\|G^U\right\|_{\mathfrak{M}^{[1]}}$, by Chernozhukov et al. (2016, Lemma 2.1), (S137) and (S138), when n is sufficiently large, $\forall\gamma\in(0,1)$,

$$\begin{aligned}&\sup_{t\in\mathbb{R}}\left|\Pr\left[\left\|S\left(\cdot\mid x;\widehat{h}\right)\right\|_{I_x}\leq t\right]-\Pr\left[\left\|G^U\right\|_{\mathfrak{M}^{[1]}}\leq t\right]\right|\\&\quad\leq\sup_{t\in\mathbb{R}}\Pr\left[\left|\left\|G^U\right\|_{\mathfrak{M}^{[1]}}-t\right|\leq C_1\left(\varepsilon_n\sqrt{\log(n)}+\upsilon_n+\sqrt{nh^5}+\frac{\log(n)^{2/3}}{\gamma^{1/3}\left(nh^3\right)^{1/6}}+\frac{\log(n)}{\gamma\sqrt{nh^3}}\right)\right]\\&\qquad+C_2\left(\gamma+\sqrt{\frac{\log(n)}{nh^3}}+\delta_n\right).\end{aligned}\tag{S141}$$

By (S139), (S140) and optimally choosing γ that gives the fastest rate of convergence of the right hand side of (S141), which should be $\gamma=\log(n)^{7/8}/\left(nh^3\right)^{1/8}$, we have (68).

Since $\left\|\bar{S}_{\mathsf{npb}}(\cdot \mid x; h)\right\|_{I_x} = \left\|\mathbb{G}_n^{U^*}\right\|_{\mathfrak{M}^{[1]}}$, by (S135), we have

$$\left\|S_{\mathsf{npb}}\left(\cdot \mid x; \widehat{h}\right)\right\|_{I_x} - \left\|\mathbb{G}_n^{U^*}\right\|_{\mathfrak{M}^{[1]}} = O_p^{\sharp}\left(\left(\frac{\log(n)}{nh^3}\right)^{1/4} + \sqrt{\log(n)\,h} + \varepsilon_n\sqrt{\log(n)}, \left(\frac{\log(n)}{nh^3}\right)^{1/4}, \sqrt{\frac{\log(n)}{nh^3}} + \delta_n\right). \tag{S142}$$

Then we apply Chernozhukov et al. (2016, Theorem 2.3) with $\mathcal{F} = \mathfrak{M}_{\pm}^{[1]}$, $B = 0$, $\sigma = \overline{\sigma}_{\mathfrak{M}^{[1]}}$, $b = F_{\mathfrak{M}^{[1]}}$ and $q = \infty$. When n is sufficiently large, for any coupling error $\gamma \in (0, 1)$, there exists a random variable $Z^*_{\mathfrak{M}_{\pm}^{[1]},\gamma}$ such that (1) $Z^*_{\mathfrak{M}_{\pm}^{[1]},\gamma}$ is independent of the data; (2) $Z^*_{\mathfrak{M}_{\pm}^{[1]},\gamma}$ has the same distribution as $\left\|G^U\right\|_{\mathfrak{M}^{[1]}}$; (3) $Z^*_{\mathfrak{M}_{\pm}^{[1]},\gamma}$ and $\left\|\mathbb{G}_n^{U^*}\right\|_{\mathfrak{M}^{[1]}}$ satisfies the deviation bound:

$$\Pr\left[\left|\left\|\mathbb{G}_n^{U^*}\right\|_{\mathfrak{M}^{[1]}} - Z^*_{\mathfrak{M}_{\pm}^{[1]},\gamma}\right| > C_1\left(\frac{\log(n)^{2/3}}{\gamma^{1/3}(nh^3)^{1/6}} + \frac{\log(n)^{3/4}}{\gamma(nh^3)^{1/4}}\right)\right] \leq C_2\left(\gamma + n^{-1}\right)$$

and by Markov's inequality,

$$\Pr\left[\Pr_{|W_1^n}\left[\left|\left\|\mathbb{G}_n^{U^*}\right\|_{\mathfrak{M}^{[1]}} - Z^*_{\mathfrak{M}_{\pm}^{[1]},\gamma}\right| > C_1\left(\frac{\log(n)^{2/3}}{\gamma^{1/3}(nh^3)^{1/6}} + \frac{\log(n)^{3/4}}{\gamma(nh^3)^{1/4}}\right)\right] > \sqrt{C_2\left(\gamma + n^{-1}\right)}\right] \leq \sqrt{C_2\left(\gamma + n^{-1}\right)}.$$

By this result and (S142), when n is sufficiently large, $\forall \gamma \in (0, 1)$,

$$\begin{aligned}\Pr\Bigg[\Pr_{|W_1^n}\Bigg[\left|\left\|S_{\mathsf{npb}}\left(\cdot \mid x; \widehat{h}\right)\right\|_{I_x} - Z^*_{\mathfrak{M}_{\pm}^{[1]},\gamma}\right| > C_1\left(\frac{\log(n)^{2/3}}{\gamma^{1/3}(nh^3)^{1/6}} + \frac{\log(n)^{3/4}}{\gamma(nh^3)^{1/4}} + \sqrt{\log(n)\,h} + \varepsilon_n\sqrt{\log(n)}\right)\Bigg] > \\ C_2\left(\sqrt{\gamma} + \left(\frac{\log(n)}{nh^3}\right)^{1/4}\right)\Bigg] \leq C_3\left(\sqrt{\gamma} + \sqrt{\frac{\log(n)}{nh^3}} + \delta_n\right).\end{aligned}$$

Then, since $Z^*_{\mathfrak{M}_{\pm}^{[1]},\gamma}$ is independent of the data and $Z^*_{\mathfrak{M}_{\pm}^{[1]},\gamma} =_d \left\|G^U\right\|_{\mathfrak{M}^{[1]}}$, by the above deviation bound and Chernozhukov et al. (2016, Lemma 2.1), with probability greater than $1 - C_3\left(\sqrt{\gamma} + \sqrt{\log(n)/(nh^3)} + \delta_n\right)$,

$$\begin{aligned}&\sup_{t\in\mathbb{R}}\left|\Pr_{|W_1^n}\left[\left\|S_{\mathsf{npb}}\left(\cdot \mid x; \widehat{h}\right)\right\|_{I_x} \leq t\right] - \Pr\left[\left\|G^U\right\|_{\mathfrak{M}^{[1]}} \leq t\right]\right| \leq \\ &\sup_{t\in\mathbb{R}}\Pr\left[\left|\left\|G^U\right\|_{\mathfrak{M}^{[1]}} - t\right| \leq C_1\left(\frac{\log(n)^{2/3}}{\gamma^{1/3}(nh^3)^{1/6}} + \frac{\log(n)^{3/4}}{\gamma(nh^3)^{1/4}} + \sqrt{\log(n)\,h} + \varepsilon_n\sqrt{\log(n)}\right)\right] + C_2\left(\sqrt{\gamma} + \left(\frac{\log(n)}{nh^3}\right)^{1/4}\right).\end{aligned}$$

By (S139) and optimally choosing $\gamma = \log(n)^{5/6}/\left(nh^3\right)^{1/6}$ (which balances $\sqrt{\gamma}$ and $\log(n)^{5/4}/\left(\gamma n^{1/4}h^{3/4}\right)$), we have (69). By repeating the arguments used to show (62), we have $\left|\Pr\left[\left\|S\left(\cdot \mid x; \widehat{h}\right)\right\|_{I_x} \leq s_{1-\alpha}^{\mathsf{npb}}\right] - (1-\alpha)\right| \leq C_1\kappa_{1,n} + C_2\kappa^*_{2,n} + C_3\kappa^*_{3,n}$. ■

S5 Additional Monte Carlo simulation results

Table S1 presents the coverage rates of two types of pointwise confidence intervals for the density $f_{\Delta}(v)$ evaluated at 1.6, 2.0 and 2.4. The method "PA" corresponds to the plug-in approach using our new standard errors and standard normal critical values (see (24)). The method "PB" corresponds to the bootstrap percentile confidence interval. The simulation design and the choice of tuning parameters are all the same as Section 5. The number of Monte Carlo replications is 1,000 and the nominal probability coverages rates are 0.90, 0.95 and 0.99. As Table S1 shows, PA produces coverage rates that are very close to the nominal levels, especially when the sample size

Table S1: Coverage rates for point-wise confidence intervals

v	n	Methods	$\gamma_0 = -0.5, \gamma_1 = 0.5$ 0.90	0.95	0.99	$\gamma_0 = -0.4, \gamma_1 = 0.6$ 0.90	0.95	0.99
1.6	2000	PA	0.890	0.933	0.981	0.878	0.924	0.973
		PB	0.936	0.975	0.996	0.943	0.979	0.999
	4000	PA	0.901	0.951	0.980	0.900	0.946	0.978
		PB	0.916	0.961	0.994	0.931	0.970	0.995
	6000	PA	0.902	0.951	0.980	0.907	0.950	0.981
		PB	0.927	0.974	0.994	0.944	0.979	0.995
2.0	2000	PA	0.900	0.940	0.977	0.874	0.914	0.954
		PB	0.935	0.977	0.997	0.939	0.982	0.997
	4000	PA	0.897	0.945	0.984	0.895	0.945	0.975
		PB	0.934	0.975	0.996	0.950	0.980	0.997
	6000	PA	0.900	0.953	0.987	0.909	0.949	0.988
		PB	0.928	0.967	0.997	0.946	0.972	0.999
2.4	2000	PA	0.874	0.915	0.950	0.832	0.880	0.922
		PB	0.934	0.972	0.991	0.940	0.974	0.997
	4000	PA	0.887	0.931	0.976	0.872	0.925	0.962
		PB	0.944	0.970	0.997	0.945	0.978	0.998
	6000	PA	0.901	0.948	0.986	0.901	0.939	0.979
		PB	0.938	0.969	0.999	0.945	0.977	0.997

is large. On the other hand, PB, though circumvents the calculation of standard errors, exhibits certain degree of over-coverage.

We then present simulation results for non-studentized bias corrected JMB and nonparametric bootstrap UCBs. The non-studentized nonparametric bootstrap UCB is defined in Appendix C. The non-studentized JMB UCB is described in Footnote 26. Tables S2 and S3 are the non-studentized version of Tables 1 and 2. As expected, the non-studentized UCBs are on average wider than the studentized ones because the non-studentized UCBs keep the same width across different values of v while the studentized versions adjust using the estimated variance at each v. In particular, the studentized versions become narrower in the region of v with a smaller value of $f_{\Delta}(v)$.

Table S2: Simultaneous coverage rates for non-studentized UCBs

		$\gamma_0 = -0.5, \gamma_1 = 0.5$						$\gamma_0 = -0.4, \gamma_1 = 0.6$					
		$v \in [0.5, 3.5]$			$v \in [0.8, 3.2]$			$v \in [0.5, 3.5]$			$v \in [0.8, 3.2]$		
n	Methods	0.90	0.95	0.99	0.90	0.95	0.99	0.90	0.95	0.99	0.90	0.95	0.99
2000	Bias corrected JMB	0.945	0.984	1.0	0.970	0.991	1.0	0.980	0.997	1.0	0.985	0.997	1.0
	Bias corrected NPB	0.922	0.968	1.0	0.960	0.991	1.0	0.956	0.990	1.0	0.976	0.996	1.0
4000	Bias corrected JMB	0.889	0.954	0.998	0.954	0.983	1.0	0.951	0.974	1.0	0.976	0.993	1.0
	Bias corrected NPB	0.884	0.946	0.991	0.945	0.985	1.0	0.933	0.977	0.998	0.968	0.992	1.0
6000	Bias corrected JMB	0.859	0.917	0.999	0.947	0.978	1.0	0.919	0.967	1.0	0.968	0.990	1.0
	Bias corrected NPB	0.859	0.924	0.986	0.958	0.982	1.0	0.889	0.944	0.995	0.966	0.987	0.999

Table S3: Average width of the 95% non-studentized UCBs relative to the interpolated pointwise CIs

	Methods	$\gamma_0 = -0.5, \gamma_1 = 0.5$		$\gamma_0 = -0.4, \gamma_1 = 0.6$	
		$v \in [0.5, 3.5]$	$v \in [0.8, 3.2]$	$v \in [0.5, 3.5]$	$v \in [0.8, 3.2]$
2000	Bias corrected JMB	1.802	1.722	1.962	1.885
	Bias corrected NPB	1.716	1.636	1.790	1.700
4000	Bias corrected JMB	1.729	1.640	1.807	1.723
	Bias corrected NPB	1.703	1.621	1.747	1.661
6000	Bias corrected JMB	1.897	1.794	1.765	1.680
	Bias corrected NPB	1.878	1.781	1.731	1.646

References

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics 113*(2), 231–263.

Chernozhukov, V., D. Chetverikov, and K. Kato (2014a). Anti-concentration and honest, adaptive confidence bands. *Annals of Statistics 42*(5), 1787–1818.

Chernozhukov, V., D. Chetverikov, and K. Kato (2014b). Gaussian approximation of suprema of empirical processes. *Annals of Statistics 42*(4), 1564 1597.

Chernozhukov, V., D. Chetverikov, and K. Kato (2016). Empirical and multiplier bootstraps for suprema of empirical processes of increasing complexity, and related gaussian couplings. *Stochastic Processes and their Applications 126*(12), 3632–3651.

Giné, E. and R. Nickl (2016). *Mathematical foundations of infinite-dimensional statistical models*, Volume 40. Cambridge University Press.

Kosorok, M. R. (2007). *Introduction to empirical processes and semiparametric inference.* Springer Science & Business Media.

Serfling, R. J. (2009). *Approximation theorems of mathematical statistics.* John Wiley & Sons.