APPENDIX: AUXILIARY RESULTS AND PROOFS

A bivariate kernel function $\kappa_2(\cdot,\cdot)$ is said to be of order (ν,ℓ) with $\nu=(\nu_1,\nu_2)$ if it satisfies

$$\int u^{\ell_1}v^{\ell_2}\kappa_2(u,v)dudv=\begin{cases}0 & 0\le \ell_1+\ell_2<\ell, \ell_1\ne\nu_1, \ell_2\ne\nu_2\\ \nu! & \ell_1=\nu_1, \ell_2=\nu_2\\ \ne 0 & \ell_1+\ell_2=\ell\end{cases} \tag{A.1}$$

and

$$\int \left|u^{\ell_1}v^{\ell_2}\kappa_2(u,v)\right|dudv<\infty,\ \text{for any } \ell_1+\ell_2=\ell, \tag{A.2}$$

where $\nu! = \nu_1!\cdot\nu_2!$. Similarly, a univariate kernel function $\kappa_1(\cdot)$ is of order (ν,ℓ) for a univariate $\nu=\nu_1$ when (A.1) and (A.2) hold for $\ell_2\equiv 0$ on the right hand side while integrating over the univariate argument u on the left.

We enforce the following technical conditions.

(i) The variable S has compact domain $\mathcal{S}$. Given $Z=z$, S has conditional density $f_{S,z}(s)$. Assume, uniformly in $z\in\mathcal{Z}$, that $\frac{\partial^\ell}{\partial s^\ell}f_{S,z}(s)$ exists and is continuous for $\ell=2$ on $\mathcal{S}$, and further $\inf_{s\in\mathcal{S}} f_{S,z}(s)>0$, analogously for T.

(ii) Denote the conditional density functions of (S,U) and (T,V) by $g_{X,z}(s,u)$ and $g_{Y,z}(t,v)$, respectively. Assume that the derivative $\dfrac{\partial^\ell}{\partial s^\ell}g_{X,z}(s,u)$ exists for all arguments (s,u), is uniformly continuous on $\mathcal{S}\times\mathbb{R}$, and is Lipschitz continuous in z, for $\ell=2$, analogously for $g_{Y,z}(t,v)$.

(iii) Denote the conditional density functions of quadruples (S_1,S_2,U_1,U_2) and (T_1,T_2,V_1,V_2) by $g_{2X,z}(s_1,s_2,u_1,u_2)$ and $g_{2Y,z}(t_1,t_2,v_1,v_2)$; for simplicity, the corresponding marginal conditional densities of (S_1,S_2) and (T_1,T_2) are also denoted by $g_{2X,z}(s_1,s_2)$ and $g_{2Y,z}(t_1,t_2)$, respectively. Denote the conditional density of (S,T,U,V) given $Z=z$ by $g_{XY,z}(s,t,u,v)$; and similarly its corresponding conditional marginal density of (S,T) by $g_{XY,z}(s,t)$. Assume that the derivatives $\dfrac{\partial^\ell}{\partial s_1^{\ell_1}\partial s_2^{\ell_2}}g_{2X,z}(s_1,s_2,u_1,u_2)$ exist for all arguments (s_1,s_2,u_1,u_2), are uniformly continuous on $\mathcal{S}^2\times\mathbb{R}^2$, and are Lipschitz continuous in z for $\ell_1+\ell_2=\ell$, $0\le\ell_1,\ell_2\le\ell=2$, analogously for $g_{2Y,z}(t_1,t_2,v_1,v_2)$ and $g_{XY,z}(s,t,u,v)$.

(iv) For every $p=1,2,\cdots,P$, $b_{X,z^{(p)}}\to 0$, $n_{z^{(p)},h}b^4_{X,z^{(p)}}\to\infty$, $n_{z^{(p)},h}b^6_{X,z^{(p)}}<\infty$, $b_{Y,z^{(p)}}\to 0$, $n_{z^{(p)},h}b^4_{Y,z^{(p)}}\to\infty$, and $n_{z^{(p)},h}b^6_{Y,z^{(p)}}<\infty$, as $n\to\infty$.

(v) For every $p = 1, 2, \cdots, P$, $h_{X,z^{(p)}} \to 0$, $n_{z^{(p)},h} h^6_{X,z^{(p)}} \to \infty$, $n_{z^{(p)},h} h^8_{X,z^{(p)}} < \infty$, $h_{Y,z^{(p)}} \to 0$, $n_{z^{(p)},h} h^6_{Y,z^{(p)}} \to \infty$, and $n_{z^{(p)},h} h^8_{Y,z^{(p)}} < \infty$, as $n \to \infty$.

(vi) For every $p = 1, 2, \cdots, P$, $h_{1,z^{(p)}}/h_{2,z^{(p)}} \to 1$, $h_{1,z^{(p)}} \to 0$, $n_{z^{(p)},h} h^6_{1,z^{(p)}} \to \infty$, and $n_{z^{(p)},h} h^8_{1,z^{(p)}} < \infty$ as $n \to \infty$.

(vii) For every $p = 1, 2, \cdots, P$, $b_{X,z^{(p)},V} \to 0$, $n_{z^{(p)},h} b^4_{X,z^{(p)},V} \to \infty$, $n_{z^{(p)},h} b^6_{X,z^{(p)},V} < \infty$, $b_{Y,z^{(p)},V} \to 0$, $n_{z^{(p)},h} b^4_{Y,z^{(p)},V} \to \infty$, and $n_{z^{(p)},h} b^6_{Y,z^{(p)},V} < \infty$, as $n \to \infty$.

(viii) Univariate kernel κ_1 and bivariate κ_2 are compactly supported, absolutely integrable, and of order $(\nu, \ell) = (0, 2)$ and $((0, 0), 2)$, respectively.

(ix) Assume $\sup_{(z,s)\in\mathcal{Z}\times\mathcal{S}} E(E(X(s) - \mu_{X,Z}(s))^4 | Z = z)) < \infty$ and analogously for Y.

(x) The slope function $\beta(z, s, t)$ is twice differentiable in z, *i.e.*, for any $(s, t) \in \mathcal{S} \times \mathcal{T}$, $\frac{\partial^2}{\partial z^2}\beta(z, s, t)$ exists and is continuous in z.

(xi) The bin width h and smoothing bandwidth b satisfy that $b/h < \infty$ as $n \to \infty$. The bin width h is chosen such that $P \propto n^{1/8}$.

Proposition 1. *For E_n defined in (3.1), under (xi) it holds that $P(E_n) \to 1$ as $n \to \infty$.*

Proof of Proposition 1. Note first that $P(\min n_{z^{(p)},h} > \tilde{n}) \geq 1 - \sum_{p=1}^{P} P(n_{z^{(p)},h} < \tilde{n})$. Consider the pth bin, denote $\pi_p = P(Z \in [z^{(p)} - \frac{h}{2}, z^{(p)} - \frac{h}{2}))$. Then $n_{z^{(p)},h}$ is asymptotically distributed as $N(n\pi_p, n\pi_p(1-\pi_p))$ due to the normal approximation to a binomial random variable. Thus $P(n_{z^{(p)},h} > \tilde{n}) \to f_{N(0,1)}(a_p)/a_p$ with $a_p = -(\tilde{n} - n\pi_p)/\sqrt{n\pi_p(1-\pi_p)}$, where $f_{N(0,1)}(\cdot)$ is the probability density function of the standard normal distribution. Due to [A1], π_p is bounded between $\underline{f_Z}/(\underline{f_Z} + (P-1)\bar{f}_Z)$ and $\bar{f}_Z/((P-1)\underline{f_Z} + \bar{f}_Z)$. It follows that $P(E_n) \to 1$ as $n \to \infty$ by noting that $\tilde{n} \propto \sqrt{n}$, $P \propto n^{1/8}$, and $f_{N(0,1)}(x)/x$ decays exponentially in x. □

We next prove the consistency of raw estimate of the mean functions of predictor and response trajectories within each bin. Consider a generic bin $[z - h/2, z + h/2)$, with bin center z and bandwidth h and let $b_{X,z}$ and $b_{Y,z}$ be smoothing bandwidths used to estimate $\mu_{X,z}(s)$ and $\mu_{Y,z}(t)$, $h_{X,z}$ and $h_{Y,z}$ for $G_{X,z}(s_1, s_2)$ and $G_{Y,z}(t_1, t_2)$, $h_{1,z}$ and $h_{2,z}$ for $C_{XY,z}(s, t)$, $b_{X,z,V}$ and $b_{Y,z,V}$ for $V_{X,z}(s) = G_{X,z}(s, s) + \sigma^2_X$ and $V_{Y,z}(t) = G_{Y,z}(t, t) + \sigma^2_Y$.

For a positive integer $l \geq 1$, let $\{\psi_p(t, v), p = 1, 2, \cdots, l\}$ be a collection of real functions $\psi_p : \mathbb{R}^2 \to \mathbb{R}$ satisfying the following conditions:

[C1.1a] The derivative functions $\dfrac{\partial^\ell}{\partial t^\ell}\psi_p(t, v)$ exist for all arguments (t, v) and are uniformly continuous on $\mathcal{T} \times I\!R$.

[C1.2a] Assume that $\int\int \psi_p^2(t,v) g_{Y,z}(t,v)dvdt < \infty$.

[C2.1a] Uniformly in $z \in \mathcal{Z}$, bandwidths $b_{Y,z}$ for one dimensional smoothers satisfy that $b_{Y,z} \to 0$, $n_{z,h} b_{Y,z}^{\nu+1} \to \infty$, and $n_{z,h} b_{Y,z}^{2\ell+2} < \infty$, as $n \to \infty$.

Define $\mu_{p\psi,z} = \mu_{p\psi,z}(t) = \frac{d^\nu}{dt^\nu} \int \psi_p(t,v) g_{Y,z}(t,v)dv$ and

$$\Psi_{pn,z} = \Psi_{pn,z}(t) = \frac{1}{n_{z,h} b_{Y,z}^{\nu+1}} \sum_{i \in \mathcal{N}_{z,h}} \frac{1}{EN} \sum_{j=1}^{N_i} \psi_p(T_{ij}, V_{ij}) \kappa_1(\frac{T_{ij}-t}{b_{Y,z}}),$$

where $g_{Y,z}(t,v)$ is the conditional density of (T,V), given $Z=z$.

Lemma 1. *Under Conditions [A0-A3], (i), (ii), (viii), [C1.1a], [C1.2a], and [C2.1a], we have* $\tau_{pn} = \sup_{(z,t)\in\mathcal{Z}\times\mathcal{T}} \mid \Psi_{pn,z}(t) - \mu_{p\psi,z}(t) \mid / (h + (\sqrt{n_{z,h}} b_{Y,z}^{\nu+1})^{-1}) = O_p(1)$.

Proof. Note that $\mid \Psi_{pn,z}(t) - \mu_{p\psi,z}(t) \mid \leq \mid \Psi_{pn,z}(t) - E\Psi_{pn,z}(t) \mid + \mid E\Psi_{pn,z}(t) - \mu_{p\psi,z}(t) \mid$ and $E \mid \tau_{pn} \mid = O(1)$ implies $\tau_{pn} = O_p(1)$. Standard conditioning techniques lead to

$$E\Psi_{pn,z}(t) = \frac{1}{b_{Y,z}^{\nu+1}} E(E(\psi_p(T_{i1}, V_{i1})\kappa_1(\frac{T_{i1}-t}{b_{Y,z}}) \mid z - \frac{h}{2} \leq Z_i < \frac{h}{2})).$$

For $Z_i = z_i \in [z-h/2, z+h/2)$, perform a Taylor expansion of order ℓ on the integrand,

$$\begin{aligned} E[\psi_p(T_{i1}, V_{i1})\kappa_1(\frac{T_{i1}-t}{b_{Y,z}})] &= \int\int \psi_p(t_1,v_1) g_{Y,z_i}(t_1,v_1) \kappa_1(\frac{t_1-t}{b_{Y,z}}) dt_1 dv_1 \\ = \int\int \left(\frac{\partial^\nu}{\partial t^\nu}(\psi_p(t,v_1) g_{Y,z_i}(t,v_1))\right) &\frac{(t_1-t)^\nu}{\nu!} \kappa_1(\frac{t_1-t}{b_{Y,z}}) dt_1 dv_1 \\ &+ \int\int \left(\frac{\partial^\ell}{\partial t^\ell}(\psi_p(t,v_1) g_{Y,z_i}(t,v_1))\right) \mid_{t=t^*} \frac{(t_1-t)^\ell}{\ell!} \kappa_1(\frac{t_1-t}{b_{Y,z}}) dt_1 dv_1, \end{aligned}$$

where t^* is between t and t_1. Hence, $\left| E[\psi_p(T_{i1}, V_{i1})\kappa_1(\frac{T_{i1}-t}{b_{Y,z}})] - \mu_{p\psi,z_i}(t) b_{Y,z}^{\nu+1} \right| \leq c_0 \frac{b_{Y,z}^{\ell+1}}{\ell!} \int \left| u^\ell \kappa_1(u) \right| du$ due to [C1.2a] and the assumption that the kernel function $\kappa_1(\cdot)$ is of type (ν,ℓ), where c_0 is bounded according to [C1.1a], $c_0 \leq \sup_{(z_i,t)\in\mathcal{Z}\times\mathcal{T}} \left| \frac{\partial^\ell}{\partial t^\ell} \int \psi_p(t,v_1) g_{Y,z_i}(t,v_1) dv_1 \right| < \infty$. Furthermore, using (ii), we may bound

$$\begin{aligned} &\sup_{t\in\mathcal{T}} \mid E\Psi_{pn,z}(t) - \mu_{p\psi,z}(t) \mid \\ \leq\ & c_0 b_{Y,z}^{\ell-\nu}/(\ell!) \int \left| u^\ell \kappa_1(u) \right| du + E\left\{ E\left[\sup_{t\in\mathcal{T}} |\mu_{p\psi,Z_i}(t) - \mu_{p\psi,z}(t)| \mid z - \frac{h}{2} \leq Z_i < \frac{h}{2} \right] \right\} \\ \leq\ & c_0 (\int \left| u^\ell \kappa_1(u) \right| du) b_{Y,z}^{\ell-\nu}/(\ell!) + c_1 h, \end{aligned} \tag{A.3}$$

where the constants do not depend on z. To bound $E\sup_{t\in\mathcal{T}} \mid \Psi_{pn,z}(t) - E\Psi_{pn,z}(t) \mid$, we denote Fourier transform of $\kappa_1(\cdot)$ by $\zeta_1(t) = \int e^{-iut}\kappa_1(u)du$ and letting $\varphi_{pn,z}(u) = \frac{1}{n_{z,h}}\sum_{m\in\mathcal{N}_{z,h}}\frac{1}{EN}\sum_{j=1}^{N_m} e^{iuT_{mj}}\psi_p(T_{mj}, Y_{mj})$,

$$\Psi_{pn,z} = \frac{1}{n_{z,h}b_{Y,z}^{\nu+1}}\sum_{m\in\mathcal{N}_{z,h}}\frac{1}{EN}\sum_{j=1}^{N_m}\kappa_1(\frac{T_{mj}-t}{b_{Y,z}})\psi_p(T_{mj}, Y_{mj}) = \frac{1}{2\pi b_{Y,z}^{\nu}}\int\varphi_{pn,z}(u)e^{-itu}\zeta_1(ub_{Y,z})du,$$

and $\sup_{t\in\mathcal{T}} \mid \Psi_{pn,z}(t) - E\Psi_{pn,z}(t) \mid\leq \frac{1}{2\pi b_{Y,z}^{\nu}}\int \mid \varphi_{pn,z}(u) - E\varphi_{pn,z}(u) \mid \cdot \mid \zeta_1(ub_{Y,z}) \mid du$.

Decomposing $\varphi_{pn,z}(\cdot)$ into real and imaginary parts,

$$\begin{aligned}\varphi_{pn,z,R}(u) &= \frac{1}{n_{z,h}}\sum_{m\in\mathcal{N}_{z,h}}\frac{1}{EN}\sum_{j=1}^{N_m}\cos(uT_{mj})\psi_p(T_{mj}, Y_{mj})\\ \varphi_{pn,z,I}(u) &= \frac{1}{n_{z,h}}\sum_{m\in\mathcal{N}_{z,h}}\frac{1}{EN}\sum_{j=1}^{N_m}\sin(uT_{mj})\psi_p(T_{mj}, Y_{mj}),\end{aligned}$$

we obtain $E \mid \varphi_{pn,z}(u) - E\varphi_{pn,z}(u) \mid= E \mid \varphi_{pn,z,R}(u) - E\varphi_{pn,z,R}(u) \mid +E \mid \varphi_{pn,z,I}(u) - E\varphi_{pn,z,I}(u) \mid$. Note the inequality $E \mid \varphi_{pn,z,R}(u) - E\varphi_{pn,z,R}(u) \mid\leq \sqrt{E \mid \varphi_{pn,z,R}(u) - E\varphi_{pn,z,R}(u) \mid^2}$ and the fact that$\{[Z_i, N_i, (T_{ij}, Y_{ij})_{j=1}^{N_i}] : i \in \mathcal{N}_{z,h}\}$ are $i.i.d.$ implies that

$$\mathrm{var}(\varphi_{pn,z,R}(u)) \leq \frac{1}{n_{z,h}}E\{E(\psi_p^2(T_{m1}, Y_{m1})|z - h/2 \leq Z_m < z + h/2)\}$$

where $m \in N_{z,h}$, analogously for the imaginary part. As a result, we have

$$E\sup_{t\in\mathcal{T}} \mid \Psi_{pn,z}(t) - E\Psi_{pn,z}(t) \mid \leq \frac{2\sqrt{E\{E(\psi_p^2(T_{m1}, Y_{m1})|z - h/2 \leq Z_m < z + h/2)\}}\int \mid \zeta_1(u) \mid du}{2\pi\sqrt{n_{z,h}}b_{Y,z}^{\nu+1}}.$$

Note that $E(\psi_p^2(T_{m1}, Y_{m1}))$ as a function of Z_m is continuous over the compact domain $\mathcal{Z}$ and as a result bounded. Denote $c_2 = 2\sup_{Z_m\in\mathcal{Z}}\sqrt{E(\psi_p^2(T_{m1}, Y_{m1}))} < \infty$. Hence we have

$$E\sup_{t\in\mathcal{T}} \mid \Psi_{pn,z}(t) - E\Psi_{pn,z}(t) \mid\leq \frac{c_2\int|\zeta_1(u)|\,du}{2\pi}(\sqrt{n_{z,h}}b_{Y,z}^{\nu+1})^{-1}, \tag{A.4}$$

where the constant $c_2(\int|\zeta_1(u)|\,du)/(2\pi)$ does not depend on z.

The result follows as Condition [A1] implies that $n_{z,h}$ goes to infinity uniformly for $z \in \mathcal{Z}$ as $n \to \infty$, $n_{z,h}b_{Y,z}^{2\ell+2} < \infty$ implies that $b_{Y,z}^{\ell-\nu} = O(1/(\sqrt{n_{z,h}}b_{Y,z}^{\nu+1}))$. We next extend Theorem 1 in Yao et al. (2005a) under some additional conditions. □

[C3] Uniformly in $z \in \mathcal{Z}$, $b_{X,z} \to 0$, $n_{z,h}b_{X,z}^4 \to \infty$, $n_{z,h}b_{X,z}^6 < \infty$, $b_{Y,z} \to 0$, $n_{z,h}b_{Y,z}^4 \to \infty$, and $n_{z,h}b_{Y,z}^6 < \infty$, as $n \to \infty$.

Lemma 2. *Under Conditions [A0-A3], (i), (ii), (viii), (ix), and [C3], we have*

$$\sup_{(z,s)\in\mathcal{Z}\times\mathcal{S}} \frac{|\ \tilde{\mu}_{X,z}(s)-\mu_{X,z}(s)\ |}{h+(\sqrt{n_{z,h}}b_{X,z})^{-1}} = O_p(1) \text{ and } \sup_{(z,t)\in\mathcal{Z}\times\mathcal{T}} \frac{|\ \tilde{\mu}_{Y,z}(t)-\mu_{Y,z}(t)\ |}{h+(\sqrt{n_{z,h}}b_{Y,z})^{-1}} = O_p(1) \quad (A.5)$$

Proof. Without loss of generality, we consider the estimation of $\mu_{Y,z}(t)$. The local linear estimator of $\tilde{\mu}_{Y,z}(t)$ is obtained by minimizing $\sum_{i\in\mathcal{N}_{z,h}} \frac{1}{EN}\sum_{j=1}^{N_i}\kappa_1(\frac{T_{ij}-t}{b_{Y,z}})(V_{ij}-\beta_{z,0}-\beta_{z,1}(T_{ij}-t))^2$ with respect to $\beta_{z,0}$ and $\beta_{z,1}$. For a kernel function $\kappa_1(\cdot)$ of order $(0,2)$, let $w_{ij}=\kappa_1((T_{ij}-t)/b_{Y,z})/(n_{z,h}b_{Y,z})$, $A_{z,h,m}=\sum_{i\in\mathcal{N}_{z,h}}\frac{1}{EN}\sum_{j=1}^{N_i}w_{ij}(T_{ij}-t)^m$, and $B_{z,h,m}=\sum_{i\in\mathcal{N}_{z,h}}\frac{1}{EN}\sum_{j=1}^{N_i}w_{ij}(T_{ij}-t)^mV_{ij}$ for $m=0,1,2$. Then $\tilde{\beta}_{z,0}(t)=B_{z,h,0}/A_{z,h,0}-\tilde{\beta}_{z,1}A_{z,h,1}/A_{z,h,0}$, where an estimator of the derivative $\mu'_{Y,z}(t)$ of $\mu_{Y,z}(t)$ as a function of t is given by $\tilde{\beta}_{z,1}(t)=(-A_{z,h,1}B_{z,h,0}+A_{z,h,0}B_{z,h,1})/(A_{z,h,0}A_{z,h,2}-A^2_{z,h,1})$.

Consider the Nadaraya-Watson estimator of $\mu_{Y,z}$ given by $\tilde{\mu}_{Y,z,NW}(t)=\tilde{\beta}_{z,0}=B_{z,h,0}/A_{z,h,0}$. One estimator of the density of T conditional on $Z=z$ is given by $\tilde{f}_{T,z}(t)=\sum_{i\in\mathcal{N}_{z,h}}\sum_{j=1}^{N_i}w_{ij}/EN=A_{z,h,0}$. Setting $\nu=0,\ell=2,l=2$, $\psi_1(t,y)=y$, and $\psi_2(t,y)\equiv 1$ in Lemma 1, we can easily see that $\tilde{\mu}_{Y,z,NW}(t)=H(\Psi_{1n,z},\Psi_{2n,z})$ with $H(a,b)=a/b$ and $\tilde{f}_{T,z}(t)=\Psi_{2n,z}$. Using the uniform version of *Slutsky*'s Theorem and Lemma 1, we have that

$$\sup_{(z,t)\in\mathcal{Z}\times\mathcal{T}}\frac{|\ \tilde{\mu}_{Y,z,NW}(t)-\mu_{Y,z}(t)\ |}{(h+(\sqrt{n_{z,h}}b_{Y,z})^{-1})}=O_p(1) \text{ and } \sup_{(z,t)\in\mathcal{Z}\times\mathcal{T}}\frac{|\ \tilde{f}_{T,z}(t)-f_{T,z}(t)\ |}{(h+(\sqrt{n_{z,h}}b_{Y,z})^{-1})}=O_p(1). \quad (A.6)$$

For the uniform consistency of $\tilde{\beta}_{z,1}$ estimating $\mu'_{Y,z}(t)$, let $\sigma^2_{\kappa_1}=\int u\kappa_1(u)du$ and use the kernel function $\tilde{\kappa}_1(t)=t\kappa_1(t)/\sigma^2_{\kappa_1}$ which is of order $(\nu,\ell)=(1,3)$. Define new $\Psi_{pn,z}$, $p=1,2,3$ with $\psi_1(u,y)=y$, $\psi_2(u,y)=1$, $\psi_3(u,y)=u-t$. Hence,

$$\begin{aligned}\Psi_{1n,z}(t) &= \frac{1}{n_{z,h}b^{\nu+1}_{Y,z}}\sum_{i\in\mathcal{N}_{z,h}}\frac{1}{EN}\sum_{j=1}^{N_i}\psi_1(T_{ij},V_{ij})\tilde{\kappa}_1(\frac{T_{ij}-t}{b_{Y,z}})\\ &= \frac{1}{n_{z,h}b^{\nu+1}_{Y,z}}\sum_{i\in\mathcal{N}_{z,h}}\frac{1}{EN}\sum_{j=1}^{N_i}V_{ij}\frac{T_{ij}-t}{b_{Y,z}}\kappa_1(\frac{T_{ij}-t}{b_{Y,z}})/\sigma^2_{\kappa_1}=B_{z,h,1}/(b^{\nu+1}_{Y,z}\sigma^2_{\kappa_1}).\end{aligned}$$

Similarly, $\Psi_{2n,z}(t)=A_{z,h,1}/(b^{\nu+1}_{Y,z}\sigma^2_{\kappa_1})$ and $\Psi_{3n,z}(t)=A_{z,h,2}/(b^{\nu+1}_{Y,z}\sigma^2_{\kappa_1})$. Correspondingly, $\mu_{1\psi,z}(t)=\mu'_{Y,z}(t)f_{T,z}(t)+\mu_{Y,z}(t)f'_{T,z}(t)$, $\mu_{2\psi,z}(t)=f'_{T,z}(t)$, and $\mu_{3\psi,z}(t)=f_{T,z}(t)$.

Observe that $\sup_{t\in\mathcal{T}}|\ \tilde{f}_{T,z}(t)-f_{T,z}(t)\ |=O_p(h+1/(\sqrt{n_{z,h}}b_{Y,z}))$. Define

$$\tilde{H}(x_1,x_2,x_3) = \frac{x_1-x_2\tilde{\mu}_{Y,z,NW}(t)}{x_3-b^2_{Y,z}\sigma^2_{\kappa_1}x_2^2/\tilde{f}_{T,z}(t)} \text{ and } H(x_1,x_2,x_3)=\frac{x_1-x_2\mu_{Y,z}(t)}{x_3}.$$

Then $H(\mu_{1\psi,z}(t), \mu_{2\psi,z}(t), \mu_{3\psi,z}(t)) = \mu'_{Y,z}(t)$ and

$$\begin{aligned}\tilde{\beta}_{z,1} &= \tilde{H}(\Psi_{1n,z}, \Psi_{2n,z}, \Psi_{3n,z}) = \Psi_{3n,z} H(\Psi_{1n,z}, \Psi_{2n,z}, \Psi_{3n,z})/(\Psi_{3n,z} + b^2_{Y,z}\sigma^2_{\kappa_1}\Psi^2_{2n,z}/\tilde{f}_{T,z}(t)) \\ &\quad + \Psi_{2n,z}(\mu_{Y|Z=z}(t) - \tilde{\mu}_{Y,z,NW}(t))/(\Psi_{3n,z} + b^2_{Y,z}\sigma^2_{\kappa_1}\Psi^2_{2n,z}/\tilde{f}_{T,z}(t)). \quad \text{(A.7)}\end{aligned}$$

So using Lemma 1 and the uniform version of *Slutsky*'s Theorem, we have

$$\sup_{(z,t)\in\mathcal{Z}\times\mathcal{T}} \frac{|\ \Psi_{pn,z} - \mu_{p\psi,z}\ |}{h + (\sqrt{n_{z,h}} b^2_{Y,z})^{-1}} = O_p(1) \text{ and } \sup_{(z,t)\in\mathcal{Z}\times\mathcal{T}} \frac{|\ H(\Psi_{1n,z}, \Psi_{2n,z}, \Psi_{3n,z}) - \mu'_{Y|Z=z}(t)\ |}{h + 1/(\sqrt{n_{z,h}} b^2_{Y,z})} = O_p(1). \quad \text{(A.8)}$$

Considering the uniform convergence of $\tilde{\beta}_{z,0}$ for $\mu_{Y,z}$, noticing that $\tilde{\beta}_{z,0}(t) = \tilde{\mu}_{Y,z,NW}(t) + b^2_{Y,z}\Psi_{2n,z}\tilde{\beta}_{z,1}/\tilde{f}_{T,z}$, and by (A.6), (A.7), and (A.8), we have $\sup_{(z,t)\in\mathcal{Z}\times\mathcal{T}} |\ \Psi_{2n,z}\tilde{\beta}_{z,1}(t)/\tilde{f}_{T,z}(t)\ | = O_p(1)$. The result follows from $n_{z,h} b^6_{Y,z} < \infty$ and (A.6). □

Our next two lemmas concern the consistency for estimating the covariance functions, based on the observations in the generic bin $[z - h/2, z + h/2)$. Let $\{\theta_p(r_1, r_2, v_1, v_2), p = 1, 2, \cdots, l\}$ be a collection of real functions $\theta_p : \mathbb{R}^4 \to \mathbb{R}$ with the following properties:

[C1.1b] The derivatives $\frac{\partial^\ell}{\partial r_1^{\ell_1} \partial r_2^{\ell_2}} \theta_p(r_1, r_2, v_1, v_2)$ exist for all arguments (r_1, r_2, v_1, v_2) and are uniformly continuous on $\mathcal{R}_1 \times \mathcal{R}_2 \times \mathbb{R}^2$, for $\ell_1 + \ell_2 = \ell$, $0 \le \ell_1, \ell_2 \le \ell$, $\ell = 0, 1, 2$.

[C1.2b] The expectation $\int\int\int\int \theta^2_p(r_1, r_2, v_1, v_2) g(r_1, r_2, v_1, v_2) dr_1 dr_2 dv_1 dv_2$ exists and is finite, uniformly bounded on $\mathcal{Z}$.

[C2.1b] Uniformly in $z \in \mathcal{Z}$, bandwidths $h_{Y,z}$ for the two-dimensional smoother satisfy $h_{Y,z} \to 0$, $n_{z,h} h^{|\nu|+2}_{Y,z} \to \infty$, $n_{z,h} h^{2\ell+4}_{Y,z} < \infty$, as $n \to \infty$.

Define $\varrho_{p\theta,z} = \varrho_{p\theta,z}(t_1, t_2) = \frac{\partial^{|\nu|}}{\partial t_1^{\nu_1} \partial t_2^{\nu_2}} \int\int \theta_p(t_1, t_2, v_1, v_2) g_{2Y,z}(t_1, t_2, v_1, v_2) dv_1 dv_2$ and

$$\Theta_{pn,z}(t_1, t_2) = \frac{1}{n_{z,h} h^{|\nu|+2}_{Y,z}} \sum_{i\in\mathcal{N}_{z,h}} \frac{1}{EN(EN-1)} \sum_{1\le j\ne k\le N_i} \theta_p(T_{ij}, T_{ik}, V_{ij}, V_{ik}) \kappa_2(\frac{T_{ij} - t_1}{h_{Y,z}}, \frac{T_{ik} - t_2}{h_{Y,z}}).$$

Lemma 3. *Under Conditions [A0-A3], (i), (ii), (iii), (viii), [C1.1b] with $\mathcal{R}_1 = \mathcal{T}$ and $\mathcal{R}_2 = \mathcal{T}$, [C1.2b] with $g(\cdot,\cdot,\cdot,\cdot) = g_{2Y,z}(\cdot,\cdot,\cdot,\cdot)$, and [C2.1b], we have*

$$\vartheta_{pn} = \sup_{(z,t_1,t_2)\in\mathcal{Z}\times\mathcal{T}\times\mathcal{T}} \frac{|\ \Theta_{pn,z} - \varrho_{p\theta,z}\ |}{h + (\sqrt{n_{z,h}} h^{|\nu|+2}_{Y,z})^{-1}} = O_p(1).$$

Proof. As in Lemma 1, use $\mid \Theta_{pn,z} - \varrho_{p\theta,z} \mid \leq \mid \Theta_{pn,z} - E\Theta_{pn,z} \mid + \mid E\Theta_{pn,z} - \varrho_{p\theta,z} \mid$, where

$$\begin{aligned} & E\Theta_{pn,z}(t_1,t_2) \\ = \; & E\frac{1}{n_{z,h}h_{Y,z}^{|\nu|+2}}\sum_{i\in\mathcal{N}_{z,h}}\frac{1}{EN(EN-1)}\sum_{1\leq j\neq k\leq N_i}\theta_p(T_{ij},T_{ik},V_{ij},V_{ik})\kappa_2(\frac{T_{ij}-t}{h_{Y,z}},\frac{T_{ik}-t}{h_{Y,z}}) \\ = \; & \frac{1}{h_{Y,z}^{|\nu|+2}}E\left\{E\left[\theta_p(T_{i1},T_{i2},V_{i1},V_{i2})\kappa_2(\frac{T_{i1}-t}{h_{Y,z}},\frac{T_{i2}-t}{h_{Y,z}}) \mid z-h/2\leq Z_i\leq z+h/2\right]\right\}. \end{aligned}$$

Conditional on $Z_i = z_i \in [z-h/2, z+h/2)$, a Taylor expansion of order ℓ yields

$$\begin{aligned} & E\theta_p(T_{i1},T_{i2},V_{i1},V_{i2})\kappa_2(\frac{T_{i1}-t_1}{h_{Y,z}},\frac{T_{i2}-t_2}{h_{Y,z}}) \\ = & \int\!\!\int\!\!\int\!\!\int\Bigg[(\sum_{m=0}^{\ell-1}\frac{1}{m!}((\tilde t_1-t_1)\frac{\partial}{\partial\tilde t_1}+(\tilde t_2-t_2)\frac{\partial}{\partial\tilde t_2})^m)\theta_p(\tilde t_1,\tilde t_2,v_1,v_2)g_{2Y,z_i}(\tilde t_1,\tilde t_2,v_1,v_2) \\ & +\frac{1}{\ell!}\sum_{j=0}^{\ell}\binom{\ell}{j}((\tilde t_1-t_1)\frac{\partial}{\partial\tilde t_1})^j((\tilde t_2-t_2)\frac{\partial}{\partial\tilde t_2})^{\ell-j}\theta_p(\tilde t_1,\tilde t_2,v_1,v_2)g_{2Y,z_i}(\tilde t_1,\tilde t_2,v_1,v_2)\mid_{\tilde t_1=t_{1j}^*,\tilde t_2=t_{2j}^*}\Bigg] \\ & \times\kappa_2(\frac{\tilde t_1-t_1}{h_{Y,z}},\frac{\tilde t_2-t_2}{h_{Y,z}})dv_1dv_2d\tilde t_1d\tilde t_2, \end{aligned}$$

where t_{ij}^* is between t and t_i for $i=1,2$ and $j=0,1,\cdots,\ell$. The assumption that $\kappa_2(\cdot,\cdot)$ is of order (ν,ℓ) implies that

$$\begin{aligned} & \sup_{z\in\mathcal{Z}}\sup_{(t_1,t_2)\in\mathcal{T}\times\mathcal{T}}|\Theta_{pn,z}(t_1,t_2)-\varrho_{p\theta,z}(t_1,t_2)|\leq\sup_{z\in\mathcal{Z}}\sup_{z_i\in[z-h/2,z+h/2)}\sup_{(t_1,t_2)\in\mathcal{T}\times\mathcal{T}}|\varrho_{p\theta,z_i}(t_1,t_2)-\varrho_{p\theta,z}(t_1,t_2)| \\ & +h_{Y,z}^{\ell-|\nu|}\sum_{j=0}^{\ell}\sup_{\tilde t_1,\tilde t_2}\left|\frac{\partial^\ell}{\partial\tilde t_1^j\partial\tilde t_2^{\ell-j}}\int\!\!\int\theta_p(\tilde t_1,\tilde t_2,v_1,v_2)g_{2Y,z_i}(\tilde t_1,\tilde t_2,v_1,v_2)dv_1dv_2\right|\int\left|u^jv^{\ell-j}\kappa_2(u,v)\right|dudv, \end{aligned}$$

where Conditions [C1.1b] and (iii) guarantee that the supremum exists and is finite, and the Lipschitz condition in (iii) implies that the first term is of order $O(h)$. Hence,

$$\sup_{z\in\mathcal{Z}}\sup_{(t_1,t_2)\in\mathcal{T}\times\mathcal{T}}\frac{|\Theta_{pn,z}(t_1,t_2)-\varrho_{p\theta,z}(t_1,t_2)|}{h+h_{Y,z}^{\ell-|\nu|}}=O_p(1).$$

We can use the same technique of inserting the Fourier transformation of $\kappa_2(\cdot,\cdot)$ as in the proof for Lemma 1 to obtain $\sup_{z\in\mathcal{Z}}\sup_{(t_1,t_2)\in\mathcal{T}\times\mathcal{T}} \mid \Theta_{pn,z}(t_1,t_2) - E\Theta_{pn,z}(t_1,t_2) \mid /(\sqrt{n_{z,h}}h_{Y,z}^{|\nu|+2}) = O_p(1)$. Then the result follows by noticing that $n_{z,h}h_{Y,z}^{2\ell+4} < \infty$ as $n \to \infty$. □

[C4] Uniformly in $z \in \mathcal{Z}$, $h_{X,z} \to 0$, $n_{z,h}h_{X,z}^6 \to \infty$, $n_{z,h}h_{X,z}^8 < \infty$, $h_{Y,z} \to 0$, $n_{z,h}h_{Y,z}^6 \to \infty$, and $n_{z,h}h_{Y,z}^8 < \infty$, as $n \to \infty$.

Lemma 4. *Under Conditions [A0-A3], (i-iii), (viii), (ix), [C3], and [C4], we have*

$$\sup_{(z,s_1,s_2)\in\mathcal{Z}\times\mathcal{S}^2} \frac{\mid \tilde{G}_{X,z}(s_1,s_2) - G_{X,z}(s_1,s_2) \mid}{(h + (\sqrt{n_{z,h}}h_{X,z}^2)^{-1})} = O_p(1) \quad \text{(A.9)}$$

$$\sup_{(z,t_1,t_2)\in\mathcal{Z}\times\mathcal{T}^2} \frac{\mid \tilde{G}_{Y,z}(t_1,t_2) - G_{Y,z}(t_1,t_2) \mid}{(h + (\sqrt{n_{z,h}}h_{Y,z}^2)^{-1})} = O_p(1) \quad \text{(A.10)}$$

Proof. Without loss of generality, we consider $G_{Y,z}(t_1,t_2)$. In the local linear estimator for the covariance function $G_{Y,z}(t_1,t_2)$, we use raw observations $G_{Y,i,z}(T_{ij},T_{ik}) = (V_{ij} - \tilde{\mu}_{Y,z}(T_{ij}))(V_{ik} - \tilde{\mu}_{Y,z}(T_{ik}))$ instead of $\check{G}_{Y,i,z}(T_{ij},T_{ik}) = (V_{ij} - \mu_{Y,z}(T_{ij}))(V_{ik} - \mu_{Y,z}(T_{ik}))$, $i \in \mathcal{N}_{z,h}$. Observe

$$G_{Y,i,z}(T_{ij},T_{ik}) = \check{G}_{Y,i,z}(T_{ij},T_{ik}) + (V_{ij} - \mu_{Y,z}(T_{ij}))(\mu_{Y,z}(T_{ik}) - \tilde{\mu}_{Y,z}(T_{ik}))$$
$$+(V_{ik} - \mu_{Y,z}(T_{ik}))(\mu_{Y,z}(T_{ij}) - \tilde{\mu}_{Y,z}(T_{ij})) + (\mu_{Y,z}(T_{ij}) - \tilde{\mu}_{Y,z}(T_{ij}))(\mu_{Y,z}(T_{ik}) - \tilde{\mu}_{Y,z}(T_{ik})).$$

Since $\sup_{(z,t)\in\mathcal{Z}\times\mathcal{T}} \mid \tilde{\mu}_{Y,z}(t) - \mu_{Y,z}(t) \mid /(h + (\sqrt{n_{z,h}}b_{Y,z})^{-1}) = O_p(1)$ by (A.5) and this rate is faster than the rate we desire in (A.9) and (A.10), the local linear estimator, $\tilde{G}_{Y,z}(t_1,t_2)$, of $G_{Y,z}(t_1,t_2)$ obtained from $G_{Y,i,z}(T_{ij},T_{ik})$ is asymptotically equivalent to that obtained from $\check{G}_{Y,i,z}(T_{ij},T_{ik})$, denoted by $\bar{G}_{Y,z}(t_1,t_2)$.

To derive the uniform convergence rate of $\bar{G}_{Y,z}(t_1,t_2)$, note that $\bar{G}_{Y,z}(t_1,t_2)$ is the minimizer of $\sum_{i\in\mathcal{N}_{z,h}} \sum_{1\le j\ne l\le N_i} \kappa_2(\frac{T_{ij}-t_1}{h_{Y,z}}, \frac{T_{il}-t_2}{h_{Y,z}}) \left[\check{G}_{Y,i,z}(T_{ij},T_{il}) - b_0 - b_{11}(T_{ij}-t_1) - b_{12}(T_{il}-t_2)\right]^2$ with respect to b_0, where minimization is carried out over b_0, b_{11}, and b_{12}.

For $l,m = 0,1,2$, define $c_{lm} = \sum_{i\in\mathcal{N}_{z,h}} \sum_{1\le j\ne k\le N_i} w_{ijl}\check{G}_{Y,i,z}(T_{ij},T_{ik})(T_{ij}-t_1)^l(T_{ik}-t_2)^m$ and $a_{lm} = \sum_{i\in\mathcal{N}_{z,h}} \sum_{1\le j\ne k\le N_i} w_{ijl}(T_{ij}-t_1)^l(T_{ik}-t_2)^m$, where $w_{ijl} = \kappa_2(\frac{T_{ij}-t_1}{h_{Y,z}}, \frac{T_{il}-t_2}{h_{Y,z}})/(n_{z,h}h_{Y,z}^2)$. Let $\mathbf{c} = (c_{00}, c_{10}, c_{01})^T$ and $A = (A_1, A_2, A_3)$ a matrix, where $A_1 = (a_{00}, a_{10}, a_{01})^T$, $A_2 = (a_{10}, a_{20}, a_{11})^T$, and $A_3 = (a_{01}, a_{11}, a_{02})^T$. Then $\bar{G}_{Y,z}(t_1,t_2) = \hat{b}_0(t_1,t_2) = \bar{G}_{Y,z,NW}(t_1,t_2) - \frac{a_{10}}{a_{00}}\hat{b}_{11} - \frac{a_{01}}{a_{00}}\hat{b}_{12}$, where $\bar{G}_{Y,z,NW}(t_1,t_2) = c_{00}/a_{00}$ is the Nadaraya-Watson estimator for $G_{Y,z}(t_1,t_2)$ based on $\check{G}_{Y,i,z}(T_{ij},T_{ik})$, and $\hat{b}_{11}$ and $\hat{b}_{12}$ are estimators for the first order partial derivatives $\frac{\partial}{\partial t_1}G_{Y,z}(t_1,t_2)$ and $\frac{\partial}{\partial t_2}G_{Y,z}(t_1,t_2)$, respectively.

First consider the uniform convergence of the Nadaraya-Watson estimator. Set $\theta_1(t_1,t_2,v_1,v_2) = 1$ and $\theta_2(t_1,t_2,v_1,v_2) = (v_1 - \mu_{Y,z}(t_1))(v_2 - \mu_{Y,z}(t_2))$. Then $\varrho_{1\theta,z} = g_{2Y,z}(t_1,t_2)$ and $\varrho_{2\theta,z} = G_{Y,z}(t_1,t_2)g_{2Y,z}(t_1,t_2)$; $\bar{G}_{Y,z,NW}(t_1,t_2) = H(\Theta_{2n,z},\Theta_{1n,z})$ for $H(a,b) = a/b$; $\tilde{g}_{2Y,z}(t_1,t_2) = \Theta_{1n,z}(t_1,t_2)$. For kernel $\kappa_2(\cdot,\cdot)$ of order $((0,0),2)$, Lemma 3 implies that $\sup_{(z,t_1,t_2)\in\mathcal{Z}\times\mathcal{T}^2} \frac{|\Theta_{pn,z} - \varrho_{p\theta,z}|}{h + (\sqrt{n_{z,h}}h_{Y,z}^2)^{-1}} =$

$O_p(1)$ and the uniform version of *Slutsky*'s theorem implies

$$\sup_{(z,t_1,t_2)\in\mathcal{Z}\times\mathcal{T}^2} \frac{\left|\bar{G}_{Y,z,NW}(t_1,t_2)-G_{Y,z}(t_1,t_2)\right|}{h+(\sqrt{n_{z,h}}h_{Y,z}^2)^{-1}} = O_p(1). \tag{A.11}$$

For symmetric kernel functions $\kappa_2(u,v)$ of order $((0,0),2)$, the modified kernel $\tilde{\kappa}_{2,u}(u,v) = u\kappa_2(u,v)/\sigma^2_{u^2,\kappa_2}$ is of order $((1,0),3)$ and $\tilde{\kappa}_{2,v}(u,v) = v\kappa_2(u,v)/\sigma^2_{v^2,\kappa_2}$ is of order $((0,1),3)$, where $\sigma^2_{u^2,\kappa_2} = \int u^2\kappa_2(u,v)dudv$ and $\sigma^2_{v^2,\kappa_2} = \int v^2\kappa_2(u,v)dudv$.

Considering the uniform convergence of $\hat{b}_{11}$ to $\frac{\partial}{\partial t_1}G_{Y,z}(t_1,t_2)$, we apply Lemma 3 with kernel $\tilde{\kappa}_{2,u}(u,v)$ on $\theta_1^{(1)}(e_1,e_2,v_1,v_2) = (v_1-\mu_{Y,z}(e_1))(v_2-\mu_{Y,z}(e_2))$, $\theta_2^{(1)}(e_1,e_2,v_1,v_2)=1$, $\theta_3^{(1)}(e_1,e_2,v_1,v_2) = e_1-t_1$, and $\theta_4^{(1)}(e_1,e_2,v_1,v_2) = e_2-t_2$, to obtain $\Theta^{(1)}_{1n,z} = \frac{1}{h^2_{Y,z}\sigma^2_{u^2,\kappa_2}}c_{10} \to \varrho^{(1)}_{1\theta,z} = \frac{\partial}{\partial t_1}\left[G_{Y,z}(t_1,t_2)g_{2Y,z}(t_1,t_2)\right]$, $\Theta^{(1)}_{2n,z} = \frac{1}{h^2_{Y,z}\sigma^2_{u^2,\kappa_2}}a_{10} \to \varrho^{(1)}_{2\theta,z} = \frac{\partial}{\partial t_1}\left[g_{2Y,z}(t_1,t_2)\right]$, $\Theta^{(1)}_{3n,z} = \frac{1}{h^2_{Y,z}\sigma^2_{u^2,\kappa_2}}a_{20} \to \varrho^{(1)}_{3\theta,z} = g_{2Y,z}(t_1,t_2)$ and $\Theta^{(1)}_{4n,z} = \frac{1}{h^2_{Y,z}\sigma^2_{u^2,\kappa_2}}a_{11} \to \varrho^{(1)}_{3\theta,z} = 0$, uniformly in (z,t_1,t_2). For kernel $\tilde{\kappa}_{2,v}(u,v)$ with $\theta_1^{(2)}(e_1,e_2,v_1,v_2) = (v_1-\mu_{Y,z}(e_1))(v_2-\mu_{Y,z}(e_2))$, $\theta_2^{(2)}(e_1,e_2,v_1,v_2) = 1$, $\theta_3^{(2)}(e_1,e_2,v_1,v_2) = e_2-t_2$, and $\theta_4^{(2)}(e_1,e_2,v_1,v_2) = e_1-t_1$, analogously we get $\Theta^{(2)}_{1n,z} = \frac{1}{h^2_{Y,z}\sigma^2_{v^2,\kappa_2}}c_{01} \to \varrho^{(2)}_{1\theta,z} = \frac{\partial}{\partial t_2}\left[G_{Y,z}(t_1,t_2)g_{2Y,z}(t_1,t_2)\right]$, $\Theta^{(2)}_{2n,z} = \frac{1}{h^2_{Y,z}\sigma^2_{v^2,\kappa_2}}a_{01} \to \varrho^{(2)}_{2\theta,z} = \frac{\partial}{\partial t_2}\left[g_{2Y,z}(t_1,t_2)\right]$, $\Theta^{(2)}_{3n,z} = \frac{1}{h^2_{Y,z}\sigma^2_{v^2,\kappa_2}}a_{02} \to \varrho^{(2)}_{3\theta,z} = g_{2Y,z}(t_1,t_2)$ and $\Theta^{(2)}_{4n,z} = \frac{1}{h^2_{Y,z}\sigma^2_{v^2,\kappa_2}}a_{11} \to 0$, uniformly in (z,t_1,t_2).

For function $H(x_1,x_2,x_3) = (x_1-x_2G_{Y,z}(t_1,t_2))/x_3$, we can easily see that $H(\Theta^{(1)}_{1n,z},\Theta^{(1)}_{2n,z},\Theta^{(1)}_{3n,z}) = (c_{10}-a_{10}G_{Y,z}(t_1,t_2))/a_{20}$ and

$$H(\Theta^{(1)}_{1n,z},\Theta^{(1)}_{2n,z},\Theta^{(1)}_{3n,z}) + \frac{a_{10}(G_{Y,z}(t_1,t_2)-\bar{G}_{Y,z,NW}(t_1,t_2))}{a_{20}} = \frac{c_{10}-a_{10}c_{00}/a_{00}}{a_{20}}.$$

Note that

$$\hat{b}_{11} = \frac{(c_{10}-a_{10}c_{00}/a_{00})/(a_{20}) + (a_{01}a_{11}c_{00}+a_{01}a_{10}c_{01}-a_{00}a_{11}c_{01}-a_{01}^2c_{10})/(a_{00}a_{20}a_{02})}{1-(a_{02}a_{10}^2-2a_{01}a_{10}a_{11}+a_{00}a_{11}^2+a_{01}^2a_{20})/(a_{00}a_{20}a_{02})},$$

$$\begin{aligned}&(a_{01}a_{11}c_{00}+a_{01}a_{10}c_{01}-a_{00}a_{11}c_{01}-a_{01}^2c_{10})/(a_{00}a_{20}a_{02})\\ =&\left(\bar{G}_{Y,z,NW}(t_1,t_2)\Theta^{(1)}_{4n,z}\Theta^{(2)}_{2n,z} - \Theta^{(1)}_{4n,z}\Theta^{(2)}_{1n,z} + h^2_{Y,z}\frac{\Theta^{(1)}_{2n,z}\Theta^{(2)}_{1n,z}\Theta^{(2)}_{2n,z}-\Theta^{(1)}_{1n,z}(\Theta^{(2)}_{2n,z})^2}{\tilde{g}_{2Y,z}(t_1,t_2)/\sigma^2_{v^2,\kappa_2}}\right)/(\Theta^{(1)}_{3n,z}\Theta^{(2)}_{3n,z}),\end{aligned}$$

$$\begin{aligned}&(a_{02}a_{10}^2-2a_{01}a_{10}a_{11}+a_{00}a_{11}^2+a_{01}^2a_{20})/(a_{00}a_{20}a_{02})\\ =&\left(\Theta^{(1)}_{4n,z}\Theta^{(2)}_{4n,z} + h^2_{Y,z}\frac{\sigma^2_{u^2,\kappa_2}[(\Theta^{(1)}_{2n,z})^2\Theta^{(2)}_{3n,z}-2\Theta^{(1)}_{2n,z}\Theta^{(2)}_{2n,z}\Theta^{(1)}_{4n,z}]+\sigma^2_{v^2,\kappa_2}\Theta^{(1)}_{3n,z}(\Theta^{(2)}_{2n,z})^2}{\tilde{g}_{2Y,z}(t_1,t_2)}\right)/(\Theta^{(1)}_{3n,z}\Theta^{(2)}_{3n,z})\end{aligned}$$

and as a result both of them converge to zero. Applying the uniform version of *Slutsky*'s theorem, we can show that $\hat{b}_{11}$ converges to $\frac{\partial}{\partial t_1} G_{Y,z}(t_1,t_2)$ uniformly in z, t_1 and t_2, and analogously for $\frac{\partial}{\partial t_2} G_{Y,z}(t_1,t_2)$.

Combining these findings leads to the uniform convergence of $\bar{G}_{Y,z}(t_1,t_2)$, as

$$\bar{G}_{Y,z}(t_1,t_2) = \hat{b}_0(t_1,t_2) \quad = \quad \bar{G}_{Y,z,NW}(t_1,t_2) - h_{Y,z}^2 \frac{\Theta_{2n,z}^{(1)} \sigma_{u^2,\kappa_2}^2}{\tilde{g}_{2Y,z}(t_1,t_2)} \hat{b}_{11} - h_{Y,z}^2 \frac{\Theta_{2n,z}^{(2)} \sigma_{v^2,\kappa_2}^2}{\tilde{g}_{2Y,z}(t_1,t_2)} \hat{b}_{12},$$

equation (A.11) and the uniform convergence of $\Theta_{2n,z}^{(i)}$, $i=1,2$, $\tilde{g}_{2Y,z}(t_1,t_2)$, $\hat{b}_{11}$, and $\hat{b}_{12}$ imply

$$\sup_{(z,t_1,t_2)\in\mathcal{Z}\times\mathcal{T}^2} \frac{\left|\bar{G}_{Y,z}(t_1,t_2) - G_{Y,z}(t_1,t_2)\right|}{h + (\sqrt{n_{z,h}} h_{Y,z}^2)^{-1}} = O_p(1), \tag{A.12}$$

considering that $n_{z,h} h_{Y,z}^8 < \infty$ as $n \to \infty$.

Since $\sup_{(z,t)\in\mathcal{Z}\times\mathcal{T}} \mid \tilde{\mu}_{Y,z}(t) - \mu_{Y,z}(t) \mid / (h + (\sqrt{n_{z,h}} b_{Y,z})^{-1}) = O_p(1)$ by (A.5), the difference between $G_{Y,i,z}(T_{ij},T_{ik})$ and $\check{G}_{Y,i,z}(T_{ij},T_{ik})$ is of order $O_p(h + (\sqrt{n_{z,h}} b_{Y,z})^{-1})$, which is faster than the convergence rate $O_p(h + (\sqrt{n_{z,h}} h_{Y,z}^2)^{-1})$ of $\bar{G}_{Y,z}(t_1,t_2)$ to $G_{Y,z}(t_1,t_2)$, in equation (A.12), and therefore $\tilde{G}_{Y,z}(t_1,t_2)$ shares the same convergence rate of $\bar{G}_{Y,z}(t_1,t_2)$, as desired in (A.10). □

To estimate variance of the measurement errors, as in Yao et al. (2005a), we first estimate $G_{X,z}(s,s) + \sigma_X^2$ (resp. $G_{Y,z}(t,t) + \sigma_Y^2$) using a local linear smoother based on $G_{X,i,z}(S_{il},S_{il})$ for $l = 1,2,\cdots,L_i$, $i \in \mathcal{N}_{z,h}$ (resp. $G_{Y,i,z}(T_{ij},T_{ij})$ for $j=1,2,\cdots,N_i$, $i \in \mathcal{N}_{z,h}$) with smoothing bandwidth $b_{X,z,V}$ (resp. $b_{Y,z,V}$) and denote the estimates by $\tilde{V}_{X,z}(s)$ (resp. $\tilde{V}_{Y,z}(t)$), removing the two ends of the interval $\mathcal{S}$ (resp. $\mathcal{T}$) to get more stable estimates of σ_X^2 (res. σ_Y^2). Denote the estimates based on the generic bin $[z-h/2, z+h/2)$ by $\tilde{\sigma}_{X,z}^2$ and $\tilde{\sigma}_{Y,z}^2$, $|\mathcal{S}|$ the length of $\mathcal{S}$ and $\mathcal{S}_1 = [\inf\{s : s\in\mathcal{S}\} + |\mathcal{S}|/4, \sup\{s : s\in\mathcal{S}\} - |\mathcal{S}|/4]$. Then

$$\tilde{\sigma}_{X,z}^2 = \frac{2}{|\mathcal{S}|} \int_{\mathcal{S}_1} [\tilde{V}_X(s) - \tilde{G}_{X,z}(s,s)] ds.$$

and analogously for $\tilde{\sigma}_{Y,z}^2$. Lemmas 2 and 4 imply the convergence of $\tilde{\sigma}_{X,z}^2$ and $\tilde{\sigma}_{Y,z}^2$, as stated in Corollary 1.

[C5] Uniformly in $z\in\mathcal{Z}$, $b_{X,z,V} \to 0$, $n_{z,h} b_{X,z,V}^4 \to \infty$, $n_{z,h} b_{X,z,V}^6 < \infty$, $b_{Y,z,V} \to 0$, $n_{z,h} b_{Y,z,V}^4 \to \infty$, and $n_{z,h} b_{Y,z,V}^6 < \infty$, as $n\to\infty$.

Corollary 1. *Under Condition [C5] and the conditions of Lemmas 2 and 4,*

$$\sup_{z\in\mathcal{Z}}\left|\tilde\sigma^2_{X,z}-\sigma^2_X\right|/(h+(\sqrt{n_{z,h}}b_{X,z,V})^{-1}+(\sqrt{n_{z,h}}h^2_{X,z})^{-1})=O_p(1) \text{ and analogously for } \tilde\sigma^2_{X,z}.$$

Proposition 2. *Under Conditions [A0-A3] in Section 2 and (i-ix), the final estimates of σ^2_X and σ^2_Y (2.8) converge in probability to their corresponding true counterparts, i.e.,*

$$\hat\sigma^2_X\xrightarrow{P}\sigma^2_X,\qquad \hat\sigma^2_Y\xrightarrow{P}\sigma^2_Y.$$

Proof of Proposition 2. The result follows straightforwardly from Corollary 1. □

While Lemma 3 implies consistency of the estimator of the variance, we also require an extension regarding estimation of the cross-covariance function. Let $\{\tilde\theta_p(s,t,u,v), p=1,2,\cdots,l\}$ be a collection of real functions $\tilde\theta_p:\mathbb{R}^4\to\mathbb{R}$.

[C2.1c] For $\ell\geq|\nu|+2$ and any pair of ℓ_1 and ℓ_2 such that $\ell=\ell_1+\ell_2$, $\ell_1\geq\nu_1+1$, $\ell_2\geq\nu_2+1$. Uniformly in $z\in\mathcal{Z}$, bandwidth $h_{1,z}$ and $h_{2,z}$ satisfy $h_{1,z}\to0$, $h_{1,z}/h_{2,z}\to1$, $n_{z,h}h_{1,z}^{|\nu|+2}\to\infty$, $n_{z,h}h_{1,z}^{2\ell+4}<\infty$, as $n\to\infty$.

Define $\varrho_{p\tilde\theta,z}=\varrho_{p\tilde\theta,z}(s,t)=\frac{\partial^{|\nu|}}{\partial s^{\nu_1}\partial t^{\nu_2}}\int\int\tilde\theta_p(s,t,u,v)g_{XY,z}(s,t,u,v)dudv$ and

$$\tilde\Theta_{pn,z}=\tilde\Theta_{pn,z}(s,t)=\frac{1}{n_{z,h}h_{1,z}^{\nu_1+1}h_{2,z}^{\nu_2+1}}\sum_{i\in\mathcal{N}_{z,h}}\frac{1}{EN}\sum_{1\leq j\leq N_i}\tilde\theta_p(S_{ij},T_{ij},U_{ij},V_{ij})\kappa_2(\frac{S_{ij}-s}{h_{1,z}},\frac{T_{ij}-t}{h_{2,z}}).$$

Lemma 5. *Under Conditions [A0-A3], (i), (ii), (iii), (viii), [C1.1b] with $\mathcal{R}_1=\mathcal{S}$ and $\mathcal{R}_2=\mathcal{T}$, [C1.2b] with $g(\cdot,\cdot,\cdot,\cdot)=g_{XY,z}(\cdot,\cdot,\cdot,\cdot)$, and [C2.1c] (with $\ell_1=\ell_2=1$ and $\nu_1=\nu_2=0$), we have $\tilde\vartheta_{pn}=\sup_{(z,s,t)\in\mathcal{Z}\times\mathcal{S}\times\mathcal{T}}|\tilde\Theta_{pn,z}(s,t)-\varrho_{p\tilde\theta,z}(s,t)|/(h+(\sqrt{n_{z,h}}h_{Y,1}^{\nu_1+1}h_{Y,2}^{\nu_2+1})^{-1})=O_p(1)$.*

Proof. The proof is analogous to that of Lemmas 1 and 3. □

[C6] Uniformly in $z\in\mathcal{Z}$, bandwidth $h_{1,z}$ and $h_{2,z}$ satisfy $h_{1,z}\to0$, $h_{1,z}/h_{2,z}\to1$, $n_{z,h}h_{1,z}^6\to\infty$, $n_{z,h}h_{1,z}^8<\infty$, as $n\to\infty$.

Lemma 6 (Convergence of convergence of the cross-covariance function between X and Y)**.** *Under Condition [A0-A3], (i), (ii), (iii), (viii), (ix), [C3], and [C6]*

$$\sup_{(z,s,t)\in\mathcal{Z}\times\mathcal{S}\times\mathcal{T}}|\tilde C_{XY,z}(s,t)-C_{XY,z}(s,t)|/(h+(\sqrt{n_{z,h}}h_{1,z}h_{2,z})^{-1})=O_p(1).$$

Proof. The proof is similar to that of Lemma 4. □

Consider the real separable Hilbert space $L^2_Y(\mathcal{T}) \equiv H_Y$ (resp. $L^2_X(\mathcal{S}) \equiv H_X$) endowed with inner product $\langle f, g\rangle_{H_Y} = \int_{\mathcal{T}} f(t)g(t)dt$ (resp. $\langle f, g\rangle_{H_X} = \int_{\mathcal{S}} f(s)g(s)ds$) and norm $\| f \|_{H_X} = \sqrt{\langle f, f\rangle_{H_X}}$ (resp. $\| f \|_{H_Y} = \sqrt{\langle f, f\rangle_{H_Y}}$) (Courant and Hilbert, 1953). Let $\mathcal{I}'_{Y,z}$ (resp. $\mathcal{I}'_{X,z}$) be the set of indices of the eigenfunctions $\phi_{z,k}(t)$ (resp. $\psi_{z,m}(s)$) corresponding to eigenvalues $\lambda_{z,k}$ (resp. $\rho_{z,m}$) of multiplicity one. We obtain the consistency of $\tilde{\lambda}_{z,k}$ (resp. $\tilde{\rho}_{z,m}$) for $\lambda_{z,k}$ (resp. $\rho_{z,m}$), the consistency of $\tilde{\phi}_{z,k}(t)$ (resp. $\tilde{\psi}_{z,m}(s)$) for $\phi_{z,k}(t)$ (resp. $\psi_{z,m}(s)$) in the L^2_Y (resp. L^2_X) norm $\| \cdot \|_{H_X}$ (resp. $\| \cdot \|_{H_Y}$) when $\lambda_{z,k}$ (resp. $\rho_{z,m}$) is of multiplicity one, and the uniform consistency of $\tilde{\phi}_{z,k}(t)$ (resp. $\tilde{\psi}_{z,m}(s)$) for $\phi_{z,k}(t)$ (resp. $\psi_{z,m}(s)$) as well.

For $f, g, h \in H_Y$, define the rank one operator $f \otimes g : h \to \langle f, h\rangle g$. Denote the separable Hilbert space of Hilbert-Schmidt operators on H_Y by $F_Y \equiv \sigma_2(H_Y)$, endowed with $\langle T_1, T_2\rangle_{F_Y} = tr(T_1 T_2^*) = \sum_j \langle T_1 u_j, T_2 u_j\rangle_{H_Y}$ and $\| T \|^2_{F_Y} = \langle T, T\rangle_{F_Y}$, where T_1, T_2, $T \in F_Y$, T_2^* is the adjoint of T_2, and $\{u_j : j \geq 1\}$ is any complete orthonormal system in H_Y. The covariance operator $\mathbf{G}_{Y,z}$ (resp. $\tilde{\mathbf{G}}_{Y,z}$) is generated by the kernel $G_{Y,z}$ (resp. $\tilde{G}_{Y,z}$), *i.e.*, $\mathbf{G}_{Y,z}(f) = \int_{\mathcal{T}} G_{Y,z}(t_1, t) f(t_1) dt_1$ (resp. $\tilde{\mathbf{G}}_{Y,z}(f) = \int_{\mathcal{T}} \tilde{G}_{Y,z}(t_1, t) f(t_1) dt_1$). Obviously, $\mathbf{G}_{Y,z}$ and $\tilde{\mathbf{G}}_{Y,z}$ are Hilbert-Schmidt operators. As a result of (A.10), we have $\sup_{z\in\mathcal{Z}} \| \tilde{\mathbf{G}}_{Y,z} - \mathbf{G}_{Y,z} \|_{F_Y} / (h + (\sqrt{n_{z,h}} h^2_{Y,z})^{-1}) = O_p(1)$.

Let $\mathcal{I}_{Y,z,i} = \{j : \lambda_{z,j} = \lambda_{z,i}\}$ and $\mathcal{I}'_{Y,z} = \{i : |\mathcal{I}_{Y,z,i}| = 1\}$, where $|\mathcal{I}_{Y,z,i}|$ denotes the number of elements in $\mathcal{I}_{Y,z,i}$. Denote $\mathbf{P}^Y_{z,j} = \sum_{k\in\mathcal{I}_{Y,z,j}} \phi_{z,k} \otimes \phi_{z,k}$ and $\tilde{\mathbf{P}}^Y_{z,j} = \sum_{k\in\mathcal{I}_{Y,z,j}} \tilde{\phi}_{z,k} \otimes \tilde{\phi}_{z,k}$ to be the true and estimated orthogonal projection operators from H_Y to the subspace spanned by $\{\phi_{z,k} : k \in \mathcal{I}_{Y,z,j}\}$. Set $\delta^Y_{z,j} = \frac{1}{2} \min\{|\lambda_{z,l} - \lambda_{z,j}| : l \notin \mathcal{I}_{Y,z,j}\}$ and $\mathbf{\Lambda}_{\delta^Y_{z,j}} = \{c \in \mathcal{C} : |c - \lambda_{z,j}| = \delta^Y_{z,j}\}$, where $\mathcal{C}$ stands for the complex numbers. Let $\mathbf{R}_{Y,z}$ (resp. $\tilde{\mathbf{R}}_{Y,z}$) to be the resolvent of $\mathbf{G}_{Y,z}$ (resp. $\tilde{\mathbf{G}}_{Y,z}$), *i.e.*, $\mathbf{R}_{Y,z}(c) = (\mathbf{G}_{Y,z} - cI)^{-1}$ (resp. $\tilde{\mathbf{R}}_{Y,z}(c) = (\tilde{\mathbf{G}}_{Y,z} - cI)^{-1}$). Denote $A_{\delta^Y_{z,j}} = \sup\{\| \mathbf{R}_{Y,z}(c) \|_{F_Y} : c \in \mathbf{\Lambda}_{\delta^Y_{z,j}}\}$ and

$$\alpha_X = \Big(\delta^X_{z,j} (A_{\delta^X_{z,j}})^2\Big) / \Big((h + (\sqrt{n_{z,h}} h^2_{X,z})^{-1})^{-1} - A_{\delta^X_{z,j}}\Big). \tag{A.13}$$

Parallel notations are made for the Y process.

Proposition 3. *Under Conditions [A0-A3] in Section 2, and Conditions (i-iii), (viii),* *(ix), [C3], [C4], and [C6], it holds that*

$$\begin{aligned} |\tilde{\rho}_{z,m} - \rho_{z,m}| &= O_p(\alpha_X) && && \text{(A.14)} \\ \| \tilde{\psi}_{z,m} - \psi_{z,m} \|_{H_X} &= O_p(\alpha_X), && m \in \mathcal{I}'_{X,z} && \text{(A.15)} \\ \sup_{s \in \mathcal{S}} \left| \tilde{\psi}_{z,m}(s) - \psi_{z,m}(s) \right| &= O_p(\alpha_X), && m \in \mathcal{I}'_{X,z} && \text{(A.16)} \end{aligned}$$

$$\begin{aligned} \left| \tilde{\lambda}_{z,k} - \lambda_{z,k} \right| &= O_p(\alpha_Y) && && \text{(A.17)} \\ \| \tilde{\phi}_{z,k} - \phi_{z,k} \|_{H_Y} &= O_p(\alpha_Y), && k \in \mathcal{I}'_{Y,z} && \text{(A.18)} \\ \sup_{t \in \mathcal{T}} \left| \tilde{\phi}_{z,k}(t) - \phi_{z,k}(t) \right| &= O_p(\alpha_Y), && k \in \mathcal{I}'_{Y,z}, && \text{(A.19)} \end{aligned}$$

$$|\tilde{\sigma}_{z,mk} - \sigma_{z,mk}| = O_p(\max(\alpha_X, \alpha_Y, h + (\sqrt{n_{z,h}} h_{1,z} h_{2,z})^{-1})), \quad \text{(A.20)}$$

where the norms on H_X and H_Y are defined on page 34, both α_X, α_Y are defined in (A.13) and converge to zero as $n \to \infty$, and the above O_p terms are uniform in $z \in \mathcal{Z}$.

Proof of Proposition 3. The proof is similar to the proof of Theorem 2 in Yao et al. (2005a). The uniformity result follows from that of Lemmas 4 and 6. □

Note that

$$\beta(z, s, t) = \sum_{k=1}^{\infty} \sum_{m=1}^{\infty} \frac{E(\zeta_{z,m} \xi_{z,k})}{E(\zeta_{z,m}^2)} \psi_{z,m}(s) \phi_{z,k}(t). \quad \text{(A.21)}$$

To define the convergence of the right hand side of (A.21) in the L_2 sense in (s, t) and uniformly in z, we require that

[A4] $\sum_{k=1}^{\infty} \sum_{m=1}^{\infty} \sigma_{z,mk}^2 / \rho_{z,m}^2 < \infty$ uniformly for $z \in \mathcal{Z}$.

The proof of the following result is straightforward.

Lemma 7. *Under Condition [A4], uniformly in $z \in \mathcal{Z}$, the right hand side of (A.21) converges in the L_2 sense.*

Proof. The orthonormality of $\{\psi_{z,m} : m \geq 1\}$ and $\{\phi_{z,k} : k \geq 1\}$ implies that $\int_{\mathcal{T}} \int_{\mathcal{S}} \beta_{MK}(z, s, t)^2 ds dt = \sum_{k=1}^{K} \sum_{m=1}^{M} \sigma_{z,mk}^2 / \rho_{z,m}^2$. Hence it is obvious that β_{MK} converges in the L_2 sense in (s, t) and uniformly in z under [A4], *i.e.*, $\sup_{z \in \mathcal{Z}} \int_{\mathcal{T}} \int_{\mathcal{S}} (\beta_{MK}(z, s, t) - \beta(z, s, t))^2 ds dt \to 0$ as $M, K \to \infty$. □

The next result requires assumptions [A4] and the following

$$[A5] \qquad \begin{array}{rcl} \sum_{m=1}^{M(n)} \frac{\delta^X_{z,m}(A_{\delta^X_{z,m}})^2}{(h+(\sqrt{n_{z,h}}h^2_{X,z})^{-1})^{-1}-A_{\delta^X_{z,m}}} & \to & 0 \\ \sum_{k=1}^{K(n)} \frac{\delta^Y_{z,k}(A_{\delta^Y_{z,k}})^2}{(h+(\sqrt{n_{z,h}}h^2_{Y,z})^{-1})^{-1}-A_{\delta^Y_{z,k}}} & \to & 0 \\ MK(h+(\sqrt{n_{z,h}}h_{1,z}h_{2,z})^{-1}) & \to & 0 \end{array} \qquad \text{uniformly in } z \in \mathcal{Z}.$$

Lemma 8. *Under conditions of Proposition 3, [A4], and [A5],*

$$\lim_{n\to\infty} \sup_{z\in\mathcal{Z}} \int_{\mathcal{S}}\int_{\mathcal{T}} [\tilde{\beta}(z,s,t)-\beta(z,s,t)]^2 = 0, \qquad \textit{in probability.} \tag{A.22}$$

Proof. The orthonormality of the eigenfunction basis implies that

$$\int_{\mathcal{T}}\int_{\mathcal{S}} [\tilde{\beta}(z,s,t)-\beta(z,s,t)]^2 dsdt = Q_{z,1}(n)+Q_{z,2}(n)+Q_{z,3}(n),$$

where $Q_{z,1}(n) = \int_{\mathcal{T}}\int_{\mathcal{S}} \left\{\sum_{k=1}^{K-1}\sum_{m=1}^{M-1}[\frac{\tilde{\sigma}_{z,mk}}{\tilde{\rho}_{z,m}}\tilde{\psi}_{z,m}(s)\tilde{\phi}_{z,k}(t) - \frac{\sigma_{z,mk}}{\rho_{z,m}}\psi_{z,m}(s)\phi_{z,k}(t)]\right\}^2 dsdt$, $Q_{z,2}(n) = \sum_{k=K}^{\infty}\sum_{m=M}^{\infty}\frac{\sigma^2_{z,mk}}{\rho^2_{z,m}}$, and

$$Q_{z,3}(n) = 2\int_{\mathcal{T}}\int_{\mathcal{S}}[\sum_{k=K}^{\infty}\sum_{m=M}^{\infty}\frac{\sigma_{z,mk}}{\rho_{z,m}}\psi_{z,m}(s)\phi_{z,k}(t)]\times\left\{\sum_{k=1}^{K-1}\sum_{m=1}^{M-1}[\frac{\tilde{\sigma}_{z,mk}}{\tilde{\rho}_{z,m}}\tilde{\psi}_{z,m}(s)\tilde{\phi}_{z,k}(t) - \frac{\sigma_{z,mk}}{\rho_{z,m}}\psi_{z,m}(s)\phi_{z,k}(t)]\right\} dsdt$$

Condition [A4] implies that $\sup_{z\in\mathcal{Z}} Q_{z,2}(n) \to 0$ as $M(n), K(n) \to \infty$. By (A.14), (A.16), (A.17), (A.19), (A.20) and [A5], the uniform version of *Slutsky*'s theorem implies that

$$\begin{aligned} Q_{z,1}(n) &= O_p(\sum_{m=1}^{M}\frac{\delta^X_{z,m}(A_{\delta^X_{z,m}})^2}{(h+(\sqrt{n_{z,h}}h^2_{X,z})^{-1})^{-1}-A_{\delta^X_{z,m}}} \\ &\quad + \sum_{k=1}^{K}\frac{\delta^Y_{z,k}(A_{\delta^Y_{z,k}})^2}{(h+(\sqrt{n_{z,h}}h^2_{Y,z})^{-1})^{-1}-A_{\delta^Y_{z,k}}} + KM(h+(\sqrt{n_{z,h}}h_{1,z}h_{2,z})^{-1})) \xrightarrow{P} 0, \end{aligned}$$

uniformly in z. Note that $|Q_{z,3}(n)| \le Q_{z,1}(n)+Q_{z,2}(n)$ due to Cauchy-Schwarz inequality implying $Q_{z,3}(n) \xrightarrow{P} 0$ uniformly in $z \in \mathcal{Z}$, and (A.22) follows. □

Proof of Theorem 1. We consider only the convergence of $\hat{\beta}(z,s,t)$. The consistency of $\hat{\mu}_{X,z}(s)$ and $\hat{\mu}_{Y,z}(t)$ is analogous. Note first that

$$\begin{aligned} &\int_{\mathcal{T}}\int_{\mathcal{S}}(\hat{\beta}(z,s,t)-\beta(z,s,t))^2 dsdt \\ \le\ & 2(2b/h+1)\sum_{p=1}^{P}\omega_{0,2}(z^{(p)},z,b)^2\int_{\mathcal{T}}\int_{\mathcal{S}}(\tilde{\beta}(z^{(p)},s,t)-\beta(z^{(p)},s,t))^2 dsdt \\ &+2\int_{\mathcal{T}}\int_{\mathcal{S}}(\sum_{p=1}^{P}\omega_{0,2}(z^{(p)},z,b)\beta(z^{(p)},s,t)-\beta(z,s,t))^2 dsdt, \end{aligned} \tag{A.23}$$

where $2b/h+1$ of the very last inequality is due to the fact that the kernel function $K(\cdot)$ is of bounded support $[-1,1]$. Denote $a(k)=\sum_{p=1}^{P}K_b(z^{(p)}-z)(z^{(p)}-z)^k$, $b(k)=\sum_{p=1}^{P}K_b(z^{(p)}-z)^2(z^{(p)}-z)^k$, $\mu_k=\int K(u)u^k du$ and $\nu_k=\int(K(u))^2u^k du$. Then we have

$$a(k)=\mu_k\frac{b^k}{h}(1+o(1)),\ \text{and}\ b(k)=\nu_k\frac{b^{k-1}}{h}(1+o(1))$$

for small h (large $P\propto 1/h$) and small b. Moreover the usual boundary techniques can be applied near the two end points. Consequently

$$\begin{aligned}&\sum_{p=1}^{P}\omega_{0,2}(z^{(p)},z,b)^2=\mathbf{e}_{1,2}^T(\mathbf{C}^T\mathbf{W}\mathbf{C})^{-1}(\mathbf{C}^T\mathbf{W}\mathbf{W}\mathbf{C})(\mathbf{C}^T\mathbf{W}\mathbf{C})^{-1}\mathbf{e}_{1,2}\\ =\ &\mathbf{e}_{1,2}^T\begin{pmatrix}a(0)&a(1)\\a(1)&a(2)\end{pmatrix}^{-1}\begin{pmatrix}b(0)&b(1)\\b(1)&b(2)\end{pmatrix}\begin{pmatrix}a(0)&a(1)\\a(1)&a(2)\end{pmatrix}^{-1}\mathbf{e}_{1,2}=(\frac{\mu_2^2\nu_0-2\mu_1\mu_2\nu_1+\mu_1^2\nu_2}{\mu_0\mu_2-\mu_1^2})(\frac{b}{h})(1+o(1)),\end{aligned}$$

due to the compactness of $\mathcal{Z}$, the above o-term is uniform in $z\in\mathcal{Z}$, implying

$$\int_{\mathcal{Z}}\sum_{p=1}^{P}\omega_{0,2}(z^{(p)},z,b)^2dz\ =\ (\frac{\mu_2^2\nu_0-2\mu_1\mu_2\nu_1+\mu_1^2\nu_2}{\mu_0\mu_2-\mu_1^2})(\frac{b}{h})\,|\mathcal{Z}|\,(1+o(1))\quad\text{(A.24)}$$

for small h and b, where $|\mathcal{Z}|$ denotes the Lebesgue measure of $\mathcal{Z}$. Hence, (A.24) and the consistency of $\tilde{\beta}(z,s,t)$ in the L_2 sense in (s,t) and uniformly in z due to (A.22) imply

$$\int_{\mathcal{Z}}\left[\sum_{p=1}^{P}\omega_{0,2}(z^{(p)},z,b)^2\int_{\mathcal{T}}\int_{\mathcal{S}}((\tilde{\beta}(z^{(p)},s,t)-\beta(z^{(p)},s,t)))^2dsdt\right]dz\xrightarrow{P}0.\quad\text{(A.25)}$$

For the second part in (A.23), applying Taylor expansion on $\beta(z^{(p)},s,t)$ at each z,

$$\begin{aligned}&\sum_{p=1}^{P}\omega_{0,2}(z^{(p)},z,b)\beta(z^{(p)},s,t)\\ =\ &\mathbf{e}_{1,2}^T\begin{pmatrix}a(0)&a(1)\\a(1)&a(2)\end{pmatrix}^{-1}\begin{pmatrix}a(0)\\a(1)\end{pmatrix}\beta(z,s,t)+\mathbf{e}_{1,2}^T\begin{pmatrix}a(0)&a(1)\\a(1)&a(2)\end{pmatrix}^{-1}\begin{pmatrix}a(1)\\a(2)\end{pmatrix}\frac{\partial}{\partial z}\beta(z,s,t)\\ &+\frac{1}{2}\mathbf{e}_{1,2}^T\begin{pmatrix}a(0)&a(1)\\a(1)&a(2)\end{pmatrix}^{-1}\begin{pmatrix}a(2)\\a(3)\end{pmatrix}\frac{\partial^2}{\partial z^2}\beta(z,s,t)+\text{higher order terms}\\ =\ &\beta(z,s,t)+\frac{1}{2}b^2\frac{\mu_2^2-\mu_1\mu_3}{\mu_0\mu_2-\mu_1^2}\frac{\partial^2}{\partial z^2}\beta(z,s,t)+\text{higher order terms.}\end{aligned}$$

Hence $\sum_{p=1}^{P}\omega_{0,2}(z^{(p)},z,b)\beta(z^{(p)},s,t)-\beta(z,s,t)=\frac{1}{2}b^2\frac{\mu_2^2-\mu_1\mu_3}{\mu_0\mu_2-\mu_1^2}\frac{\partial^2}{\partial z^2}\beta(z,s,t)(1+o(1))$, and

$$\begin{aligned}&\int_{\mathcal{Z}}\int_{\mathcal{T}}\int_{\mathcal{S}}(\sum_{p=1}^{P}\omega_{0,2}(z^{(p)},z,b)\beta(z^{(p)},s,t)-\beta(z,s,t))^2dsdtdz\\ =\ &\frac{1}{2}b^2\frac{\mu_2^2-\mu_1\mu_3}{\mu_0\mu_2-\mu_1^2}(\int_{\mathcal{Z}}\int_{\mathcal{T}}\int_{\mathcal{S}}\frac{\partial^2}{\partial z^2}\beta(z,s,t)dsdtdz)(1+o(1))\to 0.\quad\text{(A.26)}\end{aligned}$$

Combining (A.25) and (A.26) completes the proof by noting further Condition (xi). □

Proof of Theorem 2. Note that

$$\begin{aligned} Y^*(t) - \hat{Y}^*(t) &= \mu_{Y,Z^*}(t) - \hat{\mu}_{Y,Z^*}(t) + \int_{\mathcal{S}} (\beta(Z^*, s, t) - \hat{\beta}(Z^*, s, t))(X^*(s) - \mu_{X,Z^*}(s))ds \\ &\quad - \int_{\mathcal{S}} \hat{\beta}(Z^*, s, t)(\mu_{X,Z^*}(s) - \hat{\mu}_{X,Z^*}(s))ds. \end{aligned}$$

The convergence results in Theorem 1 imply that $\int_{\mathcal{T}} (Y^*(t) - \hat{Y}^*(t))^2 dt \xrightarrow{P} 0$ as desired. □