

Supplementary Material for “Online Estimation for Functional Data”

In this Supplementary Material, we first derive the asymptotic distributions of the proposed mean and covariance estimates, i.e., Theorem 1 and 2, in Section S.1 and S.2, respectively. The convergence of bandwidth selection and lower bound of relative efficiencies are shown in Section S.3 and S.4.

S.1. PROOF OF THEOREM 1

Proof. This proof is in analogy to the proof of Theorem 3.1 in Zhang and Wang (2016). Denote the current time by K and recall $e_{:,k}^{(K)}; k = 1, \dots, K$ is the pseudo-bandwidth chain at time K . We omit the superscript (K) of $e_{:,k}^{(K)}, \mathbb{E}^{(K)}$ in (18), $\mathbb{E}_{:,j}^{(K)}$ in (14) and $e^{(K)}$ in (10) in the proof when no confusion exists. The online estimate can be written as

$$e(t) = \frac{R_0 Q_2}{Q_0 Q_2} \frac{R_1 Q_1}{Q_1^2}$$

where for $r = 0, 1, 2$,

$$Q_r = \frac{1}{S_{K;1}} \prod_{k=1}^K \prod_{i=1}^k \prod_{j=1}^{k-i} W_{:,k}(T_{kij}(t)) \frac{T_{kij}(t)^r}{e_{:,k}} = \prod_k !_k Q_{r;k}$$

$$R_r = \frac{1}{S_{K;1}} \prod_{k=1}^K \prod_{i=1}^k \prod_{j=1}^{k-i} W_{:,k}(T_{kij}(t)) \frac{T_{kij}(t)^r}{e_{:,k}} Y_{kij} = \prod_k !_k R_{r;k}$$

We further denote for $r = 0, 1$,

$$\mathbb{E}_r = \frac{1}{S_{K;1}} \prod_{k=1}^K \prod_{i=1}^k \prod_{j=1}^{k-i} W_{:,k}(T_{kij}(t)) T_{kij}^r = \prod_k !_k \mathbb{E}_{r;k}$$

and define

$$e'(t) = \frac{1}{e_{:,k}} \frac{R_1}{Q_2} \frac{R_0 Q_1 = Q_0}{Q_1^2 = Q_0}$$

Note that

$$Q_1 = \prod_k !_k Q_{1;k} = \prod_k !_k \frac{\mathbb{E}_{1;k} t \mathbb{E}_{0;k}}{e_{:,k}}$$

one can derive

$$e(t) = \frac{R_0}{Q_0} \prod_k !_k \frac{\mathbb{E}_{1;k}}{Q_0} e'_k(t) + t \prod_k !_k \frac{\mathbb{E}_{0;k}}{Q_0} e'_k(t)$$

We also define

$$e(t) = \frac{R_0}{Q_0} \prod_k \frac{\mathbb{G}_{1;k}}{Q_0} f'(t) + t \prod_k \frac{\mathbb{G}_{0;k}}{Q_0} f'(t); \quad (S.1)$$

where $f'(\cdot)$ is the first derivative of $f(\cdot)$. Then

$$e(t) = \prod_k \frac{Q_{1;k}}{Q_0} \frac{Q_0(R_1 - Q_1 e_{;k} Q_2) - Q_1(R_0 - Q_0 e_{;k} Q_1)}{Q_2 Q_0 - Q_1^2}; \quad (S.2)$$

It is straightforward to show that both Q_0 and $Q_0 Q_2 - Q_1^2$ are positive and bounded away from 0 with probability tending to one,

$$Q_1 = O_p\left(\frac{1}{\sqrt{S_{K,1}}}\right);$$

$$(R_1 - Q_1 e_{;k} Q_2); (R_0 - Q_0 e_{;k} Q_1) = O_p\left(\frac{1}{\sqrt{S_{K,1}}}\right) + \frac{1}{\sqrt{S_{K,1}}} e_{;k} + \frac{1}{\sqrt{S_{K,1}}} \left(\frac{S_{K,2}}{S_{K,1}^2}\right)^{\frac{1}{2}} A;$$

Then by the Cramer-Wold device and Lyapunov condition, we can achieve the asymptotic joint normality of $(R_0, \mathbb{G}_1, \mathbb{G}_0)$, where the rate of convergence is $\min[(\frac{1}{\sqrt{S_{K,1}}})^{1=2}; f Q_2 = S_{K,1}^2 g^{1=2}]$.

Explicitly for $r, r' = 0, 1$,

$$E(\mathbb{G}_r) = t^r f(t) + \frac{1}{2} (W) f''(t) + o_p\left(\frac{1}{\sqrt{S_{K,1}}}\right);$$

$$E(R_0) = (t) f(t) + \frac{1}{2} (W) f''(t) + 2 f'(t) f'(t) + f''(t) g_{;2} + o_p\left(\frac{1}{\sqrt{S_{K,1}}}\right);$$

$$\text{Cov}(\mathbb{G}_r; \mathbb{G}_{r'}) = \frac{1}{\sqrt{S_{K,1}}} R(W) t^{r+r'} f(t) + o_p\left(\frac{1}{\sqrt{S_{K,1}}}\right);$$

$$\text{Var}(R_0) = \frac{1}{\sqrt{S_{K,1}}} R(W) (f''(t) + (t; t) f''(t) + f''(t) g_{;2}) + o_p\left(\frac{1}{\sqrt{S_{K,1}}}\right) + \frac{S_{K,2}}{S_{K,1}^2} (t; t) f(t)^2 + o_p\left(\frac{1}{\sqrt{S_{K,1}}}\right);$$

$$\text{Cov}(R_0; \mathbb{G}_r) = \frac{1}{\sqrt{S_{K,1}}} R(W) t^r (t) f(t) + o_p\left(\frac{1}{\sqrt{S_{K,1}}}\right);$$

From the delta method, $e(t)$ follows the asymptotic normality that

$$e(t) = \frac{1}{\sqrt{S_{K,1}}} (W) f''(t) g_{;2} + o_p\left(\frac{1}{\sqrt{S_{K,1}}}\right) \stackrel{d}{\rightarrow} N(0; 1);$$

where

$$= R(W) \frac{(t; t) + f''(t)}{\sqrt{S_{K,1}} f(t)} \frac{1}{\sqrt{S_{K,1}}} + (t; t) \frac{S_{K,2}}{S_{K,1}^2};$$

When $\bar{m}_{;K} = (N_K)^{1=4} \rightarrow 0$ and $\bar{\theta} = S_{K,1}^{-1=5}$, the first term in above dominates. When $\bar{m}_{;K} = (N_K)^{1=4} \rightarrow C$ and $\bar{\theta} = S_{K,1}^{-1=5}$, where $0 < C < 1$, the two terms of are of the same

order as $\mathcal{O}_p(n^{-4})$. When $m_{j,k} = (N_K)^{1-4} \rightarrow 1$ and $\mathcal{H} = S_{K,1}^{-1-5}$, the second term of \mathcal{R}_1 dominates and the bias becomes negligible. Noting (S.2), the proof of Theorem 1 is completed. \blacksquare

S.2. PROOF OF THEOREM 2 AND LEMMA 1

We omit the superscript $\setminus(K)$ of $e_{j,k}^{(K)}$, $\mathcal{H}^{(K)}$, $\mathcal{H}_j^{(K)}$ and $e^{(K)}$ in this section when no confusion exists. The proof of Lemma 1 is similar to the proof of Theorem 2, with $j: j \in Z$ replaced by \mathcal{H} . Before delineating the proof of Theorem 2, we state the following lemma which is a preparation for the proof of Theorem 2.

Lemma S. 1. *Let*

$$Q_{0R_1} = \frac{1}{S_{K,1}} \prod_{k'=1}^K \prod_{i'=1}^K \prod_{j'=1}^K W_{\cdot,k'}(T_{k'i'j'} \quad T_{kil});$$

$$\mathcal{R}_1 = S_{K,1}^{-1} S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^K \prod_{j \neq i}^K \prod_{k'=1}^K \prod_{i'=1}^K \prod_{j'=1}^K W_{\cdot,k'}(T_{kij} \quad s) W_{\cdot,k'}(T_{kil} \quad t) W_{\cdot,k'}(T_{k'i'j'} \quad T_{kil});$$

$$(Y_{kij} \quad (T_{kij})) (Y_{k'i'j'} \quad (T_{k'i'j'})) \frac{1}{2} \text{tr}(T_{kil})(T_{k'i'j'} \quad T_{kil})^2 ;$$

then

$$E(\mathcal{R}_1 = Q_{0R_1}) = O_p(S_{K,1}^{-1} S_{K,2}^{-1} S_{K,3});$$

$$\text{Var}(\mathcal{R}_1 = Q_{0R_1}) = O_p(S_{K,1}^{-2} S_{K,2}^{-2} S_{K,3}^2 + \sum_k S_{k,3} e_{j,k}^{-1} e_{i,k}^{-2} + \sum_k S_{k,4} e_{j,k}^{-1} e_{i,k}^{-1} + \sum_k S_{k,5} e_{j,k}^{-1} + S_{K,2}^{-2} \sum_k S_{k,4} e_{j,k}^4);$$

Proof. By easy calculation,

$$E \mathcal{R}_1 = S_{K,1}^{-1} S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^K \prod_{j \neq i}^K \prod_{k'=1}^K \prod_{i'=1}^K \prod_{j'=1}^K E W_{\cdot,k'}(T_{kij} \quad s) W_{\cdot,k'}(T_{kil} \quad t) W_{\cdot,k'}(T_{k'i'j'} \quad T_{kil}) (Y_{kij} \quad (T_{kij})) (Y_{k'i'j'} \quad (T_{k'i'j'}))$$

$$= S_{K,1}^{-1} S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^K \prod_{j \neq i}^K \prod_{i'=1}^K E W_{\cdot,k'}(T_{kij} \quad s) W_{\cdot,k'}(T_{kil} \quad t) W_{\cdot,k'}(T_{kij'} \quad T_{kil}) (Y_{kij} \quad (T_{kij})) (Y_{kij'} \quad (T_{kij'}))$$

$$= S_{K,1}^{-1} S_{K,2}^{-1} \prod_{k=1}^K \prod_{j \neq i}^K m_{ki}^2 (m_{ki} - 1) E W_{\cdot,k'}(T_{kij} \quad s) W_{\cdot,k'}(T_{kil} \quad t)$$

#

$$W_{k_1 i_1 j_1}(s) W_{k_1 i_1 h_1}(t) W_{k_1' i_1' j_1'}(T_{k_1 i_1 h_1}) W_{k_2 i_2 j_2}(s) W_{k_2 i_2 l_2}(t) W_{k_2' i_2' j_2'}(T_{k_2 i_2 l_2})$$

$$= O_p(S_{K,1}^{-1} S_{K,2}^{-1} S_{K,3}) ;$$

and hence $E \mathbb{R}_1 = O_{pR_1} = O_p(S_{K,1}^{-1} S_{K,2}^{-1} S_{K,3})$. Note that

$$\begin{aligned} \mathbb{R}_1^2 &= S_{K,1}^{-2} S_{K,2}^{-2} \sum_{j_1 \neq h_1, j_2 \neq l_2} W_{k_1 i_1 j_1}(s) W_{k_1 i_1 h_1}(t) W_{k_1' i_1' j_1'}(T_{k_1 i_1 h_1}) W_{k_2 i_2 j_2}(s) W_{k_2 i_2 l_2}(t) W_{k_2' i_2' j_2'}(T_{k_2 i_2 l_2}) \\ &\quad (Y_{k_1 i_1 j_1} \quad k_1 i_1 j_1) \quad (Y_{k_1' i_1' j_1'} \quad k_1' i_1' j_1') \quad \frac{1}{2} \quad \text{"} \quad k_1 i_1 h_1 (T_{k_1' i_1' j_1'} \quad T_{k_1 i_1 h_1})^2 \\ &\quad (Y_{k_2 i_2 j_2} \quad k_2 i_2 j_2) \quad (Y_{k_2' i_2' j_2'} \quad k_2' i_2' j_2') \quad \frac{1}{2} \quad \text{"} \quad k_2 i_2 l_2 (T_{k_2' i_2' j_2'} \quad T_{k_2 i_2 l_2})^2 ; \end{aligned}$$

$E \mathbb{R}_1^2$ is the summation of the following quantities:

$$\begin{aligned} (a) &= S_{K,1}^{-2} S_{K,2}^{-2} E \sum W_{k_1 i_1 j_1}(s) W_{k_1 i_1 h_1}(t) W_{k_1' i_1' j_1'}(T_{k_1 i_1 h_1}) W_{k_2 i_2 j_2}(s) W_{k_2 i_2 l_2}(t) W_{k_2' i_2' j_2'}(T_{k_2 i_2 l_2}) \\ &\quad (Y_{k_1 i_1 j_1} \quad k_1 i_1 j_1) (Y_{k_1' i_1' j_1'} \quad k_1' i_1' j_1') (Y_{k_2 i_2 j_2} \quad k_2 i_2 j_2) (Y_{k_2' i_2' j_2'} \quad k_2' i_2' j_2'); \\ (b) &= \frac{1}{2} S_{K,1}^{-2} S_{K,2}^{-2} E \sum W_{k_1 i_1 j_1}(s) W_{k_1 i_1 h_1}(t) W_{k_1' i_1' j_1'}(T_{k_1 i_1 h_1}) W_{k_2 i_2 j_2}(s) W_{k_2 i_2 l_2}(t) W_{k_2' i_2' j_2'}(T_{k_2 i_2 l_2}) \\ &\quad (Y_{k_1 i_1 j_1} \quad k_1 i_1 j_1) (Y_{k_1' i_1' j_1'} \quad k_1' i_1' j_1') (Y_{k_2 i_2 j_2} \quad k_2 i_2 j_2) \quad \text{"} \quad k_2 i_2 l_2 (T_{k_2' i_2' j_2'} \quad T_{k_2 i_2 l_2})^2; \\ (c) &= \frac{1}{2} S_{K,1}^{-2} S_{K,2}^{-2} E \sum W_{k_1 i_1 j_1}(s) W_{k_1 i_1 h_1}(t) W_{k_1' i_1' j_1'}(T_{k_1 i_1 h_1}) W_{k_2 i_2 j_2}(s) W_{k_2 i_2 l_2}(t) W_{k_2' i_2' j_2'}(T_{k_2 i_2 l_2}) \\ &\quad (Y_{k_1 i_1 j_1} \quad k_1 i_1 j_1) (Y_{k_2 i_2 j_2} \quad k_2 i_2 j_2) (Y_{k_2' i_2' j_2'} \quad k_2' i_2' j_2') \quad \text{"} \quad k_1 i_1 h_1 (T_{k_1' i_1' j_1'} \quad T_{k_1 i_1 h_1})^2; \\ (d) &= \frac{1}{4} S_{K,1}^{-2} S_{K,2}^{-2} E \sum W_{k_1 i_1 j_1}(s) W_{k_1 i_1 h_1}(t) W_{k_1' i_1' j_1'}(T_{k_1 i_1 h_1}) W_{k_2 i_2 j_2}(s) W_{k_2 i_2 l_2}(t) W_{k_2' i_2' j_2'}(T_{k_2 i_2 l_2}) \\ &\quad (Y_{k_1 i_1 j_1} \quad k_1 i_1 j_1) (Y_{k_2 i_2 j_2} \quad k_2 i_2 j_2) \quad \text{"} \quad k_1 i_1 h_1 (T_{k_1' i_1' j_1'} \quad T_{k_1 i_1 h_1})^2 \quad \text{"} \quad k_2 i_2 l_2 (T_{k_2' i_2' j_2'} \quad T_{k_2 i_2 l_2})^2; \end{aligned}$$

which are computed as follows.

- a. For (a), there are four cases according to the relationship between $(k_1; i_1)$, $(k_2; i_2)$, $(k_1'; i_1')$, $(k_2'; i_2')$ to make the quantity nonzero:

$$(B.1) \text{ Let } \overset{a}{1} = f(k_1; i_1) = (k_1'; i_1'); (k_2; i_2) = (k_2'; i_2'); (k_1; i_1) \notin (k_2; i_2); j_1 \notin h_1; j_2 \notin l_2,$$

$$\begin{aligned} (a:1) &= S_{K,1}^{-2} S_{K,2}^{-2} E \sum_{\overset{a}{1}} W_{k_1 i_1 j_1}(s) W_{k_1 i_1 h_1}(t) W_{k_1' i_1' j_1'}(T_{k_1 i_1 h_1}) W_{k_2 i_2 j_2}(s) W_{k_2 i_2 l_2}(t) \\ &\quad W_{k_2 i_2 j_2'}(T_{k_2 i_2 l_2}) (Y_{k_1 i_1 j_1} \quad k_1 i_1 j_1) (Y_{k_1' i_1' j_1'} \quad k_1' i_1' j_1') (Y_{k_2 i_2 j_2} \quad k_2 i_2 j_2) (Y_{k_2 i_2 j_2'} \quad k_2 i_2 j_2') \\ &= S_{K,1}^{-2} S_{K,2}^{-2} E \sum_{\overset{a}{1}} W_{k_1 i_1 h_1}(t) W_{k_1 i_1 j_1}(s) W_{k_1' i_1' j_1'}(T_{k_1 i_1 h_1}) (Y_{k_1 i_1 j_1} \quad k_1 i_1 j_1) (Y_{k_1' i_1' j_1'} \quad k_1' i_1' j_1') \end{aligned}$$

$$E W_{k_2 i_2 l_2}(t) W_{k_2 i_2 j_2}(s) W_{k_2 i_2 j'_2}(T_{k_2 i_2 l_2})(Y_{k_2 i_2 j_2}(Y_{k_2 i_2 j'_2}))_{k_2 i_2 j'_2} \\ = (E \mathcal{R}_1(s; t))^2:$$

(B.2) Let $\overset{a}{2} = f(k_1; i_1) = (k_2; i_2); (k'_1; i'_1) = (k'_2; i'_2); (k_1; i_1) \notin (k'_1; i'_1); j_1 \notin h_1; j_2 \notin h_2; j'_1 \notin j'_2 g$,

$$(a.2) = S_{K,1}^{-2} S_{K,2}^{-2} \overset{\times}{\underset{\overset{a}{2}}{E}} E f W_{k_1 i_1 j_1}(s) W_{i_2 j_1 k_1}(s) (Y_{k_1 i_1 j_1}(Y_{i_2 j_1 k_1}))_{i_2 j_1 k_1} g \\ E W_{k_1 i_1 h_1}(t) W_{i_2 j_1 k_1}(t) W_{k'_1 i'_1 j'_1}(T_{k_1 i_1 h_1}) W_{i'_2 j'_1 k'_1}(T_{i_2 j_1 k_1})(Y_{k'_1 i'_1 j'_1}(Y_{i'_2 j'_1 k'_1}))_{i'_2 j'_1 k'_1} \\ = S_{K,1}^{-2} S_{K,2}^{-2} \overset{\times}{\underset{\overset{a}{2}}{(s; s) (t; t) f^2(s) f^4(t) + O_p(\cdot_2 + \cdot_2)}} \\ = S_{K,1}^{-2} S_{K,2}^{-2} S_{K,3}^2 (s; s) (t; t) f^2(s) f^4(t) + O_p(\cdot_2 + \cdot_2) :$$

(B.3) Let $\overset{a}{3} = f(k_1; i_1) = (k'_2; i'_2); (k_2; i_2) = (k'_1; i'_1); (k_1; i_1) \notin (k_2; i_2); j_1 \notin h_1; j_2 \notin h_2 g$,

$$(a.3) = S_{K,1}^{-2} S_{K,2}^{-2} \overset{\times}{\underset{\overset{a}{3}}{E}} W_{k_1 i_1 j_1}(s) W_{k_1 i_1 h_1}(t) W_{k_2 i_2 j'_1}(T_{k_1 i_1 h_1}) \\ W_{k_2 i_2 j_2}(s) W_{k_2 i_2 l_2}(t) W_{k_1 i_1 j'_2}(T_{k_2 i_2 l_2}) \\ (Y_{k_1 i_1 j_1}(Y_{k_2 i_2 j'_1}(Y_{k_2 i_2 j_2}(Y_{k_1 i_1 j'_2})))_{k_1 i_1 j'_2} \\ = (E \mathcal{R}_1(s; t))^2:$$

(B.4) Let $\overset{a}{4} = f(k_1; i_1) = (k_2; i_2) = (k'_1; i'_1) = (k'_2; i'_2); j_1 \notin h_1; j_2 \notin h_2 g$ and omit the subscript ki ,

$$(a.4) = S_{K,1}^{-2} S_{K,2}^{-2} \overset{\times}{\underset{\overset{a}{4}}{E}} W_{j_1}(s) W_{i_1}(t) W_{j'_1}(T_{i_1}) W_{j_2}(s) W_{i_2}(t) W_{j'_2}(T_{i_2}) \\ (Y_{j_1}(Y_{j'_1}(Y_{j_2}(Y_{j'_2})))_{j'_2} :$$

It is the summation of the following sets,

$$\overset{a}{4.1} = f j_1 = j_2; j'_1 = j'_2; h_1 = h_2; j_1 \notin h_1; j_2 \notin h_2 g;$$

$$\overset{a}{4.2} = f j_1 = j_2; j'_1 \notin j'_2; j_1 \notin h_1; j_2 \notin h_2 g [f j_1 \notin j_2; j'_1 = j'_2; h_1 = h_2; j_1 \notin h_1; j_2 \notin h_2 g;$$

$$\overset{a}{4.3} = f j_1; j_2; j'_1; j'_2 \text{ are not equal}; j_1 \notin h_1; j_2 \notin h_2 g;$$

Correspondingly, we define

$$\begin{aligned}
 E_{1|4}(s; t) &= E \left[Y_1(T_1) Y_2(T_2) \mid T_1 = s; T_2 = t \right]; \\
 E_{2|4}(s; t) &= E \left[Y_1(T_1) Y_2(T_2) Y_3(T_3) \mid T_1 = s; T_2 = t; T_3 = t \right]; \\
 E_{3|4}(s; t) &= E \left[Y_1(T_1) Y_2(T_2) Y_3(T_3) Y_4(T_4) \mid T_1 = s; \right. \\
 &\quad \left. T_2 = t; T_3 = s; T_4 = t \right];
 \end{aligned}$$

Then by easy calculation,

$$\begin{aligned}
 (a.4) \quad a_{4.1} &\doteq S_{K,1}^{-2} S_{K,2}^{-2} \sum_{k=1}^{\infty} s_{k,3} e^{-1}_{,k} e^{-2}_{,k} R(W)^3 E_{1|4}(s; t) f(s) f^2(t); \\
 (a.4) \quad a_{4.2} &\doteq S_{K,1}^{-2} S_{K,2}^{-2} \left(\sum_{k=1}^{\infty} s_{k,5} e^{-1}_{,k} R(W) E_{2|4}(s; t) f(s) f^4(t) \right. \\
 &\quad \left. + \sum_{k=1}^{\infty} s_{k,4} e^{-1}_{,k} e^{-1}_{,k} R(W)^2 E_{2|4}(t; s) f^2(s) f^2(t) \right); \\
 (a.4) \quad a_{4.3} &\doteq S_{K,1}^{-2} S_{K,2}^{-2} S_6 E_{3|4}(s; t) f^2(s) f^4(t);
 \end{aligned}$$

b. For (b) and (c), it is required that $(k_1; i_1) = (k_2; i_2) = (k'_1; i'_1)$, define

$$E_{2|3}(s; s; t) = E \left[Y_1(T_1) Y_2(T_2) Y_3(T_3) \mid T_1 = s; T_2 = s; T_3 = t \right];$$

By easy calculation,

$$(b) + (c) \doteq S_{K,1}^{-2} S_{K,2}^{-2} (W) E_{2|3}(s; s; t) \sum_{k=1}^{\infty} s_{k,3} e^2_{,k} ;$$

c. For (d), we need $(k_1; i_1) = (k_2; i_2)$, then

$$(d) \doteq \frac{1}{4} S_{K,2}^{-2} (W) f''(t) g^2(s; s) f^2(s) f^4(t) \sum_{k=1}^{\infty} s_{k,4} e^4_{,k} ;$$

From the above arguments,

$$\begin{aligned}
 \text{Var } \hat{R}_1 &= E \hat{R}_1^2 - (E \hat{R}_1)^2 \\
 &= O_p \left[S_{K,1}^{-2} S_{K,2}^{-2} S_{K,3}^2 + \sum_k s_{k,3} e^{-1}_{,k} e^{-2}_{,k} + \sum_k s_{k,4} e^{-1}_{,k} e^{-1}_{,k} + \sum_k s_{k,5} e^{-1}_{,k} \right. \\
 &\quad \left. + S_{K,2}^{-2} \sum_k s_{k,4} e^4_{,k} \right];
 \end{aligned}$$

Note that

$$\begin{aligned} \text{Cov}(Q_{0R_1}; \hat{R}_1) &= S_{K,1}^{-2} S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^{m_{ki}} m_{ki}^2 (m_{ki} - 1) \int_{-1}^1 R(W) (s; t) f(s) f^3(t); \\ E(Q_{0R_1}) &= f(t) + \frac{1}{2} (W) f''(t) + o_p(\cdot); \\ \text{Var}(Q_{0R_1}) &= \frac{\cdot^{-1}}{S_{K,1}} R(W) f(t) + o_p \left(\frac{\cdot^{-1}}{S_{K,1}} \right); \end{aligned}$$

By the delta method,

$$\begin{aligned} \text{Var} \frac{\hat{R}_1}{Q_{0R_1}} &= \frac{\text{Var} \hat{R}_1}{E^2 Q_{0R_1}} + 2 \frac{E \hat{R}_1 \text{Cov}(Q_{0R_1}; \hat{R}_1)}{E^3 Q_{0R_1}} + \frac{E^2 \hat{R}_1 \text{Var} Q_{0R_1}}{E^4 Q_{0R_1}} \\ &= O_p \left(S_{K,1}^{-2} S_{K,2}^{-2} S_{K,3}^2 + \sum_k s_{k,3} e^{-1} e^{-2} + \sum_k s_{k,4} e^{-1} e^{-1} + \sum_k s_{k,5} e^{-1} \right) \\ &\quad + S_{K,2}^{-2} \sum_k s_{k,4} e^4; \end{aligned}$$

Now we give the proof of the asymptotic distribution of $e^{(K)}$ in Theorem 2.

Proof. Denote

$$f_1 = Q_{20} Q_{02} \quad Q_{11}^2; \quad f_2 = Q_{10} Q_{02} \quad Q_{01} Q_{11}; \quad f_3 = Q_{01} Q_{20} \quad Q_{10} Q_{11};$$

Let e be the local linear estimate based on $\hat{e}_{kijl} = (Y_{kij} - e_{kij})(Y_{kil} - e_{kil})$ which can be expressed as

$$e = \frac{f_1 R_{00} \quad f_2 R_{10} \quad f_3 R_{01}}{f_1 Q_{00} \quad f_2 Q_{10} \quad f_3 Q_{01}};$$

where for $p; q = 0; 1; 2$,

$$\begin{aligned} Q_{p;q} &= S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^{m_{ki}} \prod_{1 \leq j \neq l \leq m_{ki}} W_{\cdot;k}(T_{kij} - s) W_{\cdot;k}(T_{kil} - t) \frac{T_{kij} - s}{e_{\cdot;k}}^p \frac{T_{kil} - t}{e_{\cdot;k}}^q \\ &= \prod_k !_k Q_{pq;k}; \\ R_{p;q} &= S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^{m_{ki}} \prod_{1 \leq j \neq l \leq m_{ki}} W_{\cdot;k}(T_{kij} - s) W_{\cdot;k}(T_{kil} - t) \frac{T_{kij} - s}{e_{\cdot;k}}^p \frac{T_{kil} - t}{e_{\cdot;k}}^q e_{kijl} \\ &= \prod_k !_k R_{pq;k}; \end{aligned}$$

Let

$$\begin{aligned}
 e_{1;k} &= \frac{1}{e_{;k}} \frac{f_2(R_{10} \quad R_{00} Q_{10}=Q_{00})}{f_1 Q_{10} \quad f_2 Q_{10}^2=Q_{00} \quad f_3 Q_{01} Q_{10}=Q_{00}}; \\
 e_{2;k} &= \frac{1}{e_{;k}} \frac{f_3(R_{01} \quad R_{00} Q_{01}=Q_{00})}{f_1 Q_{01} \quad f_2 Q_{01} Q_{10}=Q_{00} \quad f_3 Q_{01}^2=Q_{00}}; \\
 e_{1;k}^* &= \frac{1}{e_{;k}} \frac{f_2(R_{10}^* \quad R_{00}^* Q_{10}=Q_{00})}{f_1 Q_{10} \quad f_2 Q_{10}^2=Q_{00} \quad f_3 Q_{01} Q_{10}=Q_{00}}; \\
 e_{2;k}^* &= \frac{1}{e_{;k}} \frac{f_3(R_{01}^* \quad R_{00}^* Q_{01}=Q_{00})}{f_1 Q_{01} \quad f_2 Q_{01} Q_{10}=Q_{00} \quad f_3 Q_{01}^2=Q_{00}}; \\
 R_{p;q}^* &= S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^K \prod_{1 \leq j \neq l \leq m_{ki}} W_{;k}(T_{kij} \quad s) W_{;k}(T_{kil} \quad t) \frac{T_{kij} \quad s^p}{e_{;k}} \frac{T_{kil} \quad t^q}{e_{;k}} C_{kijl}; \\
 \mathcal{E}_{pq} &= S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^K \prod_{1 \leq j \neq l \leq m_{ki}} W_{;k}(T_{kij} \quad s) W_{;k}(T_{kil} \quad t) T_{kij}^p T_{kil}^q = \prod_k \mathcal{E}_{pq;k}.
 \end{aligned}$$

Then we can write

$$\begin{aligned}
 e &= \frac{R_{00}}{Q_{00}} \prod_k \mathcal{E}_{10;k} \frac{s Q_{00}}{Q_{00}} \prod_k \mathcal{E}_{01;k} \frac{t Q_{00}}{Q_{00}}; \\
 - &= \frac{R_{00}}{Q_{00}} \frac{\partial}{\partial s} \frac{\mathcal{E}_{10} \quad s Q_{00}}{Q_{00}} \frac{\partial}{\partial t} \frac{\mathcal{E}_{01} \quad t Q_{00}}{Q_{00}}; \\
 -^* &= \frac{R_{00}^*}{Q_{00}} \frac{\partial}{\partial s} \frac{\mathcal{E}_{10} \quad s Q_{00}}{Q_{00}} \frac{\partial}{\partial t} \frac{\mathcal{E}_{01} \quad t Q_{00}}{Q_{00}}.
 \end{aligned}$$

Following similar arguments in the proof of Theorem 1, we can prove that

$$e^- = a_p \min_{k=1}^2 \frac{8}{e_{;k}} \frac{S_{k,3}}{e_{;k}} \frac{1}{2}; \quad S_{K,4} = S_{K,2}^2 \frac{1}{2} = 5.$$

Now we derive the asymptotic distribution of $-$. We first show that $-^* = O_p f(N_K)^{-1} g$.

(B.1) Calculate $-^*$.

From the definition of $-$ and $-^*$, we have $-^* = R_{00} = Q_{00}$. Note that

$$\begin{aligned}
 R_{00} &= S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^K \prod_{j \neq l} W_{kij}(s) W_{kil}(t) \mathcal{E}_{kijl} \\
 &= S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^K \prod_{j \neq l} W_{kij}(s) W_{kil}(t) (f(Y_{kij} \quad kij)(Y_{kil} \quad kil) + (Y_{kij} \quad kij)(e_{kil} \quad e_{kil}) \\
 &\quad + (e_{kij} \quad kij)(Y_{kil} \quad kil) + (e_{kij} \quad kij)(e_{kil} \quad kil)) g
 \end{aligned}$$

$$\begin{aligned} & \stackrel{\cdot}{=} S_{K,1}^{-1} S_{K,2}^{-1} Q_0^{-1} \prod_{k=1}^K \prod_{i=1}^k \prod_{j \neq l} \prod_{k'=1}^k \prod_{l'=1}^k \prod_{j'=1}^k W_{\cdot, k'}(T_{kij} \quad s) W_{\cdot, k'}(T_{kil} \quad t) \\ & W_{\cdot, k'}(T_{k'l'j'} \quad T_{kil})(Y_{kij} \quad (T_{kij})) Y_{k'l'j'} \quad (T_{k'l'j'}) + \\ & '(T_{kil}) \quad '(T_{k'l'j'}) (T_{k'l'j'} \quad T_{kil}) + \frac{1}{2} ''(T_{kil})(T_{k'l'j'} \quad T_{kil})^2 : \end{aligned}$$

Then take the Taylor expansion of $'(T_{k'l'j'})$ at $'(T_{kil})$ in the last bracket,

$$\begin{aligned} & S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^k \prod_{j \neq l} W_{kij}(s) W_{kil}(t) (Y_{kij} \quad kij) (\quad kil \quad e_{kil}) \\ & \stackrel{\cdot}{=} S_{K,1}^{-1} S_{K,2}^{-1} Q_0^{-1} \prod_{k=1}^K \prod_{i=1}^k \prod_{j \neq l} \prod_{k'=1}^k \prod_{l'=1}^k \prod_{j'=1}^k W_{\cdot, k'}(T_{kij} \quad s) W_{\cdot, k'}(T_{kil} \quad t) \\ & W_{\cdot, k'}(T_{k'l'j'} \quad T_{kil})(Y_{kij} \quad (T_{kij})) Y_{k'l'j'} \quad (T_{k'l'j'}) \frac{1}{2} ''(T_{kil})(T_{k'l'j'} \quad T_{kil})^2 : \end{aligned}$$

Denote

$$\begin{aligned} \Delta \mathcal{R}_1 &= S_{K,1}^{-1} S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^k \prod_{j \neq l} \prod_{k'=1}^k \prod_{l'=1}^k \prod_{j'=1}^k W_{\cdot, k'}(T_{kij} \quad s) W_{\cdot, k'}(T_{kil} \quad t) \\ & W_{\cdot, k'}(T_{k'l'j'} \quad T_{kil})(Y_{kij} \quad (T_{kij})) Y_{k'l'j'} \quad (T_{k'l'j'}) \frac{1}{2} ''(T_{kil})(T_{k'l'j'} \quad T_{kil})^2 ; \\ \Delta \mathcal{R}_2 &= S_{K,1}^{-1} S_{K,2}^{-1} \prod_{k=1}^K \prod_{i=1}^k \prod_{j \neq l} \prod_{k'=1}^k \prod_{l'=1}^k \prod_{j'=1}^k W_{\cdot, k'}(T_{kij} \quad s) W_{\cdot, k'}(T_{kil} \quad t) \\ & W_{\cdot, k'}(T_{k'l'j'} \quad T_{kij})(Y_{kil} \quad (T_{kil})) Y_{k'l'j'} \quad (T_{k'l'j'}) \frac{1}{2} ''(T_{kij})(T_{k'l'j'} \quad T_{kij})^2 ; \end{aligned}$$

$$Q_{0R1} = Q_0(T_{kil}); \quad Q_{0R2} = Q_0(T_{kij});$$

It can be written as

$$R_{00} \stackrel{\cdot}{=} R_{00}^* + \frac{\mathcal{R}_1}{Q_{0R1}} + \frac{\mathcal{R}_2}{Q_{0R2}};$$

The last terms in the right hand side are of the same order. By Lemma S.1, we have

$$\begin{aligned} E(\mathcal{R}_1 = Q_{0R1}) &= O_p(S_{K,1}^{-1} S_{K,2}^{-1} S_{K,3}); \\ \text{Var}(\mathcal{R}_1 = Q_{0R1}) &= O_p(S_{K,1}^{-2} S_{K,2}^{-2} S_{K,3}^2 (1 + S_{K,1}^{-1} \quad ; -1) + S_{K,3} \quad ; -1 \quad ; -2 \\ &+ S_4 \quad ; -1 \quad ; -1 + S_5 \quad ; -1 + S_6 + S_{K,2}^{-2} S_4 (\quad ; 2)^2) : \end{aligned}$$

Hence

$$\begin{aligned} - \quad -^* &= O_p(S_{K,1}^{-1} S_{K,2}^{-1} S_{K,3}^2 (1 + S_{K,1}^{-1} \quad ; -1) + S_{K,3} \quad ; -1 \quad ; -2 \\ &+ S_{K,4} \quad ; -1 \quad ; -1 + S_5 \quad ; -1 + S_6 \frac{1}{2} + S_{K,2}^{-1} S_4^{\frac{1}{2}} (\quad ; 2) : \end{aligned}$$

(B.2) Asymptotic distribution of R_{00}^* .

Define

$$R_{00}^* = \frac{R_{00}^*}{Q_{00}} = \frac{\mathbb{E} \mathbb{G}_{10}}{Q_{00}} + \frac{s Q_{00}}{Q_{00}} \frac{\mathbb{E} \mathbb{G}_{01}}{Q_{00}} + \frac{t Q_{00}}{Q_{00}};$$

The asymptotic normality of R_{00}^* is accomplished by the asymptotic joint normality of $(R_{00}^*, \mathbb{E} R_{00}^*, \mathbb{E} \mathbb{G}_{00}, \mathbb{E} \mathbb{G}_{10}, \mathbb{E} \mathbb{G}_{01}, \mathbb{E} \mathbb{G}_{10}, \mathbb{E} \mathbb{G}_{01})$. Specifically, let

$$E_1(s; t) = E f(Y_1 | (T_1))^2 (Y_2 | (T_2))^2 | T_1 = s; T_2 = tg;$$

$$E_2(s; t) = E f(Y_1 | (T_1))^2 (Y_2 | (T_2)) (Y_2 | (T_3)) | T_1 = s; T_2 = t; T_3 = tg;$$

By easy calculation, we have

$$\begin{aligned} E \mathbb{G}_{pq} &= s^p t^q f(s) f(t) + \frac{1}{2} (W) f s^p t^q f''(s) f(t) + s^p t^q f(s) f''(t) g^{-2} \\ &+ \frac{1}{2} (W) f^2 p f'(s) f(t) + 2 q f'(t) f(s) g^{-2} + o_p(g^{-2}); \end{aligned}$$

$$\begin{aligned} E R_{00}^* &= (s; t) f(s) f(t) + \frac{1}{2} (W)^{-2} \frac{\partial^2}{\partial s^2} (s; t) f(s) f(t) \\ &+ 2 \frac{\partial}{\partial s} (s; t) f'(s) f(t) + (s; t) f''(s) f(t) + \frac{\partial^2}{\partial t^2} (s; t) f(t) f(s) \\ &+ 2 \frac{\partial}{\partial t} (s; t) f'(t) f(s) + (s; t) f''(t) f(s) + o_p(g^{-2}); \end{aligned}$$

$$\begin{aligned} \text{Var}(R_{00}^*) &= f^2 + I(s=t) g^{-2} R(W)^2 E_1(s; t) f(s) f(t) S_{K;2}^{-2} \\ &+ R(W) f^2(s) f(t) E_2(t; s) + f(s) f(t)^2 E_2(s; t) S_{K;2}^{-1} \sum_{k=1}^{\infty} \frac{S_{k;3}}{e^{-k}} \\ &+ V_3(s; t) f^2(s) f^2(t) S_{K;2}^{-2} S_4 + o_p(S_{K;2}^{-2}) \\ &+ o_p(S_{K;2}^{-2}) \sum_{k=1}^{\infty} \frac{S_{k;3}}{e^{-k}} + o_p(S_{K;2}^{-2} S_4); \end{aligned}$$

$$\begin{aligned} \text{Cov}(\mathbb{G}_{pq}; \mathbb{G}_{p'q'}) &= f^2 + I(s=t) g^{p+p'} t^{q+q'} R(W)^2 f(s) f(t) S_{K;2}^{-2} \\ &+ R(W) f f^2(s) f(t) + f(s) f^2(t) g S_{K;2}^{-2} \sum_{k=1}^{\infty} \frac{S_{k;3}}{e^{-k}} \\ &+ o_p(S_{K;2}^{-2}) + o_p(S_{K;2}^{-2}) \sum_{k=1}^{\infty} \frac{S_{k;3}}{e^{-k}}; \end{aligned}$$

$$\text{Cov}(\mathbb{G}_{pq}; R_{00}^*) = f^2 + I(s=t) g s^p t^q R(W)^2 (s; t) f(s) f(t) S_{K;2}^{-2}$$

$$\begin{aligned}
& + R(W) (s; t) f f^2(s) f(t) + f(s) f^2(t) g S_{K;2}^{-2} \sum_{k=1}^K \frac{S_{k;3}}{e_{;k}} \\
& + o_p \left(S_{K;2}^{-2} \right) + o_p \left(S_{K;2}^{-2} \sum_{k=1}^K \frac{S_{k;3}}{e_{;k}} \right) :
\end{aligned}$$

From the delta method, $\hat{\mu}(t)$ follows the asymptotic normality that

$$\hat{\mu}^{-1} = \hat{\mu}^{-1}(s; t) - \hat{\mu}^{-1}(s; t) \frac{1}{2} \frac{\partial^2}{\partial s^2} (W) \frac{\partial^2}{\partial s^2} (s; t) + \frac{\partial^2}{\partial t^2} (s; t) + o_p \left(\frac{1}{n} \right) \stackrel{d}{\rightarrow} N(0; 1);$$

where

$$\begin{aligned}
\hat{\mu}^{-1} = & f(s) + I(s=t) g \left(S_{K;2}^{-1} \frac{V_1(s; t)}{f(s) f(t)} \right. \\
& \left. + S_{K;2}^{-2} \sum_{k=1}^K \frac{S_{k;3}}{e_{;k}} R(W) \frac{f(s) V_2(t; s) + f(t) V_2(s; t)}{f(s) f(t)} + V_3(s; t) S_{K;2}^{-2} \right)
\end{aligned}$$

The proof is completed combining the above arguments. ■

S.3. PROOF OF THEOREM 3

We mention that Theorem 3 and Theorem 4 can be extended to the standard d -dimensional local linear regression. To see this, we introduce the following notations and assumptions.

Notations and assumptions for d -dimensional local linear regression. The underlying model is $Y = m(X) + \epsilon$ where $X \in \mathbb{R}^d$, $E\epsilon = 0$ and $\text{Var } \epsilon = \sigma^2(X)$. Suppose we observe $f(Y_{ki}; X_{ki} : i = 1, \dots, n_k; k = 1, \dots, K)g$ in the k th block and $N_K = \sum_{k=1}^K n_k$, i.e., we use n_k to denote the sub-sample size of the K th data block and $N_K = \sum_{k=1}^K n_k$ to denote the full sample size up to time K , which correspond to $S_{K;1}; S_{K;1}$ in (6) for mean estimation and $S_{K;2}; S_{K;2}$ for covariance estimation, respectively. We impose the following assumptions which are in parallel to (A.1)-(A.3) and (A.6).

(B.1) Observations $f(X_{ki} : i = 1, \dots, n_k; k = 1, \dots, K)g$ are i.i.d. copies of a random variable X defined on $[0; 1]^d$ whose density $f(\cdot)$ is bounded away from 0 with bounded second derivative.

(B.2) The second and fourth derivatives of regression function $m(x)$ are bounded and continuous.

(B.3) Noises ϵ_{ki} are i.i.d. with mean 0 and variance $\sigma^2(x)$.

(B.4) The block size $K n_k = N_K \rightarrow 1$ as $K \rightarrow 1$ for $k = 1, 2, \dots, K$.

The optimal bandwidth at time K is $h_*^{(K)} = N_K^{-1/(4+d)}$. The candidate sequence is $f_i^{(K)} g_{l=1}^l$ and the pseudo-bandwidths are $\hat{f} e_k^{(K)} : k = 1, \dots, K$ for $j = 1, \dots, d$. The centroids for the candidate sequence are $f_i^{(K)} : i = 1, \dots, L$. Further define

$$f_i^{(K)} = \sum_{k=1}^K \frac{n_k}{N_K} e_k^{(K)} \quad i = 1, \dots, L$$

We also mention that the mean and covariance estimates corresponds the cases $d = 1$ and $d = 2$, respectively.

When estimating m , all observations are pooled together and the index of subjects become invalid, which is intrinsically one-dimensional local linear regression. Hence one can derive the following result for standard d -dimensional local linear regression by the same arguments in Theorem 1.

Lemma S. 2. *Suppose assumptions (B.1)-(B.3) and (A.8)-(A.9) hold. If*

$$\hat{h}_*^{(K)} = h_*^{(K)} = o_p(N_K^{-1/(4+d)}) \quad ; \quad j = 1, \dots, d; \quad (S.3)$$

then a fixed interior point $x \in (0, 1)^d$, as $K \rightarrow 1$, the online estimate $\hat{m}^{(K)}(x)$ satisfies

$$\hat{m}^{(K)}(x) - m(x) = O_p\left(\frac{1}{N_K} \sum_{j=1}^d \frac{R(W)^{d-2} f_j''(x)}{f(x)}\right) + o_p\left(\frac{1}{N_K}\right)$$

We prove in Section S.3 that (S.3) is satisfied by the proposed online method.

Denote $X = \{X_{ki} \in \mathbb{R}^d : k = 1, \dots, K; i = 1, \dots, n_k\}$, $\int_{\mathbb{R}^d} f^2 m(x) = \int_{\mathbb{R}^d} x_j^2 g f(x) dx$ and $\int_{[0,1]^d} R(W)^{d-2} f(x) dx$. Similarly, we adopt another online local cubic and online local linear to estimate $\hat{m}^{(K)}$ and $\hat{e}^{(K)}$, which are denoted as $\hat{e}^{(K)}$ and $e^{(K)}$, respectively. Then the optimal bandwidth and the online estimated bandwidth for $m(\cdot)$ are

$$h_*^{(K)} = \frac{1}{2(W)} N_K^{-1/(4+d)}; \quad \hat{h}_*^{(K)} = \frac{e^{(K)}}{2(W) \hat{e}^{(K)}} N_K^{-1/(4+d)}; \quad (S.4)$$

The candidate sequence for $\hat{m}^{(K)}$ and $\hat{e}^{(K)}$ are $f_{:,j}^{(K)} g_{l=1}^l$ and $f_{:,j}^{(K)} g_{l=1}^l$, respectively, and the pseudo-bandwidths are $\hat{f} e_{:,k}^{(K)} : k = 1, \dots, K$ and $\hat{f} e_{:,k}^{(K)} : k = 1, \dots, K$. Then we have the following conclusion.

Lemma S. 3. Suppose assumptions (B.1)-(B.4) and (A.8)-(A.9) hold. When the online bandwidths and candidate sequence for $\hat{\theta}^{(K)}$ are

$$h^{(K)} = GN_K^{-1/(6+d)}; \quad h^{(K)} = RN_K^{-1/(4+d)}; \quad 1 < G; R < 1;$$

$$\hat{\theta}_{j,1}^{(K)} < \hat{\theta}_{j,1}^{(K)} = h^{(K)}; \quad \hat{\theta}_{j,1}^{(K)} < \hat{\theta}_{j,1}^{(K)} = h^{(K)};$$

then as $K \rightarrow \infty$, $\hat{\theta}^{(K)}$ in (S.4) satisfies

$$\hat{\theta}^{(K)} \rightarrow h_*^{(K)}$$

and then $e^{(K)} = O_p(S_{K,1}^{-2-5})$. For \hat{e}_K , one can prove that the convergence rate of $e^{(K)}$ is dominated by $E_K(s)$ in step 3 of case (b) in Appendix B which equals $O_p(S_{K,2}^{-1-3})$. For dense data, we need only check the order of e_K^2 . From Zhang and Chen (2007), \hat{e}_{kij} in step 2 of (c) in Appendix B is of order $(S_{K,1})^{-2-25}$, recall that \hat{e}_{kij}^0 is a debiased version of \hat{e}_{kij} and e_K^2 is the average of \hat{e}_{kij}^0 , then $e_K^2 = O_p(S_{K,1}^{-12-25})$. In all the cases, estimates of \hat{e}_K are dominated by e_K , which completes the proof of Theorem 3.

S.4. PROOF OF THEOREM 4

Note that the asymptotic distributions of \hat{e} and e have the consistent form with the standard d -dimensional local linear regression, i.e., the bias is of $O_p(h^2)$ and the variance is of $O_p(N^{-1}h^{-d})$, where $d = 1$ for mean estimation and $d = 2$ for covariance estimation. The lower bound of the relative efficiency also holds for the standard d -dimensional local linear regression. We adopt the notations of the d -dimensional local linear in the proof for notation conciseness.

Proof. Recall that $\hat{h}^{(K)} = (h_1^{(K)}, \dots, h_L^{(K)})$, without loss of generality, we assume $h_1^{(K)} = g(l)\hat{h}^{(K)}$ where $g(l) = g(1) = 1, l = 1, \dots, L$. Note that $\hat{h}^{(K)} = h^{(K)} = o_p(N_K^{-1-(d+4)})$ and the candidate bandwidth sequence satisfies

$$\frac{e_k^{(K)}}{\hat{h}^{(K)}} = g(l) \frac{N_k}{N_K}^{-\frac{1}{d+2}} + o_p(1):$$

Write $g(l) = \tilde{f}g(l)^{1-} = g$ and note that $N_k = N_K$ grows linearly with respect to k , then \tilde{f} shall be $1-(d+2)$ and the optimal $g(l)^{1-}$ shall be linear between $(0; 1)$. Hence, the optimal \tilde{f} is as in (19).

Next we derive the asymptotic lower bound for the relative efficiency of our local linear estimator. Writing $\hat{h}^{(K)} = h_*^{(K)}(1 + O_p(N_K^{-2-(d+6)}))g$, one can derive

$$eff(\hat{e})^{-1} = \frac{d}{d+4} \sum_{k=1}^8 \frac{n_k}{N_K} \frac{e_k^{(K)}}{\hat{h}^{(K)}} + \frac{4}{d+4} \sum_{k=1}^8 \frac{n_k}{N_K} \frac{e_k^{(K)}}{\hat{h}^{(K)}} + O_p(N_K^{-\frac{2}{6+d}}) \quad (S.5)$$

As shown in Figure 1, when K tends large, there exists a breakpoint K_0 : when $k \leq K_0$, e_k equals the last candidate $(1-L)^{1-(d+4)}\hat{h}_k$ for the limit number of candidates and for $k > K_0$,

there are sufficient and close candidates to make a choice. Hence

$$\frac{e_k^{(K)}}{\mathfrak{H}(K)} = \begin{cases} \asymp L^{-\frac{1}{d+4}} \frac{N_k}{N_K}^{-\frac{1}{d+4}} + O_p(N_k^{-\frac{2}{6+d}}); & k \leq K_0 \\ \asymp 1 + \frac{N_k}{N_K} \frac{l_k}{L} O^{\frac{1}{d+4}} + O_p(N_k^{-\frac{2}{6+d}}); & k > K_0; \end{cases}$$

where $l_k = \operatorname{argmin}_{1 \leq l \leq L} jN_k = N_K \quad l=Lj$. Then when $k \leq K_0$,

$$\begin{aligned} \sum_{k=1}^{K_0} \frac{n_k}{N_K} \frac{e_k^{(K)}!^{-2}}{\mathfrak{H}(K)} &= \frac{d+4}{d+2} L^{-\frac{2}{d+4}} \frac{N_{K_0}}{N_K}^{1-\frac{2}{d+4}} + O_p(N_k^{-\frac{2}{6+d}}); \\ \sum_{k=1}^{K_0} \frac{n_k}{N_K} \frac{e_k^{(K)}!^{-d}}{\mathfrak{H}(K)} &= \frac{d+4}{2d+4} L^{\frac{d}{d+4}} \frac{N_{K_0}}{N_K}^{1+\frac{d}{d+4}} + O_p(N_k^{-\frac{2}{6+d}}); \end{aligned}$$

and when $k > K_0$,

$$\begin{aligned} \sum_{k=K_0+1}^K \frac{n_k}{N_K} \frac{e_k^{(K)}!^{-2}}{\mathfrak{H}(K)} &= 1 \frac{N_{K_0}}{N_K} + O_p(N_k^{-\frac{2}{6+d}}); \\ \sum_{k=K_0+1}^K \frac{n_k}{N_K} \frac{e_k^{(K)}!^{-1}}{\mathfrak{H}(K)} &= 1 \frac{N_{K_0}}{N_K} + O_p(N_k^{-\frac{2}{6+d}}); \end{aligned}$$

The property of breakpoint K_0 also guarantees that

$$\frac{1}{L} \sum_{k=1}^{K_0} \frac{n_k}{N_{K_0}} \frac{e_k^{(K)}!^{-1}}{\mathfrak{H}(K)} \asymp \frac{N_{K_0}}{N_K};$$

which is equivalent to

$$\frac{1}{L} \sum_{k=1}^{K_0} \frac{n_k}{N_{K_0}} \frac{e_k^{(K)}!^{-1}}{\mathfrak{H}(K)} \frac{h_*^{(K)}}{h_*^{(K)}} = \frac{1}{L} \sum_{k=1}^{K_0} \frac{n_k}{N_{K_0}} \frac{e_k^{(K)}!^{-1}}{\mathfrak{H}(K)} \frac{h_*^{(K)}}{h_*^{(K)}} + O_p(N_k^{-\frac{2}{6+d}}) \frac{N_k}{N_K} \frac{1}{d+4};$$

and under Assumption (A.5), we obtain

$$\frac{N_{K_0}}{N_K} \frac{d+4}{d+3} \frac{1}{L} + O_p(N_k^{-\frac{2}{6+d}});$$

Denote $\bar{N} = N_{K_0} = N_K$, one can derive from (S.5) that

$$\begin{aligned} \operatorname{eff}(\mathfrak{H})^{-1} &= \frac{d}{d+4} \frac{d+4}{d+2} L^{-\frac{2}{d+4}} \frac{1}{L} + 1 \\ &\quad + \frac{4}{d+4} \frac{d+4}{2d+4} L^{\frac{d}{d+4}} \frac{1}{L} + 1 + O_p(N_k^{-\frac{2}{6+d}}); \end{aligned} \quad (\text{S.6})$$

Note that (S.6) is strictly increasing with respect to $h \geq 0; f(d+4) = (d+3)g^{d+4} = L^i$. Hence we have

$$eff(\mathbb{M})^{-1} = 1 + \frac{2c_1 d + 4c_2}{d+4} L^{-1} + \frac{d}{d+4} c_1^2 L^{-2};$$

where

$$c_1 = \frac{(d+4)^{d+3}}{(d+3)^{d+2}(d+2)} \quad \frac{d+4}{d+3}^{d+4}; \quad c_2 = \frac{(d+4)^{2d+5}}{(d+3)^{2d+4}(2d+4)} \quad \frac{d+4}{d+3}^{d+4};$$

This completes the proof of Theorem 4. ■

REFERENCES

- Fan, J. and Yao, Q. (1998), 'Efficient estimation of conditional variance functions in stochastic regression', *Biometrika* **85**(3), 645{660.
- Wand, M. P. and Jones, M. C. (1994), *Kernel Smoothing*, Chapman and Hall.
- Zhang, J. and Chen, J. (2007), 'Statistical inferences for functional data', *The Annals of Statistics* **35**(3), 1052{1079.
- Zhang, X. and Wang, J.-L. (2016), 'From sparse to dense functional data and beyond', *The Annals of Statistics* **44**(5), 2281{2321.