Online Supplemental Appendix for VARYING COEFFICIENT MODELS FOR DATA WITH AUTO-CORRELATED ERROR PROCESS

Zhao Chen, Runze Li and Yan Li

Princeton University

Pennsylvania State University

eBay Inc

Denote $\mathbf{F} = (\mathbf{f}_{d+1}, \ldots, \mathbf{f}_n)^T$ with $\mathbf{f}_t = (\varepsilon_{t-1}, \ldots, \varepsilon_{t-d})^T$, and $\mathbf{E} = (\mathbf{e}_{d+1}, \ldots, \mathbf{e}_n)^T$ with $\mathbf{e}_t = (\widehat{\varepsilon}_{t-1}, \ldots, \widehat{\varepsilon}_{t-d})^T$, where $\widehat{\varepsilon}_t$ is the estimated residual in the initial step when the profile least squares method is implemented. Define $\boldsymbol{\Delta} = \mathbf{E} - \mathbf{F}$. Our proof follows a similar strategy to that used in Fan and Huang (2005) and Li and Li (2009). Note that the proof of Fan and Huang (2005) is for iid data, and the proof of Li and Li (2009) is for nonparametric regression models rather than varying coefficient models. The following conditions are imposed to facilitate the proof and are adopted from Fan and Huang (2005). They are not the weakest possible conditions.

A. The random variable $\{u_t\}$ has a bounded support Ω. Its density function $g(\cdot)$ is Lipschitz continuous with order $\gamma \geq 2$ and bounded away from 0 on its support. That is,

$$|g(x_1) - g(x_2)| \leq C|x_1 - x_2|^{\gamma},$$

for some constants $C > 0$.

B. There is an $s > 2$ such that $E||\mathbf{f}_t||^{2s} < \infty$ and $E||\mathbf{X}_t||^{2s} < \infty$ and for some $(1 - s^{-1})/2 < \xi < 2 - s^{-1}$ such that $n^{1-2s^{-1}-2\xi}h \to \infty$.

C. The $p \times p$ matrix $E(\mathbf{X}_t\mathbf{X}_t^T|\mathbf{f}_t)$ is non-singular for each $\mathbf{f}_t \in \Omega$. $E(\mathbf{X}_t\mathbf{X}_t^T|\mathbf{f}_t)$ and $E(\mathbf{X}_t\mathbf{X}_t^T|\mathbf{f}_t)^{-1}$ are all Lipschitz continuous with order $\gamma \geq 2$.

D. $\{\alpha_j(\cdot), j = 0, \ldots, p\}$ have continuous second derivatives in $u \in \Omega$.

E. The kernel function $K(\cdot)$ is a bounded symmetric density function with bounded support $[-\kappa,\kappa]$, satisfying a Lipschitz condition.

F. $nh^8 \to 0$ and $nh^2/(\log n)^2 \to \infty$.

G. $\sup_{u_t\in\Omega} |\tilde{\alpha}_j(u_t) - \alpha_j(u_t)| = o_p(n^{-\frac{1}{4}})$ for all $j = 0,\ldots,p$, where $\tilde{\alpha}_j(u_t)$ is the local linear estimator pretending that data are i.i.d..

H. The sequence of random vectors $(u_t, \mathbf{X}_t^T, \varepsilon_t)$, $t = 1, 2, \cdots$, is strictly stationary (Fan and Yao (2003)) and satisfies the following conditions for α-mixing processes (Fan and Yao (2003)):

$$\sum_l l^a[\alpha(l)]^{1-2/\delta} < \infty, \quad E|\varepsilon_1|^\delta < \infty, \quad E|\mathbf{X}_1\mathbf{X}_1^T|^\delta < \infty,$$

$$g_{u_1|\varepsilon_1}(u|\varepsilon) \le C_1 < \infty, \quad g_{u_1|\mathbf{X}_1}(u|\mathbf{X}) \le C_2 < \infty$$

with some $\delta > 2$, $a > 1 - 2/\delta$ and positive constants C_1 and C_2, where

$$\alpha(n) = \sup_{A\in\mathcal{F}_{-\infty}^0, B\in\mathcal{F}_n^\infty} |P(A)P(B) - P(AB)|. \qquad \text{(A.1)}$$

with $\mathcal{F}_{-\infty}^0$ being a σ-field generated by $\{(u_t, \mathbf{X}_t^T, \varepsilon_t) : t \le 0\}$ and $\mathcal{F}_n^\infty$ a σ-field generated by $\{(u_t, \mathbf{X}_t^T, \varepsilon_t) : t \ge n\}$.

Lemma 1. *Let $(u_1,\varepsilon_1),\ldots,(u_n,\varepsilon_n)$ be a strictly stationary sequence satisfying the mixing condition $\alpha(l) \le cl^{-\tau}$ for some $c > 0$ and $\tau > 5/2$. Assume further that for some $s > 2$ and interval $[a,b]$,*

$$E|\varepsilon_t|^s < \infty \quad \text{and} \quad \sup_{\forall x\in[a,b]} \int |\varepsilon_t|^s g(u,\varepsilon)\mathrm{d}\varepsilon < \infty,$$

where $g(\cdot,\cdot)$ denotes the joint density of (u_t,ε_t).

In addition, Condition H holds, and the conditional density $g_{u_1,u_l|\varepsilon_1,\varepsilon_l}(u_1,u_l|\varepsilon_1,\varepsilon_l) \le A_2 < \infty, \forall l \ge 1$. Let K satisfy Condition E. Then

$$\sup_{u\in[a,b]} |\frac{1}{n}\sum_{i=1}^n \{K_h(u_i-u)\varepsilon_i - E[K_h(u_i-u)\varepsilon_i]\}| = O_p(\left\{\frac{\log n}{nh}\right\}^{1/2})$$

provided that $h \to 0$, for some $\xi > 0$, $n^{1-2s^{-1}-2\xi}h \to \infty$ and $n^{(\tau+1.5)(s^{-1}+\xi)-\tau/2+5/4}h^{-\tau/2-5/4} \to$

0.

This lemma is extracted from Fan and Yao (2003) and will be used in our proof repeatedly.

Lemma 2. *Suppose that Conditions A—H hold. It follows that*

$$\frac{1}{n}\mathbf{E}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{E} \xrightarrow{P} E(\mathbf{f}\mathbf{f}^T). \tag{A.2}$$

$$\frac{1}{n}\mathbf{E}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{M} = O_p(\{h^2+\sqrt{\log(n)/nh}\}^2) \tag{A.3}$$

$$\frac{1}{n}\mathbf{E}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\boldsymbol{\Delta\beta} = o_p(n^{-1/4}\{h^2+\sqrt{\log(n)/nh}\}^2) \tag{A.4}$$

Proof To prove (A.2), we first show that

$$\frac{1}{n}\mathbf{F}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{F} \xrightarrow{P} E(\mathbf{f}\mathbf{f}^T). \tag{A.5}$$

Denote W_u to be a $(n-d)\times(n-d)$ diagonal matrix with i-th diagonal element $K_h(u_i-u)$ and

$$D_u = \begin{pmatrix} \mathbf{X}_{d+1}^T & \frac{u_{d+1}-u}{h}\mathbf{X}_{d+1}^T \\ \vdots & \vdots \\ \mathbf{X}_n^T & \frac{u_n-u}{h}\mathbf{X}_n^T \end{pmatrix},$$

Then the smoothing matrix $\mathbf{S}_h$ for the local linear regression can be expressed as

$$\mathbf{S}_h = \begin{pmatrix} [\mathbf{X}_{d+1}^T, 0]\{D_{u_{d+1}}^T W_{u_{d+1}} D_{u_{d+1}}\}^{-1} D_{u_{d+1}}^T W_{u_{d+1}} \\ \vdots \\ [\mathbf{X}_n^T, 0]\{D_{u_n}^T W_{u_n} D_{u_n}\}^{-1} D_{u_n}^T W_{u_n} \end{pmatrix},$$

where

$$D_u^T W_u D_u = \begin{pmatrix} \sum_{i=d+1}^n \mathbf{X}_i\mathbf{X}_i^T K_h(u_i-u) & \sum_{i=d+1}^n \mathbf{X}_i\mathbf{X}_i^T(\frac{u_i-u}{h})K_h(u_i-u) \\ \sum_{i=d+1}^n \mathbf{X}_i\mathbf{X}_i^T(\frac{u_i-u}{h})K_h(u_i-u) & \sum_{i=d+1}^n \mathbf{X}_i\mathbf{X}_i^T(\frac{u_i-u}{h})^2K_h(u_i-u) \end{pmatrix},$$

Each element of matrix $D_u^T W_u D_u$ is a kernel regression. Denote $G = \frac{1}{n}D_u^T W_u D_u$. Let $\mathbf{X}$ be a $p+1$-dimensional random vector whose distribution is the same as that of $\mathbf{X}_t$. With Lemma 1 and the symmetry of kernel function $K(\cdot)$, it holds uniformly in u that G equals a diagonal matrix with the 1st and 2nd diagonal elements $g(u)E(\mathbf{X}\mathbf{X}^T|u)(1+O_p(h^2+\sqrt{\log(n)/nh}))$ and $\mu_2 g(u)E(\mathbf{X}\mathbf{X}^T|u)(1+O_p(h^2+\sqrt{\log(n)/nh}))$, respectively. Thus the inverse of G is also a diagonal matrix with the 1st and 2nd diagonal elements being $[g(u)E(\mathbf{X}\mathbf{X}^T|u)]^{-1}(1+O_p(h^2+\sqrt{\frac{\log(n)}{nh}}))$ and $[\mu_2 g(u)E(\mathbf{X}\mathbf{X}^T|u)]^{-1}(1+O_p(h^2+\sqrt{\frac{\log(n)}{nh}}))$, respectively.

By the independence assumption of $(u_t,\mathbf{X}_t^T)$ and ε_t, we get $E(\mathbf{X}_t\mathbf{f}_t^T|u)=0$. Following a similar argument, we have

$$\frac{1}{n}D_u^T W_u \mathbf{F} = \begin{pmatrix} O_p(\sqrt{\frac{\log n}{nh}}) \\ O_p(\sqrt{\frac{\log n}{nh}}) \end{pmatrix}$$

holds uniformly in u. Consequently,

$$[\mathbf{X}^T,0]\{D_u^T W_u D_u\}^{-1}\{D_u^T W_u \mathbf{F}\} = \mathbf{X}^T[g(u)E(\mathbf{X}\mathbf{X}^T|u)]^{-1}O_p(\sqrt{\frac{\log n}{nh}})(1+o_p(1))$$

Substituting this result into the smoothing matrix $\mathbf{S}_h$, we have

$$\mathbf{S}_h\mathbf{F} = \begin{pmatrix} [\mathbf{X}_{d+1}^T,0]\{D_{u_{d+1}}^T W_{u_{d+1}} D_{u_{d+1}}\}^{-1} D_{u_{d+1}}^T W_{u_{d+1}}\mathbf{F} \\ \vdots \\ [\mathbf{X}_n^T,0]\{D_{u_n}^T W_{u_n} D_{u_n}\}^{-1} D_{u_n}^T W_{u_n}\mathbf{F} \end{pmatrix} = \begin{pmatrix} O_p(\sqrt{\frac{\log n}{nh}}) \\ \vdots \\ O_p(\sqrt{\frac{\log n}{nh}}) \end{pmatrix}.$$

Thus,

$$\mathbf{F}-\mathbf{S}_h\mathbf{F} = \mathbf{F}\{1+O_p(\sqrt{\frac{\log n}{nh}})\}.$$

Finally, by the stationarity and ergodicity,

$$\frac{1}{n}\mathbf{F}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{F} = \left(\frac{1}{n}\sum_{i=d+1}^{n}\mathbf{f}_i\mathbf{f}_i^T\right)\{1+O_p(\sqrt{\frac{\log n}{nh}})\} \xrightarrow{P} E(\mathbf{f}\mathbf{f}^T)$$

Thus, (A.5) holds.

Note that $\boldsymbol{\Delta}=\mathbf{E}-\mathbf{F}$, and the generic element of $\boldsymbol{\Delta}$ is of the form $\sum_{j=0}^{p}[\tilde{\alpha}_j(u_t)x_{tj}-\alpha_j(u_t)x_{tj}]$. By condition F: $\sup_{u\in\Omega}|\tilde{\alpha}_j(u_t)-\alpha_j(u_t)|=o_p(n^{-\frac{1}{4}})$ and the assumption that x_{tj} is bounded, $\boldsymbol{\Delta}$ is of order $o_p(n^{-1/4})$ uniformly in u. We observe that

$$\frac{1}{n}\mathbf{E}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{E} = \frac{1}{n}(\mathbf{F}+\boldsymbol{\Delta})^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)(\mathbf{F}+\boldsymbol{\Delta}).$$

By using an argument similar to the proof of (A.5), it can be shown that

$$\frac{1}{n}\mathbf{E}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{E} = \frac{1}{n}\mathbf{F}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{F} + o_p(n^{-1/4}).$$

Thus, (A.2) follows by (A.5).

We next show (A.3). To this end, we first show that

$$\frac{1}{n}\mathbf{F}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{M} = O_p(\{h^2+\sqrt{\log(n)/nh}\}^2). \tag{A.6}$$

It is noted that

$$\begin{aligned} &\frac{1}{n}\mathbf{F}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{M} \qquad\qquad (A.7)\\ =\;& \frac{1}{n}\sum_{i=d+1}^{n}[\mathbf{f}_i-(\mathbf{S}_h\mathbf{f})_i][\mathbf{X}_i^T\boldsymbol{\alpha}(u_i)-[\mathbf{X}_i^T,0]\{D_{u_i}^T W_{u_i} D_{u_i}\}^{-1}D_{u_i}^T W_{u_i}\mathbf{M}]. \end{aligned}$$

Similar to the argument in the proof of (A.5), we can show that

$$[\mathbf{X}^T,0]\{D_u^T W_u D_u\}^{-1}\{D_u^T W_u\mathbf{M}\} = \mathbf{X}^T\boldsymbol{\alpha}(u)(1+O_p(h^2+\sqrt{\log(n)/nh}))$$

holds uniformly in $u\in\Omega$. Plugging this into (A.7), we have $(\mathbf{f}_i-(\mathbf{S}_h\mathbf{f})_i)^T = \mathbf{f}_i^T(1+O_p(\sqrt{\log(n)/nh}))$. Hence, it follows that

$$\begin{aligned} &\frac{1}{n}\mathbf{F}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{M}\\ &=\frac{1}{n}\sum_{i=d+1}^{n}[\mathbf{f}_i-(\mathbf{S}_h\mathbf{f})_i][\mathbf{X}_i^T\boldsymbol{\alpha}(u_i)-\mathbf{X}_i^T\boldsymbol{\alpha}(u_i)(1+O_p(h^2+\sqrt{\log(n)/nh})]\\ &=\frac{1}{n}\sum_{i=d+1}^{n}\mathbf{f}_i\mathbf{X}_i^T\boldsymbol{\alpha}(u_i)[1+O_p(\sqrt{\log(n)/nh})]O_p(h^2+\sqrt{\log(n)/nh})\\ &=O_p([h^2+\sqrt{\log(n)/nh}]^2) \end{aligned}$$

Note that $E\{\mathbf{f}_i\mathbf{X}_i^T\boldsymbol{\alpha}(u_i)\}=0$ because of the independence between $\mathbf{f}_i$ and $(u_i,\mathbf{X}_i^T)$, and covariance matrix for $\{\mathbf{f}_i\mathbf{X}_i^T\boldsymbol{\alpha}(u_i)\}$ is finite. This leads to $\frac{1}{n}\sum_{i=d+1}^{n}\mathbf{f}_i\mathbf{X}_i^T\boldsymbol{\alpha}(u_i)=O_p(1/\sqrt{n})$ Thus, the last equation holds by Condition F. Thus, (A.6) holds

Since $\mathbf{E}=\mathbf{F}+\boldsymbol{\Delta}$, we can break $\frac{1}{n}\mathbf{E}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{M}$ into two terms:

$\frac{1}{n}\mathbf{F}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{M}$, which is $o_p(1)$ by (A.6), and $\frac{1}{n}\boldsymbol{\Delta}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{M}$, which is also $o_p(1)$ as $\boldsymbol{\Delta} = o_p(n^{-1/4})$. Thus, (A.3) holds.

The proof of Lemma 2 is completed since (A.4) is a direct result from the proof of (A.3).

Lemma 3. *Suppose that Conditions A—H hold. It follows that*

$$\sqrt{n}\,[\mathbf{E}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{E}]^{-1}\mathbf{E}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\boldsymbol{\eta} \xrightarrow{L} N(0,\sigma^2\{E(\mathbf{f}\mathbf{f}^T)\}^{-1})$$

<u>**Proof**</u> By Lemma 2, $\frac{1}{n}\mathbf{E}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\mathbf{E} \xrightarrow{P} E(\mathbf{f}\mathbf{f}^T)$. Using the Slutsky theorem, it suffices to show that $n^{-1/2}\mathbf{E}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\boldsymbol{\eta} \xrightarrow{L} N(0,\sigma^2\{E(\mathbf{f}\mathbf{f}^T)\})$. Since $\mathbf{E} = \mathbf{F} + \boldsymbol{\Delta}$, we may write

$$\mathbf{E}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\boldsymbol{\eta} = \mathbf{F}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\boldsymbol{\eta} + \boldsymbol{\Delta}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\boldsymbol{\eta}.$$

Note that $\boldsymbol{\Delta} = o_p(n^{-1/4})$ by Condition G, it can be shown that

$$\begin{aligned}
& n^{-1/2}\boldsymbol{\Delta}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\boldsymbol{\eta} \\
= \; & n^{-1/2}\sum_{i=d+1}^{n}\boldsymbol{\Delta}_i[\eta_i - [\mathbf{X}_i^T,0]\{D_{u_i}^T W_{u_i} D_{u_i}\}^{-1}D_{u_i}^T W_{u_i}\boldsymbol{\eta}][1+O_p(\sqrt{\frac{\log(n)}{nh}})] \\
= \; & n^{-1/2}\sum_{i=d+1}^{n}\boldsymbol{\Delta}_i\eta_i[1+O_p(\sqrt{\frac{\log(n)}{nh}})] = o_p(n^{-1/4})
\end{aligned}$$

Thus, it is enough to show that

$$\frac{1}{\sqrt{n}}\mathbf{F}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\boldsymbol{\eta} \to N(0,\sigma^2 E(\mathbf{f}\mathbf{f}^T)). \tag{A.8}$$

We observe that

$$\mathbf{F}^T(I-\mathbf{S}_h)^T(I-\mathbf{S}_h)\boldsymbol{\eta} = \sum_{i=d+1}^{n}\mathbf{f}_i[\eta_i - [\mathbf{X}_i^T,0]\{D_{u_i}^T W_{u_i} D_{u_i}\}^{-1}D_{u_i}^T W_{u_i}\boldsymbol{\eta}][1+o_p(1)]. \tag{A.9}$$

By using Lemma 1 on $\{u_i, \eta_i\}$, we can show that

$$[\mathbf{X}^T,0]\{D_u^T W_u D_u\}^{-1}\{D_u^T W_u\boldsymbol{\eta}\} = \mathbf{X}^T[g(u)E(\mathbf{X}\mathbf{X}^T|u)]^{-1}O_p\left(\sqrt{\frac{\log(n)}{nh}}\right)$$

Then $\eta_i - [\mathbf{X}_i^T, 0]\{D_{u_i}^T W_{u_i} D_{u_i}\}^{-1} D_{u_i}^T W_{u_i} \boldsymbol{\eta} = \eta_i\{1 + o_p(1)\}$. Plugging this in (A.9), we obtain that

$$\mathbf{F}^T(I - \mathbf{S}_h)^T(I - \mathbf{S}_h)\boldsymbol{\eta} = \sum_{i=d+1}^{n} \mathbf{f}_i \eta_i \{1 + o_p(1)\}.$$

Since $E(\mathbf{f}_i\eta_i) = 0$, $\mathrm{Var}(\mathbf{f}_i\eta_i) = \sigma^2\{E(\mathbf{f}\mathbf{f}^T)\} < \infty$, and $E(\mathbf{f}_i\eta_i(\mathbf{f}_j\eta_j)^T) = 0$ for $i \neq j$ since η_i is independent of $\mathbf{f}_i$. By Central Limit Theorem for strictly stationary sequence (see Theorem 2.21 of Fan and Yao, 2003),

$$\frac{1}{\sqrt{n}}\mathbf{F}^T(I - \mathbf{S}_h)^T(I - \mathbf{S}_h)\boldsymbol{\eta} \xrightarrow{L} N(0, \sigma^2\{E(\mathbf{f}\mathbf{f}^T)\}).$$

This completes the proof of Lemma 3.

Proof of Theorem 1

Let us first show the asymptotic normality of $\widehat{\boldsymbol{\beta}}$. Denote

$$\begin{aligned}
A_1 &= \sqrt{n}\,[\{\mathbf{E}^T(I - \mathbf{S}_h)^T(I - \mathbf{S}_h)\mathbf{E}\}^{-1}\mathbf{E}^T(I - \mathbf{S}_h)^T(I - \mathbf{S}_h)\mathbf{M}],\\
A_2 &= \sqrt{n}\,[\{\mathbf{E}^T(I - \mathbf{S}_h)^T(I - \mathbf{S}_h)\mathbf{E}\}^{-1}\mathbf{E}^T(I - \mathbf{S}_h)^T(I - \mathbf{S}_h)\boldsymbol{\Delta\beta}],\\
A_3 &= \sqrt{n}\,[\{\mathbf{E}^T(I - \mathbf{S}_h)^T(I - \mathbf{S}_h)\mathbf{E}\}^{-1}\mathbf{E}^T(I - \mathbf{S}_h)^T(I - \mathbf{S}_h)\boldsymbol{\eta}].
\end{aligned}$$

According to the expression in $\widehat{\boldsymbol{\beta}}$ in (2.5), we have

$$\sqrt{n}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = A_1 + A_2 + A_3.$$

Lemma 2 shows that $A_1 = o_p(1)$ and $A_2 = o_p(1)$. Moreover, Lemma 3 states that A_3 weakly converges to $N(0, \sigma^2\{E(\mathbf{f}\mathbf{f}^T)\}^{-1})$. Thus, we establish the asymptotic normality of $\widehat{\boldsymbol{\beta}}$.

Next we derive the asymptotic bias and variance of $\widehat{\alpha}_j(\cdot)$. By (2.6) and the arguments in Lemmas 2 and 3, we have

$$\widehat{\alpha}_j(u_0, \widehat{\boldsymbol{\beta}}) = e_{j+1}^T\{D_{u_0}^T W_{u_0} D_{u_0}\}^{-1} D_{u_0}^T W_{u_0}(\mathbf{y} - \mathbf{E}\widehat{\boldsymbol{\beta}}),\ j = 0, \ldots, p$$

where e_{j+1}^T is a $2(p+1) \times 1$ vector consisting of 0's except 1 at the $(j+1)^{\text{th}}$ element. By the matrix format of semiparametric varying coefficient partially linear model (2.2) and $||\widehat{\boldsymbol\beta}-\boldsymbol\beta|| = O_p(n^{-\frac{1}{2}})$, it follows

$$\widehat{\alpha}_j(u_0,\widehat{\boldsymbol\beta}) = e_{j+1}^T\{D_{u_0}^T W_{u_0} D_{u_0}\}^{-1} D_{u_0}^T W_{u_0}(\mathbf{M}+\boldsymbol\eta)\{1+o_p(1)\},\ j=0,\dots,p$$

Using techniques used in the proof of Lemma 2 and techniques to derive the asymptotic bias of local linear regression, we can show that under Conditions A—G,

$$e_{j+1}^T\{D_{u_0}^T W_{u_0} D_{u_0}\}^{-1} D_{u_0}^T W_{u_0}\mathbf{M} = \alpha_j(u_0) + \frac{1}{2}\mu_2\alpha_j''(u_0)h^2 + o_p(h^2).$$

By assumption that $\boldsymbol\eta$ is independent of $\{u_1,\cdots,u_n\}$ and $\{\mathbf{X}_1,\cdots,\mathbf{X}_n\}$, we can show that under Conditions A—G, $E[e_{j+1}^T\{D_{u_0}^T W_{u_0} D_{u_0}\}^{-1} D_{u_0}^T W_{u_0}\boldsymbol\eta] = 0$ and

$$\mathrm{Var}[e_{j+1}^T\{D_{u_0}^T W_{u_0} D_{u_0}\}^{-1} D_{u_0}^T W_{u_0}\boldsymbol\eta] = \frac{\sigma^2}{nhg(u_0)}\int K^2(u)\,du\{1+o_p(1)\}$$

by using similar techniques related to the proof of Lemma 2. Furthermore, as shown in the proof of Lemma 2, $\frac{1}{n}D_{u_0}^T W_{u_0} D_{u_0} \to g(u_0)E(\mathbf{X}\mathbf{X}^T|u_t = u_0)\otimes \mathrm{diag}\{1,\mu_2\}$ in probability, where $\otimes$ stands for the Kronecker product of two matrices and $\mathrm{diag}\{1,\mu_2\}$ is a 2×2 diagonal matrix with diagonal element 1 and μ_2. Since η_t are independent and identically distributed with mean zero and variance σ^2, it follows that the asymptotic normality of $\frac{1}{\sqrt{nh}}D_{u_0}^T W_{u_0}\boldsymbol\eta$ can be established by using the CLT for α-mixing process. The proof of Theorem 1(B) is completed by using the Slutsky theorem and noting that $\nu_0 = \int K^2(u)\,du$.

Department of Operation Research and Financial Engineering, Princeton University, Princeton, NJ 08544, USA. Email: chenzhao1985@gmail.com.

Department of Statistics and The Methodology Center, The Pennsylvania State University, University Park, Pennsylvania, 16802-2111, USA; E-mail: rzli@psu.edu

eBay Inc, 618 East Yan'an Road, Shanghai, 200001. Email: Maggie.Li.Stat@gmail.com.