

A Appendix of “Statistical Inference on Panal Data Models: A Kernel Ridge Regression Method”

This appendix contains the proofs of the main results. In Section A.1, a proof of the convergence rate in the heterogeneous model is provided (Theorem 3.1) and some auxiliary lemmas are stated. In Section A.2, we prove the FBR for the heterogeneous model (Theorem 3.2) and the joint asymptotic distributions of our estimators (Theorems 3.3 and 3.4). Section A.3 includes the proof of the convergence rate in the homogeneous model (Theorem 4.1) as well as some auxiliary lemmas. In Section A.4, proofs of the FBR in the homogeneous model (Theorem 4.2) and the corresponding asymptotic normality (Theorem 4.3) are given. We also show that the variance estimator is consistent.

A.1 Proofs in Section 3.2

In this section, we derive the rate of convergence for our estimator in the heterogeneous setting, i.e., Theorem 3.1. Before proving the results, we provide some preliminary results.

Define $\Sigma_i = E\{G_i(X_i)(G_i(X_i) - A_i(X_i))'\}$, a symmetric matrix of dimension $q_1 + d$. Following Cheng and Shang (2015), Proposition A.1 below guarantees (3.9).

Proposition A.1. *For any $u = (x, z) \in \mathcal{U}_i$ and for any $\theta = (\beta, g) \in \Theta_i$, (3.9) holds for $R_i u = (H_u^{(i)}, T_u^{(i)})$ and $P_i \theta = (H_g^{(\star i)}, T_g^{(\star i)})$, where*

$$\begin{aligned} H_u^{(i)} &= (\Omega_i + \Sigma_i)^{-1}(z - V_i(G_i, K_x^{(i)})), \\ T_u^{(i)} &= K_x^{(i)} - A_i'(\Omega_i + \Sigma_i)^{-1}(z - V_i(G_i, K_x^{(i)})), \\ H_g^{(\star i)} &= -(\Omega_i + \Sigma_i)^{-1}V_i(G_i, W_i g), \\ T_g^{(\star i)} &= W_i g + A_i'(\Omega_i + \Sigma_i)^{-1}V_i(G_i, W_i g). \end{aligned}$$

A direct application of Proposition A.1 is to calculate the Fréchet derivatives of ℓ_{i,M,η_i} . Define $U_{it} = (X_{it}, Z_t)$ for $i \in [N]$, $t \in [T]$. For $\theta = (\beta, g)$, $\Delta\theta = (\Delta\beta, \Delta g) \in \Theta_i$, we have

$$\begin{aligned} D\ell_{i,M,\eta_i}(\theta)\Delta\theta &= \left\langle -\frac{1}{T} \sum_{t=1}^T (Y_{it} - \langle R_i U_{it}, \theta \rangle_i) R_i U_{it} + P_i \theta, \Delta\theta \right\rangle_i \equiv \langle S_{i,M,\eta_i}(\theta), \Delta\theta \rangle_i, \\ DS_{i,M,\eta_i}(\theta)\Delta\theta &= \frac{1}{T} \sum_{t=1}^T \langle R_i U_{it}, \Delta\theta \rangle_i R_i U_{it} + P_i \Delta\theta, \\ D^2 S_{i,M,\eta_i}(\theta) &= 0, \end{aligned}$$

where D is the Fréchet derivative operator.

Lemma A.1. For any $\theta \in \Theta_i$, $DS_{i,M,\eta_i}^*(\theta) = id$, where $S_{i,M,\eta_i}^*(\theta) = E\{S_{i,M,\eta_i}(\theta)\}$.

Lemma A.2. There exist universal constants C_1, C_2, \dots, C_N such that,

$$\|R_i U_{it}\|_i^2 \leq C_i^2 (h_i^{-1} + Z_t' Z_t), \quad \text{for any } i \in [N], t \in [N], \quad (\text{A.1})$$

$$\|\theta\|_{i,\text{sup}} \leq C_i (1 + h_i^{-1/2}) \|\theta\|_i, \quad \text{for any } \theta \in \Theta_i. \quad (\text{A.2})$$

Proposition A.2. Under Assumption A1, as $T \rightarrow \infty$, $\max_{1 \leq t \leq T} \|Z_t\|_2 = O_P(T^{1/\alpha})$.

The following proposition holds for both (1) $T, N \rightarrow \infty$; (2) $N \rightarrow \infty$, T is fixed. That is, the result holds for $M \rightarrow \infty$.

Proposition A.3. Let Assumptions A1–A3 hold. For $i \in [N]$ and $t \in [T]$, let $p_i = p_i(M) \geq 1$ be a deterministic sequence indexed by M , and let $\psi_{i,M,t}(U_{it}; \theta)$ be a real-valued function defined on Θ_i such that $\psi_{i,M,t}(U_{it}; 0) \equiv 0$, and for any $\theta_1, \theta_2 \in \Theta_i$,

$$\|(\psi_{i,M,t}(U_{it}; \theta_1) - \psi_{i,M,t}(U_{it}; \theta_2)) R_i U_{it}\|_i \leq \|\theta_1 - \theta_2\|_{i,\text{sup}}.$$

Then there exists a universal constant $C_0 > 0$ such that, as $N \rightarrow \infty$,

$$P \left(\max_{i \in [N]} \sup_{\theta \in \mathcal{G}_i(p_i)} \frac{\sqrt{T} \|\mathbb{Z}_{iM}(\theta)\|_i}{\sqrt{T} J_i(p_i, \|\theta\|_{i,\text{sup}}) + 1} \geq C_0 \sqrt{\log N + \log \log(T J_i(p_i, 1))} \right) \rightarrow 0,$$

where

$$\mathbb{Z}_{iM}(\theta) = \frac{1}{\sqrt{T}} \sum_{t=1}^T [\psi_{i,M,t}(U_{it}; \theta) R_i U_{it} - E(\psi_{i,M,t}(U_{it}; \theta) R_i U_{it})], \quad \theta \in \Theta_i.$$

Proofs of Lemmas A.1 and A.2, Propositions A.2 and A.3 can be found in supplement document.

Proof of Theorem 3.1. Since $Y_{it} = g_{i0}(X_{it}) + Z_t' \beta_{i0} + e_{it}$, it follows that

$$S_{i,M,\eta_i}^*(\theta_{i0}) = E\{S_{i,M,\eta_i}(\theta_{i0})\} = E\left\{-\frac{1}{T} \sum_{t=1}^T e_{it} R_i U_{it} + P_i \theta_{i0}\right\}.$$

Also $e_{it} = \epsilon_{it} - \gamma'_{2i}(\bar{\Gamma}_2 \bar{\Gamma}'_2)^{-1} \bar{\Gamma}_2 \bar{v}_t = \epsilon_{it} - \Delta_i \bar{v}_t$, so we have

$$\begin{aligned}
\|S_{i,M,\eta_i}^*(\theta_{i0})\|_i &= \|E\{(\epsilon_{it} - \Delta_i \bar{v}_t)R_i U_{it} - P_i \theta_{i0}\}\|_i \\
&\leq \|E\{(\epsilon_{it} - \Delta_i \bar{v}_t)R_i U_{it}\}\|_i + \|P_i \theta_{i0}\|_i \\
&= \sup_{\|\theta\|_i=1} |\langle E\{(\epsilon_{it} - \Delta_i \bar{v}_t)R_i U_{it}\}, \theta \rangle_i| + \|P_i \theta_{i0}\|_i \\
&= \sup_{\|\theta\|_i=1} |E\{(\epsilon_{it} - \Delta_i \bar{v}_t)(g(X_{it}) + Z'_t \beta_i)\}| + \|P_i \theta_{i0}\|_i \\
&= \sup_{\|\theta\|_i=1} |E\{\Delta_i \bar{v}_t(g(X_{it}) + Z'_t \beta_i)\}| + \|P_i \theta_{i0}\|_i.
\end{aligned}$$

Since

$$|g(X_{it}) + Z'_t \beta_i| \leq (1 + \|Z_t\|_2)\|\theta\|_{i,sup} \leq C_i(1 + \|Z_t\|_2)(1 + h_i^{-1/2})\|\theta\|_i$$

and

$$E\{\Delta_i \bar{v}_t h_i^{-1/2}\} \leq E\{(\Delta_i \bar{v}_t)^2\}^{1/2} h_i^{-1/2} = O((Nh_i)^{-1/2}),$$

there exists a constant C' , such that

$$\sup_{\|\theta\|_i=1} |E\{\Delta_i \bar{v}_t(g(X_{it}) + Z'_t \beta_i)\}| \leq \frac{C'}{(Nh_i)^{1/2}}. \quad (\text{A.3})$$

In the meantime, we have

$$\|P_i \theta_{i0}\|_i = \sup_{\|\theta\|_i=1} |\langle P_i \theta_{i0}, \theta \rangle_i| = \sup_{\|\theta\|_i=1} |\eta_i \langle g_{i0}, g \rangle_i| \leq \sqrt{\eta_i} \|g_{i0}\|_{\mathcal{H}_i}, \quad i \in [N]. \quad (\text{A.4})$$

Consider an operator

$$T_{1i}(\theta) = \theta - S_{i,M,\eta_i}^*(\theta + \theta_{i0}), \quad \theta \in \Theta_i.$$

By Lemma A.1 we have for any $\theta \in \Theta_i$,

$$T_{1i}(\theta) = \theta - DS_{i,M,\eta_i}^*(\theta_{i0})\theta - S_{i,M,\eta_i}^*(\theta_{i0}) = -S_{i,M,\eta_i}^*(\theta_{i0}).$$

Since T_{1i} takes a constant value and by (A.3) and (A.4), T_{1i} is a contraction mapping from $\mathbb{B}_i(\sqrt{\eta_i} \|g_{i0}\|_{\mathcal{H}_i} + \frac{C'}{(Nh_i)^{1/2}})$ to itself, where $\mathbb{B}_i(r)$ represents the r -ball in $(\Theta_i, \|\cdot\|_i)$. By the Contraction mapping theorem, there exists a unique fixed point $\theta' \in \mathbb{B}_i(\sqrt{\eta_i} \|g_{i0}\|_{\mathcal{H}_i} + \frac{C'}{(Nh_i)^{1/2}})$ such that $T_{1i}(\theta') = \theta'$. Let $\theta_{\eta_i} = \theta' + \theta_{i0}$, then $S_{i,M,\eta_i}^*(\theta_{\eta_i}) = 0$. Obviously, $\|\theta_{\eta_i} - \theta_{i0}\|_i \leq \sqrt{\eta_i} \|g_{i0}\|_{\mathcal{H}_i} + \frac{C'}{(Nh_i)^{1/2}}$.

We fix an $i \in [N]$ and assume both T and N approach infinity. Let $\mathcal{E}_M = \{\max_{1 \leq t \leq T} \|Z_t\|_2 \leq \tilde{C}T^{1/\alpha}\}$. Proposition A.2 says that when \tilde{C} is large, \mathcal{E}_M has probability approaching one.

Write $\mathcal{E}_{M,t} = \{\|Z_t\|_2 \leq \tilde{C}T^{1/\alpha}\}$. Then $\mathcal{E}_M = \cap_{t=1}^T \mathcal{E}_{M,t}$. By Lemma A.2, $\mathcal{E}_{M,t}$ implies that $\|R_i U_{it}\|_i \leq C_i(h_i^{-1/2} + \tilde{C}T^{1/\alpha})$.

Consider another operator

$$T_{2i}(\theta) = \theta - S_{i,M,\eta_i}(\theta_{\eta_i} + \theta), \quad \theta \in \Theta_i.$$

For $i \in [N], t \in [T]$, define

$$\psi_{i,M,t}(U_{it}; \theta) = \frac{\langle R_i U_{it}, \theta \rangle_i I_{\mathcal{E}_{M,t}}}{\tilde{C}C_i T^{1/\alpha}(h_i^{-1/2} + \tilde{C}T^{1/\alpha})}, \quad \theta \in \Theta_i.$$

It is easy to see that on \mathcal{E}_M , for any $\theta_1 = (\beta_1, g_1), \theta_2 = (\beta_2, g_2) \in \Theta_i$, by Proposition A.1,

$$\begin{aligned} & \|(\psi_{i,M,t}(U_{it}; \theta_1) - \psi_{i,M,t}(U_{it}; \theta_2))R_i U_{it}\|_i \\ = & \frac{|\langle R_i U_{it}, \theta_1 - \theta_2 \rangle_i| \times \|R_i U_{it}\|_i}{C_i \tilde{C}T^{1/\alpha}(h_i^{-1/2} + \tilde{C}T^{1/\alpha})} I_{\mathcal{E}_{M,t}} \\ = & \frac{|(g_1 - g_2)(X_{it}) + Z'_t(\beta_1 - \beta_2)| \times \|R_i U_{it}\|_i}{C_i \tilde{C}T^{1/\alpha}(h_i^{-1/2} + \tilde{C}T^{1/\alpha})} I_{\mathcal{E}_{M,t}} \\ \leq & \frac{\|\theta_1 - \theta_2\|_{i,\text{sup}} \tilde{C}T^{1/\alpha} C_i (h_i^{-1/2} + \tilde{C}T^{1/\alpha})}{C_i \tilde{C}T^{1/\alpha}(h_i^{-1/2} + \tilde{C}T^{1/\alpha})} I_{\mathcal{E}_{M,t}} \leq \|\theta_1 - \theta_2\|_{i,\text{sup}}. \end{aligned} \quad (\text{A.5})$$

Notice the following decomposition:

$$\begin{aligned} T_{2i}(\theta) &= \theta - S_{i,M,\eta_i}(\theta + \theta_{\eta_i}) + S_{i,M,\eta_i}(\theta_{\eta_i}) - S_{i,M,\eta_i}(\theta_{\eta_i}) \\ &= \theta - DS_{i,M,\eta_i}(\theta_{\eta_i})\theta - S_{i,M,\eta_i}(\theta_{\eta_i}). \end{aligned}$$

We first examine $S_{i,M,\eta_i}(\theta_{\eta_i})$ as follows:

$$\begin{aligned} S_{i,M,\eta_i}(\theta_{\eta_i}) &= S_{i,M,\eta_i}(\theta_{\eta_i}) - E(S_{i,M,\eta_i}(\theta_{\eta_i})) \\ &= -\frac{1}{T} \sum_{t=1}^T [(Y_{it} - \langle R_i U_{it}, \theta_{\eta_i} \rangle_i) R_i U_{it} - E((Y_{it} - \langle R_i U_{it}, \theta_{\eta_i} \rangle_i) R_i U_{it})] \\ &= -\frac{1}{T} \sum_{t=1}^T [e_{it} R_i U_{it} - E(e_{it} R_i U_{it})] \\ &\quad + \frac{1}{T} \sum_{t=1}^T [\langle R_i U_{it}, \theta_{\eta_i} - \theta_{i0} \rangle_i R_i U_{it} - E(\langle R_i U_{it}, \theta_{\eta_i} - \theta_{i0} \rangle_i R_i U_{it})] \end{aligned}$$

Define $\xi_{it} = e_{it}R_iU_{it}$. Following [Dehling \(1983, eqn. \(3.2\)\)](#) and [Bradley \(2005, eqn. \(1.11\)\)](#),

$$\begin{aligned}
& E \left\| \sum_{t=1}^T [e_{it}R_iU_{it} - E(e_{it}R_iU_{it})] \right\|_i^2 \\
&= E \left\| \sum_{t=1}^T [\xi_{it} - E(\xi_{it})] \right\|_i^2 \\
&= \sum_{t,t'=1}^T [E(\langle \xi_{it}, \xi_{it'} \rangle_i) - \langle E(\xi_{it}), E(\xi_{it'}) \rangle_i] \\
&\leq \sum_{t,t'=1}^T 15(\phi(|t-t'|)/2)^{1-4/\alpha} E(\|\xi_{it}\|_i^{\alpha/2})^{4/\alpha}.
\end{aligned}$$

It follows from Assumption [A1 \(a\), \(c\), \(d\)](#), and Lemma [A.2](#) that

$$\begin{aligned}
E(\|\xi_{it}\|_i^{\alpha/2})^2 &= E(|e_{it}|^{\alpha/2} \|R_iU_{it}\|_i^{\alpha/2})^2 \\
&\leq E(|e_{it}|^{\alpha/2}) E(\|R_iU_{it}\|_i^\alpha) \leq c_0 h_i^{-\alpha/2},
\end{aligned}$$

where c_0 is an absolute constant. The existence of such c_0 is due to the fact $E(|e_{it}|^\alpha) < \infty$ and $E(\|Z_t\|_2^\alpha) < \infty$. Therefore, it follows from Assumption [A1 \(b\)](#) that there exists an absolute constant c_1 such that

$$E \left\| \sum_{t=1}^T [e_{it}R_iU_{it} - E(e_{it}R_iU_{it})] \right\|_i^2 \leq c_1 T h_i^{-1}.$$

Similarly, it can be shown that

$$\begin{aligned}
& E \left\| \sum_{t=1}^T [\langle R_iU_{it}, \theta_{\eta_i} - \theta_{i0} \rangle_i R_iU_{it} - E(\langle R_iU_{it}, \theta_{\eta_i} - \theta_{i0} \rangle_i R_iU_{it})] \right\|_i^2 \\
&\leq \sum_{t,t'=1}^T 15(\phi(|t-t'|)/2)^{1-4/\alpha} E(\|R_iU_{it}\|_i^\alpha)^{4/\alpha} \|\theta_{\eta_i} - \theta_{i0}\|_i^2 \\
&\leq c'_1 T h_i^{-1},
\end{aligned}$$

where c'_1 is an absolute constant. The last step follows from Proposition [A.2](#), i.e.,

$$E(\|R_iU_{it}\|_i^\alpha) = O(h_i^{-\alpha/2}),$$

and the fact $\|\theta_{\eta_i} - \theta_{i0}\|_i^2 = O(\eta_i + \frac{1}{Nh_i})$, and the condition $\eta_i + \frac{1}{Nh_i} = O(h_i)$.

Therefore, we can choose c_2 to be large such that, with the probability approaching one,

$$\|S_{i,M,\eta_i}(\theta_{\eta_i})\|_i \leq c_2(Th_i)^{-1/2}.$$

On $\mathcal{E}_{M,t}$, for any unequal $\theta_1, \theta_2 \in \Theta_i$, define

$$\theta = \frac{\theta_1 - \theta_2}{C_i(1 + h_i^{-1/2})\|\theta_1 - \theta_2\|_i}.$$

Write $\theta = (\beta, g)$. Hence, by Lemma A.2 (A.2),

$$\begin{aligned} \|\theta\|_{i,\text{sup}} &\leq 1, \\ \eta_i \|g\|_{\mathcal{H}_i}^2 &\leq \|\theta\|_i^2 = \frac{\|\theta_1 - \theta_2\|_i^2}{C_i^2(1 + h_i^{-1/2})^2\|\theta_1 - \theta_2\|_i^2} \leq C_i^{-2}h_i. \end{aligned}$$

This means that $\theta \in \mathcal{G}_i(p_i)$ with $p_i = C_i^{-1}(\eta_i^{-1}h_i)^{1/2}$. Since $\eta_i^{-1}h_i$ goes to infinity as (N, T) does, it is not of loss of generality to assume that $p_i \geq 1$. Define

$$Z_{iM}(\theta) = \frac{1}{\sqrt{T}} \sum_{t=1}^T [\psi_{i,M,t}(U_{it}; \theta) R_i U_{it} - E(\psi_{i,M,t}(U_{it}; \theta) R_i U_{it})].$$

It follows from (A.5) and Proposition A.3 that, with probability approaching one,

$$\sup_{\theta \in \mathcal{G}_i(p_i)} \frac{\sqrt{T}\|Z_{iM}(\theta)\|_i}{\sqrt{T}J_i(p_i, \|\theta\|_{i,\text{sup}}) + 1} \leq C_0 \sqrt{\log N + \log \log(TJ_i(p_i, 1))}. \quad (\text{A.6})$$

Since $h_i = o(1)$, assume that $h_i^{-1} \geq 1$. It follows from Lemma A.2 (A.1) that

$$\begin{aligned} &\|E\left(\langle R_i U_{it}, \theta \rangle_i I_{\mathcal{E}_{M,t}^c} R_i U_{it}\right)\|_i \\ &\leq E\left(|\langle R_i U_{it}, \theta \rangle_i| I_{\mathcal{E}_{M,t}^c} \|R_i U_{it}\|_i\right) \\ &\leq E\left((1 + \|Z_t\|_2) I_{\mathcal{E}_{M,t}^c} C_i(h_i^{-1/2} + \|Z_t\|_2)\right) \\ &\leq C_i h_i^{-1/2} E\left((1 + \|Z_t\|_2)^2 I_{\mathcal{E}_{M,t}^c}\right) \\ &\leq C_i h_i^{-1/2} E\left((1 + \|Z_t\|_2)^\alpha\right)^{2/\alpha} P(\mathcal{E}_{M,t}^c)^{1-2/\alpha} \\ &\leq C_i h_i^{-1/2} E\left((1 + \|Z_t\|_2)^\alpha\right)^{2/\alpha} \left(\frac{1}{\widetilde{C}^{\alpha T}} E(\|Z_t\|_2^\alpha)\right)^{1-2/\alpha}. \end{aligned} \quad (\text{A.7})$$

Consequently, with probability approaching one, for any unequal $\theta_1, \theta_2 \in \Theta_i$ On $\mathcal{E}_{M,t}$, it follows from (A.6) and (A.7) that

$$\begin{aligned}
& \|T_{2i}(\theta_1) - T_{2i}(\theta_2)\|_i \\
= & \left\| -\frac{1}{T} \sum_{t=1}^T [\langle R_i U_{it}, \theta_1 - \theta_2 \rangle_i R_i U_{it} - E(\langle R_i U_{it}, \theta_1 - \theta_2 \rangle_i R_i U_{it})] \right\|_i \\
= & \left\| -\frac{1}{T} \sum_{t=1}^T [\langle R_i U_{it}, \theta \rangle_i R_i U_{it} - E(\langle R_i U_{it}, \theta \rangle_i R_i U_{it})] \times \|\theta_1 - \theta_2\|_i C_i (1 + h_i^{-1/2}) \right\|_i \\
= & \|\theta_1 - \theta_2\|_i C_i (1 + h_i^{-1/2}) \left\| \left(-\frac{1}{T} \sum_{t=1}^T [\langle R_i U_{it}, \theta \rangle_i I_{\mathcal{E}_{M,t}} R_i U_{it} - E(\langle R_i U_{it}, \theta \rangle_i I_{\mathcal{E}_{M,t}} R_i U_{it})] \right. \right. \\
& \left. \left. + E(\langle R_i U_{it}, \theta \rangle_i I_{\mathcal{E}_{M,t}} R_i U_{it}) \right) \right\|_i \\
= & \|\theta_1 - \theta_2\|_i C_i (1 + h_i^{-1/2}) \left\| \left(-T^{-1/2} C_i \tilde{C} T^{1/\alpha} (h_i^{-1/2} + \tilde{C} T^{1/\alpha}) Z_{iM}(\theta) + E(\langle R_i U_{it}, \theta \rangle_i I_{\mathcal{E}_{M,t}} R_i U_{it}) \right) \right\|_i \\
\leq & \|\theta_1 - \theta_2\|_i C_i (1 + h_i^{-1/2}) \left(T^{-1/2} C_0 C_i \tilde{C} T^{1/\alpha} (h_i^{-1/2} + \tilde{C} T^{1/\alpha}) (J_i(p_i, 1) + T^{-1/2}) \right. \\
& \left. \times \sqrt{\log N + \log \log(T J_i(p_i, 1))} + C_i h_i^{-1/2} E((1 + \|Z_t\|_2)^\alpha)^{2/\alpha} \left(\frac{1}{\tilde{C}^\alpha T} E(\|Z_t\|_2^\alpha) \right)^{1-2/\alpha} \right) \\
\leq & c_3 \|\theta_1 - \theta_2\|_i,
\end{aligned} \tag{A.8}$$

where c_3 is a constant in $(0, 1/2)$. Note that (A.8) holds also for $\theta_1 = \theta_2$. The existence of such c_3 follows by condition $b_{N,p} = o_P(\sqrt{N}h)$.

In particular, letting $\theta_2 = 0$, one gets that for any $\theta_1 \in \mathbb{B}(2c_2(Th_i)^{-1/2})$,

$$\begin{aligned}
\|T_{2i}(\theta_1)\|_i & \leq \|T_{2i}(\theta_1) - T_{2i}(0)\|_i + \|T_{2i}(0)\|_i \\
& \leq c_3 \|\theta_1\|_i + \|S_{i,M,\eta_i}(\theta_{\eta_i})\|_i \\
& \leq 2c_2 c_3 (Th_i)^{-1/2} + c_2 (Th_i)^{-1/2} < 2c_2 (Th_i)^{-1/2}.
\end{aligned}$$

This implies that, with probability approaching one, T_{2i} is a contraction mapping from $\mathbb{B}(2c_2(Th_i)^{-1/2})$ to itself. By contraction mapping theorem, there exists uniquely a $\theta'' \in \mathbb{B}(2c_2(Th_i)^{-1/2})$ such that $T_{2i}(\theta'') = \theta''$, implying that $S_{i,M,\eta_i}(\theta_{\eta_i} + \theta'') = 0$. Thus, $\hat{\theta}_i = \theta_{\eta_i} + \theta''$ is the penalized MLE of ℓ_{i,M,η_i} . This further shows that $\|\hat{\theta}_i - \theta_{\eta_i}\|_i \leq 2c_2(Th_i)^{-1/2}$. Combined with $\|\theta_{\eta_i} - \theta_{i0}\|_i = O(\eta_i^{1/2} + (Nh_i)^{-1/2})$, we have $\|\hat{\theta}_i - \theta_{i0}\|_i = O_P((Th_i)^{-1/2} + \eta_i^{1/2} + (Nh_i)^{-1/2})$. \square

A.2 Proofs in Section 3.3

In this section, we prove Theorems 3.2, 3.3 and 3.4, and Corollary 3.6.

Proof of Theorem 3.2. Define

$$S_{i,M}(\theta) \equiv -\frac{1}{T} \sum_{t=1}^T (Y_{it} - \langle R_i U_{it}, \theta \rangle_i) R_i U_{it}$$

and

$$S_i(\theta) \equiv E\{S_{i,M}(\theta)\} = E\left\{-\frac{1}{T} \sum_{t=1}^T (Y_{it} - \langle R_i U_{it}, \theta \rangle_i) R_i U_{it}\right\}.$$

Recall $S_{i,M,\eta_i} = S_{i,M} + P_i\theta$ and $S_{i,M,\eta_i}^*(\theta) = S_i(\theta) + P_i\theta$. Denote $\theta_i = \widehat{\theta}_i - \theta_{i0}$. Since $S_{i,M,\eta_i}(\widehat{\theta}_i) = 0$, we have $S_{i,M,\eta_i}(\theta_i + \theta_{i0}) = 0$. Therefore,

$$\begin{aligned} & \|S_{i,M}(\theta_i + \theta_{i0}) - S_i(\theta_i + \theta_{i0}) - (S_{i,M}(\theta_{i0}) - S_i(\theta_{i0}))\|_i \\ &= \|S_{i,M,\eta_i}(\theta_i + \theta_{i0}) - S_{i,M,\eta_i}^*(\theta_i + \theta_{i0}) - (S_{i,M,\eta_i}(\theta_{i0}) - S_{i,M,\eta_i}^*(\theta_{i0}))\|_i \\ &= \|S_{i,M,\eta_i}^*(\theta_i + \theta_{i0}) + S_{i,M,\eta_i}(\theta_{i0}) - S_{i,M,\eta_i}^*(\theta_{i0})\|_i \\ &= \|DS_{i,M,\eta_i}^*(\theta_{i0})\theta_i + S_{i,M,\eta_i}(\theta_{i0})\|_i \\ &= \|\theta_i + S_{i,M,\eta_i}(\theta_{i0})\|_i. \end{aligned} \tag{A.9}$$

Consider an event $B_{i,M} = \{\|\theta\|_i \leq r_{i,M} \equiv C_B((Th_i)^{-1/2}) + \eta_i^{1/2} + (Nh_i)^{-1/2}\}$. For some C_B large enough, $B_{i,M}$ has probability approaching one. Let $d_{i,M} = C_i r_{i,M}(1 + h_i^{-1/2})$, where C_i is defined in lemma A.2. We have $d_{i,M} = o(1)$. For any $\theta \in \Theta_i$, we further define $\bar{\theta} = (\bar{\beta}, \bar{g}) = d_{i,M}^{-1}\theta/2$, where $\bar{\beta} = d_{i,M}^{-1}\beta/2$ and $\bar{g} = d_{i,M}^{-1}g/2$. Then, on event $B_{i,M}$, we have

$$\|\bar{\theta}\|_{i,sup} \leq C_i(1 + h_i^{-1/2})\|\bar{\theta}\|_i = C_i(1 + h_i^{-1/2})d_{i,M}^{-1}\|\theta\|_i/2 \leq \frac{1}{2}.$$

Meanwhile,

$$\|\bar{g}\|_{H_i}^2 = \frac{d_{i,M}^{-2}}{4}\eta_i^{-1}(\eta_i\|g\|_{H_i}^2) \leq \frac{d_{i,M}^{-2}}{4}\eta_i^{-1}\|\theta\|_i^2 \leq \frac{d_{i,M}^{-2}}{4}\eta_i^{-1}r_{i,M}^2 \leq C_i^{-2}h_i\eta_i^{-1}.$$

Let $p_i = C_i^{-1}(h_i\eta_i^{-1})^{1/2}$. Then $\|\bar{g}\|_{H_i} \leq p_i$. Therefore $\bar{\theta} \in \mathcal{G}_i(p_i)$. Since $(\eta_i h_i^{-1}) \rightarrow \infty$ as $(N, T) \rightarrow \infty$, $p_i > 1$ in general.

Recall $\mathcal{E}_{M,t} = \{\|Z_t\|_2 \leq \widetilde{C}T^{1/\alpha}\}$, as defined in the proof of Theorem 3.1. Let

$$\psi_{i,M,t}^d(U_{it}; \bar{\theta}) = \frac{\langle R_i U_{it}, \bar{\theta} \rangle_i I_{\mathcal{E}_{M,t}}}{2d_M \widetilde{C} C_i T^{1/\alpha} (h_i^{-1/2} + \widetilde{C}T^{1/\alpha})}, \quad \theta \in \Theta_i.$$

Following the proof of Theorem 3.1, on \mathcal{E}_M , for any $\theta_1 = (\beta_1, g_1), \theta_2 = (\beta_2, g_2) \in \Theta_i$, we have

$$\|(\psi_{i,M,t}^d(U_{it}; \bar{\theta}_1) - \psi_{i,M,t}^d(U_{it}; \bar{\theta}_2))R_i U_{it}\|_i \leq \|\bar{\theta}_1 - \bar{\theta}_2\|_{i,\text{sup}}. \quad (\text{A.10})$$

Define

$$Z_{iM}^d(\bar{\theta}) = \frac{1}{\sqrt{T}} \sum_{t=1}^T [\psi_{i,M,t}^d(U_{it}; \bar{\theta})R_i U_{it} - E(\psi_{i,M,t}^d(U_{it}; \bar{\theta})R_i U_{it})].$$

It follows from Proposition A.3 that, with probability approaching one,

$$\sup_{\theta \in \mathcal{G}_i(p_i)} \frac{\sqrt{T} \|Z_{iM}^d(\bar{\theta})\|_i}{\sqrt{T} J_i(p_i, \|\bar{\theta}\|_{i,\text{sup}}) + 1} \leq C_0 \sqrt{\log N + \log \log(T J_i(p_i, 1))}. \quad (\text{A.11})$$

Therefore

$$\begin{aligned} & \|\theta_i + S_{i,M,\eta_i}(\theta_{i0})\|_i \\ &= \|S_{i,M}(\theta_i + \theta_{i0}) - S_i(\theta_i + \theta_{i0}) - (S_{i,M}(\theta_{i0}) - S_i(\theta_{i0}))\|_i \\ &= \left\| \frac{1}{T} \sum_{t=1}^T [\langle R_i U_{it}, \theta_i \rangle_i R_i U_{it} - E\{\langle R_i U_{it}, \theta_i \rangle_i R_i U_{it}\}] \right\|_i \\ &= \left\| \frac{1}{T} \sum_{t=1}^T [\langle R_i U_{it}, \theta_i \rangle_i I_{\mathcal{E}_{M,t}} R_i U_{it} - E(\langle R_i U_{it}, \theta_i \rangle_i I_{\mathcal{E}_{M,t}} R_i U_{it})] - E(\langle R_i U_{it}, \theta_i \rangle_i I_{\mathcal{E}_{M,t}^c} R_i U_{it}) \right\|_i \\ &= \left\| 2d_M \tilde{C} C_i T^{1/\alpha-1} (h_i^{-1/2} + \tilde{C} T^{1/\alpha}) \left(\sqrt{T} Z_{iM}^d(\bar{\theta}) - E(\langle R_i U_{it}, \theta_i \rangle_i I_{\mathcal{E}_{M,t}^c} R_i U_{it}) \right) \right\|_i \\ &\leq 2d_M C_0 \tilde{C} C_i T^{1/\alpha-1} (h_i^{-1/2} + \tilde{C} T^{1/\alpha}) \left(\sqrt{T} J_i(p_i, \|\bar{\theta}\|_{i,\text{sup}}) + 1 \right) \sqrt{\log N + \log \log(T J_i(p_i, 1))} \\ &\quad + C_i h_i^{-1/2} E((1 + \|Z_t\|_2)^\alpha)^{2/\alpha} \left(\frac{1}{\tilde{C}^\alpha T} E(\|Z_t\|_2^\alpha) \right)^{1-2/\alpha}. \end{aligned} \quad (\text{A.12})$$

□

Proof of Theorem 3.3. Define $\hat{\theta}_i^h = (\hat{\beta}_i, h_i^{1/2} \hat{g}_i)$, $\theta_{i0}^{*h} = (\beta_{i0}^*, h_i^{1/2} g_{i0}^*)$, and $R_i^h u = (H_u^{(i)}, h_i^{1/2} T_u^{(i)})$, where $\theta_{i0}^* = (id - P_i)\theta_{i0}$. From Theorem 3.2, we have

$$\|\hat{\theta}_i - \theta_{i0} + S_{i,M,\eta_i}(\theta_{i0})\|_i = O_P(a_M). \quad (\text{A.13})$$

Since

$$S_{i,M,\eta_i}(\theta_{i0}) = -\frac{1}{T} \sum_{t=1}^T (Y_{it} - \langle R_i U_{it}, \theta_{i0} \rangle_i) R_i U_{it} + P_i \theta_{i0} = -\frac{1}{T} \sum_{t=1}^T e_{it} R_i U_{it} + P_i \theta_{i0},$$

Theorem 3.2 can be re-written as

$$\|\widehat{\theta}_i - \theta_{i0}^* - \frac{1}{T} \sum_{t=1}^T e_{it} R_i U_{it}\|_i = O_P(a_M). \quad (\text{A.14})$$

It implies $\|\widehat{\beta}_i - \beta_{i0}^* - \frac{1}{T} \sum_{t=1}^T e_{it} H_{it}^{(i)}\|_2 = O_P(a_M)$. Further, we define $Rem = \widehat{\theta}_i - \theta_{i0}^* - \frac{1}{T} \sum_{t=1}^T e_{it} R_i U_{it}$ and $Rem^h = \widehat{\theta}_i^h - \theta_{i0}^{*h} - \frac{1}{T} \sum_{t=1}^T e_{it} R_i^h U_{it}$. Then

$$\begin{aligned} \|Rem^h - h_i^{1/2} Rem\|_i &= \left\| \left((1 - h_i^{1/2})(\widehat{\beta}_i - \beta_{i0}^* - \frac{1}{T} \sum_{t=1}^T e_{it} H_{U_{it}}^{(i)}), 0 \right) \right\|_i \\ &\leq (1 - h_i^{1/2}) O \left(\left\| \widehat{\beta}_i - \beta_{i0}^* - \frac{1}{T} \sum_{t=1}^T e_{it} H_{U_{it}}^{(i)} \right\|_2 \right) = O_P(a_M). \end{aligned}$$

Therefore, $\|Rem^h\|_i \leq \|Rem^h - h_i^{1/2} Rem\|_i + \|h_i^{1/2} Rem\|_i = O_P(a_M)$.

The idea is to employ the Cramér-Wold device. For any z , we will obtain the limiting distribution of $T^{1/2} z'(\widehat{\beta}_i - \beta_{i0}^*) + (Th_i)^{1/2}(\widehat{g}_i(x_0) - g_{i0}^*(x_0))$, which is $T^{1/2} \langle R_i u, \widehat{\theta}_i^h - \theta_{i0}^{*h} \rangle_i$ by Proposition A.1, where $u = (x_0, z)$.

Since $T^{1/2} h^{-1/2} a_M = o(1)$, we have

$$\begin{aligned} &\left| T^{1/2} \langle R_i u, \widehat{\theta}_i^h - \theta_{i0}^{*h} - \frac{1}{T} \sum_{t=1}^T e_{it} R_i^h U_{it} \rangle_i \right| \\ &\leq T^{1/2} \|R_i u\|_i \|Rem^h\|_i \\ &= O_P(T^{1/2} h^{-1/2} a_M) = o_P(1). \end{aligned}$$

Then, to find the limiting distribution of $T^{1/2} \langle R_i u, \widehat{\theta}_i^h - \theta_{i0}^{*h} \rangle_i$, we only need to find the limiting distribution of $T^{1/2} \langle R_i u, \frac{1}{T} \sum_{t=1}^T e_{it} R_i^h U_{it} \rangle_i = T^{-1/2} \sum_{t=1}^T e_{it} (z' H_{U_{it}}^{(i)} + h^{1/2} T_{U_{it}}^{(i)}(x_0))$. Next, we use the CLT to find its limiting distribution.

Define $L(U_{it}) = z' H_{U_{it}}^{(i)} + h^{1/2} T_{U_{it}}^{(i)}(x_0)$. Since ϵ_{it} and v_{it} are i.i.d. across t , we have

$$\begin{aligned} s_T^2 &= Var \left(\sum_{t=1}^T e_{it} (z' H_{U_{it}}^{(i)} + h^{1/2} T_{U_{it}}^{(i)}(x_0)) \right) \\ &= T \cdot Var(e_{it} L(U_{it})) + \sum_{t_1 \neq t_2}^T Cov(e_{it_1} L(U_{it_1}), e_{it_2} L(U_{it_2})) \\ &= T \cdot E \{ e_{it}^2 L(U_{it})^2 \} - T \cdot E \{ e_{it} L(U_{it}) \}^2 + \sum_{t_1 \neq t_2}^T Cov(e_{it_1} L(U_{it_1}), e_{it_2} L(U_{it_2})). \quad (\text{A.15}) \end{aligned}$$

For the first term,

$$E \{e_{it}^2 L(U_{it})^2\} = E \{(\epsilon_{it} - \Delta_i \bar{v}_t)^2 L(U_{it})^2\} = \sigma_\epsilon^2 E \{L(U_{it})^2\} + E \{(\Delta_i \bar{v}_t)^2 L(U_{it})^2\}.$$

From Cauchy-Schwarz and Hölder's inequality, we can show that

$$\begin{aligned} E \{(\Delta_i \bar{v}_t)^2 L(U_{it})^2\} &\leq \|\Delta_i\|_2^2 E \{\|\bar{v}_t\|_2^2 L(U_{it})^2\} \\ &\leq \|\Delta_i\|_2^2 (E \{\|\bar{v}_t\|_2^\alpha\})^{2/\alpha} \left(E \left\{ |L(U_{it})|^{\frac{2\alpha}{\alpha-2}} \right\} \right)^{\frac{\alpha-2}{\alpha}}. \end{aligned}$$

We next will find the upper bound of $|L(U_{it})|$.

$$\begin{aligned} L(U_{it}) &= z' H_{U_{it}}^{(i)} + h^{1/2} T_{U_{it}}^{(i)}(x_0) \\ &= z' (\Omega_i + \Sigma_i)^{-1} Z_t - z' (\Omega_i + \Sigma_i)^{-1} A_i(X_{it}) + h_i^{1/2} K_{X_{it}}^{(i)}(x_0) \\ &\quad - h_i^{1/2} A_i'(x_0) (\Omega_i + \Sigma_i)^{-1} Z_t + h_i^{1/2} A_i'(x_0) (\Omega_i + \Sigma_i)^{-1} A_i(X_{it}). \end{aligned}$$

By the proof of lemma A.2, we have

$$\begin{aligned} |z' (\Omega_i + \Sigma_i)^{-1} Z_t| &\leq |c_1^{-1} z' Z_t| \leq c_{1z} \|Z_t\|_2, \\ |z' (\Omega_i + \Sigma_i)^{-1} A_i(X_{it})| &= |z' (\Omega_i + \Sigma_i)^{-1/2} (\Omega_i + \Sigma_i)^{-1/2} A_i(X_{it})| \\ &\leq \sqrt{z' (\Omega_i + \Sigma_i)^{-1} z} \sqrt{A_i'(X_{it}) (\Omega_i + \Sigma_i)^{-1} A_i(X_{it})} \\ &\leq c_{2z} h_i^{-1/2}, \\ |A_i'(x_0) (\Omega_i + \Sigma_i)^{-1} Z_t| &\leq c_{3z} h_i^{-1/2} \|Z_t\|_2, \\ |K_{X_{it}}^{(i)}(x_0)| &\leq C_{\varphi,i}^2 h_i^{-1}, \\ |A_i'(x_0) (\Omega_i + \Sigma_i)^{-1} A_i(X_{it})| &\leq c_1^{-1} C_{\varphi,i}^2 C_{G_i}^2 h_i^{-1}, \end{aligned}$$

where $c_{1z} = c_1^{-1} \|z\|_2$, $c_{2z} = c_1^{-1} C_{\varphi,i} C_{G_i} \|z\|_2$, and $c_{3z} = c_1^{-1} C_{\varphi,i} C_{G_i}$. Combine all these inequality together, we have

$$|L(U_{it})| \leq c_{4z} \|Z_t\|_2 + c_{5z} h_i^{-1/2},$$

where $c_{4z} = c_{1z} + c_{3z}$ and $c_{5z} = c_{2z} + C_{\varphi,i}^2 + c_1^{-1} C_{\varphi,i}^2 C_{G_i}^2$. Therefore, there exists a constant

c_{6z} , such that

$$\begin{aligned} E \left\{ |L(U_{it})|^{\frac{2\alpha}{\alpha-2}} \right\} &= E \left\{ \left| c_{4z} \|Z_t\|_2 + c_{5z} h_i^{-1/2} \right|^{\frac{2\alpha}{\alpha-2}} \right\} \\ &\leq c_{6z} \left(E \left\{ \|Z_t\|_2^{\frac{2\alpha}{\alpha-2}} \right\} + h_i^{-\frac{\alpha}{\alpha-2}} \right) \\ &\leq c_{6z} \left(E \left\{ \|Z_t\|_2^\alpha \right\}^{\frac{2}{\alpha-2}} + h_i^{-\frac{\alpha}{\alpha-2}} \right). \end{aligned}$$

Since $Z_t = (f'_{1t}, \bar{X}'_t)'$ and $\bar{X}_t = \bar{\Gamma}'_1 f_{1t} + \bar{\Gamma}'_2 f_{2t} + \bar{v}_t$, we have

$$\begin{aligned} E \left\{ \|Z_t\|_2^\alpha \right\} &= E \left\{ (\|f_{1t}\|_2^2 + \|\bar{X}_t\|_2^2)^{\alpha/2} \right\} \\ &\leq c_7 (E \left\{ \|f_{1t}\|_2^\alpha \right\} + E \left\{ \|\bar{X}_t\|_2^\alpha \right\}) \\ &\leq c_8 (E \left\{ \|f_{1t}\|_2^\alpha \right\} + E \left\{ \|f_{1t}\|_2^\alpha \right\} + E \left\{ \|f_{2t}\|_2^\alpha \right\} + E \left\{ \|\bar{v}_t\|_2^\alpha \right\}), \end{aligned}$$

where c_7 and c_8 are constants. By Assumption **A1**, $E \left\{ \|f_{1t}\|_2^\alpha \right\}$, $E \left\{ \|f_{2t}\|_2^\alpha \right\}$, and $E \left\{ \|v_{it}\|_2^\alpha \right\}$ are finite. Using Marcinkiewicz and Zygmund inequality in [Shao \(2003\)](#), we can show that

$$\begin{aligned} E \left\{ \|\bar{v}_t\|_2^\alpha \right\} &= \frac{1}{N^\alpha} E \left\{ \left(\sum_{l=1}^d \left(\sum_{i=1}^N v_{ilt} \right)^2 \right)^{\alpha/2} \right\} \\ &\leq c_9 \frac{1}{N^\alpha} \sum_{l=1}^d E \left\{ \left(\left(\sum_{i=1}^N v_{ilt} \right)^2 \right)^{\alpha/2} \right\} \\ &= c_9 \frac{1}{N^\alpha} \sum_{l=1}^d E \left\{ \left(\sum_{i=1}^N v_{ilt} \right)^\alpha \right\} \\ &\leq c_{10} \frac{d}{N^\alpha} \frac{1}{N^{1-\alpha/2}} \sum_{i=1}^N E \left\{ v_{ilt}^\alpha \right\} \\ &= O_P \left(\frac{1}{N^{\alpha/2}} \right), \end{aligned}$$

where c_9 and c_{10} are two constants. As $N \rightarrow \infty$, $E \left\{ \|Z_t\|_2^\alpha \right\} = O_P(1)$ and $E \left\{ |L(U_{it})|^{\frac{2\alpha}{\alpha-2}} \right\} = O_P(h_i^{-\frac{\alpha}{\alpha-2}})$. Thus, we have

$$E \left\{ (\Delta_i \bar{v}_t)^2 L(U_{it})^2 \right\} \leq \|\Delta_i\|_2^2 (E \left\{ \|\bar{v}_t\|_2^\alpha \right\})^{2/\alpha} \left(E \left\{ |L(U_{it})|^{\frac{2\alpha}{\alpha-2}} \right\} \right)^{\frac{\alpha-2}{\alpha}} = O_P \left(\frac{1}{N h_i} \right). \quad (\text{A.16})$$

By assumption $(Nh_i)^{-1} = o_P(1)$, we have $E\{(\Delta_i \bar{v}_t)^2 L(U_{it})^2\} = o_P(1)$.

Next, we will find the order of $E\{L(U_{it})^2\}$. As it is shown in [Cheng and Shang \(2015\)](#), as $\eta_i \rightarrow 0$ and $T \rightarrow \infty$, $E(|L(U_{it})|^2) \rightarrow \alpha_{x_0}^2 + 2(z + \beta_{x_0})' \Omega_i^{-1} \alpha_{x_0} + (z + \beta_{x_0})' \Omega_i^{-1} (z + \beta_{x_0})$ for any given z .

From above derivatives, we have found the leading term of the first term in Equation (A.15). Now, we turn to the second term. It is straightforward to obtain that

$$E\{e_{it}L(U_{it})\}^2 = E\{(\epsilon_{it} - \Delta_i \bar{v}_t)L(U_{it})\}^2 \leq \|\Delta_i\|_2^2 E\{\|\bar{v}_t\|_2^2\} E\{L(U_{it})^2\},$$

where $E\{\|\bar{v}_t\|_2^2\} \leq E\{\|\bar{v}_t\|_2^\alpha\}^{2/\alpha} = O_P(1/N)$ and $E\{L(U_{it})^2\} = O_P(1)$. So $E\{e_{it}L(U_{it})\}^2 = O_P(1/N)$. So the second term is of a smaller order than the first term.

For the last term of Equation (A.15), we can show that

$$\begin{aligned} & \sum_{t_1 \neq t_2}^T \text{Cov}(e_{it_1}L(U_{it_1}), e_{it_2}L(U_{it_2})) \\ & \leq \sum_{t_1 \neq t_2}^T |\text{Cov}(\Delta_i \bar{v}_{t_1}L(U_{it_1}), \Delta_i \bar{v}_{t_2}L(U_{it_2}))| \\ & \leq 8 \sum_{t_1 \neq t_2}^T \alpha(|t_1 - t_2|)^{1-4/\alpha} E\{|\Delta_i \bar{v}_{t_1}L(U_{it_1})|^{\alpha/2}\}^{4/\alpha} \\ & \leq 8 \cdot 2^{1-4/\alpha} \sum_{t_1=1}^T \sum_{t_2=1, t_2 \neq t_1}^T \phi(|t_1 - t_2|)^{1-4/\alpha} E\{|\Delta_i \bar{v}_{t_1}L(U_{it_1})|^{\alpha/2}\}^{4/\alpha}, \end{aligned}$$

where $\alpha(|t_1 - t_2|)$ and $\phi(|t_1 - t_2|)$ are the α -mixing and ϕ -mixing coefficients for $\Delta_i \bar{v}_t L(U_{it})$. We have $2 \cdot \alpha(|t_1 - t_2|) < \phi(|t_1 - t_2|)$. The second inequality is from Proposition 2.5 in [Fan and Yao \(2003\)](#). Similar to Equation (A.16), we can show that $E\{|\Delta_i \bar{v}_{t_1}L(U_{it_1})|^\alpha\} = O_P((Nh_i)^{-\alpha/2})$. So $E\{|\Delta_i \bar{v}_{t_1}L(U_{it_1})|^{\alpha/2}\}^{4/\alpha} = O_P(\frac{1}{Nh_i})$. From Assumption A1, we have $\sum_{t_2=1, t_2 \neq t_1}^\infty \phi(|t_1 - t_2|)^{1-4/\alpha} < \infty$. Thus,

$$\sum_{t_1 \neq t_2}^T \text{Cov}(e_{it_1}L(U_{it_1}), e_{it_2}L(U_{it_2})) = O_P\left(\frac{T}{Nh_i}\right).$$

Again, the third term of Equation (A.15) is of a smaller order than the first term.

To combine all above equations together, we have

$$\frac{1}{T} s_T^2 \rightarrow \sigma_s^2 \quad \text{or} \quad \frac{1}{T} s_T^2 \rightarrow (z', 1) \Psi^*(z', 1)',$$

where $\sigma_s^2 = \sigma_\epsilon^2 (\alpha_{x_0}^2 + 2(z + \beta_{x_0})' \Omega_i^{-1} \alpha_{x_0} + (z + \beta_{x_0})' \Omega_i^{-1} (z + \beta_{x_0}))$.

In order to use the CLT with mixing conditions, we need to show $E \left\{ |e_{it} L(U_{it})|^{\alpha/2} \right\}$ is finite.

$$\begin{aligned} E \left\{ |e_{it} L(U_{it})|^{\alpha/2} \right\} &\leq E \left\{ |e_{it}|^\alpha \right\}^{1/2} E \left\{ |L(U_{it})|^\alpha \right\}^{1/2} \\ &\leq c_{6z} E \left\{ |e_{it}|^\alpha \right\}^{1/2} E \left\{ \|Z_t\|_2^\alpha + h_i^{-\alpha/2} \right\}^{1/2}. \end{aligned}$$

Since $E \left\{ \|Z_t\|_2^\alpha \right\} = O_P(1)$ and $E \left\{ |e_{it}|^\alpha \right\} \leq \infty$, then $E \left\{ |e_{it} L(U_{it})|^{\alpha/2} \right\} \leq \infty$. From the above proof, we have $E \left\{ e_{it} L(U_{it}) \right\}^2 = O_P(1/N)$, which implies

$$E \left\{ e_{it} L(U_{it}) \right\} = O_P(1/\sqrt{N}). \quad (\text{A.17})$$

Then $E \left\{ |e_{it} L(U_{it}) - E \left\{ e_{it} L(U_{it}) \right\}|^{\alpha/2} \right\} < \infty$. Since ϕ -mixing condition is stronger than α -mixing, by the Theorem 2.21 of [Fan and Yao \(2003\)](#), we have

$$\frac{1}{\sqrt{T}} \left(\sum_{t=1}^T (e_{it} L(U_{it}) - E \left\{ e_{it} L(U_{it}) \right\}) \right) \xrightarrow{d} N(0, \sigma_s^2).$$

□

Proof of Theorem 3.4. Notice

$$\theta_{i0} - \theta_{i0}^* = P_i \theta_{i0} = \begin{pmatrix} -(\Omega_i + \Sigma_i)^{-1} V_i(G_i, W_i g_{i0}) \\ W_i g_{i0} + A_i' (\Omega_i + \Sigma_i)^{-1} V_i(G_i, W_i g_{i0}) \end{pmatrix}.$$

Hence the result of the theorem holds if we can show that, for any $x \in \mathcal{X}_i$,

$$\begin{aligned} \sqrt{T} E(e_{it} U_{it}^{(i)}) &= o_P(1), \\ \sqrt{T h_i} E(e_{it} H_{it}^{(i)}(x)), \\ \sqrt{T} (\Omega_i + \Sigma_i)^{-1} V_i(G_i, W_i g_{i0}) &= o_P(1), \\ \sqrt{T h_i} A_i'(x) (\Omega_i + \Sigma_i)^{-1} V_i(G_i, W_i g_{i0}) &= o_P(1), \\ \alpha_x = \beta_x &= 0, \\ \lim_{T \rightarrow \infty} \sqrt{T h_i} W_i g_{i0}(x) &= 0. \end{aligned}$$

First by (A.17), we can see that the following hold true:

$$\begin{aligned}\sqrt{T}E(e_{it}U_{it}^{(i)}) &= O_P(\sqrt{T/N}) = o_P(1), \\ \sqrt{Th_i}E(e_{it}H_{it}^{(i)}(x)) &= O_P(\sqrt{T/N}) = o_P(1).\end{aligned}$$

Similar to Shang and Cheng (2013), we have

$$W_i\varphi_\nu^{(i)} = \frac{\eta_i\rho_\nu^{(i)}}{1 + \eta_i\rho_\nu^{(i)}}\varphi_\nu^{(i)}, \quad \nu \geq 1. \quad (\text{A.18})$$

Now we see from A3 and (A.18)

$$V_i(G_{i,k}, W_i g_{i0}) = \sum_{\nu \geq 1} V_i(G_{i,k}, \varphi_\nu^{(i)}) V_i(g_{i0}, \varphi_\nu^{(i)}) \frac{\eta_i\rho_\nu^{(i)}}{1 + \eta_i\rho_\nu^{(i)}}.$$

So by Cauchy's inequality,

$$\begin{aligned}|V_i(G_{i,k}, W_i g_{i0})|^2 &\leq \sum_{\nu \geq 1} |V_i(G_{i,k}, \varphi_\nu^{(i)})|^2 \frac{\eta_i\rho_\nu^{(i)}}{1 + \eta_i\rho_\nu^{(i)}} \sum_{\nu \geq 1} |V_i(g_{i0}, \varphi_\nu^{(i)})|^2 \frac{\eta_i\rho_\nu^{(i)}}{(1 + \eta_i\rho_\nu^{(i)})} \\ &\leq \eta_i \text{const} \sum_{\nu \geq 1} |V_i(G_{i,k}, \varphi_\nu^{(i)})|^2 \frac{\eta_i\rho_\nu^{(i)}}{1 + \eta_i\rho_\nu^{(i)}} \\ &\leq \eta_i \text{const} \sum_{\nu \geq 1} |V_i(G_{i,k}, \varphi_\nu^{(i)})|^2 k_\nu \frac{\eta_i\rho_\nu^{(i)}}{(1 + \eta_i\rho_\nu^{(i)})k_\nu} \\ &\leq \eta_i \text{const},\end{aligned}$$

For $\|A_{i,k}\|_{\text{sup}}$, by definition,

$$A_{i,k}(x) = \langle A_{i,k}, K_x^{(i)} \rangle_{*,i} \leq V_i(G_{i,k}, K_x^{(i)}) = \sum_{\nu \geq 1} \frac{V_i(G_{i,k}, \varphi_\nu^{(i)})}{1 + \eta_i\rho_\nu^{(i)}} \varphi_\nu^{(i)}(x).$$

By boundedness condition of $\varphi_\nu^{(i)}$ (Assumption A3) and Cauchy's inequality, we have

$$\begin{aligned}|A_{i,k}(x)|^2 &\leq \sum_{\nu \geq 1} |V_i(G_{i,k}, \varphi_\nu^{(i)})|^2 k_\nu |\varphi_\nu^{(i)}(x)|^2 \sum_{\nu \geq 1} \frac{1}{k_\nu (1 + \eta_i\rho_\nu^{(i)})^2} \\ &\leq \text{const} \sum_{\nu \geq 1} |V_i(G_{i,k}, \varphi_\nu^{(i)})|^2 k_\nu \sum_{\nu \geq 1} \frac{1}{k_\nu} = O(1).\end{aligned}$$

The above holds uniformly for all $x \in \mathcal{X}_i$. So

$$\sqrt{T}(\Omega_i + \Sigma_i)^{-1}V_i(G_i, W_i g_{i0}) = O(\sqrt{T\eta_i}) = o(1),$$

and

$$\sqrt{Th_i}A'_i(x)(\Omega_i + \Sigma_i)^{-1}V_i(G_i, W_i g_{i0}) = O(\sqrt{Th_i\eta_i}) = o(1).$$

By the above uniform boundedness of $A_{i,k}$, we have $\beta_x = 0$. Similarly, by (A.18), we have

$$W_i A_{i,k}(x) = \sum_{\nu \geq 1} \frac{V_i(G_{i,k}, \varphi_\nu^{(i)})}{1 + \eta_i \rho_\nu^{(i)}} \eta_i \rho_\nu^{(i)} \varphi_\nu^{(i)}(x).$$

Hence

$$\begin{aligned} |W_i A_{i,k}(x)|^2 &\leq \sum_{\nu \geq 1} |V_i(G_{i,k}, \varphi_\nu^{(i)})|^2 k_\nu |\varphi_\nu^{(i)}(x)|^2 \sum_{\nu \geq 1} \frac{(\eta_i \rho_\nu^{(i)})^2}{k_\nu (1 + \eta_i \rho_\nu^{(i)})^2} \\ &\leq \text{const} \sum_{\nu \geq 1} |V_i(G_{i,k}, \varphi_\nu^{(i)})|^2 k_\nu \sum_{\nu \geq 1} \frac{1}{k_\nu} = O(1), \end{aligned}$$

and this rate is uniform for all $x \in \mathcal{X}_i$. So $\alpha_x = 0$. By definition of RKHS and W_i ,

$$\begin{aligned} |W_i g_{i0}(x)| &= |\langle W_i g_{i0}, K_x^{(i)} \rangle_{\star, i}| \\ &= |\eta_i \langle g_{i0}, K_x^{(i)} \rangle_{\mathcal{H}_i}| \\ &\leq \|g_{i0}\|_{\mathcal{H}_i} \|K_x^{(i)}\|_{\mathcal{H}_i} \\ &\leq \sqrt{\eta_i} \|g_{i0}\|_{\mathcal{H}_i} \sqrt{\langle K_x^{(i)}, K_x^{(i)} \rangle_{\star, i}} \\ &\leq \sqrt{\eta_i} \|g_{i0}\|_{\mathcal{H}_i} O(h_i^{-1/2}) \\ &= O(\sqrt{\eta_i/h_i}). \end{aligned}$$

Thus $\sqrt{Th_i}W_i g_{i0}(x) = o(1)$. □

Proof of Corollary 3.6. By (2.5),

$$Y_{it} - \widehat{g}_i(X_{it}) - Z'_t \widehat{\beta}_i = \epsilon_{it} - \Delta_i \bar{v}_t + (g_{i0}(X_{it}) - \widehat{g}_i(X_{it})) + Z'_t(\beta_{i0} - \widehat{\beta}_i).$$

Hence $\sum_{t=1}^T \{Y_{it} - \widehat{g}_i(X_{it}) - Z'_t \widehat{\beta}_i - \epsilon_{it}\}^2 / T \leq 4 \sum_{t=1}^T (A_{1t} + A_{2t} + A_{3t}) / T$, where $A_{1t} = \|\Delta_i\|_2^2 \|\bar{v}_t\|_2^2$, $A_{2t} = (g_{i0}(X_{it}) - \widehat{g}_i(X_{it}))^2$, $A_{3t} = \|Z_t\|_2^2 \|\widehat{\beta}_i - \beta_{i0}\|_2^2$. By uniform boundedness of Δ_i and i.i.d. of v_{it} in Assumption A1, we have $E(A_{t1}) \leq \|\Delta_i\|_2^2 \text{tr}(\Sigma_v) / N$, where Σ_v is

the covariance matrix of v_{it} . So $\sum_{t=1}^T A_{1t} = O_P(1/N) = o_P(1)$. Also, by Lemma A.2 and Theorem 3.1,

$$\|g_{i0} - \hat{g}_i\|_{\text{sup}} = O_P(h_i^{-1/2} \|\theta_{i0} - \hat{\theta}_i\|_i) = O_P(h_i^{-1/2} r_{i,M}).$$

So it follows that $\sum_{t=1}^T A_{2t}/T = O_P(h_i^{-1} r_{i,M}^2) = o_P(1)$. By Proposition A.2, Lemma A.2 and Theorem 3.1, we have $A_{3t} = O_P(T^{2/\alpha} r_{i,M}^2)$ uniformly for all $t = 1, 2, \dots, T$. So $\sum_{t=1}^T A_{3t}/T = O_P(T^{2/\alpha} r_{i,M}^2) = o_P(1)$. Hence $\sum_{t=1}^T \{Y_{it} - \hat{g}_i(X_{it}) - Z'_t \hat{\beta}_i\}^2/T - \sum_{t=1}^T \epsilon_{it}^2/T = o_P(1)$. The result follows by applying Law of Large Number: $\sum_{t=1}^T \epsilon_{it}^2/T \rightarrow \sigma_\epsilon^2$, in probability. \square

A.3 Proofs in Section 4.2

We prove Theorem 4.1. Let us first introduce some notation. Define

$$\begin{aligned} F'_1 &= (f_{11}, f_{12}, \dots, f_{1T}), F'_2 = (f_{21}, f_{22}, \dots, f_{2T}), \\ \epsilon_i &= (\epsilon_{i1}, \epsilon_{i2}, \dots, \epsilon_{iT})', e_i = (e_{i1}, e_{i2}, \dots, e_{iT}), \\ \bar{X} &= (\bar{X}_1, \bar{X}_2, \dots, \bar{X}_T), \bar{v} = (\bar{v}_1, \bar{v}_2, \dots, \bar{v}_T), \bar{X}^* = \bar{X} - \bar{v}, \\ P_\star &= I_T - \mathbb{Z}'_\star (\mathbb{Z}_\star \mathbb{Z}'_\star)^{-1} \mathbb{Z}_\star, \tilde{\mathbb{Z}} = \mathbb{Z} - \mathbb{Z}_\star. \end{aligned}$$

We can rewrite (2.1) as $(\mathbb{Y}_i - \langle K_{\mathbb{X}_i}, g_0 \rangle)' = \gamma'_{1i} F'_1 + \gamma'_{2i} F'_2 + \epsilon'_i$, (2.2) as $\bar{X}^* = \bar{X} - \bar{v} = \bar{\Gamma}'_1 F'_1 + \bar{\Gamma}'_2 F'_2$ and (2.5) as $\mathbb{Y}_i - \langle K_{\mathbb{X}_i}, g_0 \rangle = \mathbb{Z}' \beta_i + e_i$. Notice that we have $\mathbb{Z} = (F_1, \bar{X}')'$, $\mathbb{Z}_\star = (F_1, \bar{X}^*)'$. By definition, we have $\mathbb{Z}P = 0, \mathbb{Z}_\star P_\star = 0$. Hence $F'_1 P = 0, \bar{X} P = 0, F'_1 P_\star = 0, \bar{X}^* P_\star = 0, F'_2 P_\star = 0$. Define $S_{M,\eta}^\star(g) = E\{S_{M,\eta}(g) | \mathcal{F}_1^T\}$. By the proof of Lemma A.1, it can be easily shown that $DS_{M,\eta}^\star(g) = id$, for any $g \in \mathcal{H}$. We also have

$$\|K_x\|^2 \leq C_\varphi^2 h^{-1}, \quad \sup_{x \in \mathcal{X}} \|g(x)\| \leq C_\varphi h^{-1/2} \|g\|. \quad (\text{A.19})$$

The proof of Theorem 4.1 also relies on the following Lemmas A.3, A.4 and A.5. Proofs of these lemmas are provided in supplement document.

Lemma A.3. *Suppose that Assumptions A1, A4 and A5 hold. Then the following holds:*

$$E(\|P - P_\star\|_{op}) = O(N^{-1/2}), \quad (\text{A.20})$$

$$\max_{1 \leq i \leq N} \|E\{\gamma'_{2i} F'_2 (P - P_\star) K_{\mathbb{X}_i} | \mathcal{F}_1^T\}\| = O_P\left(\sqrt{\frac{T}{Nh}} + \frac{T}{N\sqrt{h}}\right), \quad (\text{A.21})$$

$$E(\|\tilde{\mathbb{Z}} \tilde{\mathbb{Z}}'\|_{op}) = O(T/N). \quad (\text{A.22})$$

where $F_2 = (f_{21}, \dots, f_{2T})'$.

Lemma A.4. *Suppose that Assumptions A1, A4 and A5 hold. Let $p = p(M) \geq 1$ be an \mathcal{F}_1^T -measurable sequence indexed by M and let $\psi(\mathbb{X}, g) : \mathbb{R}^T \times \mathcal{H} \rightarrow \mathbb{R}^T$ be a measurable function satisfying $\psi(\mathbb{X}, 0) \equiv 0$ and the following Lipschitz condition:*

$$\|\psi(\mathbb{X}, g_1) - \psi(\mathbb{X}, g_2)\|_2 \leq L\sqrt{h/T}\|g_1 - g_2\|_{\text{sup}}, \quad \text{for any } g_1, g_2 \in \mathcal{H},$$

where $L > 0$ is a constant. Then with $M \rightarrow \infty$,

$$\sup_{g \in \mathcal{G}(p)} \|Z_M(g)\| = O_P\left(1 + \sqrt{\log \log(NJ(p, 1))}(J(p, 1) + N^{-1/2})\right),$$

where

$$Z_M(g) = \frac{1}{\sqrt{N}} \sum_{i=1}^N [\psi(\mathbb{X}_i, g)' PK_{\mathbb{X}_i} - E\{\psi(\mathbb{X}_i, g)' PK_{\mathbb{X}_i} | \mathcal{F}_1^T\}].$$

Lemma A.5. *Under conditions in Theorem 4.1,*

$$\|S_{M,\eta}(g_\eta)\| = O_P\left(\frac{1}{\sqrt{NT}h} + \frac{1}{N\sqrt{h}}\right) + o_P(\sqrt{\eta}).$$

Proof of Theorem 4.1. The proof consists of two parts.

Part one: Define $T_1(g) = g - S_{M,\eta}^*(g + g_0)$. So

$$T_1(g) = g - DS_{M,\eta}^*(g_0)g - S_{M,\eta}^*(g_0) = -S_{M,\eta}^*(g_0).$$

So we have

$$\begin{aligned} \|S_{M,\eta}^*(g_0)\| &= \|E\{-\frac{1}{NT} \sum_{i=1}^N (\mathbb{Y}_i - \langle K_{\mathbb{X}_i}, g_0 \rangle)' PK_{\mathbb{X}_i} + W_\eta g_0 | \mathcal{F}_1^T\}\| \\ &= \|E\{-\frac{1}{NT} \sum_{i=1}^N e_i' PK_{\mathbb{X}_i} - W_\eta g_0 | \mathcal{F}_1^T\}\| \\ &= \|E\{-\frac{1}{NT} \sum_{i=1}^N (\gamma'_{1i} F'_1 + \gamma'_{2,i} F'_2 + \epsilon'_i) PK_{\mathbb{X}_i} - W_\eta g_0 | \mathcal{F}_1^T\}\| \\ &= \|E\{-\frac{1}{NT} \sum_{i=1}^N \gamma'_{2,i} F'_2 PK_{\mathbb{X}_i} - W_\eta g_0 | \mathcal{F}_1^T\}\|, \\ &= \|E\{-\frac{1}{NT} \sum_{i=1}^N \gamma'_{2,i} F'_2 (P - P_\star) K_{\mathbb{X}_i} - W_\eta g_0 | \mathcal{F}_1^T\}\|, \end{aligned}$$

where the second last equation is using independence of ϵ_i and \mathbb{X}_i, F_1, F_2 . By directly calculations,

$$\|W_\eta g_0\| = \sup_{\|g\|=1} |\langle W_\eta g_0, g \rangle| = \sup_{\|g\|=1} |\eta \langle g_0, g \rangle_{\mathcal{H}}| \leq \sup_{\|g\|=1} \sqrt{\eta} \|g_0\|_{\mathcal{H}} \sqrt{\eta} \|g\|_{\mathcal{H}} \leq \sqrt{\eta} \|g_0\|_{\mathcal{H}}.$$

For the first term, by Lemma A.3,

$$\|E\{-\frac{1}{NT} \sum_{i=1}^N \gamma'_{2,i} F'_2(P - P_\star) K_{\mathbb{X}_i} | \mathcal{F}_1^T\}\| = O_P(\frac{1}{\sqrt{NT}h} + \frac{1}{N\sqrt{h}}).$$

As a consequence, $\|S_{M,\eta}^*(g_0)\| = O_P(1/\sqrt{NT}h + 1/(N\sqrt{h}) + \sqrt{\eta})$. Hence with probability approaching one, we have

$$\|T_1(g)\| \leq C(\frac{1}{\sqrt{NT}h} + \frac{1}{N\sqrt{h}} + \sqrt{\eta}) \equiv r,$$

for some constant $C > 0$. Notice that T_1 is also a contraction mapping from $\bar{B}(0, r)$ to $\bar{B}(0, r)$, so there exists a $\tilde{g}_1 \in \bar{B}(0, r) \subset \mathcal{H}$ such that, $T_1(\tilde{g}_1) = \tilde{g}_1$. By Taylor expansion,

$$\tilde{g}_1 = T_1(\tilde{g}_1) = \tilde{g}_1 - S_{M,\eta}^*(\tilde{g}_1 + g_0),$$

and hence it follows that $S_{M,\eta}^*(\tilde{g}_1 + g_0) = 0$. Let $g_\eta = \tilde{g}_1 + g_0$, and with the probability approaching one we have,

$$\|g_\eta - g_0\| \leq C(\frac{1}{\sqrt{NT}h} + \frac{1}{N\sqrt{h}} + \sqrt{\eta}), \quad S_{M,\eta}^*(g_\eta) = 0. \quad (\text{A.23})$$

Part two: Define $T_2(g) = g - S_{M,\eta}(g_\eta + g)$, for any $g, \Delta g \in \mathcal{H}$. Since

$$\begin{aligned} \Delta g &= DS_{M,\eta}^*(g) \Delta g \\ &= E(DS_{M,\eta}(g) \Delta g | \mathcal{F}_1^T) = \frac{1}{NT} \sum_{i=1}^N E\{\tau_i \Delta g' P K_{\mathbb{X}_i} + W_\eta \Delta g | \mathcal{F}_1^T\}, \end{aligned}$$

we have that

$$\begin{aligned}
\|T_2(g_1) - T_2(g_2)\| &= \|(g_1 - g_2) - (S_{M,\eta}(g_\eta + g_1) - S_{M,\eta}(g_\eta + g_2))\| \\
&= \|DS_{M,\eta}^*(g)(g_1 - g_2) - (S_{M,\eta}(g_\eta + g_1) - S_{M,\eta}(g_\eta + g_2))\| \\
&= \left\| \frac{1}{NT} \sum_{i=1}^N \{ \tau_i(g_1 - g_2)' PK_{\mathbb{X}_i} + W_\eta(g_1 - g_2) \right. \\
&\quad \left. - E(\tau_i(g_1 - g_2)' PK_{\mathbb{X}_i} + W_\eta(g_1 - g_2) | \mathcal{F}_1^T) \right\| \\
&= \left\| \frac{1}{NT} \sum_{i=1}^N \{ \tau_i(g_1 - g_2)' PK_{\mathbb{X}_i} - E(\tau_i(g_1 - g_2)' PK_{\mathbb{X}_i} | \mathcal{F}_1^T) \right\| \\
&\equiv \|\kappa(g_1 - g_2)\|,
\end{aligned}$$

where $\kappa(g) = \sum_{i=1}^N (\tau_i g' PK_{\mathbb{X}_i} - E\{\tau_i g' PK_{\mathbb{X}_i} | \mathcal{F}_1^T\}) / (NT)$. Let $\Psi(\mathbb{X}_i, g) = \sqrt{h} \tau_i g / T$. It follows that, for any $g_1, g_2 \in \mathcal{H}$,

$$\begin{aligned}
\|\Psi(\mathbb{X}_i, g_1) - \Psi(\mathbb{X}_i, g_2)\|_2 &\leq \frac{\sqrt{h}}{T} \|\tau_i(g_1 - g_2)\|_2 \\
&\leq \frac{\sqrt{h}}{T} \sqrt{\sum_{t=1}^T |g_1(X_{it}) - g_2(X_{it})|^2} \leq \sqrt{\frac{h}{T}} \|g_1 - g_2\|_{\text{sup}}.
\end{aligned}$$

Let $\tilde{g} \equiv (g_1 - g_2) / (c_\varphi h^{-1/2} \|g_1 - g_2\|)$, then

$$\|\tilde{g}\|_{\text{sup}} = \frac{\|g_1 - g_2\|_{\text{sup}}}{c_\varphi h^{-1/2} \|g_1 - g_2\|} \leq 1,$$

and

$$\|\tilde{g}_\eta\|_{\mathcal{H}} = \frac{\|g_1 - g_2\|_{\mathcal{H}}}{c_\varphi h^{-1/2} \|g_1 - g_2\|} \leq \frac{1}{c_\varphi \sqrt{h^{-1} \eta}}.$$

So $\tilde{g} \in \mathcal{G}(p)$, with $p = 1 / (c_\varphi \sqrt{h^{-1} \eta})$. By Lemma A.4,

$$\frac{\|Z_M(g_1 - g_2)\|}{c_\varphi h^{-1/2} \|g_1 - g_2\|} = \|Z_M(\tilde{g})\| = O_P \left(1 + \sqrt{\log \log (NJ(p, 1))} (J(p, 1) + N^{-1/2}) \right) = O_P(b_{N,p}).$$

So by assumption $b_{N,p} = o_P(\sqrt{N}h)$, we have

$$\begin{aligned}
\|\kappa(g_1 - g_2)\| &= \left\| \frac{1}{\sqrt{N}h} Z_M(g_1 - g_2) \right\| \\
&= c_\varphi \frac{1}{\sqrt{N}h} \|g_1 - g_2\| O_P(b_{N,p}) = O_P\left(\frac{b_{N,p}}{\sqrt{N}h}\right) \|g_1 - g_2\| = o_P(1) \|g_1 - g_2\|,
\end{aligned} \tag{A.24}$$

where the terms O_P, o_P do not depend on g_1, g_2 . Hence with probability approaching one, uniformly for any g_1, g_2 , $\|T_2(g_1) - T_2(g_2)\| \leq \frac{1}{2}\|g_1 - g_2\|$. Also with probability approaching one, uniformly for g ,

$$\|T_2(g)\| \leq \|T_2(g) - T_2(0)\| + \|T_2(0)\| \leq \frac{1}{2}\|g\| + \|S_{M,\eta}(g_\eta)\|.$$

By Lemma A.5, with probability approaching one,

$$\|S_{M,\eta}(g_\eta)\| \leq C \left(\frac{1}{\sqrt{NT}h} + \frac{1}{N\sqrt{h}} + \sqrt{\eta} \right) \equiv R/2,$$

for some $C > 0$. Hence with probability approaching one, it follows that $\sup_{\|g\| \leq R} \|T_2(g)\| \leq R/2 + R/2 = R$. The above implies that T_2 is a contraction mapping from $\bar{B}(0, R)$ to itself. By contraction mapping theorem, there exists $\tilde{g}_2 \in \bar{B}(0, R)$ such that $\tilde{g}_2 = T_2(\tilde{g}_2) = \tilde{g}_2 - S_{M,\eta}(g_\eta + \tilde{g}_2)$. Hence $S_{M,\eta}(g_\eta + \tilde{g}_2) = 0$. So $\hat{g} = g_\eta + \tilde{g}_2$. Therefore,

$$\|\hat{g} - g_0\| \leq \|g_\eta - g_0\| + \|\tilde{g}_2\| = O_P \left(\frac{1}{\sqrt{NT}h} + \frac{1}{N\sqrt{h}} + \sqrt{\eta} \right).$$

□

A.4 Proofs in Section 4.3

We will prove Theorems 4.2 and 4.3, and Proposition 4.5. Let us introduce some additional notation and preliminaries. Let m be the increasing sequence of integers provided in Theorems 4.2 and 4.3. For any fixed $x_0 \in \mathcal{X}$, define $V_{NT} = \frac{1}{NT} \sum_{i=1}^N K_{\mathbb{X}_i}(x_0)' P K_{\mathbb{X}_i}(x_0)$, $A_{NT} = V_{NT}^{-1/2}$. Let $V_{NTm} = \sum_{i=1}^N \phi_m' \Phi_i' P \Phi_i \phi_m / (NT)$, $A_{NTm} = V_{NTm}^{-1/2}$ and $H_{NTm} = \sum_{i=1}^N \Phi_i' P \Phi_i / (NT)$, where

$$\Phi_i = \begin{pmatrix} \varphi_1(X_{i1}) & \varphi_2(X_{i1}) & \cdots & \varphi_m(X_{i1}) \\ \varphi_1(X_{i2}) & \varphi_2(X_{i2}) & \cdots & \varphi_m(X_{i2}) \\ & & \cdots & \\ \varphi_1(X_{iT}) & \varphi_2(X_{iT}) & \cdots & \varphi_m(X_{iT}) \end{pmatrix}, \quad \phi_m = \left(\frac{\varphi_1(x_0)}{1 + \eta\rho_1}, \frac{\varphi_2(x_0)}{1 + \eta\rho_2}, \dots, \frac{\varphi_m(x_0)}{1 + \eta\rho_m} \right)'$$

The proof of Theorems 4.2 and 4.3 rely on the following Lemmas A.6, A.7 and A.8. Proofs of these lemmas can be found in supplement document.

Lemma A.6. *Under Assumptions A1, A4 and A5, suppose $m = o(\sqrt{N})$, then $\|H_{NTm} - I_m\|_F = o_P(1)$, $\lambda_{\min}^{-1}(H_{NTm}) = O_P(1)$ and $\lambda_{\max}(H_{NTm}) = O_P(1)$.*

Lemma A.7. Under Assumptions A1, A4 and A5, suppose $h^{-1} = o_P(\sqrt{N})$, $D_m = o_P(1)$, then for any $x_0 \in \mathcal{X}$, $A_{NT} = O_P(1)$.

Lemma A.8. Under Assumptions A1, A4 and A5, suppose $D_m = o_P(\sqrt{N})$, $m = o_P(\sqrt{N})$, for any $x_0 \in \mathcal{X}$, we have

$$\sqrt{NT}A_{NTm} \left(\frac{1}{NT} \sum_{i=1}^N \phi'_m \Phi'_i P \epsilon_i \right) \xrightarrow{d} N(0, \sigma_\epsilon^2).$$

Proof of Theorem 4.2. Let $g = \hat{g} - g_0$. By the fact $DS_{M,\eta}^*(g_0) = id$ (see Section A.3) and $S_{M,\eta}(g + g_0) = 0$, we have

$$\begin{aligned} \|g + S_{M,\eta}(g_0)\| &= \|DS_{M,\eta}^*(g_0)g + S_{M,\eta}(g_0)\| \\ &= \|S_{M,\eta}^*(g + g_0) - S_{M,\eta}^*(g_0) - S_{M,\eta}(g + g_0) + S_{M,\eta}(g_0)\| \\ &= \left\| \frac{1}{NT} \sum_{i=1}^N (\tau_i g' P K_{\mathbb{X}_i} - E(\tau_i g' P K_{\mathbb{X}_i})) + W_\eta g - E(W_\eta g | \mathcal{F}_1^T) \right\| \\ &\leq \|\kappa(g)\|, \end{aligned}$$

where $\kappa(g)$ is defined in Part two of the proof of Theorem 4.1. By (A.24) with $g_1 - g_2$ therein replaced by g we have $\|\kappa(g)\| = O_P\left(\frac{b_{N,p}}{\sqrt{N}h}\right)\|g\|$. Hence

$$\|g + S_{M,\eta}(g_0)\| = O_P\left(\frac{b_{N,p}}{\sqrt{N}h}\right)\|g\| = O_P\left(\frac{b_{N,p}}{\sqrt{N}h} \left(\frac{1}{\sqrt{NT}h} + \frac{1}{N\sqrt{h}} + \sqrt{\eta}\right)\right).$$

For fixed $x_0 \in \mathcal{X}$,

$$\begin{aligned} |A_{NT}|g(x_0) + S_{M,\eta}(g_0)(x_0)| &\leq A_{NT}\|g + S_{M,\eta}(g_0)\|_{\text{sup}} \\ &\leq c_\varphi h^{-1/2}\|g + S_{M,\eta}(g_0)\| \\ &= O_P\left(\frac{b_{N,p}}{\sqrt{N}h} \left(\frac{1}{\sqrt{NT}h} + \frac{1}{Nh} + \sqrt{\frac{\eta}{h}}\right)\right). \end{aligned}$$

Since $g_0 \in \mathcal{H}$ is fixed,

$$\|W_\eta g_0\| = \sup_{\|\tilde{g}\|=1} \langle W_\eta g_0, \tilde{g} \rangle = \sup_{\|\tilde{g}\|=1} \eta \langle g_0, \tilde{g} \rangle_{\mathcal{H}} \leq \sup_{\|\tilde{g}\|=1} \eta \|g_0\|_{\mathcal{H}} \|\tilde{g}\|_{\mathcal{H}} \leq \sqrt{\eta} \|g_0\|_{\mathcal{H}}. \quad (\text{A.25})$$

It follows from (A.25) that $\|W_\eta g_0\| \leq \sqrt{\eta} \|g_0\|_{\mathcal{H}} = O_P(\sqrt{\eta})$, and hence

$$|W_\eta g_0(x_0)| \leq c_\varphi h^{-1/2} \|W_\eta g_0\| = O_P(\sqrt{\eta/h}).$$

Hence, we have

$$A_{NT}|g(x_0) + S_{M,\eta}(g_0)(x_0) - W_\eta g_0(x_0)| = O_P \left(\frac{b_{N,p}}{\sqrt{Nh}} \left(\frac{1}{\sqrt{NT}h} + \frac{1}{Nh} + \sqrt{\frac{\eta}{h}} \right) \right).$$

□

Proof of Theorem 4.3. By Theorem 4.2, the asymptotic distribution of $\sqrt{NT}A_{NT}(\widehat{g}(x_0) - g_0(x_0) + W_\eta g_0)$ is the same as $\sqrt{NT}A_{NT}[-S_{M,\eta}(g_0)(x_0) + W_\eta g_0(x_0)] = \sqrt{NT}A_{NT}[\frac{1}{NT} \sum_{i=1}^N (\mathbb{Y}_i - \tau_i g_0)' P K_{\mathbb{X}_i}(x_0)]$.

By $\mathbb{Y}_i = \tau_i g_0 + \mathbb{Z}'\beta_i + e_i$ (4.2), $\mathbb{Z}P = 0$ (see Section A.3) and $e_i = \epsilon_i - \bar{v}'\Delta'_i$, it yields that

$$\begin{aligned} & A_{NT} \left[\frac{1}{NT} \sum_{i=1}^N (\mathbb{Y}_i - \tau_i g_0)' P K_{\mathbb{X}_i}(x_0) \right] \\ &= A_{NT} \left(\frac{1}{NT} \sum_{i=1}^N K'_{\mathbb{X}_i}(x_0) P \epsilon_i \right) - A_{NT} \left(\frac{1}{NT} \sum_{i=1}^N K'_{\mathbb{X}_i}(x_0) P \bar{v}' \Delta'_i \right) \\ &\equiv A_{NT} \zeta - A_{NT} \xi. \end{aligned}$$

Let $\zeta_m = \sum_{i=1}^N \phi'_m \Phi'_i P \epsilon_i / (NT)$ and $\xi_m = \sum_{i=1}^N \phi'_m \Phi'_i P \bar{v}' \Delta'_i / (NT)$. To prove the result of the theorem, it is sufficient to prove the following:

$$A_{NT} - A_{NTm} = o_P(1), \tag{A.26}$$

$$A_{NT} = O_P(1), \tag{A.27}$$

$$\sqrt{NT} A_{NTm} \xi_m = o_P(1), \tag{A.28}$$

$$\sqrt{NT} (\xi - \xi_m) = o_P(1), \tag{A.29}$$

$$\sqrt{NT} (\zeta - \zeta_m) = o_P(1), \tag{A.30}$$

$$\sqrt{NT} A_{NTm} \zeta_m \xrightarrow{d} N(0, 1). \tag{A.31}$$

(A.26) and (A.27) are guaranteed by Lemma A.7; (A.31) follows from Lemma A.8. For

(A.28), by the expression of A_{NTm} and Lemma A.6, we get,

$$\begin{aligned}
|A_{NTm}\xi_m| &= |A_{NTm}\phi'_m \frac{1}{NT} \sum_{i=1}^N \Phi'_i P \bar{v}' \Delta'_i| \\
&\leq \|A_{NTm}\phi_m\|_2 \times \left\| \frac{1}{NT} \sum_{i=1}^N \Phi'_i P \bar{v}' \Delta'_i \right\|_2 \\
&\leq \sqrt{\frac{\|\phi_m\|_2^2}{\phi'_m H_{NTm} \phi_m}} \left\| \frac{1}{NT} \sum_{i=1}^N \Phi'_i P \bar{v}' \Delta'_i \right\|_2 \\
&\leq \lambda_{\min}^{-1/2}(H_{NTm}) \left\| \frac{1}{NT} \sum_{i=1}^N \Phi'_i P \bar{v}' \Delta'_i \right\|_2.
\end{aligned}$$

Recall $\Delta_i = \gamma'_{2i}(\bar{\Gamma}_2 \bar{\Gamma}'_2)^{-1} \bar{\Gamma}_2 \equiv \gamma'_{2i} \widetilde{M}$ (see Assumption A1(d)), which leads to

$$\left\| \sum_{i=1}^N \Phi'_i P \bar{v}' \Delta'_i \right\|_2^2 = \text{Tr} \left\{ \left(\sum_{i=1}^N \Phi'_i P \bar{v}' \Delta'_i \right) \left(\sum_{i=1}^N \Delta_i \bar{v} P \Phi_i \right) \right\}.$$

In matrix form, the above becomes

$$\begin{aligned}
&\text{Tr} \left\{ (\Phi'_1, \Phi'_2, \dots, \Phi'_N) P_N \bar{v}'_N \widetilde{M}'_N \gamma_2 \gamma_2' \widetilde{M}_N \bar{v}_N P_N (\Phi'_1, \Phi'_2, \dots, \Phi'_N)' \right\} \\
&\leq \lambda_{\max}(\gamma_2 \gamma_2') \lambda_{\max}(\widetilde{M}'_N \widetilde{M}_N) \lambda_{\max}(\bar{v}'_N \bar{v}_N) \text{Tr} \left\{ \sum_{i=1}^N \Phi'_i P \Phi_i \right\},
\end{aligned}$$

where P_N is defined in Section 4.1, $\widetilde{M}_N = I_N \otimes \widetilde{M}$, $\bar{v}_N = I_N \otimes \bar{v}$.

By Lemma A.3, Assumption A1 and Lemma A.6, we have

$$\lambda_{\max}(\bar{v}'_N \bar{v}_N) = \|\widetilde{Z}\widetilde{Z}'\|_{\text{op}} = O_P(T/N), \quad \lambda_{\max}(\widetilde{M}'_N \widetilde{M}_N) = \lambda_{\max}((\bar{\Gamma}_2 \bar{\Gamma}'_2)^{-1}) = O_P(1),$$

and

$$\text{Tr} \left\{ \sum_{i=1}^N \Phi'_i P \Phi_i \right\} = NT \text{Tr}(H_{NTm}) = O_P(NTm).$$

So it can be seen that,

$$|A_{NTm}\xi_m| = O_P \left(\frac{\sqrt{\lambda_{\max}(\gamma_2 \gamma_2') m}}{N} \right),$$

thus

$$\sqrt{NT} A_{NTm} \xi_m = O_P \left(\sqrt{\frac{\lambda_{\max}(\gamma_2 \gamma_2') m T}{N}} \right) = o_P(1),$$

and hence (A.28) is true.

Similar to (S.20) in the proof of Lemma A.5 (see supplement document), it can be shown that

$$|\zeta - \xi_m| = O_P\left(\frac{D_m}{\sqrt{NT}h} + \frac{D_m}{N\sqrt{h}}\right). \quad (\text{A.32})$$

More explicitly, the proof of (A.32) follows by replacing $K_{\mathbb{X}_i}$ in the expression of T_2 with $\Phi_i\phi_m$, and by a line-by-line check. Therefore, we have

$$\sqrt{NT}|\zeta - \xi_m| = O_P\left(\frac{D_m}{\sqrt{h}} + \frac{D_m\sqrt{T}}{\sqrt{Nh}}\right) = o_P(1),$$

i.e., (A.29) holds.

Let $\mathcal{D}_1^T = \sigma(f_{1t}, f_{2t}, X_{it} : t \in [T], i \in [N])$. By independence of ϵ_i and \mathcal{D}_1^T , we have $E(|\zeta - \zeta_m|^2 | \mathcal{D}_1^T) = \sum_{i=1}^N R_i' P R_i / (NT)$, where $R_i = (K_{\mathbb{X}_i}(x_0) - \Psi_i\phi_m)' P \epsilon_i$. Let $R_{x_0}(\cdot) = \sum_{\nu \geq m+1} \frac{\varphi_\nu(x_0)\varphi_\nu(\cdot)}{1+\eta\rho_\nu}$, since $\mathcal{F}_1^T \subset \mathcal{D}_1^T$, it follows that,

$$\begin{aligned} E(|\zeta - \zeta_m|^2 | \mathcal{F}_1^T) &= E\left(\frac{1}{N^2 T^2} \sum_{i=1}^N R_i' P R_i | \mathcal{F}_1^T\right) \\ &= \frac{1}{NT} V(R_{x_0}, R_{x_0}) \\ &= \frac{1}{NT} \sum_{\nu=m+1}^{\infty} \frac{\varphi_\nu^2(x_0)}{(1+\eta\rho_\nu)^2} \\ &\leq \frac{1}{NT} c_\varphi^2 D_m \\ &= O_P\left(\frac{D_m}{NT}\right), \end{aligned}$$

so $\zeta - \zeta_m = O_P(\sqrt{D_m/(NT)})$ which implies that (A.30) is valid. Proof is completed. \square

Proof of Proposition 4.5. By $Y_i = \tau_i g_0 + \mathbb{Z}'\beta_i + e_i$, $\mathbb{Z}P = 0$, $e_i = \epsilon_i - \bar{v}'\Delta'_i$, we have

$$\begin{aligned} &\frac{1}{NT} \sum_{i=1}^N (\mathbb{Y}_i - \tau_i \hat{g})' P (\mathbb{Y}_i - \tau_i \hat{g}) \\ &= \frac{1}{NT} \sum_{i=1}^N (\tau_i (g_0 - \hat{g}))' P (\tau_i (g_0 - \hat{g})) + \frac{2}{NT} \sum_{i=1}^N (\tau_i (g_0 - \hat{g}))' P e_i + \frac{1}{NT} \sum_{i=1}^N e_i' P e_i \\ &\equiv T_1 + T_2 + T_3. \end{aligned}$$

By Theorem 4.1,

$$\|\widehat{g} - g_0\|_{\text{sup}} \leq c_\varphi h^{-1/2} \|\widehat{g} - g_0\| = O_P\left(\frac{1}{\sqrt{NT}h} + \frac{1}{Nh} + \sqrt{\frac{\eta}{h}}\right) = o_P(1).$$

So $|T_1| \leq \sum_{i=1}^N \sum_{t=1}^T |g_0(X_{it} - \widehat{g}(X_{it}))|^2 / (NT) \leq \|\widehat{g} - g_0\|_{\text{sup}}^2 = o_P(1)$.

By the definitions of \bar{v} and \widetilde{Z} and by Lemma A.3, we have $E(\|\bar{v}\bar{v}'\|_{\text{op}}) = E(\|\widetilde{Z}\widetilde{Z}'\|_{\text{op}}) = O(T/N)$. Hence it holds that

$$\begin{aligned} E(\|e_i\|_2^2) &= E(\epsilon'_i \epsilon_i) + E(\Delta'_i \bar{v} \bar{v}' \Delta_i) - 2E(\epsilon'_i \bar{v}' \Delta_i) \\ &\leq T\sigma_\epsilon^2 + E(\|\bar{v}\bar{v}'\|_{\text{op}}) \sup_{1 \leq i \leq N} \|\Delta_i\|_2^2 = O(T). \end{aligned}$$

It then follows from Cauchy inequality that

$$\begin{aligned} |T_2| &\leq \frac{2}{NT} \sum_{i=1}^N \|\tau_i(g_0 - \widehat{g})\|_2 \|e_i\|_2 \\ &\leq \frac{2}{NT} \sum_{i=1}^N \|e_i\|_2 \sqrt{T \|\widehat{g} - g_0\|_{\text{sup}}} \\ &= O_P(1) \sqrt{\|\widehat{g} - g_0\|_{\text{sup}}} = o_P(1). \end{aligned}$$

Meanwhile, the following decomposition holds

$$\begin{aligned} T_3 &= \frac{1}{NT} \sum_{i=1}^N \epsilon'_i P \epsilon_i + \frac{1}{NT} \sum_{i=1}^N \Delta'_i \bar{v} P \bar{v}' \Delta_i - \frac{2}{NT} \sum_{i=1}^N \Delta'_i \bar{v} P \epsilon_i \\ &= \frac{1}{NT} \sum_{i=1}^N \epsilon'_i P_\star \epsilon_i + \frac{1}{NT} \sum_{i=1}^N \epsilon'_i (P - P_\star) \epsilon_i + \frac{1}{NT} \sum_{i=1}^N \Delta'_i \bar{v} P \bar{v}' \Delta_i - \frac{2}{NT} \sum_{i=1}^N \Delta'_i \bar{v} P \epsilon_i \\ &\equiv T_{31} + T_{32} + T_{33} - T_{34}. \end{aligned}$$

We handle the above terms $T_{31}, T_{32}, T_{33}, T_{34}$ respectively. By Lemma A.3, it follows that

$$|T_{32}| \leq \frac{1}{NT} \|P - P_\star\|_{\text{op}} \sum_{i=1}^N \epsilon'_i \epsilon_i = O_P(N^{-1/2}) = o_P(1).$$

In the meantime,

$$|T_{33}| \leq \frac{1}{NT} \sup_{1 \leq i \leq N} \|\Delta_i\|_2^2 \sum_{i=1}^N \|\bar{v}\bar{v}'\|_{\text{op}} = O_P\left(\frac{1}{N}\right) = o_P(1),$$

and

$$|T_{34}| \leq \frac{2}{NT} \|\epsilon_i\|_2 \|\bar{v}' \Delta_i\| \leq 2 \sqrt{\frac{\sum_{i=1}^N \epsilon'_i \epsilon_i}{NT}} \sqrt{\frac{\sum_{i=1}^N \Delta'_i \bar{v} \bar{v}' \Delta_i}{NT}} = O_P\left(\frac{1}{\sqrt{N}}\right) = o_P(1).$$

Next we look at T_{31} . By direct examinations,

$$E(T_{31} | \mathcal{F}_1^T) = \frac{1}{T} \text{Tr}\{P_\star E(\epsilon_i \epsilon'_i | \mathcal{F}_1^T)\} = \sigma_\epsilon^2 \frac{1}{T} \text{Tr}(P_\star) = \sigma_\epsilon^2 \frac{T - (q_1 + d)}{T},$$

and by Chebyshev's inequality, for any $\epsilon > 0$,

$$\begin{aligned} P(|T_{31} - E(T_{31} | \mathcal{F}_1^T)| > \epsilon | \mathcal{F}_1^T) &\leq \frac{1}{N} \frac{E\{|\epsilon'_1 P_\star \epsilon_1 / T - E(T_{31} | \mathcal{F}_1^T)|^2\} | \mathcal{F}_1^T}{\epsilon^2} \\ &\leq \frac{1}{N \epsilon^2} E(|\epsilon'_1 P_\star \epsilon_1 / T|^2 | \mathcal{F}_1^T) \\ &\leq \frac{1}{NT^2 \epsilon^2} E(|\epsilon'_1 \epsilon_1|^2 | \mathcal{F}_1^T) \\ &\leq \frac{1}{NT^2 \epsilon^2} (TE(\epsilon_{11}^4) + T^2 \sigma_\epsilon^4) \\ &= o(1). \end{aligned}$$

So $T_{31} \rightarrow \sigma_\epsilon^2(T - q_1 - d)/T$ in probability. Since T_{32}, T_{33}, T_{34} are all $o_P(1)$ as shown in the above, we have $T_3 \rightarrow \sigma_\epsilon^2(T - q_1 - d)/T$ in probability. Proof is completed. \square

B Appendix of “Statistical Inference on Panal Data Models: A Kernel Ridge Regression Method” – Additional Application Results

Our real data application is based on a balanced panel data of 48 countries for the period 1950–2014, the names of these countries are reported in Table B.1. Furthermore, the heterogeneous effects of government spending are summarized in Figure B.1, and the heterogeneous effects of exports are summarized in Figure B.2.

Table B.1: Countries in the application

1	Argentina	11	Congo	21	Ireland	31	Norway	41	Thailand
2	Australia	12	Denmark	22	Israel	32	Pakistan	42	Trinidad and Tobago
3	Austria	13	Ecuador	23	Italy	33	Peru	43	Turkey
4	Belgium	14	Egypt	24	Japan	34	Philippines	44	Uganda
5	Bolivia	15	Finland	25	Kenya	35	Portugal	45	United Kingdom
6	Brazil	16	France	26	Luxembourg	36	South Africa	46	United States
7	Canada	17	Germany	27	Mexico	37	Spain	47	Uruguay
8	Colombia	18	Guatemala	28	Morocco	38	Sri Lanka	48	Venezuela
9	Costa Rica	19	Iceland	29	Netherlands	39	Sweden		
10	Cyprus	20	India	30	New Zealand	40	Switzerland		

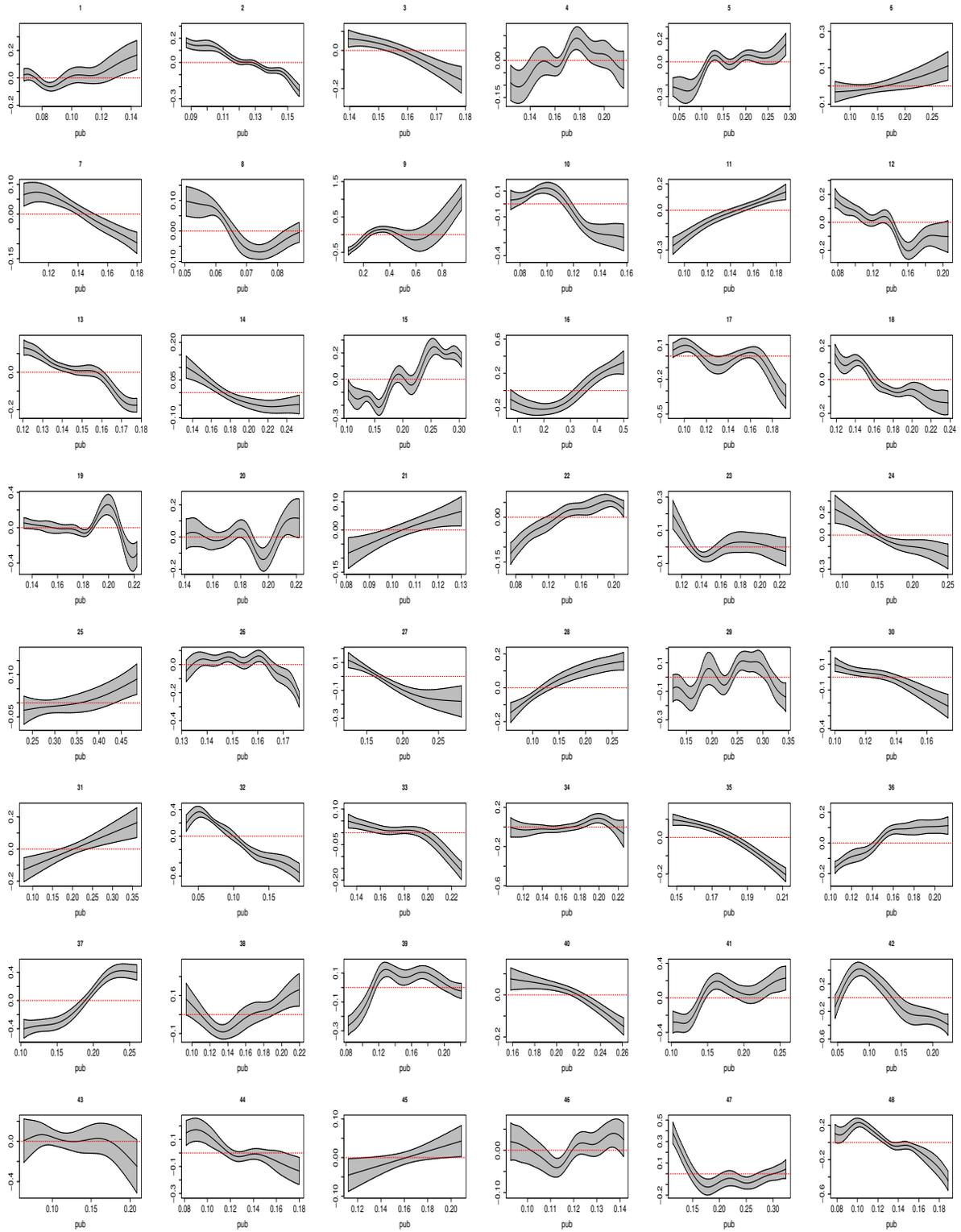


Figure B.1: Heterogeneous effects of government spending

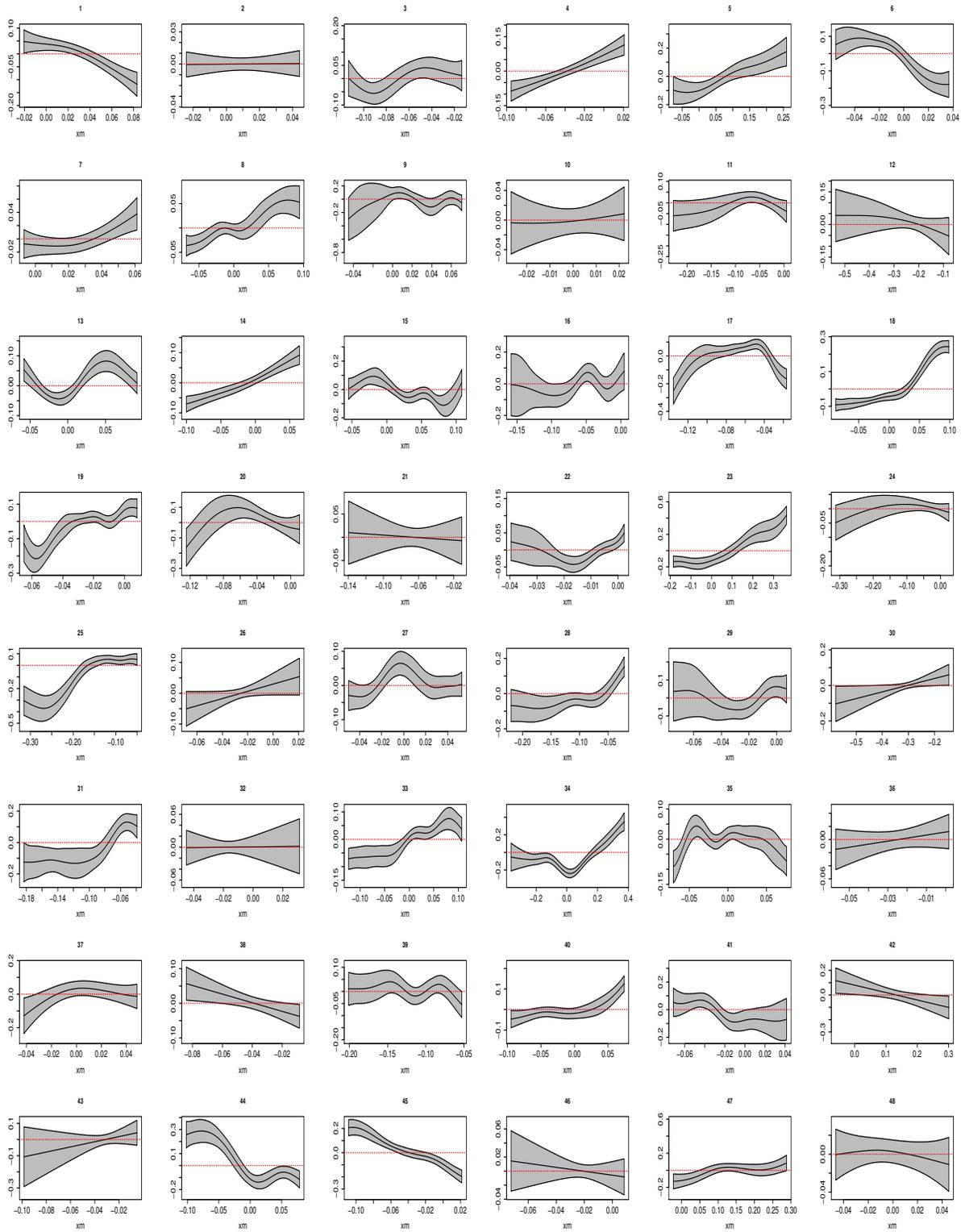


Figure B.2: Heterogeneous effects of exports

References

- Bradley, R. C. (2005). Basic properties of strong mixing conditions. a survey and some open questions. *Probability Surveys*, 2(2):107–144.
- Cheng, G. and Shang, Z. (2015). Joint asymptotics for semi-nonparametric regression models with partially linear structure. *The Annals of Statistics*, 43(3):1351–1390.
- Dehling, H. (1983). Limit theorems for sums of weakly dependent banach space valued random variables. *Probability Fields and Related Fields*, 63(3):393–432.
- Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer Science Business Media, LLC.
- Shang, Z. and Cheng, G. (2013). Local and global asymptotic inference in smoothing spline models. *The Annals of Statistics*, 41(5):2608–2638.
- Shao, J. (2003). *Mathematical Statistics*. Springer: New York.