Online Appendix to "SPECIFICATION TESTS FOR TIME-VARYING COEFFICIENT PANEL DATA MODELS"

Alev Atak[a], Thomas Tao Yang[b], Yonghui Zhang[c*], Qiankun Zhou[d]

[a]Department of Economics, Middle East Technical University, Turkey
[b]Research School of Economics, Australian National University, Australia
[c]School of Economics, Renmin University of China, China
[d]Department of Economics, Louisiana State University, USA

In this online appendix, we provide all proofs for lemmas, sketch proofs for main results in Section 4, the equivalence of the one-step procedure and our two-step procedure, and additional results for simulations and application.

C PROOFS FOR LEMMAS

Proof for Lemma A.1. Let $\mathbf{g}_{(2)} = (g_1, \ldots, g_d)'$ and $\mathbf{g} = (g_0, \mathbf{g}_{(2)}')'$. By Assumption 1(iv), all the eigenvalues of $E\left(\tilde{X}_{it}\tilde{X}_{it}'\right)$ are bounded away from zero and above from ∞. Then we have

$$\begin{aligned}
\|\mathbf{g}\|_i^2 &= E\left\{\frac{1}{T}\sum_{t=1}^{T}\left[\mathbf{g}'(\tau_t)\tilde{X}_{it}\right]\left[\tilde{X}_{it}'\mathbf{g}(\tau_t)\right]\right\} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{g}'(\tau_t)E\left(\tilde{X}_{it}\tilde{X}_{it}'\right)\mathbf{g}(\tau_t) \\
&\asymp \frac{1}{T}\sum_{t=1}^{T}\mathbf{g}_{(2)}'(\tau_t)\mathbf{g}_{(2)}(\tau_t) + \frac{1}{T}\sum_{t=1}^{T} g_0^2(\tau_t) \\
&= \sum_{l=1}^{d}\theta_l'\left[\frac{1}{T}\sum_{t=1}^{T}B^K(\tau_t)B^K(\tau_t)'\right]\theta_l + \theta_0'\left[\frac{1}{T}\sum_{t=1}^{T}B_{-1}^K(\tau_t)B_{-1}^K(\tau_t)'\right]\theta_0 \\
&= \sum_{l=1}^{d}\theta_l'\theta_l + \theta_0'\theta_0 + o(1) = \|\theta\|^2 + o(1)
\end{aligned}$$

by Assumption 1(iv) and the fact $T^{-1}\sum_{t=1}^{T}B^K(\tau_t)B^K(\tau_t)' = I_K + o(K/T)$ (see Lemma C.4.(i) in Dong and Linton (2018)). ■

Proof for Lemma A.2. The proofs of (i) and (ii) are analogous to that of Lemma A.2(i)-(ii) in Su et al. (2019), which uses the argument of Lemma A.2 in Huang et al. (2004). The only difference is that we use cosine functions as basis function, but it does not affect the results because cosine functions are bounded as well. We list the outline of the proof in the following, and the details can be found in Huang et al. (2004) and Su et al. (2019).

For (i), first, due to the uniform boundedness of cosine functions, we can obtain

$$P\left(\sup_{\mathbf{g}\in\mathcal{G}_{-1}\times\mathcal{G}^{\otimes d}}\left|\frac{T^{-1}\sum_{t=1}^{T}\left[\mathbf{g}'(\tau_t)\tilde{X}_{it}\right]^s}{T^{-1}\sum_{t=1}^{T}E\left[\mathbf{g}'(\tau_t)\tilde{X}_{it}\right]^s}-1\right|>\epsilon\right)\le C_1K^2\exp\left(-C_2\frac{T}{K}\frac{\epsilon^2}{1+\epsilon}\right), \tag{A.1}$$

for some positive constant C_1 and C_2 (see the proof of Lemma A.2 in Huang et al. (2004) for details), and for $s = 1, 2$.

Second,

$$
\begin{aligned}
&P\left(\max_i \sup_{\mathbf{g}\in\mathcal{G}_{-1}\times\mathcal{G}^{\otimes d}}\left|\frac{T^{-1}\sum_{t=1}^T\left[\mathbf{g}'(\tau_t)\tilde{X}_{it}\right]^s}{T^{-1}\sum_{t=1}^T E\left[\mathbf{g}'(\tau_t)\tilde{X}_{it}\right]^s}-1\right|>\epsilon\right)\\
&\leq \sum_{i=1}^N P\left(\sup_{\mathbf{g}\in\mathcal{G}_{-1}\times\mathcal{G}^{\otimes d}}\left|\frac{T^{-1}\sum_{t=1}^T\left[\mathbf{g}'(\tau_t)\tilde{X}_{it}\right]^s}{T^{-1}\sum_{t=1}^T E\left[\mathbf{g}'(\tau_t)\tilde{X}_{it}\right]^s}-1\right|>\epsilon\right)\\
&\leq C_1 N K^2 \exp\left(-C_2\frac{T}{K}\frac{\epsilon^2}{1+\epsilon}\right)
\end{aligned}
$$

using (A.1). By Assumption 2, $\frac{T}{K} \gg N^C$ for some fixed positive C, thus $C_1NK^2\exp\left(-C_2\frac{T}{K}\frac{\epsilon^2}{1+\epsilon}\right) = o\left(N^{-1}\right)$ as desired.

Similarly, we can show (ii). ■

Proof of Lemma A.3. We can follow the proofs of Lemma S4.3 (i) in Lu and Su (2022) to show (i), (ii), (iv) and (v). The only difference is that the dimension in our case is increasing as sample size (N, T) increases. The details are omitted. Lastly, (iii) is a direct result of (i) and (ii). ■

Proof for Lemma A.4. We first prove (i). Recall that $Z_{it} = (B_{-1,t}, B_t \otimes X_{it})'$ and $\dot{Z}_{it} = Z_{it} - \bar{Z}_i$. Write

$$
\hat{Q}_{i,\dot{z}\dot{z}} = \frac{1}{T}\sum_{t=1}^T Z_{it}Z_{it}' - \bar{Z}_i\bar{Z}_i' \equiv \hat{Q}_{i,\dot{z}\dot{z}}^{(1)} - \hat{Q}_{i,\dot{z}\dot{z}}^{(2)}, \text{ say.}
$$

Let $\varpi = (\varpi_0', \varpi_1', \ldots, \varpi_d')' = \left(\varpi_0', \varpi^{(2)\prime}\right)'$ with $\varpi_0 \in \mathbb{R}^{K-1}$ and $\varpi_l \in \mathbb{R}^K$ for $l = 1, \ldots, d$, and $\|\varpi\| \leq C \leq \infty$. Let $g_l(\tau, \varpi_l) = \varpi_l' B^K(\tau)$ and $g_0(\tau, \varpi_0) = \varpi_0' B_{-1}^K(\tau)$. Let $\mathbf{g}_\varpi = \left(g_0(\tau, \varpi_0), \mathbf{g}_{\varpi^{(2)}}'\right)'$, where $\mathbf{g}_{\varpi^{(2)}} = (g_1(\tau, \varpi_1), \ldots, g_d(\tau, \varpi_d))'$.

First, we show that $\lambda_{\max}(\hat{Q}_{i,\dot{z}\dot{z}})$ is bounded by some positive number uniformly in i. By Lemmas A.1 and A.4, we have that uniformly in i and ϖ,

$$
\varpi'\hat{Q}_{i,\dot{z}\dot{z}}^{(1)}\varpi = \frac{1}{T}\sum_{t=1}^T\left[\mathbf{g}_\varpi'(\tau_t)\tilde{X}_{it}\right]^2 = \frac{1}{T}\sum_{t=1}^T E\left[\mathbf{g}_\varpi'(\tau_t)\tilde{X}_{it}\right]^2(1+o_p(1)) \asymp \|\varpi\|^2.
$$

Then the largest eigenvalue of $\hat{Q}_{i,\dot{z}\dot{z}}^{(1)}$ is bounded above by some positive number $\bar{c}_{\dot{z}}$ uniformly in i with probability $1 - o(N^{-1})$. Noting that $\lambda_{\max}(\hat{Q}_{i,\dot{z}\dot{z}}) \leq \lambda_{\max}(\hat{Q}_{i,\dot{z}\dot{z}}^{(1)})$, we have $\lambda_{\max}(\hat{Q}_{i,\dot{z}\dot{z}}) \leq \bar{c}_{\dot{z}} < \infty$ uniformly in i with probability $1 - o(N^{-1})$.

Second, we prove that $\lambda_{\min}(\hat{Q}_{i,\dot{z}\dot{z}})$ is bounded away from zero uniformly in i. By Lemma A.2, $\varpi'\hat{Q}_{i,\dot{z}\dot{z}}^{(2)}\varpi = [T^{-1}\sum_{t=1}^T \mathbf{g}_\varpi'(\tau_t)\tilde{X}_{it}]^2 = [T^{-1}\sum_{t=1}^T \mathbf{g}_\varpi(\tau_t)' E\tilde{X}_{it}]^2(1+o(1))$ uniformly in i and ϖ. By Cauchy-Schwarz inequality, we have $[T^{-1}\sum_{t=1}^T \mathbf{g}_\varpi'(\tau_t)E\tilde{X}_{it}]^2 \leq T^{-1}\sum_{t=1}^T \left\|E\tilde{X}_{it}\right\|^2 \times$

$T^{-1}\sum_{t=1}^{T}\|\mathbf{g}_{\varpi}(\tau_t)\|^2 \leq C\|\varpi\|^2 < \infty$ uniformly in i and ϖ because of $T^{-1}\sum_{t=1}^{T}\|\mathbf{g}_{\varpi}(\tau_t)\|^2 = \|\varpi\|^2(1+o(1))$ (see the proof of Lemma A.1). It follows that

$$\varpi'\hat{Q}_{i,\dot{z}\dot{z}}\varpi = \frac{1}{T}\sum_{t=1}^{T}E\left\{[\mathbf{g}'_{\varpi}(\tau_t)\tilde{X}_{it}]^2\right\} - \left[\frac{1}{T}\sum_{t=1}^{T}\mathbf{g}'_{\varpi}(\tau_t)E\tilde{X}_{it}\right]^2 + o_p(1) \equiv A_{i,\varpi} + o_p(1).$$

We want to show that $A_{i,\varpi} \geq C\|\varpi\|^2$ for some positive constant C. Recall that $\mu_i^{(x)}(\tau_t) = EX_{it}$. Let $\Omega_i^{(x)}(\tau_t) \equiv \text{Var}(X_{it}) = \Xi_i(\tau_t) - \mu_i^{(x)}(\tau_t)\mu_i^{(x)}(\tau_t)'$ and $\tilde{\mu}_i(\tau_t) \equiv E(\tilde{X}_{it}) = \begin{pmatrix} 1 \\ \mu_i^{(x)}(\tau_t) \end{pmatrix}$, $\tilde{\Xi}_i(\tau_t) \equiv E(\tilde{X}_{it}\tilde{X}'_{it}) = \begin{pmatrix} 1 & \mu_i^{(x)}(\tau_t)' \\ \mu_i^{(x)}(\tau_t) & \Xi_i(\tau_t) \end{pmatrix}$, and $\tilde{\Omega}_i(\tau_t) \equiv \text{Var}\left(\tilde{X}_{it}\right) = \begin{pmatrix} 0 & \mathbf{0}_{d\times 1} \\ \mathbf{0}_{d\times 1} & \Omega_i^{(x)}(\tau_t) \end{pmatrix}$.
Then we have

$$\begin{aligned} A_{i,\varpi} &= \int_0^1 \mathbf{g}'_{\varpi}(\tau)\tilde{\Xi}_i(\tau)\mathbf{g}_{\varpi}(\tau)d\tau - \left\{\int_0^1 \mathbf{g}'_{\varpi}(\tau)\tilde{\mu}_i(\tau)d\tau\right\}^2 + o(1) \\ &= \int_0^1 \mathbf{g}'_{\varpi^{(2)}}(\tau)\Omega_i^{(x)}(\tau)\mathbf{g}_{\varpi^{(2)}}(\tau)d\tau + \left[\int_0^1\left[\mathbf{g}'_{\varpi}(\tau)\tilde{\mu}_i(\tau)\right]^2 d\tau - \left(\int_0^1 \mathbf{g}'_{\varpi}(\tau)\tilde{\mu}_i(\tau)d\tau\right)^2\right] + o(1) \\ &\equiv A_{i,\varpi}^{(1)} + A_{i,\varpi}^{(2)} + o(1). \end{aligned}$$

For the first term, we have

$$A_{i,\varpi}^{(1)} = \int_0^1 \mathbf{g}'_{\varpi^{(2)}}(\tau)\Omega_i^{(x)}(\tau)\mathbf{g}_{\varpi^{(2)}}(\tau)d\tau = \varpi'\begin{pmatrix} \mathbf{0}_{(K-1)\times(K-1)} & \mathbf{0}_{(K-1)\times dK} \\ \mathbf{0}_{dK\times(K-1)} & \int_0^1\left(\Omega_i^{(x)}(\tau)\otimes B(\tau)B(\tau)'\right)d\tau \end{pmatrix}\varpi.$$

Let $\underline{\mu}_i(\tau) = (B(\tau)\otimes\mu_i^{(x)}(\tau))'$ and $\underline{\mu}_i^{(c)}(\tau) = \underline{\mu}_i(\tau) - \int_0^1\underline{\mu}_i(\tau)d\tau$. Define

$$\mathbb{B}_i = \begin{pmatrix} \int_0^1 B_{-1}(\tau)B_{-1}(\tau)'d\tau & \int_0^1 B_{-1}(\tau)\underline{\mu}_i^{(c)}(\tau)'d\tau \\ \int_0^1\underline{\mu}_i^{(c)}(\tau)B_{-1}(\tau)'d\tau & \int_0^1\underline{\mu}_i^{(c)}(\tau)\underline{\mu}_i^{(c)}(\tau)'d\tau \end{pmatrix}.$$

Then for the second term, we have $A_{i,\varpi}^{(2)} = \varpi'\mathbb{B}_i\varpi$. Since $\int_0^1 B_{-1}(\tau)B_{-1}(\tau)'d\tau = I_{K-1}$, it follows that

$$\begin{aligned} &A_{i,\varpi}^{(1)} + A_{i,\varpi}^{(2)} \\ &= \varpi'\begin{pmatrix} I_{K-1} & \int_0^1 B_{-1}(\tau)\underline{\mu}_i^{(c)}(\tau)'d\tau \\ \int_0^1\underline{\mu}_i^{(c)}(\tau)B_{-1}(\tau)'d\tau & \int_0^1\underline{\mu}_i^{(c)}(\tau)\underline{\mu}_i^{(c)}(\tau)'d\tau + \int_0^1\left(\Omega_i^{(x)}(\tau)\otimes B(\tau)B(\tau)'\right)d\tau \end{pmatrix}\varpi \\ &= \varpi' D_{1i}\begin{pmatrix} I_{K-1} & \mathbf{0}_{(K-1)\times dK} \\ \mathbf{0}_{dK\times(K-1)} & D_{0i} \end{pmatrix}D'_{1i}\varpi \end{aligned}$$

where $D_{1i} = \begin{pmatrix} I_{K-1} & \mathbf{0} \\ -\int_0^1 \underline{\mu}_i^{(c)}(\tau) B_{-1}(\tau)' d\tau & I_{Kd} \end{pmatrix}$, $D_{0i} = \int_0^1 \left(\Omega_i(\tau) \otimes B(\tau) B(\tau)'\right) d\tau + \bar{D}_{0i}$, and

$$\bar{D}_{0i} = \int_0^1 \underline{\mu}_i^{(c)}(\tau) \underline{\mu}_i^{(c)}(\tau)' d\tau - \int_0^1 \underline{\mu}_i^{(c)}(\tau) B_{-1}(\tau)' d\tau \int_0^1 B_{-1}(\tau) \underline{\mu}_i^{(c)}(\tau)' d\tau.$$

Noting that $D_{1i} D_{1i}' = I$, we have $A_{i,\varpi}^{(1)} + A_{i,\varpi}^{(2)} \geq \lambda_{\min}(D_{0i}) \varpi' D_{1i} D_{1i}' \varpi = \lambda_{\min}(D_{0i}) \|\varpi\|^2$

$$\begin{aligned} A_{i,\varpi}^{(1)} + A_{i,\varpi}^{(2)} &\geq \lambda_{\min}(D_{0i}) \varpi' D_{1i} D_{1i}' \varpi = \lambda_{\min}(D_{0i}) \|\varpi\|^2 \\ &\geq \lambda_{\min}\left(\bar{D}_{0i}\right) \|\varpi\|^2 + \lambda_{\min}\left[\int_0^1 \left(\Omega_i^{(x)}(\tau) \otimes B(\tau) B(\tau)'\right) d\tau\right] \|\varpi\|^2 \end{aligned}$$

by Weyl inequality. Noting that

$$\begin{aligned} &\lambda_{\min}\left[\int_0^1 \left(\Omega_i^{(x)}(\tau) \otimes B(\tau) B(\tau)'\right) d\tau\right] \\ &= \inf_{\|C\|=1, C \in \mathbb{R}^{d \times K}} \int_0^1 \operatorname{vec}(C)' \left(\Omega_i^{(x)}(\tau) \otimes B(\tau) B(\tau)'\right) \operatorname{vec}(C) d\tau \\ &= \inf_{\|C\|=1} \int_0^1 B(\tau)' C' \Omega_i^{(x)}(\tau) C B(\tau) d\tau \\ &\geq \lambda_{\min}\left(\Omega_i^{(x)}(\tau)\right) \int_0^1 \operatorname{tr}\left[B(\tau)' C' C B(\tau)\right] d\tau \\ &= \lambda_{\min}\left(\Omega_i^{(x)}(\tau)\right) \operatorname{tr}\left[C' C \left(\int_0^1 B(\tau) B(\tau)' d\tau\right)\right] \\ &= \lambda_{\min}\left(\Omega_i^{(x)}(\tau)\right) \operatorname{tr}\left(C' C\right) \\ &= \|C\|^2 \lambda_{\min}\left(\Omega_i^{(x)}(\tau)\right) = \lambda_{\min}\left(\Omega_i^{(x)}(\tau)\right) \geq \min_i \lambda_{\min}\left(\Omega_i^{(x)}(\tau)\right) \end{aligned}$$

we are left to show that $\bar{D}_{0i}$ is positive semi-definite (p.s.d.). Define

$$\underline{\mu}_{i,P}^{(c)}(\tau) = \int_0^1 \underline{\mu}_i^{(c)}(\tau) B_{-1}(\tau)' d\tau \left\{\int_0^1 B_{-1}(\tau) B_{-1}(\tau)' d\tau\right\}^{-1} B_{-1}(\tau).$$

Clearly, by the fact that $\int_0^1 B_{-1}(\tau) B_{-1}(\tau)' d\tau = I_{K-1}$, we have

$$\begin{aligned} \underline{\mu}_{i,P}^{(c)}(\tau) &= \int_0^1 \underline{\mu}_i^{(c)}(\tau) B_{-1}(\tau)' d\tau B_{-1}(\tau), \\ \int_0^1 \underline{\mu}_{i,P}^{(c)}(\tau) \underline{\mu}_{i,P}^{(c)}(\tau)' d\tau &= \int_0^1 \underline{\mu}_i^{(c)}(\tau) B_{-1}(\tau)' d\tau \left(\int_0^1 B_{-1}(\tau) B_{-1}(\tau)' d\tau\right) \int_0^1 B_{-1}(\tau) \underline{\mu}_i^{(c)}(\tau)' d\tau \\ &= \int_0^1 \underline{\mu}_i^{(c)}(\tau) B_{-1}(\tau)' d\tau \int_0^1 B_{-1}(\tau) \underline{\mu}_i^{(c)}(\tau)' d\tau. \end{aligned}$$

Observing that

$$\int_0^1 \underline{\mu}_i^{(c)}(\tau)\,\underline{\mu}_{i,P}^{(c)}(\tau)'\,d\tau = \int_0^1 \underline{\mu}_i^{(c)}(\tau)\,B_{-1}(\tau)'\,d\tau \int_0^1 B_{-1}(\tau)\,\underline{\mu}_i^{(c)}(\tau)'\,d\tau = \int_0^1 \underline{\mu}_{i,P}^{(c)}(\tau)\,\underline{\mu}_{i,P}^{(c)}(\tau)'\,d\tau$$

we can write $\bar{D}_{0i}$ as

$$\bar{D}_{0i} = \int_0^1 \left[\underline{\mu}_i^{(c)}(\tau) - \underline{\mu}_{i,P}^{(c)}(\tau)\right]\left[\underline{\mu}_i^{(c)}(\tau) - \underline{\mu}_{i,P}^{(c)}(\tau)\right]' d\tau.$$

Clearly, $\bar{D}_{0i}$ is p.s.d. and $\lambda_{\min}\left(\bar{D}_{0i}\right) \geq 0$.

(ii) The proof of (ii) is much simpler than (i). It is omitted here.

(iii)) The proof is analogous to (i) and thus is omitted. We can replace X_{it} by $\varepsilon_{it}X_{it}$ and apply Assumption 1(vi) in place of Assumption 1(v). Noting that $\mathrm{Var}(\varepsilon_{it}X_{it}) = E(\varepsilon_{it}^2 X_{it}X_{it}')$. Assumption 1(v) and moment conditions on $\varepsilon_{it}X_{it}$ suffice to the proof of (v). ■

Proof for Lemma A.5. Note that

$$\begin{aligned}
\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T r_{g,it}^2 &= \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \left(r_{f,it} + X_{it}'r_{\beta,it}\right)^2 \\
&\leq \frac{2}{NT}\sum_{i=1}^N\sum_{t=1}^T X_{it}'r_{\beta,it}r_{\beta,it}'X_{it} + \frac{2}{NT}\sum_{i=1}^N\sum_{t=1}^T r_{f,it}^2 \\
&\leq \max_i \sup_{\tau\in[0,1]} r_{f,i}^2(\tau) + \max_i \sup_{\tau\in[0,1]} \|r_{\beta,i}(\tau)\|^2 \frac{2}{NT}\sum_{i=1}^N\sum_{t=1}^T \|X_{it}\|^2 \\
&= O\left(K^{-2\kappa}\right) + O_p\left(K^{-2\kappa}\right)O_p(1) = O_p\left(K^{-2\kappa}\right)
\end{aligned}$$

by Assumption 3 in Newey (1997). ■

Proof for Lemma A.6. (i) By the repeated use of $n\lambda_{\min}(A) \leq \mathrm{tr}(A) \leq n\lambda_{\max}(A)$ and $\lambda_{\min}(A)\mathrm{tr}(B) \leq \mathrm{tr}(AB) \leq \lambda_{\max}(A)\mathrm{tr}(B)$ for p.s.d. $n\times n$ matrix A and symmetric matrix B, we have

$$\mathbb{V}_{NT} = \frac{2}{N}\sum_{i=1}^N E\left[\mathrm{tr}\left(\mathbb{Q}_i\mathring{\Omega}_i\mathbb{Q}_i\mathring{\Omega}_i\right)\right] \leq CK\max_i \lambda_{\max}^2(\mathbb{Q}_i)\max_i \lambda_{\max}^2\left(\mathring{\Omega}_i\right) = O_p(K),$$

and

$$\mathbb{V}_{NT} = \frac{2}{N}\sum_{i=1}^N E\left[\mathrm{tr}\left(\mathbb{Q}_i\mathring{\Omega}_i\mathbb{Q}_i\mathring{\Omega}_i\right)\right] \geq CK\min_i \lambda_{\min}^2(\mathbb{Q}_i)\min_i \lambda_{\min}^2\left(\mathring{\Omega}_i\right) = O_p(K).$$

It follows that $\mathbb{V}_{NT} \asymp K$.

(ii) Note that $\mathring{\mathcal{K}}_{i,tt} = \mathring{Z}_{it}'\mathbb{Q}_i\mathring{Z}_{it} \leq \lambda_{\max}(\mathbb{Q}_i)||\mathring{Z}_{it}||^2 \leq C||\mathring{Z}_{it}||^2$ uniformly in i and t. Similarly, $\mathring{\mathcal{K}}_{i,tt} \geq \lambda_{\min}^2(\mathbb{Q}_i)||\mathring{Z}_{it}||^2 \geq C||\mathring{Z}_{it}||^2$ uniformly in i and t. It follows

$$\mathbb{B}_{NT} \asymp \frac{1}{\sqrt{N}T}\sum_{i=1}^N\sum_{t=1}^T E(||\mathring{Z}_{it}||^2\varepsilon_{it}^2) \asymp KN^{1/2}.$$

■

Proof for Lemma A.7. (i) Recall that $\mathcal{K}_i = \dot{Z}_i \hat{\mathbb{Q}}_i \dot{Z}_i'$ and $\mathring{\mathcal{K}}_i = \mathring{Z}_i \mathbb{Q}_i \mathring{Z}_i'$. First, note that $\dot{Z}_{it} = \mathring{Z}_{it} + \bar{\xi}_{z,i}^{(1)}$, where $\bar{\xi}_{z,i}^{(1)} = \bar{Z}_i - E\bar{Z}_i = T^{-1}\sum_{t=1}^T \xi_{z,it}^{(1)}$ and $\xi_{z,it}^{(1)} = Z_{it} - E(Z_{it})$. As the proof of Lemma A.3, we can show that $\max_i ||\bar{\xi}_{z,i}^{(1)}|| = O_p[(K\ln N/T)^{1/2}]$; Second, by Taylor expansion and keeping the first order terms, we have $\hat{\mathbb{Q}}_i - \mathbb{Q}_i = \bar{\xi}_{\mathbb{Q},i} + \mathbb{Q}_i^{(R)}$, where $\bar{\xi}_{\mathbb{Q},i} = T^{-1}\sum_{t=1}^T \xi_{\mathbb{Q},it}$, $\xi_{\mathbb{Q},it} = Q_{i,\dot{z}\dot{z}}^{-1}\xi_{\dot{z},it}\mathbb{Q}_i + \mathbb{Q}_i \xi_{\dot{z},it} Q_{i,\dot{z}\dot{z}}^{-1} + Q_{i,\dot{z}\dot{z}}^{-1}\xi_{z,it}^{(2)} Q_{i,\dot{z}\dot{z}}^{-1}$ with

$$\xi_{z,it}^{(2)} = Z_{it}Z_{it}' - E\left(Z_{it}Z_{it}'\right) \text{ and } \xi_{\dot{z},it} = \xi_{z,it}^{(2)} - \xi_{z,it}^{(1)} E\left(\bar{Z}_i'\right) - E\left(\bar{Z}_i\right)\xi_{z,it}^{(1)\prime},$$

and $\mathbb{Q}_i^{(R)}$ comes from the higher order (≥ 2) terms and $\max_i ||\mathbb{Q}_i^{(R)}|| = O_p(K^2 \ln N/T)$. It follows that

$$\begin{aligned}
&\frac{1}{NT^2}\sum_{i=1}^N \varepsilon_i'\left(\mathcal{K}_i - \mathring{\mathcal{K}}_i\right)\varepsilon_i \\
&= \frac{1}{NT^2}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T\left[\left(\mathring{Z}_{it} + \bar{\xi}_{z,i}^{(1)}\right)'\left(\mathbb{Q}_i + \bar{\xi}_{\mathbb{Q},i} + \mathbb{Q}_i^{(R)}\right)\left(\mathring{Z}_{it} + \bar{\xi}_{z,i}^{(1)}\right) - \mathring{\mathcal{K}}_{i,ts}\right]\varepsilon_{is}\varepsilon_{it} \\
&= \frac{1}{NT^3}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T\left[\mathring{Z}_{it}'\mathbb{Q}_i\bar{\xi}_{z,i}^{(1)} + \mathring{Z}_{it}'\bar{\xi}_{\mathbb{Q},i}\mathring{Z}_{it} + \mathring{Z}_{it}'\bar{\xi}_{\mathbb{Q},i}\bar{\xi}_{z,i}^{(1)} + \mathring{Z}_{it}'\mathbb{Q}_i^{(R)}\mathring{Z}_{it} + \mathring{Z}_{it}'\mathbb{Q}_i^{(R)}\bar{\xi}_{z,i}^{(1)}\right. \\
&\left.+\bar{\xi}_{z,i}^{(1)\prime}\mathbb{Q}_i\mathring{Z}_{it} + \bar{\xi}_{z,i}^{(1)\prime}\mathbb{Q}_i\bar{\xi}_{z,i}^{(1)} + \bar{\xi}_{z,i}^{(1)\prime}\bar{\xi}_{\mathbb{Q},i}\mathring{Z}_{it} + \bar{\xi}_{z,i}^{(1)\prime}\bar{\xi}_{\mathbb{Q},i}\bar{\xi}_{z,i}^{(1)} + \bar{\xi}_{z,i}^{(1)\prime}\mathbb{Q}_i^{(R)}\mathring{Z}_{it} + \bar{\xi}_{z,i}^{(1)\prime}\mathbb{Q}_i^{(R)}\bar{\xi}_{z,i}^{(1)}\right]\varepsilon_{is}\varepsilon_{it} \\
&= \sum\nolimits_{l=1}^{11} E_{NT,l}, \text{ say.}
\end{aligned}$$

It is easy to show that $E_{NT,l} = O_p\left(K^{3/2}/\left(N^{1/2}T^{3/2}\right)\right) = o\left(N^{-1/2}T^{-1}K^{1/2}\right)$ for $l = 1$, 2, 6, by Chebyshev inequality and $E_{NT,l} = O_p\left(K^2\ln N/T\right)O_p\left(K/T\right) = O_p\left(K^3\ln N/T^2\right) = o_p\left(N^{-1/2}T^{-1}K^{1/2}\right)$ for other l's.

(ii) The proof is simpler than that of (i). We omit the details to save space. ■

Proof of Lemma A.8. See the proof of Theorem 4.1 in Shao and Yu (1996). ■

Proof of Lemma A.9. Recall that $\mathring{Z}_{it} = Z_{it} - E\left(\bar{Z}_i\right)$. Define $\check{X}_{it} = X_{it} - E\left(X_{it}\right)$. Let $\mathring{Z}_{it,l}$ and $\check{X}_{it,l}$ be the l-th element of $\mathring{Z}_{it}$ and $\check{X}_{it}$, respectively. By the moment conditions and Assumption 1, the following holds:

$$\max_{1\leq i\leq N}\lambda_{\max}\left(T^{-1}\textstyle\sum_{t=1}^T E[\mathring{Z}_{it}\mathring{Z}_{it}']\right) < \infty,$$

$$0 < \underline{c}_{zz\varepsilon^2} = \min_{1\leq i\leq N}\lambda_{\min}\left(T^{-1}\textstyle\sum_{t=1}^T E[\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2]\right) \leq \max_{1\leq i\leq N}\lambda_{\max}\left(T^{-1}\textstyle\sum_{t=1}^T E[\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2]\right) = \overline{c}_{zz\varepsilon^2} < \infty,$$

$$\max_{1\leq i\leq N}\max_{1\leq l\leq d}\lambda_{\max}\left(T^{-1}\textstyle\sum_{t=1}^T E[\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}^2]\right) < \infty.$$

For a matrix or vector A, we denote $\|A\|_\infty = \max_{i,j}|A_{ij}|$ where A_{ij} denotes (i,j)-th element of A.

Before the main proof of the lemma, we first make the following 7 claims (claims 0-6) whose proofs are deferred after the main proof.

Claim 0. For $n \times n$ matrices A and B,

$$\max \{|\lambda_{\min}(A) - \lambda_{\min}(B)|, |\lambda_{\max}(A) - \lambda_{\max}(B)|\} \leq n \|A - B\|_{\infty}. \tag{A.2}$$

Claim 1. Let $K^* = K_1 + dK$.

$$\left(\bigcap_{i=1}^{N} \{\|\bar{X}_i - E(\bar{X}_i)\| < \epsilon/K^*\}\right) \bigcap \left(\bigcap_{i=1}^{N} \{|\bar{\varepsilon}_i| < \epsilon/K\}\right) \bigcap \left(\left\|\beta_P - \hat{\beta}_{FE}\right\| < \epsilon/K^*\right)$$

holds almost surely after some T_1^*.

Claim 2.

$$\bigcap_{i=1}^{N} \left\{\left\|T^{-1}\sum_{t=1}^{T} X_{it}\varepsilon_{it}\right\| < \epsilon/K^*\right\} \bigcap \left(\bigcap_{i=1}^{N} \left\|T^{-1}\sum_{t=1}^{T} \left[\check{X}_{it}\check{X}_{it}' - E(\check{X}_{it}\check{X}_{it}')\right]\right\|_{\infty} < \epsilon/K^*\right)$$

holds almost surely after some T_2^*.

Claim 3.

$$\bigcap_{i=1}^{N} \left\{\left\|T^{-1}\sum_{t=1}^{T} [Z_{it} - E(Z_{it})]\right\|_{\infty} < \epsilon/K^*\right\}$$

holds almost surely after some T_3^*.

Claim 4.

$$\bigcap_{i=1}^{N} \left(\left\{\left\|T^{-1}\sum_{t=1}^{T} [\mathring{Z}_{it}\mathring{Z}_{it}' - E(\mathring{Z}_{it}\mathring{Z}_{it}')]\right\|_{\infty} < \epsilon/K^*\right\}\right),$$

$$\bigcap_{i=1}^{N} \left(\left\{\left\|T^{-1}\sum_{t=1}^{T} [\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2 - E(\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2)]\right\|_{\infty} < \epsilon/K^*\right\}\right),$$

$$\bigcap_{i=1}^{N} \left(\bigcap_{l=1}^{d} \left\{\left\|T^{-1}\sum_{t=1}^{T} [\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l} - E(\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l})]\right\|_{\infty} < \epsilon/K^*\right\}\right), \text{ and}$$

$$\bigcap_{i=1}^{N} \left(\bigcap_{l=1}^{d}\bigcap_{l'=1}^{d} \left\{\left\|T^{-1}\sum_{t=1}^{T} [\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it,l'} - E(\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it,l'})]\right\|_{\infty} < \epsilon/K^*\right\}\right)$$

hold almost surely after some T_4^*.

Claim 5. We can decompose $\hat{\Omega}_i$ as:

$$\hat{\Omega}_i = T^{-1}\sum_{t=1}^{T} \mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it} + \tilde{f}_{it})^2 + \varepsilon_{it}^2] + R_i,$$

where

$$\check{\beta}_{it} = \beta_{it} - \beta_P \text{ and } \tilde{f}_{it} = E(X_{it})'\check{\beta}_{it} - T^{-1}\sum_{s=1}^{T} E(X_{is})'\check{\beta}_{is} + f_{it} - \bar{f}_i,$$

and R_i satisfies that $\bigcap_{i=1}^{N} \{\|R_i\|_{\infty} \leq \epsilon/K^*\}$ holds almost surely after some T_5^*.

Claim 6.

$$\min_{1 \leq i \leq N} \lambda_{\min}\left(T^{-1}\sum_{t=1}^{T} E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it} + \tilde{f}_{it})^2 + \varepsilon_{it}^2]\right]\right) \geq \underline{c}_{zz\varepsilon^2}, \text{ and}$$

$$\max_{1 \leq i \leq N} \lambda_{\max}\left(T^{-1}\sum_{t=1}^{T} E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it} + \tilde{f}_{it})^2 + \varepsilon_{it}^2]\right]\right) \leq C^{\dagger} < \infty$$

for a positive $C^{\dagger}$.

Main proof. Using the decomposition in Claim 5,

$$\begin{aligned} &\hat{\Omega}_i - T^{-1}\sum_{t=1}^{T} E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it} + \tilde{f}_{it})^2 + \varepsilon_{it}^2]\right] \\ &= T^{-1}\sum_{t=1}^{T} \left\{\mathring{Z}_{it}\mathring{Z}_{it}'\left[(\check{X}_{it}'\check{\beta}_{it} + \tilde{f}_{it})^2 + \varepsilon_{it}^2\right] - E\left[\mathring{Z}_{it}\mathring{Z}_{it}'((\check{X}_{it}'\check{\beta}_{it} + \tilde{f}_{it})^2 + \varepsilon_{it}^2)\right]\right\} + R_i. \end{aligned} \tag{A.3}$$

The difference of the first two terms in the above is bounded by a linear combination of finite number of (because d is fixed) $T^{-1}\sum_{t=1}^{T}[\mathring{Z}_{it}\mathring{Z}_{it}' - E(\mathring{Z}_{it}\mathring{Z}_{it}')]$, $T^{-1}\sum_{t=1}^{T}[\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l} - E(\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l})]$, $T^{-1}\sum_{t=1}^{T}[\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it,l'} - E(\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it,l'})]$, and $\sum_{t=1}^{T}[\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2 - E(\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2)]$. By the results in Claim 4 and uniform boundedness of $\|\beta_{it}\|$ and f_{it}, for any $\epsilon > 0$,

$$\bigcap_{i=1}^{N}\left\{\left\|T^{-1}\sum_{t=1}^{T}\left\{\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2] - E[\mathring{Z}_{it}\mathring{Z}_{it}'((\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2)]\right\}\right\|_{\infty} \le \frac{\epsilon}{K^*}\right\} \tag{A.4}$$

holds almost surely after some T_7^*. Together, equation (A.3), the result in equation (A.4), and the result in Claim 5 (on R_i) imply that

$$\bigcap_{i=1}^{N}\left\{\left\|\hat{\Omega}_i - T^{-1}\sum_{t=1}^{T}E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2]\right]\right\|_{\infty} \le \frac{\epsilon}{K^*}\right\} \tag{A.5}$$

holds almost surely after some T_8^*.

Apply the result in equation (A.2) in Claim 0,

$$\begin{aligned}
&\left|\lambda_{\min}\left(\hat{\Omega}_i\right) - \lambda_{\min}\left(T^{-1}\sum_{t=1}^{T}E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2]\right]\right)\right| \\
&\le K^*\left\|\hat{\Omega}_i - T^{-1}\sum_{t=1}^{T}E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2]\right]\right\|_{\infty}, \text{ and} \\
&\left|\lambda_{\max}\left(\hat{\Omega}_i\right) - \lambda_{\max}\left(T^{-1}\sum_{t=1}^{T}E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2]\right]\right)\right| \\
&\le K^*\left\|\hat{\Omega}_i - T^{-1}\sum_{t=1}^{T}E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2]\right]\right\|_{\infty}.
\end{aligned}$$

Therefore, the event in equation (A.5) implies that

$$\begin{aligned}
&\left|\lambda_{\min}\left(\hat{\Omega}_i\right) - \lambda_{\min}\left(T^{-1}\sum_{t=1}^{T}E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2]\right]\right)\right| \le \epsilon, \text{ and} \\
&\left|\lambda_{\max}\left(\hat{\Omega}_i\right) - \lambda_{\max}\left(T^{-1}\sum_{t=1}^{T}E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2]\right]\right)\right| \le \epsilon,
\end{aligned}$$

for all $i = 1, 2, ..., N$. By Claim 6 and setting ϵ small enough, the above guarantees that

$$0 < \frac{1}{2}c_{zz\varepsilon^2} \le \min_{1\le i\le N}\lambda_{\min}\left(\hat{\Omega}_i\right) \le \max_{1\le i\le N}\lambda_{\max}\left(\hat{\Omega}_i\right) \le \frac{3}{2}C^* < \infty. \tag{A.6}$$

Note we show that the event in equation (A.5) holds almost surely after some T_8^*. Since (A.5) implies (A.6), we can say the event in equation (A.6) holds almost surely after some T_8^*, as desired.

We now show those Claims we made. For the proof below, we use T^* to denote a large number that may vary from line to line, and ϵ to denote some arbitrary small constant that may vary across lines.

Proof of Claim 0. First, for $n \times n$ matrices A and B, we note

$$\lambda_{\max}(A+B) \le \lambda_{\max}(A) + \lambda_{\max}(B), \tag{A.7}$$

$$\max\{\lambda_{\min}(A), \lambda_{\min}(B)\} \le \lambda_{\min}(A+B) \text{ (for p.d. } A \text{ and p.d. } B\text{)},$$

$$\max\{|\lambda_{\min}(A)|, |\lambda_{\max}(A)|\} \le n\|A\|_{\infty}, \text{ and}$$

$$|\lambda_{\min}(A) - \lambda_{\min}(B)| \le \max\{|\lambda_{\min}(A-B)|, |\lambda_{\min}(B-A)|\},$$

which can be seen from the proof of Lemma 5 in Fan et al. (2011). The fourth line in the above implies (by taking $A = -A$ and $B = -B$)

$$|\lambda_{\max}(A) - \lambda_{\max}(B)| \leq \max\{|\lambda_{\max}(B-A)|, |\lambda_{\max}(A-B)|\}. \tag{A.8}$$

Applying the third line in equation (A.7) on $B - A$ and $A - B$**,** we have

$$\max\{|\lambda_{\min}(A-B)|, |\lambda_{\min}(B-A)|, |\lambda_{\max}(A-B)|, |\lambda_{\max}(B-A)|\} \leq n\|A-B\|_{\infty}. \tag{A.9}$$

Using the inequalities in the fourth line of equation (A.7) and the inequalities in (A.8) and (A.9), we obtain

$$\max\{|\lambda_{\min}(A) - \lambda_{\min}(B)|, |\lambda_{\max}(A) - \lambda_{\max}(B)|\} \leq n\|A-B\|_{\infty}.$$

Proof of Claims 1-3. These can be proved similarly as we do for the results in Claim 4, and they require less strict conditions. We omit the proof for brevity.

Proof of Claim 4. We only show the second and the fourth equations and others can be proved similarly. We show the second equation first. We define event $\mathcal{B}^{(1)}_{iT,(j,k)}$ as

$$\mathcal{B}^{(1)}_{iT,(j,k)} = \left\{\left|T^{-1}\textstyle\sum_{t=1}^{T}\mathring{Z}_{it,j}\mathring{Z}_{it,k}\varepsilon_{it}^2 - E\left(\mathring{Z}_{it,j}\mathring{Z}_{it,k}\varepsilon_{it}^2\right)\right| \geq \epsilon/K^*\right\},$$

which is regarding the (j,k)-th element of $T^{-1}\sum_{t=1}^{T}[\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2 - E(\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2)]$. Note that $Z_{it} = \left(B_{-1,t}, (X_{it}\otimes B_t)'\right)'$ and B_t are uniformly bounded. Therefore $E(\mathring{Z}_{it,j}^{16+\eta})$ is bounded. As a result, we are able to apply Lemma A.8 with $r = 4 + 9\eta/40$ and $\delta = \eta/40$ on $\mathcal{B}^{(1)}_{iT,(j,k)}$, and we obtain

$$\begin{aligned} P\left(\mathcal{B}^{(1)}_{iT,(j,k)}\right) &\leq \frac{K^{*(4+9\eta)/40}}{\epsilon^{(4+9\eta)/40}T^{4+9\eta/40}}E\left[\left|\textstyle\sum_{t=1}^{T}\left[\mathring{Z}_{it,j}\mathring{Z}_{it,k}'\varepsilon_{it}^2 - E(\mathring{Z}_{it,j}\mathring{Z}_{it,k}'\varepsilon_{it}^2)\right]\right|^{(4+9\eta)/40}\right] \\ &\leq \frac{C_1K^{*4+9\eta/40}T^{2+9\eta/80}}{\epsilon^{(4+9\eta)/40}T^{4+9\eta/40}} = \frac{C_1K^{*4+9\eta/40}}{\epsilon^{4+9\eta/40}T^{2+9\eta/80}} \end{aligned} \tag{A.10}$$

for some $C_1 > 0$.

Define

$$\mathcal{B}^{(1)}_{iT,(\cdot,\cdot)} = \left\{\left\|T^{-1}\textstyle\sum_{t=1}^{T}\left[\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2 - E(\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2)\right]\right\|_{\infty} \geq \epsilon/K^*\right\} = \textstyle\bigcup_{j=1}^{K^*}\bigcup_{l=1}^{K^*}\mathcal{B}^{(1)}_{iT,(j,k)},$$

and

$$\mathcal{B}^{(1)*}_{T} = \textstyle\bigcup_{i=1}^{N}\mathcal{B}^{(1)}_{iT,(\cdot,\cdot)}. \tag{A.11}$$

Using the bound in equation (A.10), the probability bound for $\mathcal{B}^{(1)*}_{T}$ can be calculated as

$$\begin{aligned} P\left(\mathcal{B}^{(1)*}_{T}\right) &= P\left(\textstyle\bigcup_{i=1}^{N}\bigcup_{j=1}^{K^*}\bigcup_{l=1}^{K^*}\mathcal{B}^{(1)}_{iT,(j,k)}\right) \\ &\leq \sum_{i=1}^{N}\sum_{j=1}^{K^*}\sum_{l=1}^{K^*}P\left(\mathcal{B}^{(1)}_{iT,(j,k)}\right) \leq \frac{C_1NK^{*6+9\eta/40}}{\epsilon^{4+9\eta/40}T^{2+9\eta/80}} \leq \frac{C_2}{T(\ln T)^2}, \end{aligned} \tag{A.12}$$

for a positive C_2, where the last inequality holds by $\frac{NK^{*(6+9\eta)/40}}{T^{(1+9\eta)/80}} = O[(\ln T)^{-2}]$. For the sequence of events $\{\mathcal{B}_T^{(1)*}\}_{T=3}^{\infty}$,[13] using the result in equation (A.12), one can obtain

$$\sum_{T=3}^{\infty} P\left(\mathcal{B}_T^{(1)*}\right) \leq \sum_{T=3}^{\infty} \frac{C_2}{T(\ln T)^2} \leq C_2 \int_2^{\infty} \frac{1}{x(\ln x)^2} dx$$
$$= C_2 \int_2^{\infty} \frac{1}{(\ln x)^2} d\ln(x) = C_2 \int_{\ln 2}^{\infty} \frac{1}{u^2} du < \infty.$$

With the above, we are able to apply the Borel-Cantelli Lemma which implies that there exists a large T^* such that

$$P\left(\bigcup_{T=T^*}^{\infty} \mathcal{B}_T^{(1)*}\right) = 0.$$

Thus,

$$\begin{aligned}
1 &= P\left(\left(\bigcup_{T=T^*}^{\infty} \mathcal{B}_T^{(1)*}\right)^c\right) = P\left(\bigcap_{T=T^*}^{\infty} \mathcal{B}_T^{(1)*c}\right) \\
&= P\left(\bigcap_{T=T^*}^{\infty} \bigcap_{i=1}^{N} \bigcap_{j=1}^{K^*} \bigcap_{l=1}^{K^*} \mathcal{B}_{iT,(j,k)}^{(1)c}\right) \\
&= P\left(\bigcap_{T=T^*}^{\infty} \bigcap_{i=1}^{N} \bigcap_{j=1}^{K^*} \bigcap_{l=1}^{K^*} \left\{\left|T^{-1}\sum_{t=1}^{T}\left[\mathring{Z}_{it,j}\mathring{Z}_{it,k}\varepsilon_{it}^2 - E(\mathring{Z}_{it,j}\mathring{Z}_{it,k}\varepsilon_{it}^2)\right]\right| < \epsilon/K^*\right\}\right) \\
&= P\left(\bigcap_{T=T^*}^{\infty} \bigcap_{i=1}^{N} \left\{\left\|T^{-1}\sum_{t=1}^{T}\left[\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2 - E(\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2)\right]\right\|_{\infty} \leq \epsilon/K^*\right\}\right),
\end{aligned}$$

where the last equality holds by the definition in equation (A.11). Note the above is equivalent to say the event in the second equation of Claim 4 holds almost surely after T^*.

The result on the fourth equation can be proved similarly. Define event

$$\mathcal{B}_{iT,(j,k,l,l')}^{(2)} = \left\{\left|T^{-1}\sum_{t=1}^{T}\left[\mathring{Z}_{it,j}\mathring{Z}_{it,k}\check{X}_{it,l}\check{X}_{it,l'} - E(\mathring{Z}_{it,j}\mathring{Z}_{it,k}\check{X}_{it,l}\check{X}_{it,l'})\right]\right| \geq \epsilon/K^*\right\},$$

which is regarding the (j,k)-th element of $T^{-1}\sum_{t=1}^{T}\left[\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it,l'} - E(\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it,l'})\right]$. For the same reason as we obtain equation (A.10), we can obtain the probability bound

$$\begin{aligned}
&P\left(\mathcal{B}_{iT,(j,k,l,l')}^{(2)}\right) \\
&\leq \frac{K^{*(4+9\eta)/40}}{\epsilon^{(4+9\eta)/40}T^{(4+9\eta)/40}} E\left[\left|\sum_{t=1}^{T}\left[\mathring{Z}_{it,j}\mathring{Z}_{it,k}\check{X}_{it,l}\check{X}_{it,l'} - E(\mathring{Z}_{it,j}\mathring{Z}_{it,k}\check{X}_{it,l}\check{X}_{it,l'})\right]\right|^{(4+9\eta)/40}\right] \\
&\leq \frac{C_3 K^{*4+9\eta/40}}{\epsilon^{4+9\eta/40}T^{2+9\eta/80}}. \qquad \text{(A.13)}
\end{aligned}$$

Similarly, we define

$$\begin{aligned}
\mathcal{B}_{iT,(\cdot,\cdot,\cdot,\cdot)}^{(2)} &= \bigcup_{l=1}^{d}\bigcup_{l'=1}^{d}\left\{\left\|T^{-1}\sum_{t=1}^{T}\left[\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it,l'} - E(\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it,l'})\right]\right\|_{\infty} \geq \epsilon/K^*\right\} \\
&= \bigcup_{l=1}^{d}\bigcup_{l'=1}^{d}\left\{\bigcup_{j=1}^{K^*}\bigcup_{k=1}^{K^*}\mathcal{B}_{iT,(j,k,l,l')}^{(2)}\right\},
\end{aligned}$$

[13] To study whether the sequence of events $\left\{\mathcal{B}_T^{(1)*}\right\}_{T=1}^{\infty}$ happen infinitely often, we can ignore a finite number of $\mathcal{B}_T^{(1)*}$. We drop $\mathcal{B}_1^{(1)*}$ and $\mathcal{B}_2^{(1)*}$ for technical convenience.

and

$$\mathcal{B}_T^{(2)*} = \bigcup_{i=1}^N \mathcal{B}_{iT,(\cdot,\cdot,\cdot,\cdot)}^{(2)}.$$

For the sequence of events $\{\mathcal{B}_T^{(2)*}\}_{T=3}^{\infty}$, note

$$P\left(\mathcal{B}_T^{(2)*}\right) = P\left(\bigcup_{i=1}^N \bigcup_{l=1}^d \bigcup_{l'=1}^d \bigcup_{j=1}^{K^*} \bigcup_{l=1}^{K^*} \mathcal{B}_{iT,(j,k,l,l')}^{(2)}\right) \leq \sum_{i=1}^{N}\sum_{l=1}^{d}\sum_{l'=1}^{d}\sum_{j=1}^{K^*}\sum_{k=1}^{K^*} P\left(\mathcal{B}_{iT,(j,k,l,l')}^{(2)}\right)$$
$$\leq \frac{C_3 N d^2 K^{*6+9\eta/40}}{\epsilon^{4+9\eta/40} T^{2+9\eta/80}} \leq \frac{C_4 N K^{*6+9\eta/40}}{\epsilon^{4+9\eta/40} T^{2+9\eta/80}} \leq \frac{C_5}{\epsilon^{4+9\eta/40} T (\ln T)^2},$$

for some positive C_4 and C_5, where the second line holds by using the bound in equation (A.13) and the fact that d is a fixed number, and the last line holds by $\frac{NK^{*6+9\eta/40}}{T^{1+9\eta/80}} = O[(\ln T)^{-2}]$. So we have shown a similar result as in equation (A.12). Continuing with the same logic as we do for the event in the second equation in Claim 4, we can say that the event

$$\bigcap_{i=1}^N \left(\bigcap_{l=1}^d \bigcap_{l'=1}^d \left\{\left\| T^{-1}\sum_{t=1}^T \left[\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it,l'} - E(\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it,l'})\right]\right\|_\infty < \epsilon/K^*\right\}\right)$$

holds almost surely after some large T^*, as desired.

Proof of Claim 5. Note that the model is

$$Y_{it} = X_{it}'\beta_{it} + f_{it} + \alpha_i + \varepsilon_{it}.$$

Recall that $\beta_P = [\sum_{i=1}^N E(X_i' M_T X_i)]^{-1} \sum_{i=1}^N E(X_i' M_T Y_i)$ and $\hat{\Omega}_i = T^{-1}\sum_{t=1}^T \dot{Z}_{it}\dot{Z}_{it}'\hat{\varepsilon}_{r,it}^2$. We divide the analysis into four parts. The first part is on $\hat{\varepsilon}_{r,it}^2$. The second part is on $\dot{Z}_{it}\dot{Z}_{it}'$. The third part puts the results in parts 1 and 2 together, and shows the decomposition in the claim. The fourth part shows that $\bigcap_{i=1}^N \{\|R_i\|_\infty \leq \epsilon/K^*\}$ holds almost surely after some large T^*.

Part 1. Note that $\hat{\varepsilon}_{r,it} = \hat{u}_{it} - \overline{\hat{u}}_i$. The population version of $\hat{u}_{it}$ is

$$u_{it} \equiv Y_{it} - X_{it}'\beta_P = X_{it}'(\beta_{it} - \beta_P) + f_{it} + \alpha_i + \varepsilon_{it},$$

and the population version of $\hat{u}_{it} - \overline{\hat{u}}_i$ is

$$\begin{aligned} u_{it} - \overline{u}_i &= X_{it}'(\beta_{it} - \beta_P) - T^{-1}\sum_{t=1}^T X_{it}'(\beta_{it} - \beta_P) + f_{it} - \bar{f}_i + \varepsilon_{it} - \bar{\varepsilon}_i \\ &= [X_{it} - E(X_{it})]'(\beta_{it} - \beta_P) - T^{-1}\sum_{t=1}^T [X_{it} - E(X_{it})]'(\beta_{it} - \beta_P) + \varepsilon_{it} \\ &\quad + E(X_{it})'(\beta_{it} - \beta_P) - T^{-1}\sum_{t=1}^T E(X_{it})'(\beta_{it} - \beta_P) + f_{it} - \bar{f}_i - \bar{\varepsilon}_i. \end{aligned}$$

Using the expression of u_{it} and $\bar{u}_i$, $\hat{u}_{it}$ and $\hat{u}_{it} - \bar{u}_i$ can be written as

$$\begin{aligned} \hat{u}_{it} &= u_{it} + X_{it}'\left(\beta_P - \hat{\beta}_{FE}\right) = X_{it}'(\beta_{it} - \beta_P) + f_{it} + \alpha_i + \varepsilon_{it} + X_{it}'\left(\beta_P - \hat{\beta}_{FE}\right), \\ \overline{\hat{u}}_i &= \bar{u}_i + \bar{X}_i'\left(\beta_P - \hat{\beta}_{FE}\right). \end{aligned}$$

As a result,

$$
\begin{aligned}
\hat{u}_{it}-\overline{\hat{u}}_i &= u_{it}-\bar{u}_i+\left(X_{it}-\bar{X}_i\right)'\left(\beta_P-\hat{\beta}_{FE}\right)\\
&=\left[X_{it}-E\left(X_{it}\right)\right]'\left(\beta_{it}-\beta_P\right)-T^{-1}{\textstyle\sum_{t=1}^{T}}\left[X_{it}-E\left(X_{it}\right)\right]'\left(\beta_{it}-\beta_P\right)+\varepsilon_{it}\\
&+E\left(X_{it}\right)'\left(\beta_{it}-\beta_P\right)-T^{-1}{\textstyle\sum_{t=1}^{T}}E\left(X_{it}\right)'\left(\beta_{it}-\beta_P\right)+f_{it}-\bar{f}_i\\
&+\left(X_{it}-\bar{X}_i\right)'\left(\beta_P-\hat{\beta}_{FE}\right)-\bar{\varepsilon}_i\\
&=\check{X}_{it}'\left(\beta_{it}-\beta_P\right)\\
&+\left[E\left(X_{it}\right)'\left(\beta_{it}-\beta_P\right)-T^{-1}{\textstyle\sum_{t=1}^{T}}E\left(X_{it}\right)'\left(\beta_{it}-\beta_P\right)+f_{it}-\bar{f}_i\right]+\varepsilon_{it}\\
&+\left(X_{it}-\bar{X}_i\right)'\left(\beta_P-\hat{\beta}_{FE}\right)-\bar{\varepsilon}_i\\
&\equiv\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it}+\varepsilon_{it}+\left[X_{it}-E\left(X_{it}\right)\right]'\varrho_{1,NT}+\varrho_{i2}\\
&=\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it}+\varepsilon_{it}+\check{X}_{it}'\varrho_{1,NT}+\varrho_{i2},
\end{aligned}
$$

where

$$
\begin{aligned}
&\tilde{f}_{it}\equiv E\left(X_{it}\right)'\check{\beta}_{it}-T^{-1}{\textstyle\sum_{t=1}^{T}}E\left(X_{it}\right)'\check{\beta}_{it}+f_{it}-\bar{f}_i,\ \check{\beta}_{it}\equiv\beta_{it}-\beta_P,\\
&\varrho_{1,NT}=\beta_P-\hat{\beta}_{FE},\ \varrho_{i2}=\left[\bar{X}_i-E\left(\bar{X}_i\right)\right]'\left(\beta_P-\hat{\beta}_{FE}\right)-\bar{\varepsilon}_i.
\end{aligned}
$$

We have for any $\epsilon>0$,

$$
\left\{\left\|\varrho_{1,NT}\right\|<\frac{\epsilon}{K^*}\right\}\text{ and }{\textstyle\bigcap_{i=1}^{N}}\left\{\left|\varrho_{i2}\right|<\frac{\epsilon}{K}\right\}
$$

holds almost surely after some large T^* due to the results in Claims 1, 2, and 3. $\tilde{f}_i\left(t\right)$ is uniformly bounded due to the uniform boundedness of β_{it} and f_{it}. Recall $\hat{\varepsilon}_{r,it}=\hat{u}_{it}-\overline{\hat{u}}_i$, therefore,

$$
\begin{aligned}
\hat{\varepsilon}_{r,it}^2&=\left(\hat{u}_{it}-\overline{\hat{u}}_i\right)^2\\
&=(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it}+\varepsilon_{it}+\check{X}_{it}'\varrho_{1,NT}+\varrho_{i2})^2\\
&=(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2+2(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})\varepsilon_{it}\\
&+(\check{X}_{it}'\varrho_{1,NT}+\varrho_{i2})^2+2(\check{X}_{it}'\varrho_{1,NT}+\varrho_{i2})\varepsilon_{it}+2(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})(\check{X}_{it}'\varrho_{1,NT}+\varrho_{i2}).
\end{aligned}
\tag{A.14}
$$

Part 2. We similarly write

$$
\begin{aligned}
\dot{Z}_{it}\dot{Z}_{it}'&=\left(\mathring{Z}_{it}+\left[E\left(\bar{Z}_i\right)-\bar{Z}_i\right]\right)\left(\mathring{Z}_{it}+\left[E\left(\bar{Z}_i\right)-\bar{Z}_i\right]\right)'\\
&=\mathring{Z}_{it}\mathring{Z}_{it}'+\varrho_{i4}-\mathring{Z}_{it}\varrho_{i3}'-\varrho_{i3}\mathring{Z}_{it}'
\end{aligned}
\tag{A.15}
$$

where $\varrho_{i3} = \bar{Z}_i - E\left(\bar{Z}_i\right)$ is a $K \times 1$ vector and $\varrho_{i4} = \varrho_{i3}\varrho_{i3}'$ is a $K \times K$ matrix such that for any $\epsilon > 0$

$$\bigcap_{i=1}^{N} \left\{ \|\varrho_{i3}\|_\infty < \frac{\epsilon}{K} \right\} \text{ and } \bigcap_{i=1}^{N} \left\{ \|\varrho_{i4}\|_\infty < \frac{\epsilon}{K} \right\}$$

holds almost surely after some large T^* due to the results in Claim 3.

Part 3. Put the results in equations (A.14) and (A.15) together, we get

$$\begin{aligned}\dot{Z}_{it}\dot{Z}_{it}'\hat{\varepsilon}_{it}^2 = &\left[\mathring{Z}_{it}\mathring{Z}_{it}' + \varrho_{i4} - \mathring{Z}_{it}\varrho_{i3}' - \varrho_{i3}\mathring{Z}_{it}'\right]\Big[(\check{X}_{it}'\check{\beta}_{it} + \tilde{f}_{it})^2 + \varepsilon_{it}^2 + 2(\check{X}_{it}'\check{\beta}_{it} + \tilde{f}_{it})\varepsilon_{it} \\ &+(\check{X}_{it}'\varrho_{1,NT} + \varrho_{i2})^2 + 2\left(\check{X}_{it}'\varrho_{1,NT} + \varrho_{i2}\right)\varepsilon_{it} + 2(\check{X}_{it}'\check{\beta}_{it} + \tilde{f}_{it})\left(\check{X}_{it}'\varrho_{1,NT} + \varrho_{i2}\right)\Big].\end{aligned}$$

Therefore,

$$\hat{\Omega}_i = T^{-1}\sum_{t=1}^{T}\mathring{Z}_{it}\mathring{Z}_{it}'\left[(\check{X}_{it}'\check{\beta}_{it} + \tilde{f}_{it})^2 + \varepsilon_{it}^2\right] + R_{1i} + R_{2i},$$

where

$$R_{1i} = 2T^{-1}\sum_{t=1}^{T}\mathring{Z}_{it}\mathring{Z}_{it}'(\check{X}_{it}'\check{\beta}_{it} + \tilde{f}_{it})\varepsilon_{it}, \text{ and}$$

$$R_{2i} = \text{Terms involved with } \varrho_{1,NT}, \varrho_{i2}, \varrho_{i3}, \varrho_{i4}.$$

Apparently, $R_i = R_{1i} + R_{2i}$.

Part 4. In this part, we show that

$$\bigcap_{i=1}^{N}\left\{\|R_i\|_\infty \leq \epsilon/K^*\right\},$$

holds almost surely after some large T^*. One sufficient condition to the above is that

$$\bigcap_{i=1}^{N}\left\{\|R_{1i}\|_\infty \leq \epsilon/K^*\right\} \text{ and } \bigcap_{i=1}^{N}\left\{\|R_{2i}\|_\infty \leq \epsilon/K^*\right\}$$

holds almost surely after some large T^*. We first show the results on R_{1i} and then move to R_{2i}.

For R_{1i}, it is a linear combination of finite terms of

$$T^{-1}\sum_{t=1}^{T}\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it} \text{ and } T^{-1}\sum_{t=1}^{T}\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\varepsilon_{it},\ l = 1, 2, ..., d,$$

and the coefficients before those terms are uniformly bounded by our assumptions on β_{it} and f_{it}. Each element in those matrices are mean 0. Thus we can apply the same logic as in the proof of Claim 4, and show that

$$\bigcap_{i=1}^{N}\left\{\|R_{1i}\|_\infty \leq \epsilon/K^*\right\}$$

holds almost surely after some large T^*.

We turn to R_{2i}. The term inside R_{2i} that demands the most strict condition is

$$T^{-1}\sum_{t=1}^{T}\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it}'\varrho_{1,NT}.$$

Note we have shown that $\|\varrho_{1,NT}\| < \epsilon/K^*$ holds almost surely after some large T^*. Thus we only need to show that

$$\bigcap_{i=1}^{N}\left\{\left\|T^{-1}\sum_{t=1}^{T}\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it}'\right\|_\infty < C\right\}$$

holds almost surely after some T^* for some generic C. By the moment conditions we imposed on X, we can similarly apply the same logic as in the proof of Claim 4 and show the above statement. Then

$$\bigcap_{i=1}^{N}\left\{\left\|T^{-1}\sum_{t=1}^{T}\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it}'\varrho_{1,NT}\right\|_{\infty}<C\epsilon/K^*\right\}$$

holds almost surely after some T^*. Further R_{2i} is a linear combination of finite terms similar to $T^{-1}\sum_{t=1}^{T}\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}\check{X}_{it}'\varrho_{1,NT}$ with uniformly bounded coefficients. Consequently,

$$\bigcap_{i=1}^{N}\left\{\|R_{2i}\|_{\infty}\leq\epsilon/K^*\right\}$$

holds almost surely after some large T^*.

Proof of Claim 6. This is a direct result from inequalities in (A.7). Apply the second inequality in (A.7), we have

$$\lambda_{\min}\left(T^{-1}\sum_{t=1}^{T}E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2]\right]\right)\geq\lambda_{\min}\left(T^{-1}\sum_{t=1}^{T}E(\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2)\right).$$

Then

$$\min_{1\leq i\leq N}\lambda_{\min}\left(T^{-1}\sum_{t=1}^{T}E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2]\right]\right)\geq\min_{1\leq i\leq N}\lambda_{\min}\left(T^{-1}\sum_{t=1}^{T}E(\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2)\right)\geq\underline{c}_{zz\varepsilon^2},$$

where the last inequality holds by one of the assumptions we imposed. For the second part in this claim, we apply the first inequality in (A.7) and use the condition that $\check{\beta}_{it}$ and f_{it} are uniformly bounded to get

$$\begin{aligned}&\lambda_{\max}\left(T^{-1}\sum_{t=1}^{T}E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2]\right]\right)\\&\leq C_6\sum_{l=1}^{d}\lambda_{\max}\left(T^{-1}\sum_{t=1}^{T}E[\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}^2]\right)+C_7\lambda_{\max}\left(T^{-1}\sum_{t=1}^{T}E[\mathring{Z}_{it}\mathring{Z}_{it}']\right)\\&+\lambda_{\max}\left(T^{-1}\sum_{t=1}^{T}E[\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2]\right)\end{aligned}$$

for some positive C_6 and C_7. Note that by Assumptions 1-2, $\lambda_{\max}\left(T^{-1}\sum_{t=1}^{T}E[\mathring{Z}_{it}\mathring{Z}_{it}'\check{X}_{it,l}^2]\right)$, $\lambda_{\max}\left(T^{-1}\sum_{t=1}^{T}E[\mathring{Z}_{it}\mathring{Z}_{it}']\right)$, $\lambda_{\max}\left(T^{-1}\sum_{t=1}^{T}E[\mathring{Z}_{it}\mathring{Z}_{it}'\varepsilon_{it}^2]\right)$ are all uniformly bounded. As a result

$$\max_{1\leq i\leq N}\lambda_{\max}\left(T^{-1}\sum_{t=1}^{T}E\left[\mathring{Z}_{it}\mathring{Z}_{it}'[(\check{X}_{it}'\check{\beta}_{it}+\tilde{f}_{it})^2+\varepsilon_{it}^2]\right]\right)\leq C^{\dagger}<\infty,$$

for a positive $C^{\dagger}$. ■

Additional References:

Fan, J., Y. Feng, & R. Song (2011) Nonparametric independence screening in sparse ultra-high-dimensional additive models. *Journal of the American Statistical Association* 106, 544–557.

Huang, J.Z., C.O. Wu, & L. Zhou (2004) Polynomial spline estimation and inference for varying coefficient models with longitudinal data. *Statistica Sinica* 14, 763–788.

D THE SKETCH OF PROOFS FOR MAIN RESULTS IN SECTION 4

In this section, we give some additional assumptions for the tests for stability of heterogeneous coefficients and for *homogeneity* of time-varying coefficients. Since the proofs for Theorems 4.3 and 4.1 are similar to that of Theorem 3.1, we provide the sketch of the proofs.

D.1 Test for the Stability of Heterogeneous Coefficients

To start, we first study the behavior of $\hat{u}_{it}$ under $\mathbb{H}_{s1,\gamma_{NT}}$. By the definition of $\bar{\beta}_{P,i}$, we still have we have $\bar{\beta}_{P,i}=\beta_i$ under $\mathbb{H}_{s1,\gamma_{NT}}$ and

$$\begin{aligned}\hat{\beta}_i-\beta_i &= \gamma_{NT}\left(X_i'M_TX_i\right)^{-1}X_i'M_Tg_{\Delta,i}+\left(X_i'M_TX_i\right)^{-1}X_i'M_T\varepsilon_i\\ &= \gamma_{NT}\bar{\beta}_{\Delta i}+\gamma_{NT}\nu_{\Delta i,T}+\nu_{i,T},\end{aligned}$$

where $\bar{\beta}_{\Delta i}\equiv[E\left(X_i'M_TX_i\right)]^{-1}E\left(X_i'M_Tg_{\Delta,i}\right)$, $\nu_{\Delta i,T}\equiv\hat{\beta}_{\Delta i,T}-\bar{\beta}_{\Delta i}$ with $\hat{\beta}_{\Delta i,T}\equiv\left(X_i'M_TX_i\right)^{-1}$ $\times X_i'M_Tg_{\Delta,i}$ and $\nu_{i,T}\equiv\left(X_i'M_TX_i\right)^{-1}X_i'M_T\varepsilon_i$. Then

$$\hat{u}_{it}=\vec{\varepsilon}_{it}+\alpha_i+\gamma_{NT}\breve{g}_{\Delta,it}-\gamma_{NT}X_{it}'\nu_{\Delta i,T} \text{ and} \tag{A.1}$$

$$\hat{u}_i=\vec{\varepsilon}_i+\alpha_i\iota_T+\gamma_{NT}\breve{g}_{\Delta,i}-\gamma_{NT}X_i\nu_{\Delta i,T} \tag{A.2}$$

where $\vec{\varepsilon}_{it}\equiv\varepsilon_{it}-X_{it}'\nu_{i,T}$, $\vec{\varepsilon}_i\equiv\left(\vec{\varepsilon}_{i1},\ldots,\vec{\varepsilon}_{iT}\right)'=\varepsilon_i-X_i\left(X_i'M_TX_i\right)^{-1}X_i'M_T\varepsilon_i$, $\breve{g}_{\Delta,it}\equiv g_{\Delta,it}-X_{it}'\bar{\beta}_{\Delta i}$, and $\breve{g}_{\Delta,i}\equiv\left(\breve{g}_{\Delta,i1},\ldots,\breve{g}_{\Delta,iT}\right)'$.

Now we give the sketch of the proof of Theorem 4.1.

The Sketch of proof for Theorem 4.1. We only give the sketch proof for (ii) because (i) can be seen as a special case of (ii) with $\gamma_{NT}=0$. Using (A.2), we can decompose Γ_{NT} as follows

$$\Gamma_{NT}=\frac{1}{NT^2}\sum_{i=1}^{N}\left(\vec{\varepsilon}_i+\gamma_{NT}\breve{g}_{\Delta,i}-\gamma_{NT}X_i\nu_{\Delta i,T}\right)'\mathcal{K}_i\left(\vec{\varepsilon}_i+\gamma_{NT}\breve{g}_{\Delta,i}-\gamma_{NT}X_i\nu_{\Delta i,T}\right)\equiv\sum_{s=1}^{6}\Gamma_{NT}^{(s)},\text{ say}$$

where $\Gamma_{NT}^{(1)}=\frac{1}{NT^2}\sum_{i=1}^{N}\vec{\varepsilon}_i'\mathcal{K}_i\vec{\varepsilon}_i$, $\Gamma_{NT}^{(2)}=\frac{\gamma_{NT}^2}{NT^2}\sum_{i=1}^{N}\breve{g}_{\Delta,i}'\mathcal{K}_i\breve{g}_{\Delta,i}$, $\Gamma_{NT}^{(3)}=\frac{\gamma_{NT}^2}{NT^2}\sum_{i=1}^{N}\nu_{\Delta i,T}'X_i'\mathcal{K}_iX_i\nu_{\Delta i,T}$, $\Gamma_{NT}^{(4)}=\frac{2\gamma_{NT}}{NT^2}\sum_{i=1}^{N}\vec{\varepsilon}_i'\mathcal{K}_i\breve{g}_{\Delta,i}$, $\Gamma_{NT}^{(5)}=\frac{-2\gamma_{NT}}{NT^2}\sum_{i=1}^{N}\vec{\varepsilon}_i'\mathcal{K}_iX_i\nu_{\Delta i,T}$, and $\Gamma_{NT}^{(6)}=\frac{-2\gamma_{NT}^2}{NT^2}\sum_{i=1}^{N}\breve{g}_{\Delta,i}'\mathcal{K}_iX_i\nu_{\Delta i,T}$. With the decomposition, we have

$$\hat{J}_{NT}=\frac{N^{1/2}T\Gamma_{NT}-\hat{\mathbb{B}}_{NT}}{\hat{\mathbb{V}}_{NT}^{1/2}}=\left(J_{NT}+\sum_{s=2}^{6}\frac{N^{1/2}T\Gamma_{NT}^{(s)}}{\mathbb{V}_{NT}^{1/2}}+\frac{\mathbb{B}_{NT}-\hat{\mathbb{B}}_{NT}}{\mathbb{V}_{NT}^{1/2}}\right)\frac{\mathbb{V}_{NT}^{1/2}}{\hat{\mathbb{V}}_{NT}^{1/2}}$$

where $J_{NT}=(N^{1/2}T\Gamma_{NT}^{(1)}-\mathbb{B}_{NT})/\mathbb{V}_{NT}^{1/2}$. We can complete the proof by showing that (i) $J_{NT}\xrightarrow{d}N\left(0,1\right)$; (ii) $N^{1/2}T\Gamma_{NT}^{(2)}/\mathbb{V}_{NT}^{1/2}=\Phi_\Delta+o_p\left(1\right)$, where $\Phi_\Delta=\text{plim}_{(N,T)\to\infty}\Phi_{\Delta,NT}$ with

$\Phi_{\Delta,NT} = N^{-1}T^{-2}\sum_{i=1}^{N}\breve{g}_{\Delta,it}^{2}$; (iii) $N^{1/2}T\Gamma_{NT}^{(s)}/\mathbb{V}_{NT}^{1/2} = o_p(1)$ for $s = 3,\ldots,6$; (iv) $\hat{\mathbb{B}}_{NT} - \mathbb{B}_{NT} = o_p(K^{1/2})$; (v) $\hat{\mathbb{V}}_{NT}/\mathbb{V}_{NT} = 1 + o_p(1)$.

First, it is straightforward to show (i)-(ii), (iv)-(v) by modifying the corresponding proofs for Theorem 3.1. For (iii), following the proof of (iii) in Theorem 3.1, we can show that $\Gamma_{NT}^{(3)} = o_p\left(\gamma_{NT}^2\right)$, $\Gamma_{NT}^{(4)} = \gamma_{NT}O_p([K/(NT)]^{1/2}) = o_p\left(\gamma_{NT}^2\right)$, $\Gamma_{NT}^{(5)} = \gamma_{NT}O_p([K/(NT)]^{1/2})o_p(1) = o_p\left(\gamma_{NT}^2\right)$, and $\Gamma_{NT}^{(6)} = o_p\left(\gamma_{NT}^2\right)$. ■

Proof of Corollary 4.2. We can follow the proof of Theorem 3.2 to show the corollary. The details are omitted here. ■

D.2 Test the Homogeneity of Time-Varying Coefficients

We first study the behavior of $\hat{u}_{it}$ and $\hat{g}_{it}$ under the local alternative. There exist $\Pi_\beta^0 \in \mathbb{R}^{d\times L}$ and $\Pi_f^0 \in \mathbb{R}^{L-1}$ such that $\beta_0(\cdot) \approx \Pi_\beta^0 B^L(\cdot)$ and $f_0(\cdot) \approx \Pi_f^{0\prime} B_{-1}^L(\cdot)$. Let $g_{it} \equiv g_{0,it} + \gamma_{NT}g_{\Delta,it}$, where $g_{0,it} \equiv X_{it}'\beta_0(\tau_t) + f_0(\tau_t)$. Given $Z_{it}^L \equiv (B_{-1t}^{L\prime}, \left(X_{it}\otimes B_t^L\right)')'$, denote $r_{g_0,it} \equiv g_{0,it} - Z_{it}^{L\prime}\Pi^0$, where $\Pi^0 \equiv (\Pi_f^{0\prime}, \text{vec}(\Pi_\beta^{0\prime}))'$. Let $S_{\dot{Z}\dot{Z}} \equiv \sum_{i=1}^{N}\dot{Z}_i^{L\prime}\dot{Z}_i^L$, $\hat{\Pi}_{\Delta,NT} \equiv S_{\dot{Z}\dot{Z}}^{-1}\sum_{i=1}^{N}\dot{Z}_i^{L\prime}g_{\Delta,i}$, $\bar{\Pi}_\Delta \equiv \left[E\left(S_{\dot{Z}\dot{Z}}\right)\right]^{-1}\sum_{i=1}^{N}E\left(\dot{Z}_i^{L\prime}g_{\Delta,i}\right)$, $R_{g_0,i} \equiv (r_{g_0,i1},\ldots,r_{g_0,iT})'$ and $g_{\Delta,i} \equiv (g_{\Delta,i1},\ldots,g_{\Delta,iT})'$. Then we have

$$
\begin{aligned}
\hat{\Pi}_{FE} - \Pi^0 &= S_{\dot{Z}\dot{Z}}^{-1}\sum_{i=1}^{N}\dot{Z}_i^{L\prime}R_{g_0,i} + \gamma_{NT}\bar{\Pi}_\Delta + \gamma_{NT}\left[\hat{\Pi}_{\Delta,NT} - \bar{\Pi}_\Delta\right] + S_{\dot{Z}\dot{Z}}^{-1}\sum_{i=1}^{N}\dot{Z}_i^{L\prime}\varepsilon_i \\
&\equiv R_{g_0,NT} + \gamma_{NT}\bar{\Pi}_\Delta + \gamma_{NT}\nu_{\Pi_\Delta,NT} + \nu_{L,NT},
\end{aligned}
$$

where $R_{g_0,NT} \sim S_{\dot{Z}\dot{Z}}^{-1}\sum_{i=1}^{N}\dot{Z}_i^{L\prime}R_{g_0,i}$, $\nu_{\Pi_\Delta,NT} \equiv \hat{\Pi}_{\Delta,NT} - \bar{\Pi}_\Delta$ and $\nu_{L,NT} \equiv S_{\dot{Z}\dot{Z}}^{-1}\sum_{i=1}^{N}\dot{Z}_i^{L\prime}\varepsilon_i$. Let $\breve{g}_{\Delta,it} \equiv g_{\Delta,it} - Z_{it}^{L\prime}\bar{\Pi}_\Delta$ and $\breve{\nu}_{L,NT} \equiv \gamma_{NT}\nu_{\Pi_\Delta,NT} + \nu_{L,NT}$. We can write

$$
\begin{aligned}
g_{it} - \hat{g}_{it} &= (g_{0,it} + \gamma_{NT}g_{\Delta,it}) - Z_{it}^{L\prime}\left(\Pi^0 + R_{g_0,NT} + \gamma_{NT}\bar{\Pi}_\Delta + \breve{\nu}_{L,NT}\right) \\
&= \left(g_{0,it} - Z_{it}^{L\prime}\Pi^0\right) + \gamma_{NT}\left(g_{\Delta,it} - Z_{it}^{L\prime}\bar{\Pi}_\Delta\right) - Z_{it}^{L\prime}R_{g_0,NT} - Z_{it}^{L\prime}\breve{\nu}_{L,NT} \\
&= \left(r_{g_0,it} - Z_{it}^{L\prime}R_{g_0,NT}\right) + \gamma_{NT}\breve{g}_{\Delta,it} - Z_{it}^{L\prime}\breve{\nu}_{L,NT} \\
&= \breve{r}_{g_0,it} + \gamma_{NT}\breve{g}_{\Delta,it} - Z_{it}^{L\prime}\breve{\nu}_{L,NT}
\end{aligned}
$$

where $\breve{r}_{g_0,it} \equiv r_{g_0,it} - Z_{it}^{L\prime}R_{g_0,NT}$. Let $\breve{R}_{g_0,i} = (\breve{r}_{g_0,i1},\ldots,\breve{r}_{g_0,iT})'$ and $\breve{g}_{\Delta,i} = (\breve{g}_{\Delta,i1},\ldots,\breve{g}_{\Delta,iT})'$. Then we have

$$\hat{u}_{it} = \varepsilon_{it} + \alpha_i + \gamma_{NT}\breve{g}_{\Delta,it} + \breve{r}_{g_0,it} - Z_{it}^{L\prime}\breve{\nu}_{L,NT} \text{ and} \tag{A.3}$$

$$\hat{u}_i = \varepsilon_i + \alpha_i\iota_T + \gamma_{NT}\breve{g}_{\Delta,i} + \breve{R}_{g_0,i} - Z_i^L\breve{\nu}_{L,NT}. \tag{A.4}$$

To establish the asymptotic distribution of $\hat{J}_{NT}$, we need the following assumptions.

Assumption 3*. (i) $f(\cdot)$ and $\beta_{0,l}(\cdot)$ for $l = 1,\ldots,d$ are all continuously differentiable up to κ-th order for some $\kappa > 0$; (ii) For each i, $\Delta_{\beta,il}(\cdot)$ for $l = 1,\ldots,d$, and $\Delta_{f,i}(\cdot)$ are all continuously differentiable up to κ-th order for some $\kappa > 0$.

Assumption 4.** As $(N,T) \to \infty$, $\Phi_\Delta = \text{plim}_{(N,T)\to\infty} \Phi_{\Delta,NT} > 0$ under $\mathbb{H}_{h1,\gamma_{NT}}$.
Assumption 5. As $(N,T) \to \infty$, $L \to \infty$, $L^2/T \to 0$, and $K/L \to 0$.

Now we give the sketch for the proof of Theorem 4.3.

Sketch of Proof for Theorem 4.3. We only give the sketch proof for (ii) because (i) can be seen as a special case of (ii) with $\gamma_{NT} = 0$. Using (A.4) and $\Gamma_{NT} = \frac{1}{NT^2}\sum_{i=1}^N \hat{u}_i' \mathcal{K}_i \hat{u}_i$, we have $\Gamma_{NT} \equiv \sum_{s=1}^{10} \Gamma_{NT}^{(s)}$, where $\Gamma_{NT}^{(1)} \equiv \frac{1}{NT^2}\sum_{i=1}^N \varepsilon_i' \mathcal{K}_i \varepsilon_i$, $\Gamma_{NT}^{(2)} \equiv \frac{\gamma_{NT}^2}{NT^2}\sum_{i=1}^N \breve{g}_{\Delta,i}' \mathcal{K}_i \breve{g}_{\Delta,i}$, $\Gamma_{NT}^{(3)} \equiv \frac{1}{NT^2}\sum_{i=1}^N \breve{R}_{g0,i}' \mathcal{K}_i \breve{R}_{g0,i}$, $\Gamma_{NT}^{(4)} \equiv \frac{1}{NT^2}\sum_{i=1}^N \breve{\nu}_{L,NT}' Z_i^{L\prime} \mathcal{K}_i Z_i^L \breve{\nu}_{L,NT}$, $\Gamma_{NT}^{(5)} \equiv \frac{2\gamma_{NT}}{NT^2}\sum_{i=1}^N \varepsilon_i' \mathcal{K}_i \breve{g}_{\Delta,i}$, $\Gamma_{NT}^{(6)} \equiv \frac{2}{NT^2}\sum_{i=1}^N \varepsilon_i' \mathcal{K}_i \breve{R}_{g0,i}$, $\Gamma_{NT}^{(7)} \equiv \frac{-2}{NT^2}\sum_{i=1}^N \varepsilon_i' \mathcal{K}_i Z_i^L \breve{\nu}_{L,NT}$, $\Gamma_{NT}^{(8)} \equiv \frac{2\gamma_{NT}}{NT^2}\sum_{i=1}^N \breve{g}_{\Delta,i}' \mathcal{K}_i \breve{R}_{g0,i}$, $\Gamma_{NT}^{(9)} \equiv \frac{-2\gamma_{NT}}{NT^2}\sum_{i=1}^N \breve{g}_{\Delta,i}' \mathcal{K}_i Z_i^L \breve{\nu}_{L,NT}$, and $\Gamma_{NT}^{(10)} \equiv \frac{-2}{NT^2}\sum_{i=1}^N \breve{R}_{g0,i}' \mathcal{K}_i Z_i^L \breve{\nu}_{L,NT}$. Then $\hat{J}_{NT}$ can be decomposed as follows

$$\hat{J}_{NT} = \frac{N^{1/2}T\Gamma_{NT} - \hat{\mathbb{B}}_{NT}}{\hat{\mathbb{V}}_{NT}^{1/2}} = \left(J_{NT} + \sum_{s=2}^{10} \frac{N^{1/2}T\Gamma_{NT}^{(s)}}{\mathbb{V}_{NT}^{1/2}} + \frac{\mathbb{B}_{NT} - \hat{\mathbb{B}}_{NT}}{\mathbb{V}_{NT}^{1/2}} \right) \frac{\mathbb{V}_{NT}^{1/2}}{\hat{\mathbb{V}}_{NT}^{1/2}}.$$

We complete the proof by showing that: (i) $J_{NT} = (N^{1/2}T\Gamma_{NT}^{(1)} - \mathbb{B}_{NT})/\mathbb{V}_{NT}^{1/2} \xrightarrow{d} N(0,1)$; (ii) $J_{NT}^{(2)} \equiv N^{1/2}T\Gamma_{NT}^{(2)}/\mathbb{V}_{NT}^{1/2} = \Phi_\Delta + o_p(1)$, where $\Phi_\Delta = \text{plim}_{(N,T)\to\infty} \Phi_{\Delta,NT}$ with $\Phi_{\Delta,NT} = N^{-1}T^{-1}\sum_{i=1}^N \sum_{t=1}^T \breve{g}_{\Delta,it}^2$; (iii) $J_{NT}^{(s)} \equiv N^{1/2}T\Gamma_{NT}^{(s)}/\mathbb{V}_{NT}^{1/2} = o_p(1)$ for $s = 3, \ldots, 10$; (iv) $\hat{\mathbb{B}}_{NT} - \mathbb{B}_{NT} = o_p(K^{1/2})$; (v) $\hat{\mathbb{V}}_{NT}/\mathbb{V}_{NT} = 1 + o_p(1)$.

First, by modifying the proof of Theorem 3.1, we can easily show that (i), (ii), (iv) and (v). Second, we can follow the proofs of (iii) for Theorem 3.1 to show that $\Gamma_{NT}^{(3)} = O_p\left(L^{-2\kappa}\right) = o_p(\gamma_{NT}^2)$, $\Gamma_{NT}^{(4)} = o_p\left(\gamma_{NT}^2\right) + O_p\left(L/(NT)\right) = o_p(\gamma_{NT}^2)$, $\Gamma_{NT}^{(5)} = O_p(\gamma_{NT}[K/(NT)]^{1/2}) = o_p(\gamma_{NT}^2)$, $\Gamma_{NT}^{(6)} = O_p(L^{-\kappa}[K/(NT)]^{1/2}) = o_p(\gamma_{NT}^2)$, $\Gamma_{NT}^{(7)} = O_p([K/(NT)]^{1/2})\left[o_p(\gamma_{NT}) + O_p([L/(NT)]^{1/2})\right] = o_p(\gamma_{NT}^2)$, $\Gamma_{NT}^{(8)} = o_p(\gamma_{NT} L^{-\kappa}) = o_p(\gamma_{NT}^2)$, $\Gamma_{NT}^{(9)} = o_p\left(\gamma_{NT}^2\right) + O_p(\gamma_{NT}[L/(NT)]^{1/2}) = o_p(\gamma_{NT}^2)$, $\Gamma_{NT}^{(10)} = O_p\left(L^{-\kappa}\right)\left[o_p(\gamma_{NT}) + O_p([L/(NT)]^{1/2}\right) = o_p(\gamma_{NT}^2)$. ■

Proof for Corollary 4.4. We can follow the proof of Theorem 3.2 to show the corollary. The details are omitted here. ■

E A COMPARISON BETWEEN ONE-STAGE AND TWO-STAGE PROCEDURES

We compare the sizes and powers of the two procedures in this section. We first show below the relationship between the one-stage test statistic and our two-stage test statistic. After that, we show that they are asymptotically equivalent under $\mathbb{H}_0$. In the end, we show that they can detect the deviation from $\mathbb{H}_0$ in a very similar way, or in other words, they have similar power under $\mathbb{H}_1$. Due to the page limit, we only outline the proof and leave out the tedious technical details. For an easier illustration, we suppose X is a scalar.

Step 1: Relationship. The model can be written as:

$$\begin{aligned} Y_{it} &= X_{it}\beta_i(\tau_t) + f_i(\tau_t) + \alpha_i + \varepsilon_{it} \\ &\approx X_{it}\sum_{j=0}^{K}\tilde{\vartheta}_{\beta,i,j}B_j(\tau_t) + \sum_{j=1}^{K}\tilde{\vartheta}_{f,i,j}B_j(\tau_t) + \alpha_i + \varepsilon_{it} \\ &\equiv Z_{it}'\tilde{\vartheta}_i + \alpha_i + \varepsilon_{it} \\ &= X_{it}\tilde{\vartheta}_{\beta,i,0} + Z_{-1,it}'\tilde{\vartheta}_{-1,i} + \alpha_i + \varepsilon_{it}, \end{aligned} \tag{A.1}$$

where we note $B_0(\tau_t) = 1$ and $X_{it}\tilde{\vartheta}_{\beta,i,0}B_0(\tau_t) = X_{it}\tilde{\vartheta}_{\beta,i,0}$,

$$\begin{aligned} Z_{it} &= (X_{it}, X_{it}B_1(\tau_t)..., X_{it}B_K(\tau_t), B_1(\tau_t), ..., B_K(\tau_t))', \\ \tilde{\vartheta}_i &= \left(\tilde{\vartheta}_{\beta,i,0}, \tilde{\vartheta}_{\beta,i,1}, ..., \tilde{\vartheta}_{\beta,i,K}, \tilde{\vartheta}_{f,i,1}, ..., \tilde{\vartheta}_{f,i,K}\right)' \end{aligned}$$

and $Z_{-1,it}$ denote the Z_{it} without the first element.

The test statistic formed using only one-stage is:

$$\tilde{\Gamma}_{NT} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[X_{it}\left(\widehat{\tilde{\vartheta}}_{\beta,i,0} - \frac{1}{N}\sum_{i=1}^{N}\widehat{\tilde{\vartheta}}_{\beta,i,0}\right) + Z_{-1,it}'\widehat{\tilde{\vartheta}}_{-1,i}\right]^2, \tag{A.2}$$

where

$$\widehat{\tilde{\vartheta}}_i = \left(Z_i'M_{\iota_T}Z_i\right)^{-1}\left(Z_i'M_{\iota_T}Y_i\right), \tag{A.3}$$

$Z_i = (Z_{i1}, Z_{i2}, ..., Z_{iT})'$, $\widehat{\tilde{\vartheta}}_{\beta,i,0}$ is the first element of $\widehat{\tilde{\vartheta}}_i$, and $\widehat{\tilde{\vartheta}}_{-1,i}$ is $\widehat{\tilde{\vartheta}}_i$ without its first element. Write $Y_{it} = Z_{it}'\widehat{\tilde{\vartheta}}_i + \widehat{\tilde{u}}_{it}$, where $\widehat{\tilde{u}}_{it}$ is the generalized residual.

We turn to the two-stage statistics. $\hat{u}_{it}$ is obtained from

$$\begin{aligned} \hat{u}_{it} &= Y_{it} - X_{it}\hat{\beta}_{FE} \\ &\approx X_{it}\left(\widehat{\tilde{\vartheta}}_{\beta,i,0} - \hat{\beta}_{FE}\right) + Z_{-1,it}'\hat{\tilde{\vartheta}}_{-1,i} + \widehat{\tilde{u}}_{it} \end{aligned} \tag{A.4}$$

using the last line of equation (A.1). In the matrix form, we have

$$\begin{aligned}\hat{u}_i &\approx X_i\left(\tilde{\vartheta}_{\beta,i,0}-\hat{\beta}_{FE}\right)+Z_{-1,i}\tilde{\vartheta}_{-1,i}+\widehat{\tilde{u}}_i\\ &\approx Z_i\left((\tilde{\vartheta}_{\beta,i,0}-\hat{\beta}_{FE})',\tilde{\vartheta}'_{-1,i}\right)'+\widehat{\tilde{u}}_i.\end{aligned}$$

Note the estimator of ϑ_i is

$$\hat{\vartheta}_i=\left(Z_i'M_{\iota_T}Z_i\right)^{-1}\left(Z_i'M_{\iota_T}\hat{u}_i\right)=\left((\tilde{\vartheta}_{\beta,i,0}-\hat{\beta}_{FE})',\tilde{\vartheta}'_{-1,i}\right)'+\left(Z_i'M_{\iota_T}Z_i\right)^{-1}Z_i'M_{\iota_T}\widehat{\tilde{u}}_i. \tag{A.5}$$

We can show the above second term is $o_p(1)$ and the following relationships hold asymptotically:

$$\widehat{\tilde{\vartheta}}_{\beta,i,0}-\hat{\beta}_{FE}=\hat{\vartheta}_{\beta,i,0}+o_p(1)\ \text{ and }\ \widehat{\tilde{\vartheta}}_{-1,i}=\hat{\vartheta}_{-1,i}+o_p(1)\,.$$

Using the above identities, the two-stage test statistic can be written as

$$\begin{aligned}\Gamma_{NT}&=\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(Z_{it}'\hat{\vartheta}_i\right)^2\\ &=\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[X_{it}\hat{\vartheta}_{\beta,i,0}+Z'_{-1,it}\hat{\vartheta}_{-1,i}\right]^2\\ &=\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[X_{it}\left(\widehat{\tilde{\vartheta}}_{\beta,i,0}-\frac{1}{N}\sum_{i=1}^{N}\widehat{\tilde{\vartheta}}_{\beta,i,0}\right)+Z'_{-1,it}\widehat{\tilde{\vartheta}}_{-1,i}+X_{it}'\bar{D}_{NT}\right]^2,\end{aligned} \tag{A.6}$$

where $\bar{D}_{NT}=\frac{1}{N}\sum_{i=1}^{N}\widehat{\tilde{\vartheta}}_{\beta,i,0}-\hat{\beta}_{FE}$. Comparing equations (A.2) and (A.6), we can see the additional term $X_{it}'\bar{D}_{NT}$, which comes from different centers are used.

We compare $\tilde{\Gamma}_{NT}$ and Γ_{NT} based on the results in equations (A.2) and (A.6).

Step 2: Comparison under $\mathbb{H}_0$. Under $\mathbb{H}_0$, it is not hard to see that $\frac{1}{N}\sum_{i=1}^{N}\widehat{\tilde{\vartheta}}_{\beta,i,0}$ and $\hat{\beta}_{FE}$ converge to the same true β in probability and at the same rate ($O_p[(NT)^{-1/2}]$). Then the effect of $X_{it}'\bar{D}_{NT}$ is asymptotically negligible. Therefore, $\tilde{\Gamma}_{NT}$ converges to zero at the same rate as Γ_{NT} does. Under some proper normalization (with some work), $\tilde{\Gamma}_{NT}$ will converge to the same standard normal distribution. In view of this observation, we conclude that these two procedures are asymptotic equivalent.

Step 3: Power comparison under $\mathbb{H}_1$. We discuss the powers under local alternatives in three cases. In the first case, we compare the two procedures when only f deviates from the null. In the second case, only β deviates from the null. In the last case, both f and β differ from the null.

Case 1: $\mathbb{H}_1$ with deviation in f only. Specifically, we suppose $\beta_{it}=\beta_0$ and $f_{it}=\gamma_{NT}\Delta_{f,it}$; $\Delta_{f,it}$ was defined in Section 3.3. γ_{NT} measures the magnitude of the deviation, and we assume γ_{NT} is much bigger than the estimation error so that the deviation is significant enough for the sample size.

- For the one-step procedure, $\widehat{\tilde{\vartheta}}_{\beta,i,0}\xrightarrow{P}\beta_0$. Thus $\frac{1}{N}\sum_{i=1}^{N}\widehat{\tilde{\vartheta}}_{\beta,i,0}\xrightarrow{P}\beta_0$.

- For the two-step procedure, it is not hard to see that $\hat{\beta}_{FE} = \beta_0 + O_P(\gamma_{NT})$ because $f_{it} = O_P(\gamma_{NT})$.

- Comparing the above two, we can see that $\frac{1}{N}\sum_{i=1}^{N}\widehat{\tilde{\vartheta}}_{\beta,i,0} - \hat{\beta}_{FE} = O_P(\gamma_{NT})$. At the same time, $Z'_{-1,it}\widehat{\tilde{\vartheta}}_{-1,i} \approx f_{it}$, and thus $E\left[\left(Z'_{-1,it}\widehat{\tilde{\vartheta}}_{-1,i}\right)^2\right] \propto \gamma_{NT}^2$. Collecting all the results, we can see that $\tilde{\Gamma}_{NT} \propto \gamma_{NT}^2$ and $\Gamma_{NT} \propto \gamma_{NT}^2$ in probability, which means that these two procedures have the same power (roughly) .

Case 2: $\mathbb{H}_1$ with deviation in β only. Specifically, we suppose $\beta_{it} = \beta_0 + \gamma_{NT}\Delta_{\beta,it}$, $f_{it} = 0$, and $\Delta_{\beta,it}$ was defined in Section 3.3.

- For the one-step procedure, it is not hard to see that $\widehat{\tilde{\vartheta}}_{\beta,i,0} = \beta_0 + O_P(\gamma_{NT})$ because $\beta_{it} - \beta_0 = O(\gamma_{NT})$.

- For the two-step procedure, $\hat{\beta}_{FE} = \beta_0 + O_P(\gamma_{NT})$ for the same reason as above.

- Thus, $\frac{1}{N}\sum_{i=1}^{N}\widehat{\tilde{\vartheta}}_{\beta,i,0} - \hat{\beta}_{FE} = O_P(\gamma_{NT})$.

Continuing from the last line for $\tilde{\Gamma}_{NT}$ in (A.2),

$$\begin{aligned}
\tilde{\Gamma}_{NT} &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[X_{it}\left(\widehat{\tilde{\vartheta}}_{\beta,i,0} - \frac{1}{N}\sum_{i=1}^{N}\widehat{\tilde{\vartheta}}_{\beta,i,0}\right) + Z'_{-1,it}\widehat{\tilde{\vartheta}}_{-1,i}\right]^2 \\
&= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[X_{it}\left(\widehat{\tilde{\vartheta}}_{\beta,i,0} - \beta_0\right) + Z'_{-1,it}\widehat{\tilde{\vartheta}}_{-1,i} + X_{it}\left(\beta_0 - \frac{1}{N}\sum_{i=1}^{N}\widehat{\tilde{\vartheta}}_{\beta,i,0}\right)\right]^2 \\
&\propto_P \gamma_{NT}^2,
\end{aligned}$$

due to $X_{it}\left(\widehat{\tilde{\vartheta}}_{\beta,i,0} - \beta_0\right) + Z'_{-1,it}\widehat{\tilde{\vartheta}}_{-1,i} \approx \beta_{it} - \beta_0 = \gamma_{NT}\Delta_{\beta,it}$, and $\beta_0 - \frac{1}{N}\sum_{i=1}^{N}\widehat{\tilde{\vartheta}}_{\beta,i,0} = O_P(\gamma_{NT})$, with some regular conditions.

Similarly, continuing from the last line for Γ_{NT} in (A.6),

$$\begin{aligned}
\Gamma_{NT} &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[X_{it}\left(\widehat{\tilde{\vartheta}}_{\beta,i,0} - \hat{\beta}_{FE}\right) + Z'_{-1,it}\widehat{\tilde{\vartheta}}_{-1,i}\right]^2 \\
&= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[X_{it}\left(\widehat{\tilde{\vartheta}}_{\beta,i,0} - \beta_0\right) + Z'_{-1,it}\widehat{\tilde{\vartheta}}_{-1,i} + X_{it}\left(\beta_0 - \hat{\beta}_{FE}\right)\right]^2 \\
&\propto_P \gamma_{NT}^2,
\end{aligned}$$

due to $X_{it}\left(\widehat{\tilde{\vartheta}}_{\beta,i,0} - \beta_0\right) + Z'_{-1,it}\widehat{\tilde{\vartheta}}_{-1,i} \approx \gamma_{NT}\Delta_{\beta,it}$, and $\beta_0 - \hat{\beta}_{FE} = O_P(\gamma_{NT})$, with some regularity conditions.

Therefore, both procedures have the same local power for Case 2.

Case 3: $\mathbb{H}_1$ with deviations in β and f. Specifically, we suppose $\beta_{it} = \beta_0 + \gamma_{NT,1}\Delta_{\beta,it}$ and $f_{it} = \gamma_{NT,2}\Delta_{f,it}$. Using the results from Cases 1 and 2, it is not hard to see that

$$\widehat{\tilde{\vartheta}}_{\beta,i,0} = \beta_0 + O_P\left(\max\left\{\gamma_{NT,1}, \gamma_{NT,2}\right\}\right), \ \hat{\beta}_{FE} = \beta_0 + O_P\left(\max\left\{\gamma_{NT,1}, \gamma_{NT,2}\right\}\right).$$

Therefore,

$$\widehat{\tilde{\vartheta}}_{\beta,i,0} - \hat{\beta}_{FE} = O_P\left(\max\left\{\gamma_{NT,1}, \gamma_{NT,2}\right\}\right).$$

Using it, we apply similar analysis in Cases 1 and 2 here, and we can obtain

$$\tilde{\Gamma}_{NT}, \Gamma_{NT} \propto_P \left(\max\left\{\gamma_{NT,1}^2, \gamma_{NT,2}^2\right\}\right),$$

which implies that both procedures have the same local power.

To summarize, both procedures have the same power in general.

As for the global power, note that

$$\bar{D}_{NT} = \frac{1}{N}\sum_{i=1}^{N}\widehat{\tilde{\vartheta}}_{\beta,i,0} - \hat{\beta}_{FE} \to_p \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\tilde{\vartheta}_{\beta,i,0} - \beta_P = c \neq 0.$$

It is unclear whether the additional term $X_{it}\bar{D}_{NT} \approx X_{it}c$ in Γ_{NT} makes Γ_{NT} greater or smaller than $\tilde{\Gamma}_{NT}$. Neither is dominant in global power performance. It should depend on the data generating process, that is, the heterogeneity in β_{it} and f_{it} and the regressor X_{it}. But they (normalized version) tend to ∞ at the same rate $\left(O_P\left(TN^{1/2}/K\right)\right)$ as sample size increases to ∞.

Conclusion: In terms of sizes and powers, both procedures are asymptotically equivalent. We note that compared with the two-step procedure, the one-step procedure is more intuitive and measures the deviation more precisely when $\beta_{it} = \beta_0$ but f_{it} being different across i and t. Due to the page limit, we leave the full investigation of the one-step procedure to the future.

F ADDITIONAL SIMULATION AND APPLICATION RESULTS

In this section, we present the size results using normal critical values for $\mathbb{H}_0$ and the testing results for the three tests discussed in Section 4. We show the sieve estimates for the application under $\mathbb{H}_{h0}$ in Figure 1.

Table 4: Size performance for testing $\mathbb{H}_0$ using normal critical values

DGP	T	N	K_1 10%	K_1 5%	K_1 1%	K_2 10%	K_2 5%	K_2 1%	K_3 10%	K_3 5%	K_3 1%	K_{cv} 10%	K_{cv} 5%	K_{cv} 1%
1	25	25	0.086	0.052	0.024	0.095	0.062	0.026	0.095	0.071	0.028	0.091	0.053	0.016
		50	0.111	0.072	0.033	0.130	0.085	0.034	0.096	0.057	0.027	0.116	0.077	0.027
	50	25	0.071	0.049	0.013	0.084	0.047	0.018	0.084	0.057	0.025	0.067	0.041	0.017
		50	0.078	0.041	0.011	0.082	0.047	0.012	0.101	0.067	0.027	0.070	0.048	0.011
	100	25	0.055	0.038	0.011	0.058	0.035	0.012	0.074	0.042	0.013	0.062	0.038	0.010
		50	0.060	0.030	0.009	0.060	0.042	0.014	0.080	0.048	0.020	0.057	0.036	0.014

Table 5: Power sensitivity studies: DGP 2-6 using different K

DGP	T	N	K_1 10%	K_1 5%	K_1 1%	K_2 10%	K_2 5%	K_2 1%	K_3 10%	K_3 5%	K_3 1%
2	25	25	0.369	0.237	0.080	0.246	0.145	0.046	0.182	0.101	0.028
		50	0.525	0.374	0.149	0.343	0.220	0.063	0.206	0.122	0.028
	50	25	0.923	0.813	0.520	0.855	0.727	0.388	0.794	0.615	0.293
		50	0.992	0.973	0.854	0.979	0.932	0.746	0.944	0.866	0.571
3	25	25	0.594	0.483	0.252	0.484	0.326	0.147	0.401	0.285	0.125
		50	0.754	0.626	0.390	0.655	0.509	0.269	0.521	0.387	0.162
	50	25	0.962	0.910	0.762	0.934	0.887	0.704	0.914	0.845	0.609
		50	0.995	0.991	0.954	0.995	0.977	0.908	0.980	0.959	0.857
4	25	25	0.559	0.423	0.203	0.437	0.318	0.145	0.341	0.215	0.093
		50	0.690	0.569	0.336	0.583	0.457	0.244	0.454	0.296	0.137
	50	25	0.944	0.901	0.709	0.929	0.847	0.611	0.888	0.771	0.521
		50	0.999	0.987	0.938	0.998	0.981	0.881	0.977	0.943	0.799
5	25	25	0.516	0.360	0.142	0.362	0.234	0.088	0.256	0.148	0.044
		50	0.664	0.500	0.249	0.492	0.330	0.127	0.333	0.203	0.077
	50	25	0.944	0.884	0.653	0.914	0.812	0.552	0.867	0.731	0.478
		50	0.999	0.984	0.931	0.989	0.966	0.858	0.979	0.934	0.764
6	25	25	0.471	0.332	0.125	0.364	0.217	0.069	0.212	0.131	0.045
		50	0.662	0.497	0.204	0.492	0.311	0.090	0.261	0.138	0.042
	50	25	0.988	0.964	0.831	0.980	0.949	0.755	0.965	0.893	0.632
		50	1.000	0.999	0.991	1.000	0.997	0.961	0.998	0.989	0.904

Note: $K_1 = \lceil T^{1/5} \rceil$, $K_2 = \lceil 1.5T^{1/5} \rceil$, $K_3 = \lceil 2T^{1/5} \rceil$

Table 6: Simulation results for testing $\mathbb{H}_{s0}$

			K_1			K_2			K_3			K_{cv}		
DGP	T	N	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
1	25	25	0.087	0.044	0.007	0.076	0.038	0.006	0.092	0.044	0.013	0.073	0.032	0.007
		50	0.074	0.039	0.010	0.091	0.043	0.009	0.106	0.051	0.014	0.063	0.028	0.002
	50	25	0.078	0.034	0.007	0.090	0.037	0.005	0.103	0.047	0.011	0.073	0.032	0.006
		50	0.085	0.040	0.010	0.087	0.036	0.004	0.087	0.039	0.010	0.078	0.040	0.004
2	25	25	0.767	0.606	0.320	0.589	0.410	0.152	0.330	0.192	0.049	0.950	0.877	0.647
		50	0.912	0.839	0.591	0.764	0.608	0.298	0.476	0.293	0.088	0.998	0.986	0.889
	50	25	0.999	0.994	0.952	0.997	0.985	0.890	0.977	0.937	0.764	1.000	1.000	0.999
		50	1.000	1.000	0.999	1.000	1.000	0.992	0.999	0.998	0.964	1.000	1.000	1.000
3	25	25	0.094	0.037	0.006	0.081	0.045	0.007	0.083	0.039	0.005	0.081	0.032	0.003
		50	0.076	0.041	0.004	0.103	0.047	0.005	0.095	0.038	0.004	0.073	0.032	0.005
	50	25	0.094	0.043	0.009	0.103	0.047	0.011	0.098	0.049	0.008	0.083	0.043	0.008
		50	0.091	0.041	0.007	0.084	0.031	0.004	0.098	0.035	0.003	0.072	0.033	0.005
4	25	25	0.141	0.073	0.019	0.131	0.059	0.013	0.108	0.048	0.011	0.168	0.092	0.018
		50	0.177	0.093	0.018	0.142	0.077	0.016	0.101	0.048	0.010	0.151	0.082	0.022
	50	25	0.313	0.187	0.056	0.269	0.154	0.047	0.219	0.123	0.022	0.364	0.232	0.057
		50	0.432	0.300	0.100	0.333	0.215	0.057	0.286	0.175	0.052	0.431	0.289	0.115
5	25	25	0.520	0.354	0.119	0.333	0.199	0.060	0.241	0.128	0.034	0.761	0.624	0.311
		50	0.692	0.545	0.237	0.496	0.324	0.109	0.330	0.204	0.048	0.904	0.814	0.556
	50	25	0.963	0.913	0.694	0.884	0.794	0.506	0.773	0.654	0.370	0.999	0.992	0.931
		50	0.999	0.994	0.958	0.987	0.966	0.832	0.974	0.917	0.706	1.000	1.000	0.998
6	25	25	0.947	0.856	0.612	0.783	0.640	0.333	0.542	0.375	0.143	0.993	0.982	0.894
		50	0.994	0.980	0.885	0.933	0.865	0.626	0.704	0.549	0.255	1.000	1.000	0.994
	50	25	1.000	1.000	0.998	1.000	1.000	0.995	0.999	0.998	0.976	1.000	1.000	1.000
		50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000

Note: 1. $K_C = \left\lceil CT^{1/5} \right\rceil$, $C = 1, 1.5, 2$, K_{cv} refers to the number of sieve terms by LOOCV;
2. DGPs 1 and 3 are for the size study and DGPs 2 and 4-6 are for power comparison.

Table 7: Simulation results for testing $\mathbb{H}_{h0}$

			K_1			K_2			K_3			K_{cv}		
DGP	T	N	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
1	25	25	0.104	0.063	0.023	0.104	0.044	0.012	0.106	0.051	0.015	0.110	0.056	0.011
		50	0.099	0.045	0.012	0.100	0.047	0.012	0.114	0.054	0.015	0.095	0.053	0.014
	50	25	0.093	0.056	0.016	0.098	0.057	0.009	0.121	0.066	0.019	0.100	0.051	0.012
		50	0.099	0.045	0.012	0.106	0.054	0.016	0.105	0.052	0.011	0.120	0.060	0.017
2	25	25	0.110	0.065	0.016	0.103	0.055	0.009	0.112	0.063	0.013	0.110	0.061	0.013
		50	0.100	0.054	0.012	0.118	0.065	0.022	0.088	0.046	0.013	0.094	0.046	0.012
	50	25	0.104	0.056	0.015	0.108	0.055	0.008	0.106	0.058	0.019	0.107	0.058	0.009
		50	0.093	0.055	0.015	0.099	0.048	0.013	0.106	0.058	0.013	0.122	0.059	0.017
3	25	25	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.990	1.000	1.000	0.999
		50	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.999	1.000	1.000	1.000
	50	25	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	25	25	0.545	0.406	0.187	0.422	0.287	0.116	0.373	0.230	0.100	0.650	0.506	0.283
		50	0.720	0.587	0.333	0.584	0.441	0.198	0.426	0.283	0.121	0.825	0.726	0.472
	50	25	0.933	0.867	0.656	0.904	0.820	0.583	0.869	0.764	0.510	0.978	0.942	0.826
		50	0.994	0.986	0.897	0.986	0.955	0.846	0.973	0.943	0.808	0.999	0.994	0.978
5	25	25	0.313	0.206	0.072	0.257	0.156	0.055	0.224	0.132	0.049	0.361	0.255	0.099
		50	0.407	0.292	0.105	0.326	0.202	0.080	0.238	0.143	0.054	0.495	0.369	0.173
	50	25	0.605	0.490	0.239	0.577	0.441	0.218	0.506	0.358	0.149	0.732	0.609	0.363
		50	0.790	0.689	0.431	0.720	0.589	0.364	0.724	0.586	0.341	0.910	0.829	0.586
6	25	25	0.098	0.052	0.010	0.098	0.051	0.012	0.096	0.047	0.010	0.095	0.056	0.024
		50	0.090	0.047	0.014	0.076	0.048	0.005	0.115	0.070	0.018	0.081	0.043	0.009
	50	25	0.082	0.036	0.007	0.092	0.048	0.015	0.090	0.052	0.023	0.094	0.048	0.008
		50	0.102	0.051	0.010	0.101	0.050	0.013	0.081	0.039	0.006	0.096	0.050	0.009

Note: 1. $K_C = \left\lceil CT^{1/5} \right\rceil$, $C = 1, 1.5, 2$, K_{cv} refers to the number of sieve terms by LOOCV;

2. DGPs 1-2 are for the size study. and DGPs 3-5 are for power comparison.

3. DGP 6 satisfies $\mathbb{H}_{h0}$ but comes with non-smooth f_0 and β_0.

Table 8: Simulation results for testing $\mathbb{H}_0$ with both individual and time fixed effects

			K_1			K_2			K_3			K_{cv}		
DGP	T	N	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
1	25	25	0.119	0.058	0.014	0.115	0.050	0.010	0.104	0.058	0.016	0.119	0.065	0.011
		50	0.092	0.046	0.006	0.097	0.045	0.009	0.108	0.047	0.012	0.089	0.050	0.005
	50	25	0.102	0.055	0.012	0.102	0.056	0.008	0.104	0.053	0.005	0.089	0.057	0.016
		50	0.104	0.053	0.014	0.102	0.057	0.014	0.104	0.052	0.014	0.093	0.044	0.011
2	25	25	0.603	0.474	0.230	0.451	0.323	0.134	0.408	0.278	0.100	0.740	0.629	0.386
		50	0.804	0.707	0.461	0.674	0.518	0.279	0.532	0.395	0.199	0.925	0.844	0.669
	50	25	0.915	0.835	0.658	0.825	0.733	0.496	0.807	0.695	0.427	0.987	0.974	0.921
		50	0.996	0.983	0.934	0.976	0.960	0.843	0.958	0.919	0.743	1.000	0.999	0.999
3	25	25	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	50	25	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	25	25	0.728	0.596	0.351	0.610	0.471	0.255	0.507	0.368	0.176	0.837	0.744	0.517
		50	0.927	0.866	0.669	0.782	0.675	0.438	0.720	0.571	0.339	0.970	0.935	0.821
	50	25	0.975	0.947	0.823	0.946	0.890	0.717	0.892	0.815	0.575	0.997	0.992	0.965
		50	1.000	0.999	0.984	0.994	0.990	0.947	0.987	0.964	0.879	1.000	1.000	1.000
5	25	25	0.675	0.557	0.295	0.581	0.459	0.222	0.498	0.359	0.161	0.821	0.732	0.509
		50	0.898	0.818	0.589	0.788	0.671	0.408	0.685	0.554	0.294	0.965	0.939	0.819
	50	25	0.953	0.920	0.790	0.900	0.851	0.666	0.889	0.810	0.599	0.995	0.992	0.956
		50	0.998	0.995	0.975	0.998	0.985	0.930	0.989	0.971	0.873	1.000	1.000	1.000
6	25	25	0.930	0.865	0.673	0.872	0.802	0.595	0.808	0.712	0.453	0.978	0.967	0.878
		50	0.995	0.988	0.947	0.985	0.958	0.873	0.953	0.915	0.774	1.000	1.000	0.992
	50	25	1.000	1.000	1.000	1.000	0.999	0.987	0.999	0.997	0.973	1.000	1.000	1.000
		50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Note: 1. $K_C = \left\lceil CT^{1/5} \right\rceil$, $C = 1, 1.5, 2$, K_{cv} refers to the number of sieve terms by LOOCV;

2. DGP 1 is for the size study and DGPs 2-6 are for power comparison.

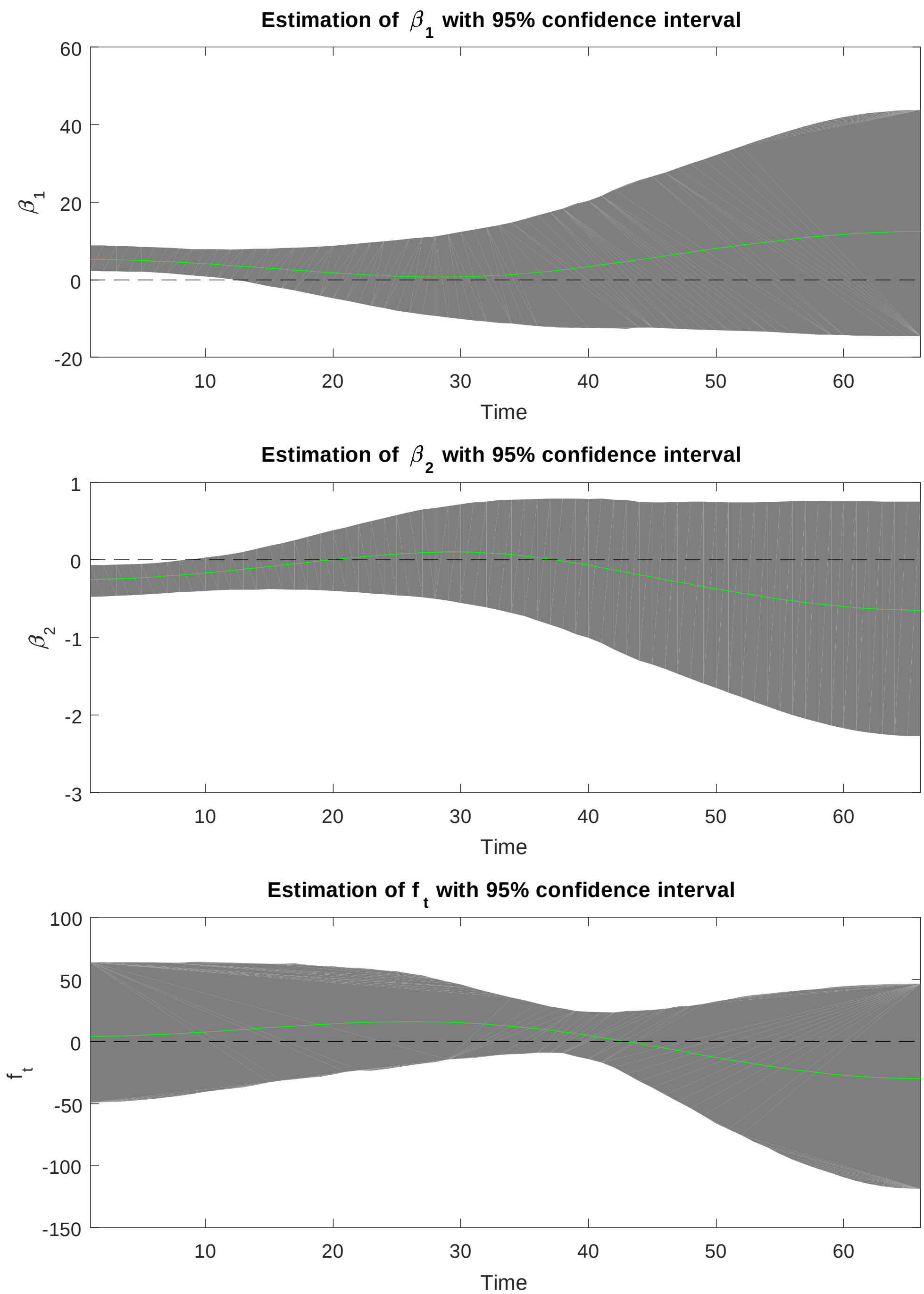

Figure 1: Estimated homogeneous TVCs under $\mathbb{H}_{h0}$