Online Appendices to “Interactive Effects Panel Data Models with General Factors and Regressors”

Bin Peng* Liangjun Su† Joakim Westerlund‡

Monash University Tsinghua University Lund University

and

Deakin University

Yanrong Yang§

Australian National University

July 31, 2023

Abstract

We begin this appendix by laying out the notation that will be used throughout. This is done in Section A. Section B then proves the asymptotic results presented in Section 4 of the main paper. Section C discusses conditions that ensure that the asymptotic distribution of $\widehat{\boldsymbol{\beta}}$ is unbiased. The appendix is concluded with Section D, which provides an empirical application to the long-run relationship between US house prices and income.

A Notation

The matrices $\boldsymbol{\Sigma}_{F^0}$ and $\boldsymbol{\Sigma}_{\Gamma^0}$ have been defined in Assumption 1. In this appendix, we use $\boldsymbol{\Sigma}_{F_g^0}$ and $\boldsymbol{\Sigma}_{\Gamma_g^0}$ to denote the sub-matrices of $\boldsymbol{\Sigma}_{F^0}$ and $\boldsymbol{\Sigma}_{\Gamma^0}$ corresponding to $T^{-\nu_g}\mathbf{F}_g^{0\prime}\mathbf{F}_g^0$ and $N^{-1}\boldsymbol{\Gamma}_g^{0\prime}\boldsymbol{\Gamma}_g^0$, respectively, for $g = 1,\ldots,G$. We also define $\mathbf{V}_g = \mathrm{diag}(\widehat{\lambda}_{g,1},\ldots,\widehat{\lambda}_{g,d_{max}})$, where $\widehat{\lambda}_{g,d}$ has been defined in Step 2 of the IPC estimation procedure. We partition $\mathbf{F}^0 = (\mathbf{F}_1^0,\ldots,\mathbf{F}_g^0,\mathbf{F}_{+g}^0)$ and $\mathbf{C}_T = \mathrm{diag}(T^{-\nu_1/2}\mathbf{I}_{d_1},\ldots,T^{-\nu_g/2}\mathbf{I}_{d_g},\mathbf{C}_{+g,T})$, where $\mathbf{F}_{+g}^0 = (\mathbf{F}_{g+1}^0,\ldots,\mathbf{F}_G^0)$ is $T\times(d_{g+1}+\cdots+d_G)$, and $\mathbf{C}_{+g,T} = \mathrm{diag}(T^{-\nu_{g+1}/2}\mathbf{I}_{d_{g+1}},\ldots,T^{-\nu_G/2}\mathbf{I}_{d_G})$ is $(d_{g+1}+\cdots+d_G)\times(d_{g+1}+\cdots+d_G)$. We partition $\widehat{\mathbf{F}}$, $\boldsymbol{\gamma}_i^0$ and $\boldsymbol{\Gamma}^0$ conformably as $\widehat{\mathbf{F}} = (\widehat{\mathbf{F}}_1,\ldots,\widehat{\mathbf{F}}_g,\widehat{\mathbf{F}}_{+g})$, $\boldsymbol{\gamma}_i^0 = (\boldsymbol{\gamma}_{1,i}^{0\prime},\ldots,\boldsymbol{\gamma}_{g,i}^{0\prime},\boldsymbol{\gamma}_{+g,i}^{0\prime})'$ and $\boldsymbol{\Gamma}^0 = (\boldsymbol{\Gamma}_1^0,\ldots,\boldsymbol{\Gamma}_g^0,\boldsymbol{\Gamma}_{+g}^0)$, respectively.

*E-mail address: bin.peng@monash.edu.

†E-mail address: sulj@sem.tsinghua.edu.cn.

‡E-mail address: joakim.westerlund@nek.lu.se.

§E-mail address: yanrong.yang@anu.edu.au.

We introduce $\lambda_{g,d} = T^{-\nu_g}\mathbf{h}^{0\prime}_{g,d}\mathbf{F}^{0\prime}_g\mathbf{\Sigma}^0_g\mathbf{F}^0_g\mathbf{h}^0_{g,d}$, where $\mathbf{\Sigma}^0_g = N^{-1}\mathbf{F}^0_g\mathbf{\Gamma}^{0\prime}_g\mathbf{\Gamma}^0_g\mathbf{F}^{0\prime}_g$ and $\mathbf{h}^0_{g,d}$ is the d-th column of $\mathbf{H}^0_g = N^{-1}T^{(\nu_g-\delta)/2}\mathbf{\Gamma}^{0\prime}_g\mathbf{\Gamma}^0_g\mathbf{F}^{0\prime}_g\widehat{\mathbf{F}}^0_g(\mathbf{V}^0_g)^{-1}$ with $\widehat{\mathbf{F}}^0_g$ being the $T \times d_g$ matrix consisting of the first d_g columns of $\widehat{\mathbf{F}}_g$ and $\mathbf{V}^0_g$ being the leading $d_g \times d_g$ principal submatrix of $\mathbf{V}_g$. In other words, $\widehat{\mathbf{F}}^0_g$ and $\mathbf{V}^0_g$ are $\widehat{\mathbf{F}}_g$ and $\mathbf{V}_g$ based on treating the number of factors for each group g, d_g, as known. We also define $\mathbf{H}_g = T^{-(\nu_g-\delta)/2}\mathbf{H}^0_g$. In order to appreciate the implication of the difference in normalization with respect to T, let us consider $\mathbf{H}^0_g$. By Assumption 1, $N^{-1}\mathbf{\Gamma}^{0\prime}_g\mathbf{\Gamma}^0_g$ is asymptotically of full rank, and hence $\|N^{-1}\mathbf{\Gamma}^{0\prime}_g\mathbf{\Gamma}^0_g\| = O_P(1)$. Hence, since

$$\|T^{-(\nu_g+\delta)/2}\mathbf{F}^{0\prime}_g\widehat{\mathbf{F}}_g\|^2 \leq T^{-\delta}\|\widehat{\mathbf{F}}_g\|^2_2 T^{-\nu_g}\|\mathbf{F}^0_g\|^2 = T^{-\nu_g}\|\mathbf{F}^0_g\|^2 = O_P(1) \tag{A.1}$$

and $\|(T^{-\nu_1}\mathbf{V}^0_1)^{-1}\| = O_P(1)$ as explained under (B.28), we can show that

$$\|\mathbf{H}^0_g\| \leq \|N^{-1}\mathbf{\Gamma}^{0\prime}_g\mathbf{\Gamma}^0_g\|\|T^{-(\nu_g+\delta)/2}\mathbf{F}^{0\prime}_g\widehat{\mathbf{F}}^0_g\|\|(T^{-\nu_g}\mathbf{V}^0_g)^{-1}\| = O_P(1), \tag{A.2}$$

which in turn implies

$$\|\mathbf{H}_g\| = T^{-(\nu_g-\delta)/2}\|\mathbf{H}^0_g\| = O_P(T^{-(\nu_g-\delta)/2}). \tag{A.3}$$

We further use $\widehat{\mathbf{F}}^0_{g,d}$ to refer to the d-th column of $\widehat{\mathbf{F}}^0_g$. In this notation, $\widehat{\lambda}_{g,d} = T^{-\delta}\widehat{\mathbf{F}}^{0\prime}_{g,d}\widehat{\mathbf{\Sigma}}_g\widehat{\mathbf{F}}^0_{g,d}$ for $d = 1, \ldots, d_g$.

We also partition $\mathbf{X}_i$ as $\mathbf{X}_i = (\mathbf{X}_{1,i}, \ldots, \mathbf{X}_{d_x,i})$ with $\mathbf{X}_{j,i}$ being the j-th column of $\mathbf{X}_i$. The j-th column of $\mathbf{X}_i\mathbf{D}_T$ is therefore given by $T^{-\kappa_j/2}\mathbf{X}_{j,i}$. Moreover, $\operatorname{vec}\mathbf{A}$, $\operatorname{rank}\mathbf{A}$, $\operatorname{span}\mathbf{A}$ and $\lambda(\mathbf{A})$ denote the vectorized version, rank, span and eigenvalues of $\mathbf{A}$, respectively, $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$. For two random variables x and y, $x \asymp y$ means that $x = O_P(y)$ and $y = O_P(x)$.

B Proofs

B.1 Outline

In this section, we describe the outline of the proofs of this section. We assume throughout that $G \geq 2$. The proofs for the cases when $G \in \{0, 1\}$ are much simpler, and can be obtained by manipulating the proofs for $G \geq 2$.

Sections B.2 and B.3 contain a number of auxiliary lemmas that will be used later in Section B.4 to prove the results reported in the main paper and their proofs. The proofing is done stepwise, starting with Step 1 of the IPC estimation procedure. Lemma B.1 is a widely used result for the eigenvalues of large dimensional matrices (Lam et al., 2011), and is presented here for convenience. Lemma B.2 is more novel. It presents two order results that will be used repeatedly in the proofs. Given Lemma B.2, we are able to establish Lemma 1 of the main text, which provides a lower bound on the rate of convergence of the initial Step 1 estimator $\widehat{\boldsymbol{\beta}}_0$. As a first step towards establishing the consistency of the Step 2 estimators of $(d_1, \ldots, d_G)$, in Lemma B.3 we study limiting behavior of the eigenvalues of $\widehat{\boldsymbol{\Sigma}}_1$ in equation (6) of the main paper. Lemmas B.4 and B.5 enable consistent estimation of $(d_2, \ldots, d_G)$, and are thus key in proving Lemma 2 of the main text. Sections B.2 and B.3 also establish the rate of convergence of $\widehat{\boldsymbol{\beta}}_0$ under the conditions of Theorem 1. This rate is provided as a part of Lemma B.6.

As already mentioned, Section B.4 provides the proofs of Lemmas 1 and 2, and Theorem 1 of the main paper. Corollary 1 is an immediate consequence of Theorem 1.

B.2 Auxiliary lemmas

Lemma B.1. *Suppose that* $\mathbf{A}$ *and* $\mathbf{A}+\mathbf{E}$ *are* $n \times n$ *symmetric matrices and that* $\mathbf{Q} = (\mathbf{Q}_1, \mathbf{Q}_2)$, *where* $\mathbf{Q}_1$ *is* $n \times r$ *and* $\mathbf{Q}_2$ *is* $n \times (n-r)$, *is an orthogonal matrix such that* span $\mathbf{Q}_1$ *is an invariant subspace for* $\mathbf{A}$. *Decompose* $\mathbf{Q}'\mathbf{A}\mathbf{Q}$ *and* $\mathbf{Q}'\mathbf{E}\mathbf{Q}$ *as* $\mathbf{Q}'\mathbf{A}\mathbf{Q} = \operatorname{diag}(\mathbf{D}_1, \mathbf{D}_2)$ *and*

$$\mathbf{Q}'\mathbf{E}\mathbf{Q} = \begin{pmatrix} \mathbf{E}_{11} & \mathbf{E}_{12} \\ \mathbf{E}_{21} & \mathbf{E}_{22} \end{pmatrix}.$$

Let $\operatorname{sep}(\mathbf{D}_1, \mathbf{D}_2) = \min_{\lambda_1 \in \lambda(\mathbf{D}_1),\ \lambda_2 \in \lambda(\mathbf{D}_2)} |\lambda_1 - \lambda_2|$. *If* $\operatorname{sep}(\mathbf{D}_1, \mathbf{D}_2) > 0$ *and* $\|\mathbf{E}\|_2 \leq \operatorname{sep}(\mathbf{D}_1, \mathbf{D}_2)/5$, *then there exists a* $(n-r) \times r$ *matrix* $\mathbf{P}$ *with* $\|\mathbf{P}\|_2 \leq 4\|\mathbf{E}_{21}\|_2/\operatorname{sep}(\mathbf{D}_1, \mathbf{D}_2)$, *such that the columns of* $\mathbf{Q}_1^0 = (\mathbf{Q}_1 + \mathbf{Q}_2\mathbf{P})(\mathbf{I}_r + \mathbf{P}'\mathbf{P})^{-1/2}$ *define an orthonormal basis for a subspace that is invariant for* $\mathbf{A}+\mathbf{E}$.

Lemma B.2. *Under Assumption 1, as* $N, T \to \infty$,

(a) $\sup_{\mathbf{F} \in \mathbb{D}_F} (NT)^{-1} \sum_{i=1}^N \boldsymbol{\varepsilon}_i' \mathbf{P}_F \boldsymbol{\varepsilon}_i = O_P(N^{-1} \vee T^{-1})$;

(b) $\sup_{\mathbf{F} \in \mathbb{D}_F} \|(NT)^{-1} \sum_{i=1}^N \mathbf{D}_T \mathbf{X}_i' \mathbf{P}_F \boldsymbol{\varepsilon}_i\| = O_P(N^{-1/2} \vee T^{-1/2})$.

If Assumption 3 also holds, then

(c) $(NT)^{-1}\|\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\| = O_P(N^{-1/2} \vee T^{-1/2})$ *and* $(NT)^{-1}\|\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\| = O_P(N^{-1/2} \vee T^{-1/2})$;

(d) $\|\boldsymbol{\Gamma}^{0\prime}\boldsymbol{\varepsilon}\| = O_P(\sqrt{NT})$.

Lemma B.3. *Let Assumptions 1–4 hold. Then, as* $N, T \to \infty$,

(a) $T^{-\nu_1}|\widehat{\lambda}_{1,d} - \lambda_{1,d}| = O_P(T^{-(\nu_1-\nu_2)/2})$ *for* $d = 1, \ldots, d_1$;

(b) $T^{-\nu_1}|\widehat{\lambda}_{1,d}| = O_P(T^{-(\nu_1-\nu_2)})$ *for* $d = d_1+1, \ldots, d_{max}$.

Lemma B.4. *As* $N, T \to \infty$, *the following results hold under Assumptions 1–4:*

(a) $T^{-\delta}\|\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_1\| = O_P(T^{-(\delta+\nu_1-\nu_2)/2} + N^{-1/2}T^{-(\delta+\nu_1-\nu_2-1)/2} + N^{-(1-p)}T^{-(\delta+\nu_1-2\nu_2)/2} + T^{-(\nu_1+\delta)/2-q})$;

(b) $\sum_{i=1}^N \|\mathbf{F}_1^{0\prime}\boldsymbol{\gamma}_{1,i}^0 - \widehat{\mathbf{F}}_1\widehat{\boldsymbol{\gamma}}_{1,i}\|^2 = O_P(N \vee T + NT^{2q-\nu_1} + N^{-(1-2p)}T^{\nu_2})$.

Lemma B.5. *Let $\tau_{NT} = N^{-1/2}T^{-(\nu_2-1)/2} + T^{-\nu_2/2} + T^{-(\nu_2-\nu_3)} + T^{q-(\nu_1+\nu_2)/2} + N^{-(1-p)}$. As $N, T \to \infty$, the following results hold under the conditions of Lemma B.4:*

(a) $T^{-\nu_2}|\widehat{\lambda}_{2,d} - \lambda_{2,d}| = O_P(\tau_{NT})$ *for* $d = 1, \ldots, d_2$;

(b) $T^{-\nu_2}|\widehat{\lambda}_{2,d}| = O_P(\tau_{NT}^2)$ *for* $d = d_2 + 1, \ldots, d_{max}$.

Lemma B.6. *Let Assumptions 1–4 hold. In addition, let $NT^{-\nu_G} < \infty$, as $N, T \to \infty$,*

(a) $\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\| = O_P((NT)^{-1/2} \vee \|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\|)$;

(b) $\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\| = O_P((NT)^{-1/2})$.

B.3 Proofs of auxiliary lemmas

Proof of Lemma B.1.

This is Lemma 3 of Lam et al. (2011). The proof is therefore omitted. ■

Proof of Lemma B.2.

Consider (a). We have

$$\sup_{\mathbf{F}\in\mathbb{D}_F} \frac{1}{NT}\sum_{i=1}^N \boldsymbol{\varepsilon}_i'\mathbf{P}_F\boldsymbol{\varepsilon}_i = \sup_{\mathbf{F}\in\mathbb{D}_F}(NT)^{-1}\mathrm{tr}(\mathbf{P}_F\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}) \le O(1)\sup_{\mathbf{F}\in\mathbb{D}_F}(NT)^{-1}\|\mathbf{P}_F\|_2\|\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\|_2$$
$$\le O(1)\sup_{\mathbf{F}\in\mathbb{D}_F}(NT)^{-1}\|\boldsymbol{\varepsilon}\|_2^2 = O_P(N^{-1} \vee T^{-1}), \tag{B.1}$$

where the first inequality follows from the fact that $|\mathrm{tr}\,\mathbf{A}| \le \mathrm{rank}\,\mathbf{A}\,\|\mathbf{A}\|_2$, the second inequality follows from the fact that $\|\mathbf{P}_F\|_2 = 1$, and the second equality holds by Assumption 1 (b).

The result in (b) is due to

$$\sup_{\mathbf{F}\in\mathbb{D}_F}\left\|\frac{1}{NT}\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{P}_F\boldsymbol{\varepsilon}_i\right\| \le \frac{1}{NT}\sum_{j=1}^{d_x}\sup_{\mathbf{F}\in\mathbb{D}_F}\left|\mathrm{tr}\,(T^{-\kappa_j/2}\mathbf{X}_j\mathbf{P}_F\boldsymbol{\varepsilon}')\right|$$
$$\le O(1)\frac{1}{NT}\sum_{j=1}^{d_x}\sup_{\mathbf{F}\in\mathbb{D}_F}\|T^{-\kappa_j/2}\mathbf{X}_j\|_2\|\mathbf{P}_F\|_2\|\boldsymbol{\varepsilon}\|_2$$
$$= O(1)(NT)^{-1}O_P(\sqrt{NT})O_P(\sqrt{N} \vee \sqrt{T})$$
$$= O_P(N^{-1/2} \vee T^{-1/2}), \tag{B.2}$$

where, with a slight abuse of notation and in this proof only, $\mathbf{X}_j = (\mathbf{X}_{j,1}, \ldots, \mathbf{X}_{j,N})'$, the second inequality follows from $|\mathrm{tr}\, \mathbf{A}| \leq \mathrm{rank}\, \mathbf{A}\, \|\mathbf{A}\|_2$, while the first equality is due to Assumption 1.

For (c), we use

$$
\begin{aligned}
(NT)^{-2}\mathbb{E}\|\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\|^2 &= \frac{1}{(NT)^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\sum_{i=1}^{N}\mathbb{E}[\varepsilon_{i,t}^2\varepsilon_{i,s}^2] + \sum_{i=1}^{N}\sum_{j\neq i}\mathbb{E}[\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{j,t}\varepsilon_{j,s}]\right) \\
&= \frac{1}{(NT)^2}\sum_{t=1}^{T}\left(\sum_{i=1}^{N}\mathbb{E}(\varepsilon_{i,t}^4) + \sum_{i=1}^{N}\sum_{j\neq i}\mathbb{E}[(\varepsilon_{i,t}\varepsilon_{j,t} - \sigma_{\varepsilon,ij})^2]\right) \\
&+ \frac{1}{(NT)^2}\sum_{t=1}^{T}\sum_{s\neq t}\left(\sum_{i=1}^{N}\mathbb{E}[\varepsilon_{i,t}^2\varepsilon_{i,s}^2] + \sum_{i=1}^{N}\sum_{j\neq i}\mathbb{E}[(\varepsilon_{i,t}\varepsilon_{j,t} - \sigma_{\varepsilon,ij})(\varepsilon_{i,s}\varepsilon_{j,s} - \sigma_{\varepsilon,ij})]\right) \\
&+ \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\sigma_{\varepsilon,ij}^2 = O(N^{-1}) + O(T^{-1}), \qquad \text{(B.3)}
\end{aligned}
$$

where the third equality follows from using the mixing condition on $\varepsilon_{i,t}\varepsilon_{j,t}$ across t. The above result implies that $(NT)^{-1}\|\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\| = O_P(N^{-1/2}) + O_P(T^{-1/2})$ as required for the first result in (c). The second follows from

$$
\begin{aligned}
(NT)^{-2}\mathbb{E}\|\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\|^2 &= \frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}[\varepsilon_{i,t}\varepsilon_{j,t}\varepsilon_{i,s}\varepsilon_{j,s}] \\
&= \frac{1}{(NT)^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\sum_{i=1}^{N}\mathbb{E}[\varepsilon_{i,t}^2\varepsilon_{i,s}^2] + \sum_{i=1}^{N}\sum_{j\neq i}\mathbb{E}[\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{j,t}\varepsilon_{j,s})\right) \\
&= O(N^{-1}) + O(T^{-1}), \qquad \text{(B.4)}
\end{aligned}
$$

where the last step follows by the same arguments used to establish the first result of (c).

It remains to prove (d), which is a direct consequence of Assumptions 1 and 3, as seen from

$$
\mathbb{E}\|\boldsymbol{\Gamma}^{0\prime}\boldsymbol{\varepsilon}\|^2 = \sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}[\boldsymbol{\gamma}_i^{0\prime}\boldsymbol{\gamma}_j^0\varepsilon_{i,t}\varepsilon_{j,t}] \leq O(T)\sum_{i=1}^{N}\sum_{j=1}^{N}|\sigma_{\varepsilon,ij}| = O(NT). \qquad \text{(B.5)}
$$

This establishes (d) and hence the proof of the lemma is complete. ■

Proof of Lemma B.3.

Consider (a). As in Appendix A, decompose $\mathbf{F}^0 = (\mathbf{F}_1^0, \mathbf{F}_{+1}^0)$ and $\mathbf{C}_T = \mathrm{diag}(T^{-\nu_1/2}\mathbf{I}_{d_1}, \mathbf{C}_{+1,T})$, where $\mathbf{F}_{+1}^0 = (\mathbf{F}_2^0, \ldots, \mathbf{F}_G^0)$ is $T\times(d_f - d_1)$ and $\mathbf{C}_{+1,T} = \mathrm{diag}(T^{-\nu_2/2}\mathbf{I}_{d_2}, \ldots, T^{-\nu_G/2}\mathbf{I}_{d_G})$ is $(d_f - d_1)\times(d_f - d_1)$. We partition $\widehat{\mathbf{F}}$, $\boldsymbol{\gamma}_i^0$ and $\boldsymbol{\Gamma}^0$ conformably as $\widehat{\mathbf{F}} = (\widehat{\mathbf{F}}_1, \widehat{\mathbf{F}}_{+1})$, $\boldsymbol{\gamma}_i^0 = (\boldsymbol{\gamma}_{1,i}^{0\prime}, \boldsymbol{\gamma}_{+1,i}^{0\prime})'$ and $\boldsymbol{\Gamma}^0 = (\boldsymbol{\Gamma}_1^0, \boldsymbol{\Gamma}_{+1}^0)$, respectively.

By the definition of the eigenvectors and eigenvalues, $\widehat{\boldsymbol{\Sigma}}_1\widehat{\mathbf{F}}_1 = \widehat{\mathbf{F}}_1\mathbf{V}_1$. By using this and the definition of $\widehat{\mathbf{\Sigma}}_1$,

$$\begin{aligned}
&T^{-(\nu_1+\delta)/2}\widehat{\mathbf{F}}_1\mathbf{V}_1\\
&=\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\mathbf{X}_i(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\widehat{\mathbf{F}}_1\\
&+\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\mathbf{X}_i(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\boldsymbol{\gamma}_{1,i}^{0\prime}\mathbf{F}_1^{0\prime}\widehat{\mathbf{F}}_1+\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\widehat{\mathbf{F}}_1\\
&+\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\mathbf{X}_i(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0+\boldsymbol{\varepsilon}_i)'\widehat{\mathbf{F}}_1\\
&+\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0+\boldsymbol{\varepsilon}_i)(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\widehat{\mathbf{F}}_1\\
&+\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0+\boldsymbol{\varepsilon}_i)(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0+\boldsymbol{\varepsilon}_i)'\widehat{\mathbf{F}}_1\\
&+\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0+\boldsymbol{\varepsilon}_i)'\widehat{\mathbf{F}}_1+\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0+\boldsymbol{\varepsilon}_i)\boldsymbol{\gamma}_{1,i}^{0\prime}\mathbf{F}_1^{0\prime}\widehat{\mathbf{F}}_1\\
&+\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0\boldsymbol{\gamma}_{1,i}^{0\prime}\mathbf{F}_1^{0\prime}\widehat{\mathbf{F}}_1\\
&=\sum_{j=1}^{9}\mathbf{J}_j, \qquad \text{(B.6)}
\end{aligned}$$

with implicit definitions of $\mathbf{J}_1,\dots,\mathbf{J}_9$. Note that

$$\mathbf{J}_9=\mathbf{F}_1^0(N^{-1}\mathbf{\Gamma}_1^{0\prime}\mathbf{\Gamma}_1^0)(T^{-(\nu_1+\delta)/2}\mathbf{F}_1^{0\prime}\widehat{\mathbf{F}}_1). \qquad \text{(B.7)}$$

Hence, moving this term over to the left-hand side, the above expression for $T^{-(\nu_1+\delta)/2}\widehat{\mathbf{F}}_1\mathbf{V}_1$ becomes

$$T^{-(\nu_1+\delta)/2}\widehat{\mathbf{F}}_1\mathbf{V}_1-\mathbf{J}_9=\sum_{j=1}^{8}\mathbf{J}_j. \qquad \text{(B.8)}$$

We now evaluate each of the terms on the right-hand side.

Because $T^{-\delta}\|\widehat{\mathbf{F}}_1\|^2=d_{max}$ and $(NT)^{-1}\sum_{i=1}^N\|\mathbf{X}_i\mathbf{D}_T\|^2=O_p(1)$ by Assumption 1, the order of $\mathbf{J}_1$ is given by

$$\begin{aligned}
T^{-\delta/2}\|\mathbf{J}_1\|&\le\left\|\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\mathbf{X}_i(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\right\|T^{-\delta/2}\|\widehat{\mathbf{F}}_1\|\\
&\le O(1)\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\|\mathbf{X}_i\mathbf{D}_T\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|^2\\
&=O_P(T^{1-(\nu_1+\delta)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|^2). \qquad \text{(B.9)}
\end{aligned}$$

Moreover, since

$$\frac{1}{NT^{\nu_1}}\sum_{i=1}^{N}\|\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0\|^2 \le \frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{\gamma}_{1,i}^0\|^2 T^{-\nu_1}\|\mathbf{F}_1^0\|^2 = O_P(1) \tag{B.10}$$

by Assumption 1, we can show that

$$\begin{aligned}
T^{-\delta/2}\|\mathbf{J}_2\| &\le O(1)\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\|\mathbf{X}_i(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\boldsymbol{\gamma}_{1,i}^{0\prime}\mathbf{F}_1^{0\prime}\| \\
&\le O(1)\left(\frac{1}{NT^{\delta}}\sum_{i=1}^{N}\|\mathbf{X}_i\mathbf{D}_T\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|^2\right)^{1/2}\left(\frac{1}{NT^{\nu_1}}\sum_{i=1}^{N}\|\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0\|^2\right)^{1/2} \\
&= O_P(T^{(1-\delta)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|),
\end{aligned} \tag{B.11}$$

and by exactly the same arguments,

$$T^{-\delta/2}\|\mathbf{J}_3\| = O_P(T^{(1-\delta)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|). \tag{B.12}$$

For $\mathbf{J}_4$, we use

$$\begin{aligned}
&\frac{1}{NT^{(\nu_1+\delta)/2}}\left\|\sum_{i=1}^{N}\mathbf{X}_i\mathbf{D}_T\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0)'\right\| \\
&\le \left(\frac{1}{NT^{\delta}}\sum_{i=1}^{N}\|\mathbf{X}_i\mathbf{D}_T\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|^2\right)^{1/2}\left(\frac{1}{NT^{\nu_1}}\sum_{i=1}^{N}\|\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0\|^2\right)^{1/2} \\
&\le O_P(T^{(1-\delta)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|)O_P(T^{-\nu_1/2}\|\mathbf{C}_{+1,T}^{-1}\|) \\
&= O_P(T^{(1-\delta-\nu_1+\nu_2)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|),
\end{aligned} \tag{B.13}$$

where the last equality makes use of the fact that $\|\mathbf{C}_{+1,T}^{-1}\| = O(T^{\nu_2/2})$, as $\nu_2 > \cdots > \nu_G$ by Assumption 1. We can further show that

$$\begin{aligned}
&\frac{1}{NT^{(\nu_1+\delta)/2}}\left\|\sum_{i=1}^{N}\mathbf{X}_i\mathbf{D}_T\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\boldsymbol{\varepsilon}_i'\right\| \\
&= \left\|\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\mathrm{vec}\left[\mathbf{X}_i\mathbf{D}_T\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\boldsymbol{\varepsilon}_i'\right]\right\| \\
&\le \left\|\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}(\boldsymbol{\varepsilon}_i\otimes\mathbf{X}_i\mathbf{D}_T)\right\|\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\| \\
&= O_P(N^{-1/2}T^{(2-\nu_1-\delta)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|),
\end{aligned} \tag{B.14}$$

where the last equality holds, because by Assumption 3 and $(\mathrm{tr}\,\mathbf{A}'\mathbf{B})^2 \leq (\mathrm{tr}\,\mathbf{A}'\mathbf{A})(\mathrm{tr}\,\mathbf{B}'\mathbf{B})$, we have

$$\begin{aligned}
&\mathbb{E}\left\|\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}(\boldsymbol{\varepsilon}_i\otimes\mathbf{X}_i\mathbf{D}_T)\right\|^2\\
&= \frac{1}{N^2T^{\nu_1+\delta}}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}(\mathbf{x}_{j,t}'\mathbf{D}_T^2\mathbf{x}_{i,t}\varepsilon_{i,s}\varepsilon_{j,s}) \leq \frac{1}{N^2T^{\nu_1+\delta}}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}|\mathrm{tr}(\mathbf{D}_T\mathbf{X}_j'\mathbf{X}_i\mathbf{D}_T)||\sigma_{\varepsilon,ij}|\\
&\leq \frac{1}{N^2T^{\nu_1+\delta-2}}\sum_{i=1}^{N}\sum_{j=1}^{N}\sqrt{T^{-1}\mathbb{E}\|\mathbf{D}_T\mathbf{X}_j\|^2}\sqrt{T^{-1}\mathbb{E}\|\mathbf{X}_i\mathbf{D}_T\|^2}|\sigma_{\varepsilon,ij}|\\
&= O(1)\frac{1}{N^2T^{\nu_1+\delta-2}}\sum_{i=1}^{N}\sum_{j=1}^{N}|\sigma_{\varepsilon,ij}| = O(N^{-1}T^{2-\nu_1-\delta}). \qquad \text{(B.15)}
\end{aligned}$$

Hence, by adding the results,

$$\begin{aligned}
T^{-\delta/2}\|\mathbf{J}_4\| &\leq O(1)\frac{1}{NT^{(\nu_1+\delta)/2}}\left\|\sum_{i=1}^{N}\mathbf{X}_i(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0+\boldsymbol{\varepsilon}_i)'\right\|\\
&= O_P(T^{(1-\delta-\nu_1+\nu_2)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\| + N^{-1/2}T^{(2-\nu_1-\delta)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|)\\
&= o_P(T^{-(1-\delta)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|), \qquad \text{(B.16)}
\end{aligned}$$

where we have used Assumptions 1 and 4 ($T/N^2 = O(1)$ under $\nu_G < 1$) to show that $T^{(1-\delta-\nu_1+\nu_2)/2}$ and $N^{-1/2}T^{(2-\nu_1-\delta)/2}$ are $o(1)$. The same steps can be used to show that

$$T^{-\delta/2}\|\mathbf{J}_5\| \leq o_P(T^{-(1-\delta)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|). \qquad \text{(B.17)}$$

For $\mathbf{J}_6$, we use

$$\begin{aligned}
T^{-\delta/2}\|\mathbf{J}_6\| &\leq O(1)\left\|\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0+\boldsymbol{\varepsilon}_i)(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0+\boldsymbol{\varepsilon}_i)'\right\|\\
&\leq O(1)(N^{-1}T^{-(\nu_1+\delta)/2}\|\mathbf{F}_{+1}^0\boldsymbol{\Gamma}_{+1}^{0\prime}\boldsymbol{\Gamma}_{+1}^0\mathbf{F}_{+1}^{0\prime}\| + N^{-1}T^{-(\nu_1+\delta)/2}\|\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\|\\
&\quad + 2N^{-1}T^{-(\nu_1+\delta)/2}\|\mathbf{F}_{+1}^0\boldsymbol{\Gamma}_{+1}^{0\prime}\boldsymbol{\varepsilon}\|). \qquad \text{(B.18)}
\end{aligned}$$

By Assumption 1,

$$N^{-1}T^{-(\nu_1+\delta)/2}\|\mathbf{F}_{+1}^0\boldsymbol{\Gamma}_{+1}^{0\prime}\boldsymbol{\Gamma}_{+1}^0\mathbf{F}_{+1}^{0\prime}\| = O_P(T^{\nu_2-(\nu_1+\delta)/2}). \qquad \text{(B.19)}$$

Another application of Assumption 1 and Lemma B.2 gives

$$\begin{aligned}
N^{-1}T^{-(\nu_1+\delta)/2}\|\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\| &= O_P(T^{1-(\nu_1+\delta)/2}(N^{-1/2}\vee T^{-1/2})), \qquad \text{(B.20)}\\
N^{-1}T^{-(\nu_1+\delta)/2}\|\mathbf{F}_{+1}^0\boldsymbol{\Gamma}_{+1}^{0\prime}\boldsymbol{\varepsilon}\| &\leq N^{-1/2}T^{-(\nu_1+\delta-\nu_2-1)/2}T^{-\nu_2/2}\|\mathbf{F}_{+1}^0\|(NT)^{-1/2}\|\boldsymbol{\Gamma}_{+1}^{0\prime}\boldsymbol{\varepsilon}\|
\end{aligned}$$

$$= O_P(N^{-1/2}T^{-(\nu_1+\delta-\nu_2-1)/2}). \tag{B.21}$$

These results can be inserted into the expression for $T^{-\delta/2}\|\mathbf{J}_6\|$, giving

$$T^{-\delta/2}\|\mathbf{J}_6\| = O_P(T^{\nu_2-(\nu_1+\delta)/2} + N^{-1/2}T^{1-(\nu_1+\delta)/2}). \tag{B.22}$$

Next up is $\mathbf{J}_7$. By using Assumptions 1 and 4, and Lemma B.2, and the arguments use in evaluating $\mathbf{J}_6$,

$$\begin{aligned}
T^{-\delta/2}\|\mathbf{J}_7\| &\leq O(1)\left\|\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0+\boldsymbol{\varepsilon}_i)'\right\| \\
&\leq O(1)N^{-1}T^{-(\nu_1+\delta)/2}\|\mathbf{F}_1^0\boldsymbol{\Gamma}_1^{0\prime}\boldsymbol{\Gamma}_{+1}^0\mathbf{F}_{+1}^{0\prime}\| + O(1)N^{-1}T^{-(\nu_1+\delta)/2}\|\mathbf{F}_1^0\boldsymbol{\Gamma}_1^{0\prime}\boldsymbol{\varepsilon}\| \\
&= O_P(T^{(\nu_2-\delta)/2} + N^{-1/2}T^{(1-\delta)/2}) = O_P(T^{(\nu_2-\delta)/2}),
\end{aligned} \tag{B.23}$$

and we can similarly show that

$$T^{-\delta/2}\|\mathbf{J}_8\| = O_P(T^{(\nu_2-\delta)/2}). \tag{B.24}$$

By putting everything together, (B.8) becomes

$$\begin{aligned}
T^{-\delta/2}\|T^{-(\nu_1+\delta)/2}\widehat{\mathbf{F}}_1\mathbf{V}_1 - \mathbf{J}_9\| &= O_P(T^{-(\delta-1)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|) \\
&\quad + O_P(N^{-1/2}T^{1-(\nu_1+\delta)/2}) + O_P(T^{-(\delta-\nu_2)/2}).
\end{aligned} \tag{B.25}$$

We now left multiply (B.8) by $T^{-(\nu_1+\delta)/2}\widehat{\mathbf{F}}_1'$ to obtain that

$$\begin{aligned}
&T^{-\nu_1}\mathbf{V}_1 - (T^{-(\nu_1+\delta)/2}\widehat{\mathbf{F}}_1'\mathbf{F}_1^0)(N^{-1}\boldsymbol{\Gamma}_1^{0\prime}\boldsymbol{\Gamma}_1^0)(T^{-(\nu_1+\delta)/2}\mathbf{F}_1^{0\prime}\widehat{\mathbf{F}}_1) \\
&= T^{-(\nu_1-\delta)/2}O_P(T^{-(\delta-1)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\| + N^{-1/2}T^{1-(\nu_1+\delta)/2} + T^{-(\delta-\nu_2)/2}) \\
&= O_P(T^{-(\nu_1-1)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|) + O_P(T^{-(\nu_1-\nu_2)/2}) \\
&= O_P(T^{-(\nu_1-\nu_2)/2}),
\end{aligned} \tag{B.26}$$

where the third equality follows from Assumption 4 and Lemma 1. This implies that $\mathbf{V}_1$ is at most of rank d_1. Similarly, we can left-multiply (B.8) by $T^{-\nu_1}\mathbf{F}_1^{0\prime}$ to obtain

$$\|T^{-(\nu_1+\delta)/2}\mathbf{F}_1^{0\prime}\widehat{\mathbf{F}}_1(T^{-\nu_1}\mathbf{V}_1) - T^{-\nu_1}\mathbf{F}_1^{0\prime}\mathbf{J}_9\| = O_P(T^{-(\nu_1-\nu_2)/2}), \tag{B.27}$$

which in turn implies that

$$\boldsymbol{\Sigma}_{F_1^0}\boldsymbol{\Sigma}_{\Gamma_1^0}(T^{-(\nu_1+\delta)/2}\mathbf{F}_1^{0\prime}\widehat{\mathbf{F}}) = (T^{-(\nu_1+\delta)/2}\mathbf{F}_1^{0\prime}\widehat{\mathbf{F}})(T^{-\nu_1}\mathbf{V}_1) + o_P(1). \tag{B.28}$$

Note that $T^{-(\nu_1+\delta)/2}\mathbf{F}_1^{0\prime}\widehat{\mathbf{F}}$ is of rank d_1, which then further indicates that $\mathbf{V}_1$ has at least d_1 non-zero elements on the main diagonal which converge to the eigenvalues of $\mathbf{\Sigma}_{F_1^0}\mathbf{\Sigma}_{\Gamma_1^0}$. We now can conclude that $\mathbf{V}_1$ is of rank d_1 in limit.

We are now ready to investigate $\widehat{\lambda}_{1,d}$ for $d \leq d_1$. Because here $d \leq d_1$, we then focus on $\widehat{\mathbf{F}}_1^0$ and $\mathbf{V}_1^0$, which are defined in Appendix A. Let us write (B.8) as follows:

$$T^{-(\nu_1+\delta)/2}\widehat{\mathbf{F}}_1^0\mathbf{V}_1^0 - \mathbf{J}_9^0 = (T^{(\nu_1-\delta)/2}\widehat{\mathbf{F}}_1^0 - \mathbf{F}_1^0\mathbf{H}_1^0)(T^{-\nu_1}\mathbf{V}_1^0), \tag{B.29}$$

where $\mathbf{H}_1^0$ is defined in Appendix A. The above expression implies that (B.8) can be written as

$$T^{(\nu_1-\delta)/2}\widehat{\mathbf{F}}_1^0 - \mathbf{F}_1^0\mathbf{H}_1^0 = \sum_{j=1}^{8}\mathbf{J}_j^0(T^{-\nu_1}\mathbf{V}_1^0)^{-1}, \tag{B.30}$$

where $\mathbf{J}_1^0,\ldots,\mathbf{J}_9^0$ are $\mathbf{J}_1,\ldots,\mathbf{J}_9$ as defined in (B.6), except that now d_1 is taken as known. Note that by the above development

$$\sum_{j=1}^{8}T^{-\nu_1/2}\|\mathbf{J}_j^0\| = T^{-(\nu_1-\delta)/2}\sum_{j=1}^{8}T^{-\delta/2}\|\mathbf{J}_j^0\| = O_P(T^{-(\nu_1-\nu_2)/2}). \tag{B.31}$$

Moreover, since $T^{-\nu_1}\mathbf{V}_1^0$ converges to a full rank matrix by the argument under (B.28), we have $\|(T^{-\nu_1}\mathbf{V}_1^0)^{-1}\| = O_P(1)$, which in turn implies

$$T^{-\nu_1/2}\|T^{(\nu_1-\delta)/2}\widehat{\mathbf{F}}_1^0 - \mathbf{F}_1^0\mathbf{H}_1^0\| \leq \sum_{j=1}^{8}T^{-\nu_1/2}\|\mathbf{J}_j^0\|\|(T^{-\nu_1}\mathbf{V}_1^0)^{-1}\| = O_P(T^{-(\nu_1-\nu_2)/2}). \tag{B.32}$$

This is an important result and in what follows we will use it frequently.

Let us now consider $T^{-\nu_1}(\widehat{\lambda}_{1,d} - \lambda_{1,d})$. By the definitions of $\lambda_{1,d}$ and $\widehat{\lambda}_{1,d}$ given in Section A,

$$\begin{aligned}
T^{-\nu_1}(\widehat{\lambda}_{1,d} - \lambda_{1,d}) &= (T^{-\delta/2}\widehat{\mathbf{F}}_{1,d}^0 - T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0 + T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0)'^{-\nu_1}(\widehat{\mathbf{\Sigma}}_1 - \mathbf{\Sigma}_1^0 + \mathbf{\Sigma}_1^0)\\
&\quad\times(T^{-\delta/2}\widehat{\mathbf{F}}_{1,d}^0 - T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0 + T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0) - T^{-2\nu_1}\mathbf{h}_{1,d}^{0\prime}\mathbf{F}_1^{0\prime}\mathbf{\Sigma}_1^0\mathbf{F}_1^0\mathbf{h}_{1,d}^0\\
&= (T^{-\delta/2}\widehat{\mathbf{F}}_{1,d}^0 - T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0)'^{-\nu_1}(\widehat{\mathbf{\Sigma}}_1 - \mathbf{\Sigma}_1^0)(T^{-\delta/2}\widehat{\mathbf{F}}_{1,d}^0 - T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0)\\
&\quad + 2(T^{-\delta/2}\widehat{\mathbf{F}}_{1,d}^0 - T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0)'^{-\nu_1}(\widehat{\mathbf{\Sigma}}_1 - \mathbf{\Sigma}_1^0)T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0\\
&\quad + (T^{-\delta/2}\widehat{\mathbf{F}}_{1,d}^0 - T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0)'^{-\nu_1}\mathbf{\Sigma}_1^0(T^{-\delta/2}\widehat{\mathbf{F}}_{1,d}^0 - T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0)\\
&\quad + 2(T^{-\delta/2}\widehat{\mathbf{F}}_{1,d}^0 - T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0)'^{-3\nu_1/3}\mathbf{\Sigma}_1^0\mathbf{F}_1^0\mathbf{h}_{1,d}^0\\
&\quad + T^{-2\nu_1}\mathbf{h}_{1,d}^{0\prime}\mathbf{F}_1^{0\prime}(\widehat{\mathbf{\Sigma}}_1 - \mathbf{\Sigma}_1^0)\mathbf{F}_1^0\mathbf{h}_{1,d}^0\\
&= J_1 + 2J_2 + J_3 + 2J_4 + J_5,
\end{aligned} \tag{B.33}$$

with obvious definitions of $J_1, \ldots, J_5$. From (B.32),

$$T^{-\nu_1/2}\|T^{(\nu_1-\delta)/2}\widehat{\mathbf{F}}_{1,d}^0 - \mathbf{F}_1^0\mathbf{h}_{1,d}^0\| = \|T^{-\delta/2}\widehat{\mathbf{F}}_{1,d}^0 - T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0\| = O_P(T^{-(\nu_1-\nu_2)/2}), \tag{B.34}$$

which is $o_P(1)$ under Assumption 1. This implies $|J_1| = o_P(|J_5|)$, $|J_2| = o_P(|J_5|)$ and $|J_3| = o_P(|J_4|)$. It remains to consider J_4 and J_5. The order of the first of these terms is given by

$$\begin{aligned}
|J_4| &\leq \|T^{-\delta/2}\widehat{\mathbf{F}}_{1,d}^0 - T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0\|T^{-3\nu_1/2}\|\boldsymbol{\Sigma}_1^0\mathbf{F}_1^0\mathbf{h}_{1,d}^0\| \\
&\leq \|T^{-\delta/2}\widehat{\mathbf{F}}_{1,d}^0 - T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{h}_{1,d}^0\|T^{-\nu_1/2}\|\mathbf{F}_1^0\|\|(N^{-1}\boldsymbol{\Gamma}_1^{0\prime}\boldsymbol{\Gamma}_1^0)\|T^{-\nu_1}\|\mathbf{F}_1^{0\prime}\mathbf{F}_1^0\|\|\mathbf{h}_{1,d}^0\| \\
&= O_P(T^{-(\nu_1-\nu_2)/2}),
\end{aligned} \tag{B.35}$$

where we have made use of the fact that $\|\mathbf{H}_g^0\| = O_P(1)$ (see Appendix A), which implies that $\|\mathbf{h}_{1,d}^0\|$ is of the same order.

The order of J_5 is the same as that of J_4. In order to appreciate this, we begin by noting

$$\begin{aligned}
T^{-\nu_1}(\widehat{\boldsymbol{\Sigma}}_1 - \boldsymbol{\Sigma}_1^0) &= \frac{1}{NT^{\nu_1}}\sum_{i=1}^N \mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i' \\
&+ \frac{1}{NT^{\nu_1}}\sum_{i=1}^N \mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\boldsymbol{\gamma}_{1,i}^{0\prime}\mathbf{F}_1^{0\prime} + \frac{1}{NT^{\nu_1}}\sum_{i=1}^N \mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i' \\
&+ \frac{1}{NT^{\nu_1}}\sum_{i=1}^N \mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0 + \boldsymbol{\varepsilon}_i)' \\
&+ \frac{1}{NT^{\nu_1}}\sum_{i=1}^N (\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0 + \boldsymbol{\varepsilon}_i)(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i' \\
&+ \frac{1}{NT^{\nu_1}}\sum_{i=1}^N (\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0 + \boldsymbol{\varepsilon}_i)(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0 + \boldsymbol{\varepsilon}_i)' \\
&+ \frac{1}{NT^{\nu_1}}\sum_{i=1}^N \mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0 + \boldsymbol{\varepsilon}_i)' + \frac{1}{NT^{\nu_1}}\sum_{i=1}^N (\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0 + \boldsymbol{\varepsilon}_i)\boldsymbol{\gamma}_{1,i}^{0\prime}\mathbf{F}_1^{0\prime}.
\end{aligned} \tag{B.36}$$

By the proof for each term of (B.6), it is easy to know that

$$T^{-\nu_1}\|\widehat{\boldsymbol{\Sigma}}_1 - \boldsymbol{\Sigma}_1^0\| = O_P(T^{-(\nu_1-\nu_2)/2}), \tag{B.37}$$

and so

$$|J_5| \leq \|\mathbf{h}_{1,d}^0\|^2 T^{-\nu_1}\|\mathbf{F}_1^0\|^2 T^{-\nu_1}\|\widehat{\boldsymbol{\Sigma}}_1 - \boldsymbol{\Sigma}_1^0\| = O_P(T^{-(\nu_1-\nu_2)/2}). \tag{B.38}$$

Hence, by putting everything together,

$$T^{-\nu_1}|\widehat{\lambda}_{1,d} - \lambda_{1,d}| \leq |J_1| + 2|J_2| + |J_3| + 2|J_4| + |J_5| = O_P(T^{-(\nu_1-\nu_2)/2}), \tag{B.39}$$

which establishes (a).

Consider (b). This proof is based on Lemma B.1. We therefore start by introducing some notation in order to make the problem here fit the one in Lemma B.1. Let us therefore denote by $\mathbf{F}_1^\perp$ a $T \times (d_{max} - d_1)$ matrix such that $T^{-\nu_1}(\mathbf{F}_1^\perp, \mathbf{F}_1^0\mathbf{R})'(\mathbf{F}_1^\perp, \mathbf{F}_1^0\mathbf{R}) = \text{diag}(\mathbf{I}_{d_{max}-d_1}, \mathbf{I}_{d_1})$, where $\mathbf{R}$ is a $d_1 \times d_1$ rotation matrix. The matrices $T^{-\nu_1/2}\mathbf{F}_1^\perp$, $T^{-\nu_1/2}\mathbf{F}_1^0\mathbf{R}$, $\mathbf{\Sigma}_1^0$ and $\widehat{\mathbf{\Sigma}}_1 - \mathbf{\Sigma}_1^0$ correspond to $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{A}$ and $\mathbf{E}$ of Lemma B.1. Our counterpart of the matrix $\mathbf{Q}_1^0$ appearing in this other lemma is thus given by

$$\widehat{\mathbf{F}}^\perp = T^{-\nu_1/2}(\mathbf{F}_1^\perp + \mathbf{F}_1^0\mathbf{R}\mathbf{P})(\mathbf{I}_{d_{max}-d_1} + \mathbf{P}'\mathbf{P})^{-1/2}, \tag{B.40}$$

where

$$\begin{aligned}\|\mathbf{P}\|_2 &\leq \frac{4}{\text{sep}(0, T^{-2\nu_1}\mathbf{F}_1^{0\prime}\mathbf{\Sigma}_1^0\mathbf{F}_1^0)} T^{-\nu_1}\|\widehat{\mathbf{\Sigma}}_1 - \mathbf{\Sigma}_1^0\| \leq O_P(1)T^{-\nu_1}\|\widehat{\mathbf{\Sigma}}_1 - \mathbf{\Sigma}_1^0\| \\ &= O_P(T^{-(\nu_1-\nu_2)/2}).\end{aligned} \tag{B.41}$$

Since $\widehat{\mathbf{F}}^\perp$ is an orthonormal basis for a subspace that is invariant for $\widehat{\mathbf{\Sigma}}_1$, we have $\widehat{\lambda}_{1,d_1+d} = \widehat{\mathbf{F}}_d^{\perp\prime}\widehat{\mathbf{\Sigma}}_1\widehat{\mathbf{F}}_d^\perp$, where $d = 1, \ldots, d_{max} - d_1$ and $\widehat{\mathbf{F}}_d^\perp$ is the d-th column of $\widehat{\mathbf{F}}^\perp$. Consider $\|\widehat{\mathbf{F}}^\perp - T^{-\nu_1/2}\mathbf{F}_1^\perp\|_2$. By the definition of $\widehat{\mathbf{F}}^\perp$,

$$\begin{aligned}&\|\widehat{\mathbf{F}}^\perp - T^{-\nu_1/2}\mathbf{F}_1^\perp\|_2 \\ &= T^{-\nu_1/2}\|[\mathbf{F}_1^\perp + \mathbf{F}_1^0\mathbf{R}\mathbf{P} - \mathbf{F}_1^\perp(\mathbf{I}_{d_{max}-d_1} + \mathbf{P}'\mathbf{P})^{1/2}](\mathbf{I}_{d_{max}-d_1} + \mathbf{P}'\mathbf{P})^{-1/2}\|_2 \\ &\leq T^{-\nu_1/2}\|\mathbf{F}_1^\perp(\mathbf{I}_{d_{max}-d_1} - (\mathbf{I}_{d_{max}-d_1} + \mathbf{P}'\mathbf{P})^{1/2})(\mathbf{I}_{d_{max}-d_1} + \mathbf{P}'\mathbf{P})^{-1/2}\|_2 \\ &+ T^{-\nu_1/2}\|\mathbf{F}_1^0\mathbf{R}\mathbf{P}(\mathbf{I}_{d_{max}-d_1} + \mathbf{P}'\mathbf{P})^{-1/2}\|_2 \\ &\leq \|(\mathbf{I}_{d_{max}-d_1} - (\mathbf{I}_{d_{max}-d_1} + \mathbf{P}'\mathbf{P})^{1/2})(\mathbf{I}_{d_{max}-d_1} + \mathbf{P}'\mathbf{P})^{-1/2}\|_2 + \|\mathbf{P}(\mathbf{I}_{d_{max}-d_1} + \mathbf{P}'\mathbf{P})^{-1/2}\|_2 \\ &\leq \|\mathbf{I}_{d_{max}-d_1} - (\mathbf{I}_{d_{max}-d_1} + \mathbf{P}'\mathbf{P})^{1/2}\|_2 + \|\mathbf{P}\|_2 \leq 2\|\mathbf{P}\|_2 \\ &= O_P(T^{-(\nu_1-\nu_2)/2}),\end{aligned} \tag{B.42}$$

where the second and third inequalities follow from (Magnus and Neudecker, 2007, Exercise 1 on page 231). This last result can be used to show that

$$\begin{aligned}&T^{-\nu_1}|\widehat{\lambda}_{1,d_1+d}| = |\widehat{\mathbf{F}}_d^{\perp\prime}(T^{-\nu_1}\widehat{\mathbf{\Sigma}}_1)\widehat{\mathbf{F}}_d^\perp| \\ &= |(\widehat{\mathbf{F}}_d^\perp - T^{-\nu_1/2}\mathbf{F}_{1,d}^\perp + T^{-\nu_1/2}\mathbf{F}_{1,d}^\perp)'^{-\nu_1}(\widehat{\mathbf{\Sigma}}_1 - \mathbf{\Sigma}_1^0 + \mathbf{\Sigma}_1^0)(\widehat{\mathbf{F}}_d^\perp - T^{-\nu_1/2}\mathbf{F}_{1,d}^\perp + T^{-\nu_1/2}\mathbf{F}_{1,d}^\perp)| \\ &\leq \|\widehat{\mathbf{F}}_d^\perp - T^{-\nu_1/2}\mathbf{F}_{1,d}^\perp\|^2 T^{-\nu_1}\|\widehat{\mathbf{\Sigma}}_1 - \mathbf{\Sigma}_1^0\| \\ &+ 2\|\widehat{\mathbf{F}}_d^\perp - T^{-\nu_1/2}\mathbf{F}_{1,d}^\perp\| T^{-\nu_1}\|\widehat{\mathbf{\Sigma}}_1 - \mathbf{\Sigma}_1^0\| T^{-\nu_1/2}\|\mathbf{F}_{1,d}^\perp\| + \|\widehat{\mathbf{F}}_d^\perp - T^{-\nu_1/2}\mathbf{F}_{1,d}^\perp\|^2 T^{-\nu_1}\|\mathbf{\Sigma}_1^0\| \\ &= O_P(T^{-(\nu_1-\nu_2)}),\end{aligned} \tag{B.43}$$

where $\mathbf{F}_{1,d}^{\perp}$ is the d-th column of $\mathbf{F}_1^{\perp}$, and the last equality follows from (B.42) and the proof of part (a). This completes the proof of the lemma. ■

Proof of Lemma B.4.

For (a), we take the same starting point as in the proof of part (a) in Lemma B.3, which is (B.8) with $\widehat{\mathbf{F}}_1$ and $\mathbf{V}_1$ based on treating d_1 as known. The rationale for doing so is, as already explained in Appendix A, that $\widehat{d}_1$ is consistent. Pre-multiplying this equation through by $T^{-(\nu_1+\delta)/2}\mathbf{F}_2^{0\prime}$ gives

$$
\begin{aligned}
&T^{-(\nu_1+\delta)}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_1\mathbf{V}_1 - T^{-(\nu_1+\delta)/2}\mathbf{F}_2^{0\prime}\mathbf{J}_9 \\
&= T^{-(\nu_1+\delta)}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_1\mathbf{V}_1 - T^{-(\nu_1+\delta)/2}\mathbf{F}_2^{0\prime}\mathbf{F}_1^0(N^{-1}\mathbf{\Gamma}_1^{0\prime}\mathbf{\Gamma}_1^0)(T^{-(\nu_1+\delta)/2}\mathbf{F}_1^{0\prime}\widehat{\mathbf{F}}_1) \\
&= \sum_{j=1}^{8} T^{-(\nu_1+\delta)/2}\mathbf{F}_2^{0\prime}\mathbf{J}_j,
\end{aligned}
\tag{B.44}
$$

or

$$
\begin{aligned}
&T^{-\delta}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_1 - T^{-(\nu_1+\delta)/2}\mathbf{F}_2^{0\prime}\mathbf{J}_9(T^{-\nu_1}\mathbf{V}_1)^{-1} \\
&= T^{-\delta}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_1 - T^{-(\nu_1+\delta)/2}\mathbf{F}_2^{0\prime}\mathbf{F}_1^0(N^{-1}\mathbf{\Gamma}_1^{0\prime}\mathbf{\Gamma}_1^0)(T^{-(\nu_1+\delta)/2}\mathbf{F}_1^{0\prime}\widehat{\mathbf{F}}_1)(T^{-\nu_1}\mathbf{V}_1)^{-1} \\
&= T^{-(\nu_1-\nu_2)/2}\sum_{j=1}^{8} T^{-(\delta+\nu_2)/2}\mathbf{F}_2^{0\prime}\mathbf{J}_j(T^{-\nu_1}\mathbf{V}_1)^{-1}.
\end{aligned}
\tag{B.45}
$$

Under Assumption 4, the orders of $\mathbf{J}_1,\ldots,\mathbf{J}_6$ are the same as in those in the proof of the first result of Lemma B.3. The stated orders of $\mathbf{J}_7$ and $\mathbf{J}_8$ are, however, not sharp and can be improved upon. The order of $T^{-(\delta+\nu_2)/2}\|\mathbf{F}_2^{0\prime}\mathbf{J}_7\|$ is given by

$$
\begin{aligned}
T^{-(\delta+\nu_2)/2}\|\mathbf{F}_2^{0\prime}\mathbf{J}_7\| &\leq O_P(1)T^{-\nu_2/2}\left\|\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\mathbf{F}_2^{0\prime}\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0(\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0+\boldsymbol{\varepsilon}_i)'\right\| \\
&\leq O_P(1)\frac{1}{NT^{(\nu_2+\nu_1+\delta)/2}}\|\mathbf{F}_2^{0\prime}\mathbf{F}_1^0\mathbf{\Gamma}_1^{0\prime}\mathbf{\Gamma}_{+1}^0\mathbf{F}_{+1}^{0\prime}\| \\
&\quad + O_P(1)\frac{1}{NT^{(\nu_2+\nu_1+\delta)/2}}\|\mathbf{F}_2^{0\prime}\mathbf{F}_1^0\mathbf{\Gamma}_1^{0\prime}\boldsymbol{\varepsilon}\| \\
&= O_P(N^{-(1-p)}T^{q-(\nu_1+\delta)/2}) + O_P(N^{-1/2}T^{q-(\nu_2+\nu_1+\delta-1)/2}),
\end{aligned}
\tag{B.46}
$$

where the last equality follows from Assumption 4 and Lemma B.2. We can similarly show that

$$
\begin{aligned}
T^{-(\delta+\nu_2)/2}\|\mathbf{F}_2^{0\prime}\mathbf{J}_8\| &\leq O_P(1)T^{-\nu_2/2}\left\|\frac{1}{NT^{(\nu_1+\delta)/2}}\sum_{i=1}^{N}\mathbf{F}_2^{0\prime}(\mathbf{F}_2^0\boldsymbol{\gamma}_{2,i}^0+\mathbf{F}_{+2}^0\boldsymbol{\gamma}_{+2,i}^0+\boldsymbol{\varepsilon}_i)\boldsymbol{\gamma}_{1,i}^{0\prime}\mathbf{F}_1^{0\prime}\right\| \\
&\leq O_P(1)\frac{1}{NT^{(\nu_2+\nu_1+\delta)/2}}\|\mathbf{F}_2^{0\prime}\mathbf{F}_2^0\mathbf{\Gamma}_2^{0\prime}\mathbf{\Gamma}_1^0\mathbf{F}_1^{0\prime}\| \\
&\quad + O_P(1)N^{-1}T^{-(\nu_2+\nu_1+\delta)/2}\|\mathbf{F}_2^{0\prime}\mathbf{F}_{+2}^0\mathbf{\Gamma}_{+2}^{0\prime}\mathbf{\Gamma}_1^0\mathbf{F}_1^{0\prime}\| \\
&\quad + O_P(1)N^{-1}T^{-(\nu_2+\nu_1+\delta)/2}\|\mathbf{F}_2^{0\prime}\boldsymbol{\varepsilon}'\mathbf{\Gamma}_1^0\mathbf{F}_1^{0\prime}\|
\end{aligned}
$$

$$= O_P(N^{-(1-p)}T^{-(\delta-\nu_2)/2}) + O_P(N^{-1/2}T^{-(\delta-1)/2}), \tag{B.47}$$

where $\mathbf{F}^0_{+2}$ and $\mathbf{\Gamma}^0_{+2}$ are defined analogously to $\mathbf{F}^0_{+1}$ and $\mathbf{\Gamma}^0_{+1}$ in the proof of Lemma B.3. By using these last two results together with the orders of $\mathbf{J}_1, \dots, \mathbf{J}_6$ given in the proof of the first result of Lemma B.3,

$$\begin{aligned}
&\left\| \sum_{j=1}^{8} T^{-(\delta+\nu_2)/2} \mathbf{F}_2^{0\prime} \mathbf{J}_j \right\| \\
&\leq O_P(T^{-(\delta-1)/2} \|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|) + O_P(T^{\nu_2-(\nu_1+\delta)/2}) + O_P(N^{-1/2}T^{1-(\nu_1+\delta)/2}) \\
&+ O_P(N^{-(1-p)}T^{-(\delta-\nu_2)/2}) + O_P(N^{-1/2}T^{-(\delta-1)/2}) \\
&= O_P(T^{-\delta/2} + N^{-1/2}T^{-(\delta-1)/2} + N^{-(1-p)}T^{-(\delta-\nu_2)/2}),
\end{aligned} \tag{B.48}$$

where we have used $\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\| = O_P(N^{-1/2} \vee T^{-1/2})$ of Lemma 1, and Assumptions 3 and 4. Hence,

$$\begin{aligned}
&\|T^{-\delta}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_1 - T^{-(\nu_1+\delta)/2}\mathbf{F}_2^{0\prime}\mathbf{J}_9(T^{-\nu_1}\mathbf{V}_1)^{-1}\| \\
&\leq T^{-(\nu_1-\nu_2)/2} \left\| \sum_{j=1}^{8} T^{-(\delta+\nu_2)/2} \mathbf{F}_2^{0\prime} \mathbf{J}_j \right\| \|(T^{-\nu_1}\mathbf{V}_1)^{-1}\| \\
&= T^{-(\nu_1-\nu_2)/2}[O_P(T^{-\delta/2} + N^{-1/2}T^{-(\delta-1)/2} + N^{-(1-p)}T^{-(\delta-\nu_2)/2})] \\
&= O_P(T^{-(\delta+\nu_1-\nu_2)/2} + N^{-1/2}T^{-(\delta+\nu_1-\nu_2-1)/2} + N^{-(1-p)}T^{-(\delta+\nu_1-2\nu_2)/2}),
\end{aligned} \tag{B.49}$$

which together with Assumption 4 yields

$$\begin{aligned}
\|T^{-\delta}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_1\| &\leq \|T^{-(\nu_1+\delta)/2}\mathbf{F}_2^{0\prime}\mathbf{J}_9\| \|(T^{-\nu_1}\mathbf{V}_1)^{-1}\| \\
&+ T^{-(\nu_1-\nu_2)/2} \left\| \sum_{j=1}^{8} T^{-(\delta+\nu_2)/2} \mathbf{F}_2^{0\prime} \mathbf{J}_j \right\| \|(T^{-\nu_1}\mathbf{V}_1)^{-1}\| \\
&= O_P(T^{q-(\nu_1+\delta)/2} + T^{-(\delta+\nu_1-\nu_2)/2} + N^{-1/2}T^{-(\delta+\nu_1-\nu_2-1)/2} \\
&+ N^{-(1-p)}T^{-(\delta+\nu_1-2\nu_2)/2}),
\end{aligned} \tag{B.50}$$

as was to be shown for (a).

Let us now consider (b). Analogously to the proof of (a), by invoking Assumption 4 we can improve the orders of $\mathbf{J}_7$ and $\mathbf{J}_8$. For $\mathbf{J}_7$,

$$\begin{aligned}
T^{-\delta/2}\|\mathbf{J}_7\| &\leq O_P(1)N^{-1}T^{-(\nu_1+\delta)/2}\|\mathbf{F}_1^0\mathbf{\Gamma}_1^{0\prime}\mathbf{\Gamma}_{+1}^0\mathbf{F}_{+1}^{0\prime}\| + O_P(1)N^{-1}T^{-(\nu_1+\delta)/2}\|\mathbf{F}_1^0\mathbf{\Gamma}_1^{0\prime}\boldsymbol{\varepsilon}\| \\
&= O_P(N^{-(1-p)}T^{-(\delta-\nu_2)/2}) + O_P(N^{-1/2}T^{-(\delta-1)/2}),
\end{aligned} \tag{B.51}$$

where the equality follows from Assumption 4 and Lemma B.2. For $\mathbf{J}_8$,

$$T^{-\delta/2}\|\mathbf{J}_8\| \le O_P(N^{-(1-p)}T^{-(\delta-\nu_2)/2}) + O_P(N^{-1/2}T^{-(\delta-1)/2}). \tag{B.52}$$

This implies that the result in (B.32) changes to (after replacing $T^{-\nu_1/2}$ by $T^{-\delta/2}$)

$$\begin{aligned} T^{-\delta/2}\|T^{(\nu_1-\delta)/2}\widehat{\mathbf{F}}_1 - \mathbf{F}_1^0\mathbf{H}_1^0\| &\le \sum_{j=1}^{8} T^{-\delta/2}\|\mathbf{J}_j\|\|(T^{-\nu_1}\mathbf{V}_1)^{-1}\| \\ &= O_P(T^{-\delta/2} + N^{-1/2}T^{-(\delta-1)/2} + N^{-(1-p)}T^{-(\delta-\nu_2)/2}). \end{aligned} \tag{B.53}$$

Hence, since $\mathbf{H}_1^0$ is invertible with $\|\mathbf{H}_1^0\| = O_P(1)$ and $\mathbf{H}_1 = T^{-(\nu_1-\delta)/2}\mathbf{H}_1^0$,

$$T^{-\delta/2}\|\widehat{\mathbf{F}}_1\mathbf{H}_1^{-1} - \mathbf{F}_1^0\| = O_P(T^{-\delta/2} + N^{-1/2}T^{-(\delta-1)/2} + N^{-(1-p)}T^{-(\delta-\nu_2)/2}). \tag{B.54}$$

We are now ready to consider $\sum_{i=1}^N \|\mathbf{F}_1^{0\prime}\boldsymbol{\gamma}_{1,i}^0 - \widehat{\mathbf{F}}_1\widehat{\boldsymbol{\gamma}}_{1,i}\|^2$.

$$\begin{aligned} &\sum_{i=1}^N \|\mathbf{F}_1^{0\prime}\boldsymbol{\gamma}_{1,i}^0 - \widehat{\mathbf{F}}_1\widehat{\boldsymbol{\gamma}}_{1,i}\|^2 = \sum_{i=1}^N \|\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0 - T^{-\delta}\widehat{\mathbf{F}}_1\widehat{\mathbf{F}}_1'(\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}_0)\|^2 \\ &\le O(1)\sum_{i=1}^N [\|\mathbf{P}_{\widehat{F}_1}\mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|^2 + \|\mathbf{M}_{\widehat{F}_1}\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0\|^2 + \|\mathbf{P}_{\widehat{F}_1}\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0\|^2 + \|\mathbf{P}_{\widehat{F}_1}\boldsymbol{\varepsilon}_i\|^2]. \end{aligned} \tag{B.55}$$

We now evaluate each of the terms on the right-hand side one by one. Making use of Assumption 1 and Lemma 1, we get

$$\begin{aligned} \sum_{i=1}^N \|\mathbf{P}_{\widehat{F}_1}\mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|^2 &\le \sum_{i=1}^N \|\mathbf{X}_i\mathbf{D}_T\|^2\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|^2 \\ &= O_P(NT)O_P(N^{-1} \vee T^{-1}) = O_P(N \vee T), \end{aligned} \tag{B.56}$$

and by another application of Lemma B.2,

$$\sum_{i=1}^N \|\mathbf{P}_{\widehat{F}_1}\boldsymbol{\varepsilon}_i\|^2 \le \sum_{i=1}^N \|\boldsymbol{\varepsilon}_i\|^2 = O_P(N \vee T). \tag{B.57}$$

For $\sum_{i=1}^N \|\mathbf{M}_{\widehat{F}_1}\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0\|^2$, we use (B.54) from which it follows that

$$\begin{aligned} \|\mathbf{M}_{\widehat{F}_1}\mathbf{F}_1^0\|^2 &= \|\mathbf{M}_{\widehat{F}_1}(\mathbf{F}_1^0 - \widehat{\mathbf{F}}_1\mathbf{H}_1^{-1})\|^2 \le T^{\delta}(T^{-\delta}\|\mathbf{F}_1^0 - \widehat{\mathbf{F}}_1\mathbf{H}_1^{-1}\|^2) \\ &= O_P(1) + O_P(N^{-1}T) + O_P(N^{-2(1-p)}T^{\nu_2}), \end{aligned} \tag{B.58}$$

which in turn implies

$$\sum_{i=1}^{N} \|\mathbf{M}_{\widehat{F}_1}\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0\|^2 = O_P(N) + O_P(T) + O_P(N^{-(1-2p)}T^{\nu_2}). \tag{B.59}$$

For $\sum_{i=1}^{N} \|\mathbf{P}_{\widehat{F}_1}\mathbf{F}_{+1}^0\boldsymbol{\gamma}_{+1,i}^0\|^2$, we use the result given in part (a), giving

$$\begin{aligned}\sum_{i=1}^{N} \|\mathbf{P}_{\widehat{F}_1}\mathbf{F}_2^0\boldsymbol{\gamma}_{2,i}^0\|^2 &= \sum_{i=1}^{N} \|T^{-\delta}\widehat{\mathbf{F}}_1\widehat{\mathbf{F}}_1'\mathbf{F}_2^0\boldsymbol{\gamma}_{2,i}^0\|^2 = O_P(NT^{\delta})\|T^{-\delta}\widehat{\mathbf{F}}_1'\mathbf{F}_2^0\|^2 \\ &= O_P(NT^{\delta})[O_P(T^{q-(\nu_1+\delta)/2}) + O_P(T^{-(\delta+\nu_1-\nu_2)/2}) \\ &+ O_P(N^{-1/2}T^{-(\delta+\nu_1-\nu_2-1)/2}) + O_P(N^{-(1-p)}T^{-(\delta+\nu_1-2\nu_2)/2})]^2 \\ &= O_P(N)[O_P(T^{-(\nu_1-\nu_2)/2}) + O_P(N^{-1/2}T^{-(\nu_1-\nu_2-1)/2}) \\ &+ O_P(N^{-(1-p)}T^{-(\nu_1-2\nu_2)/2}) + O_P(T^{q-\nu_1/2})]^2. \end{aligned} \tag{B.60}$$

The order of $\sum_{i=1}^{N} \|\mathbf{P}_{\widehat{F}_1}\mathbf{F}_{+2}^0\boldsymbol{\gamma}_{+2,i}^0\|^2$ is the same. Hence, by adding the above results, (b) follows after simple algebra. The proof is now complete. ■

Proof of Lemma B.5.

Let $\mathbf{U}_i = \mathbf{F}_{+2}^0\boldsymbol{\gamma}_{+2,i}^0 + \mathbf{F}_1^{0\prime}\boldsymbol{\gamma}_{1,i}^0 - \widehat{\mathbf{F}}_1\widehat{\boldsymbol{\gamma}}_{1,i}$. In this notation,

$$\begin{aligned}&T^{-(\nu_2+\delta)/2}\widehat{\mathbf{F}}_2\mathbf{V}_2 \\ &= \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}[(\mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0) + \mathbf{F}_2^0\boldsymbol{\gamma}_{2,i}^0 + \mathbf{U}_i + \boldsymbol{\varepsilon}_i) \\ &\times (\mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0) + \mathbf{F}_2^0\boldsymbol{\gamma}_{2,i}^0 + \mathbf{U}_i + \boldsymbol{\varepsilon}_i)']\widehat{\mathbf{F}}_2 \\ &= \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\widehat{\mathbf{F}}_2 \\ &+ \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\boldsymbol{\gamma}_{2,i}^{0\prime}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_2 + \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{F}_2^0\boldsymbol{\gamma}_{2,i}^0(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\widehat{\mathbf{F}}_2 \\ &+ \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\mathbf{U}_i'\widehat{\mathbf{F}}_2 + \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{U}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\widehat{\mathbf{F}}_2 \\ &+ \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\boldsymbol{\varepsilon}_i'\widehat{\mathbf{F}}_2 + \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\widehat{\mathbf{F}}_2 \\ &+ \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\widehat{\mathbf{F}}_2 + \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{F}_2^0\boldsymbol{\gamma}_{2,i}^0\boldsymbol{\varepsilon}_i'\widehat{\mathbf{F}}_2 + \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i\boldsymbol{\gamma}_{2,i}^{0\prime}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_2 \\ &+ \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i\mathbf{U}_i'\widehat{\mathbf{F}}_2 + \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{U}_i\boldsymbol{\varepsilon}_i'\widehat{\mathbf{F}}_2 + \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{U}_i\boldsymbol{\gamma}_{2,i}^{0\prime}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_2 \\ &+ \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{F}_2^0\boldsymbol{\gamma}_{2,i}^0\mathbf{U}_i'\widehat{\mathbf{F}}_2 + \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{U}_i\mathbf{U}_i'\widehat{\mathbf{F}}_2 + \frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{F}_2^0\boldsymbol{\gamma}_{2,i}^0\boldsymbol{\gamma}_{2,i}^{0\prime}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_2\end{aligned}$$

$$= \sum_{j=1}^{16} \mathbf{K}_j, \tag{B.61}$$

where $\mathbf{K}_1, \ldots, \mathbf{K}_{16}$ are implicitly defined. Analogously to the proof of the first result of Lemma B.3 we move $\mathbf{K}_{16} = \mathbf{F}_2^0(N^{-1}\mathbf{\Gamma}_2^{0\prime}\mathbf{\Gamma}_2^0)(T^{-(\nu_2+\delta)/2}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_2)$ over to the left, giving

$$T^{-(\nu_2+\delta)/2}\widehat{\mathbf{F}}_2\mathbf{V}_2 - \mathbf{F}_2^0(N^{-1}\mathbf{\Gamma}_2^{0\prime}\mathbf{\Gamma}_2^0)(T^{-(\nu_2+\delta)/2}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_2) = \sum_{j=1}^{15} \mathbf{K}_j. \tag{B.62}$$

By using the same steps employed in the proof of the first result of Lemma B.3, we can show that

$$T^{-\delta/2}\|\mathbf{K}_1 + \mathbf{K}_2 + \mathbf{K}_3\| = O_P(T^{-(\delta-1)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|). \tag{B.63}$$

For $\mathbf{K}_4$,

$$\begin{aligned}
T^{-\delta/2}\|\mathbf{K}_4\| &= T^{-\delta/2}\left\| \frac{1}{NT^{(\nu_2+\delta)/2}} \sum_{i=1}^N \mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\mathbf{U}_i'\widehat{\mathbf{F}}_2 \right\| \\
&\le O(1)\left(\frac{1}{NT^\delta} \sum_{i=1}^N \|\mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|^2 \right)^{1/2} \\
&\times \left(\frac{1}{NT^{\nu_2}} \sum_{i=1}^N \|\mathbf{F}_{+2}^0\boldsymbol{\gamma}_{2,i}^0\|^2 + \frac{1}{NT^{\nu_2}} \sum_{i=1}^N \|\mathbf{F}_1^{0\prime}\boldsymbol{\gamma}_{1,i}^0 - \widehat{\mathbf{F}}_1\widehat{\boldsymbol{\gamma}}_{1,i}\|^2 \right)^{1/2} \\
&= O_P(1)T^{-(\delta-1)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|[O_P(T^{-(\nu_2-\nu_3)}) + O_P(T^{-\nu_2}) + O_P(N^{-1}T^{-(\nu_2-1)}) \\
&+ O_P(T^{-(\nu_1-\nu_2)}) + O_P(T^{-(\nu_2+\nu_1-2q)}) + O_P(N^{-1}T^{-(\nu_1+\nu_2-2)}) + O_P(N^{-2(1-p)})] \\
&= o_P(T^{-(\delta-1)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|),
\end{aligned} \tag{B.64}$$

where the second equality follows Lemma B.4, and the third follows from Assumptions 1, 3, and 4. The same arguments can be used to show that

$$T^{-\delta/2}\|\mathbf{K}_5\| = o_P(T^{-(\delta-1)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|). \tag{B.65}$$

For $\mathbf{K}_6$,

$$\begin{aligned}
T^{-\delta/2}\|\mathbf{K}_6\| &\le T^{-\delta/2}\|\widehat{\mathbf{F}}_2\| \left\| \frac{1}{NT^{(\nu_2+\delta)/2}} \sum_{i=1}^N \mathbf{X}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\boldsymbol{\varepsilon}_i' \right\| \\
&= O_P(N^{-1/2}T^{-(\delta+\nu_2-2)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|) \\
&= o_P(T^{-(\delta-1)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|),
\end{aligned} \tag{B.66}$$

where the development is similar to (B.14). The order of $T^{-\delta/2}\|\mathbf{K}_7\|$ is the same.

For $\mathbf{K}_8$,

$$T^{-\delta/2}\|\mathbf{K}_8\| = T^{-\delta/2}\left\|\frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\widehat{\mathbf{F}}_2\right\| \leq T^{-\delta/2}\|\widehat{\mathbf{F}}_2\|N^{-1}T^{-(\nu_2+\delta)/2}\|\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\|$$
$$= O_P(T^{1-(\nu_2+\delta)/2}(N^{-1/2}\vee T^{-1/2})), \quad \text{(B.67)}$$

where the last equality holds by Lemma B.2.

Further use of Lemma B.2 gives

$$T^{-\delta/2}\|\mathbf{K}_9\| = T^{-\delta/2}\left\|\frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\mathbf{F}_2^0\boldsymbol{\gamma}_{2,i}^0\boldsymbol{\varepsilon}_i'\widehat{\mathbf{F}}_2\right\| \leq N^{-1}T^{-(\nu_2+\delta)/2}\|\mathbf{F}_2^0\boldsymbol{\Gamma}_2^{0\prime}\boldsymbol{\varepsilon}\|T^{-\delta/2}\|\widehat{\mathbf{F}}_2\|$$
$$= O_P(N^{-1/2}T^{-(\delta-1)/2}), \quad \text{(B.68)}$$

and we can show that $T^{-\delta/2}\|\mathbf{K}_{10}\|$ is of the same order.

$\mathbf{K}_{11}$ requires more work. We begin by expanding it in the following way:

$$T^{-\delta/2}\|\mathbf{K}_{11}\| = T^{-\delta/2}\left\|\frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i(\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0 - \widehat{\mathbf{F}}_1\widehat{\boldsymbol{\gamma}}_{1,i} + \mathbf{F}_{+2}^0\boldsymbol{\gamma}_{2,i}^0)'\widehat{\mathbf{F}}_2\right\|$$
$$\leq \left\|\frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i(\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0 - \widehat{\mathbf{F}}_1T^{-\delta}\widehat{\mathbf{F}}_1'(\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}_0) + \mathbf{F}_{+2}^0\boldsymbol{\gamma}_{2,i}^0)'\right\|T^{-\delta/2}\|\widehat{\mathbf{F}}_2\|$$
$$\leq O(1)\left\|\frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\mathbf{P}_{\widehat{F}_1}\right\|$$
$$+ O(1)\left\|\frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i\boldsymbol{\gamma}_{1,i}^{0\prime}\mathbf{F}_1^{0\prime}\mathbf{M}_{\widehat{F}_1}\right\| + O(1)\left\|\frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i\boldsymbol{\gamma}_{+1,i}^{0\prime}\mathbf{F}_{+1}^{0\prime}\mathbf{P}_{\widehat{F}_1}\right\|$$
$$+ O(1)\left\|\frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{P}_{\widehat{F}_1}\right\| + O(1)\left\|\frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i\boldsymbol{\gamma}_{+2,i}^{0\prime}\mathbf{F}_{+2}^{0\prime}\right\|$$
$$= O(1)(K_{111} + K_{112} + K_{113} + K_{114} + K_{115}), \quad \text{(B.69)}$$

where, similarly to the analysis of $\mathbf{K}_6$ and using $\|\mathbf{P}_{\widehat{F}_1}\|_2 = 1$,

$$K_{111} \leq \left\|\frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\right\| = O_P(N^{-1/2}T^{-(\delta+\nu_2-2)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|)$$
$$= o_P(T^{-(\delta-1)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|). \quad \text{(B.70)}$$

Also, making use of (B.58), we can show that

$$K_{112} \leq N^{-1}T^{-(\nu_2+\delta)/2}\|\boldsymbol{\Gamma}_1^{0\prime}\boldsymbol{\varepsilon}\|\|\mathbf{M}_{\widehat{F}_1}\mathbf{F}_1^0\|$$
$$= O_P(N^{-1/2}T^{-(\nu_2+\delta-1)/2}) + O_P(N^{-1}T^{1-(\nu_2+\delta)/2}) + O_P(N^{-(3/2-p)}T^{-(\delta-1)/2}). \quad \text{(B.71)}$$

Similarly, for K_{113}, we can show that

$$K_{113} = o_P(K_{112}) + O_P(N^{-1/2}T^{q-(\nu_1+\nu_2+\delta-1)/2}), \tag{B.72}$$

where we have used Lemmas B.2 and B.4, and Assumption 1. Further use of Lemma B.2 shows that K_{114} is of the following order:

$$K_{114} \leq N^{-1}T^{-(\nu_2+\delta)/2}\|\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\| = O_P(T^{-(\nu_2+\delta-2)/2}(N^{-1/2} \vee T^{-1/2})), \tag{B.73}$$

while the order of K_{115} is

$$K_{115} = N^{-1}T^{-(\nu_2+\delta)/2}\|\boldsymbol{\varepsilon}'\boldsymbol{\Gamma}^0_{+2}\mathbf{F}^{0\prime}_{+2}\| = O_P(N^{-1/2}T^{-(\nu_2+\delta-1-\nu_3)/2}). \tag{B.74}$$

By inserting the above results into (B.69), we obtain

$$\begin{aligned}
T^{-\delta/2}\|\mathbf{K}_{11}\| &= O_P(N^{-(3/2-p)}T^{-(\delta-1)/2}) + O_P(N^{-1/2}T^{q-(\nu_1+\nu_2+\delta-1)/2}) \\
&\quad + O_P(T^{-(\nu_2+\delta-2)/2}(N^{-1/2} \vee T^{-1/2})) + O_P(N^{-1/2}T^{-(\nu_2+\delta-1-\nu_3)/2}) \\
&\quad + o_P(T^{-(\delta-1)/2}\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\|).
\end{aligned} \tag{B.75}$$

The order of $T^{-\delta/2}\|\mathbf{K}_{12}\|$ is the same, which can be shown using the above steps.

We move on to $\mathbf{K}_{13}$, whose order is given by

$$\begin{aligned}
T^{-\delta/2}\|\mathbf{K}_{13}\| &\leq \left\| \frac{1}{NT^{(\nu_2+\delta)/2}} \sum_{i=1}^{N} (\mathbf{F}^0_{+2}\boldsymbol{\gamma}^0_{2,i} + \mathbf{F}^0_1\boldsymbol{\gamma}^0_{1,i} - \widehat{\mathbf{F}}_1\widehat{\boldsymbol{\gamma}}_{1,i})\boldsymbol{\gamma}^{0\prime}_{2,i}\mathbf{F}^{0\prime}_2 \right\| T^{-\delta/2}\|\widehat{\mathbf{F}}_2\| \\
&\leq O(1)N^{-1}T^{-(\nu_2+\delta)/2}\|\mathbf{F}^0_{+2}\boldsymbol{\Gamma}^{0\prime}_{+2}\boldsymbol{\Gamma}^0_2\mathbf{F}^{0\prime}_2\| \\
&\quad + O(1)\left\| \frac{1}{NT^{(\nu_2+\delta)/2}} \sum_{i=1}^{N} (\mathbf{F}^{0\prime}_1\boldsymbol{\gamma}^0_{1,i} - \widehat{\mathbf{F}}_1\widehat{\boldsymbol{\gamma}}_{1,i})\boldsymbol{\gamma}^{0\prime}_{2,i}\mathbf{F}^{0\prime}_2 \right\| \\
&= O_P(N^{-(1-p)}T^{-(\delta-\nu_3)/2}) + \left(\frac{1}{NT^{\delta}} \sum_{i=1}^{N} \|\mathbf{F}^0_1\boldsymbol{\gamma}^0_{1,i} - \widehat{\mathbf{F}}_1\widehat{\boldsymbol{\gamma}}_{1,i}\|^2 \right)^{1/2} \left(\frac{1}{NT^{\nu_2}} \sum_{i=1}^{N} \|\mathbf{F}^0_2\boldsymbol{\gamma}^0_{2,i}\|^2 \right)^{1/2} \\
&= O_P(N^{-1/2}T^{-(\delta-1)/2}) + O_P(T^{-\delta/2}) + O_P(T^{q-(\delta+\nu_1)/2}) + O_P(N^{-(1-p)}T^{-(\delta-\nu_2)/2}),
\end{aligned} \tag{B.76}$$

where the second equality follows from Lemma B.4 and Assumption 1. The order of $T^{-\delta/2}\|\mathbf{K}_{14}\|$ is the same.

For $\mathbf{K}_{15}$,

$$T^{-\delta/2}\|\mathbf{K}_{15}\| = T^{-\delta/2}\left\| \frac{1}{NT^{(\nu_2+\delta)/2}} \sum_{i=1}^{N} \mathbf{U}_i\mathbf{U}_i'\widehat{\mathbf{F}}_2 \right\|$$

$$
\begin{aligned}
&\leq O_P(1)\frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\|\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0-\widehat{\mathbf{F}}_1\widehat{\boldsymbol{\gamma}}_{1,i}\|^2+O_P(1)\frac{1}{NT^{(\nu_2+\delta)/2}}\sum_{i=1}^{N}\|\mathbf{F}_{+2}^0\boldsymbol{\gamma}_{+2,i}^0\|^2\\
&=O_P(T^{-(\nu_2+\delta)/2})+O_P(N^{-1}T^{1-(\nu_2+\delta)/2})+O_P(T^{2q-\nu_1-(\nu_2+\delta)/2})\\
&+O_P(N^{-2(1-p)}T^{-(\delta-\nu_2)/2})+O_P(T^{\nu_3-(\nu_2+\delta)/2}),
\end{aligned}
\tag{B.77}
$$

where the second equality follows from Lemma B.4.

We now insert the above results for $\mathbf{K}_1,\ldots,\mathbf{K}_{15}$ into (B.62). But first we left multiply by $T^{-\nu_2}\mathbf{F}_2^{0\prime}$ and $T^{-(\nu_2+\delta)/2}\widehat{\mathbf{F}}_2$ respectively. It then gives

$$
\begin{aligned}
&\|T^{-(\nu_2+\delta)/2}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_2(T^{-\nu_2}\mathbf{V}_2)-(T^{-\nu_2}\mathbf{F}_2^{0\prime}\mathbf{F}_2^0)(N^{-1}\boldsymbol{\Gamma}_2^{0\prime}\boldsymbol{\Gamma}_2^0)(T^{-(\nu_2+\delta)/2}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_2)\|\\
&=O_P(N^{-1/2}T^{-(\nu_2-1)/2})+O_P(T^{-\nu_2/2})+O_P(T^{-(\nu_2-\nu_3)})\\
&+O_P(T^{q-(\nu_1+\nu_2)/2})+O_P(N^{-(1-p)})
\end{aligned}
\tag{B.78}
$$

and

$$
\begin{aligned}
&\|T^{-\nu_2}\mathbf{V}_2-(T^{-(\nu_2+\delta)/2}\widehat{\mathbf{F}}_2'\mathbf{F}_2^0)(N^{-1}\boldsymbol{\Gamma}_2^{0\prime}\boldsymbol{\Gamma}_2^0)(T^{-(\nu_2+\delta)/2}\mathbf{F}_2^{0\prime}\widehat{\mathbf{F}}_2)\|\\
&=O_P(N^{-1/2}T^{-(\nu_2-1)/2})+O_P(T^{-\nu_2/2})+O_P(T^{-(\nu_2-\nu_3)})\\
&+O_P(T^{q-(\nu_1+\nu_2)/2})+O_P(N^{-(1-p)}).
\end{aligned}
\tag{B.79}
$$

The rest of the proofs of (a) and (b) follows from the same arguments used in Lemma B.3. It is therefore omitted. ■

Before continuing onto the proof of Lemma B.6, we note that the rates given in Lemma B.5 can actually be improved upon. Suppose that d_2 in known. Consider the next term which is a part of $\mathbf{K}_{15}$:

$$
\begin{aligned}
&N^{-1}T^{-(\nu_2/2+\delta)}\|\mathbf{F}_{+2}^0\boldsymbol{\Gamma}_{+2}^{0\prime}\boldsymbol{\Gamma}_{+2}^0\mathbf{F}_{+2}^{0\prime}\widehat{\mathbf{F}}_2\|\\
&=N^{-1}T^{-(\nu_2/2+\delta)}\|\mathbf{F}_{+2}^0\boldsymbol{\Gamma}_{+2}^{0\prime}\boldsymbol{\Gamma}_{+2}^0\mathbf{F}_{+2}^{0\prime}(\widehat{\mathbf{F}}_2\mathbf{V}_2^0-\mathbf{F}_2^0\mathbf{H}_2\mathbf{V}_2^0+\mathbf{F}_2^0\mathbf{H}_2\mathbf{V}_2^0)(\mathbf{V}_2^0)^{-1}\|\\
&\leq N^{-1}T^{-(\nu_2/2+\delta)}\|\mathbf{F}_{+2}^0\boldsymbol{\Gamma}_{+2}^{0\prime}\boldsymbol{\Gamma}_{+2}^0\mathbf{F}_{+2}^{0\prime}(\widehat{\mathbf{F}}_2\mathbf{V}_2^0-\mathbf{F}_2^0\mathbf{H}_2\mathbf{V}_2^0)(\mathbf{V}_2^0)^{-1}\|\\
&+N^{-1}T^{-(\nu_2/2+\delta)}\|\mathbf{F}_{+2}^0\boldsymbol{\Gamma}_{+2}^{0\prime}\boldsymbol{\Gamma}_{+2}^0\mathbf{F}_{+2}^{0\prime}\mathbf{F}_2^0\mathbf{H}_2\|\\
&\leq O_P(1)T^{\nu_3-\nu_2}T^{-(\nu_2/2+\delta)}\|\widehat{\mathbf{F}}_2\mathbf{V}_2^0-\mathbf{F}_2^0\mathbf{H}_2\mathbf{V}_2^0\|+O_P(1)T^{\nu_3/2-(\nu_2+\delta/2)}\|\mathbf{F}_{+2}^{0\prime}\mathbf{F}_2^0\|,
\end{aligned}
\tag{B.80}
$$

where $\mathbf{V}_2^0$ and $\mathbf{H}_2$ are defined in Appendix A. Simple algebra shows that the term $T^{-(\nu_2-\nu_3)}$ in τ_{NT} of Lemma B.5 can be dropped.

Proof of Lemma B.6.

Consider (a). Let us assume without loss of generality that $\boldsymbol{\beta}^0=\mathbf{0}_{d_x\times 1}$, as in Bai (2009). It then follows that

$$
\begin{aligned}
&(NT)^{-1}[\mathrm{SSR}(\widehat{\boldsymbol{\beta}}_1, \widehat{\mathbf{F}}) - \mathrm{SSR}(\widehat{\boldsymbol{\beta}}_0, \widehat{\mathbf{F}})] \\
&= \frac{1}{NT}\sum_{i=1}^{N}(-\mathbf{X}_i\widehat{\boldsymbol{\beta}}_1 + \mathbf{F}^0\boldsymbol{\gamma}_i^0)'\mathbf{M}_{\widehat{F}}(-\mathbf{X}_i\widehat{\boldsymbol{\beta}}_1 + \mathbf{F}^0\boldsymbol{\gamma}_i^0) \\
&- \widehat{\boldsymbol{\beta}}_1'\frac{2}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_i + \frac{2}{NT}\sum_{i=1}^{N}\boldsymbol{\gamma}_i^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_i + \frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_i \\
&- \frac{1}{NT}\sum_{i=1}^{N}(-\mathbf{X}_i\widehat{\boldsymbol{\beta}}_0 + \mathbf{F}^0\boldsymbol{\gamma}_i^0)'\mathbf{M}_{\widehat{F}}(-\mathbf{X}_i\widehat{\boldsymbol{\beta}}_0 + \mathbf{F}^0\boldsymbol{\gamma}_i^0) \\
&+ \widehat{\boldsymbol{\beta}}_0'\frac{2}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_i - \frac{2}{NT}\sum_{i=1}^{N}\boldsymbol{\gamma}_i^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_i - \frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_i \\
&\geq \frac{1}{NT}\sum_{i=1}^{N}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_i(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0) \\
&- 2\left(\frac{1}{NT}\sum_{i=1}^{N}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_i(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\right)^{1/2}\left(\frac{1}{NT}\sum_{i=1}^{N}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)'\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_i(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\right)^{1/2} \\
&- \frac{2}{NT}\sum_{i=1}^{N}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{F}^0\boldsymbol{\gamma}_i^0 - (\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)'\frac{2}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_i. \qquad \text{(B.81)}
\end{aligned}
$$

Note that by expanding the term $\mathbf{M}_{\widehat{F}}\mathbf{F}^0$ as in the proof for the second result of this lemma, we can obtain that

$$
\begin{aligned}
&\left\|\frac{1}{NT}\sum_{i=1}^{N}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{F}^0\boldsymbol{\gamma}_i^0\right\| \\
&\leq \|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\|\left\|\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{F}^0\boldsymbol{\gamma}_i^0\right\| \\
&= O_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\|)[O_P((NT)^{-1/2}) + o_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\|)]. \qquad \text{(B.82)}
\end{aligned}
$$

It follows that

$$
\begin{aligned}
0 &\geq (NT)^{-1}[\mathrm{SSR}(\widehat{\boldsymbol{\beta}}_1, \widehat{\mathbf{F}}) - \mathrm{SSR}(\widehat{\boldsymbol{\beta}}_0, \widehat{\mathbf{F}})] \\
&= \frac{1}{NT}\sum_{i=1}^{N}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_i(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0) \\
&+ O_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\|)[O_P((NT)^{-1/2}) + o_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\|)] \\
&\geq \lambda_{min}\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_i\mathbf{D}_T\right)\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\|^2 \\
&+ O_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\|)[O_P((NT)^{-1/2}) + o_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\|)], \qquad \text{(B.83)}
\end{aligned}
$$

where the first term on the right is quadratic in $\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\|$. Hence, to ensure the right-hand side is non-positive, $\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1-\widehat{\boldsymbol{\beta}}_0)\|$ cannot converge to zero at a rate faster than $(NT)^{-1/2}\vee\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0-\boldsymbol{\beta}^0)\|$. It follows that

$$\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1-\widehat{\boldsymbol{\beta}}_0)\| = O_P((NT)^{-1/2}\vee\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0-\boldsymbol{\beta}^0)\|), \tag{B.84}$$

as was to be shown.

We now turn to (b). Note that

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_1-\boldsymbol{\beta}^0 &= \mathbf{D}_T\left(\frac{1}{NT}\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_i\mathbf{D}_T\right)^{-1}\\ &\quad\times\left(\frac{1}{NT}\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{F}^0\boldsymbol{\gamma}_i^0+\frac{1}{NT}\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_i\right)\\ &= \mathbf{D}_T\mathbf{B}^{-1}\left(\mathbf{L}+\frac{1}{NT}\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_i\right),\end{aligned} \tag{B.85}$$

with obvious definitions of $\mathbf{L}$ and $\mathbf{B}$.

Consider $\mathbf{L}$. Let $\mathbf{Q}_g=(N^{-1}\boldsymbol{\Gamma}_g^{0\prime}\boldsymbol{\Gamma}_g^0)(T^{-(\nu_g+\delta)/2}\mathbf{F}_g^{0\prime}\widehat{\mathbf{F}}_g)$, such that $\mathbf{H}_g^{-1}=T^{-(\nu_g+\delta)/2}\mathbf{V}_g^0\mathbf{Q}_g^{-1}$. We also introduce $\mathbf{e}_{g,i}$, which is defined to be zero for $g=1$ and $\mathbf{e}_{g,i}=\sum_{j=1}^{g-1}(\mathbf{F}_j^0\gamma_{j,i}^0-\widehat{\mathbf{F}}_j\widehat{\gamma}_{j,i})$ for $g=2,\dots,G$. In this notation,

$$\begin{aligned}\mathbf{L}=\frac{1}{NT}\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{F}^0\boldsymbol{\gamma}_i^0 &= -\frac{1}{NT}\sum_{g=1}^G\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}(\widehat{\mathbf{F}}_g\mathbf{H}_g^{-1}-\mathbf{F}_g^0)\boldsymbol{\gamma}_{g,i}^0\\ &= -(\mathbf{L}_1+\cdots+\mathbf{L}_{15}),\end{aligned} \tag{B.86}$$

where

$$\mathbf{L}_1=\frac{1}{NT}\sum_{g=1}^G\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^N \mathbf{X}_j(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_2=\frac{1}{NT}\sum_{g=1}^G\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^N \mathbf{X}_j(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\boldsymbol{\gamma}_{g,j}^{0\prime}\mathbf{F}_g^{0\prime}\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_3=\frac{1}{NT}\sum_{g=1}^G\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^N \mathbf{F}_g^0\boldsymbol{\gamma}_{g,j}^0(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_4=\frac{1}{NT}\sum_{g=1}^G\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^N \mathbf{X}_j(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\mathbf{e}_{g,j}'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_5=\frac{1}{NT}\sum_{g=1}^G\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^N \mathbf{e}_{g,j}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_6=\frac{1}{NT}\sum_{g=1}^G\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^N \mathbf{X}_j(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\boldsymbol{\varepsilon}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_7 = \frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\boldsymbol{\varepsilon}_j(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_8 = \frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\boldsymbol{\varepsilon}_j\boldsymbol{\varepsilon}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_9 = \frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\mathbf{F}_g^0\boldsymbol{\gamma}_{g,j}^0\boldsymbol{\varepsilon}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_{10} = \frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\boldsymbol{\varepsilon}_j\boldsymbol{\gamma}_{g,j}^{0\prime}\mathbf{F}_g^{0\prime}\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_{11} = \frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\boldsymbol{\varepsilon}_j\mathbf{e}_{g,j}'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_{12} = \frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\mathbf{e}_{g,j}\boldsymbol{\varepsilon}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_{13} = \frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\mathbf{e}_{g,j}\boldsymbol{\gamma}_{g,j}^{0\prime}\mathbf{F}_g^{0\prime}\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_{14} = \frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\mathbf{F}_g^0\boldsymbol{\gamma}_{g,j}^0\mathbf{e}_{g,j}'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,$$

$$\mathbf{L}_{15} = \frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\mathbf{e}_{g,j}\mathbf{e}_{g,j}'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0.$$

We now evaluate each of these terms. From the analysis of $\mathbf{L}_2$ below, it is easy to show that $\|\mathbf{L}_1\| = o_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0-\boldsymbol{\beta}^0)\|)$. We therefore start from $\mathbf{L}_2$, which we write as

$$\begin{aligned}
\mathbf{L}_2 &= \frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\mathbf{X}_j\mathbf{D}_T\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)(\mathbf{F}_g^0\boldsymbol{\gamma}_{g,j}^0)'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0 \\
&= \frac{1}{N^2T}\sum_{g=1}^{G}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_T\boldsymbol{\gamma}_{g,j}^{0\prime}(T^{-(\nu_g+\delta)/2}\mathbf{F}_g^{0\prime}\widehat{\mathbf{F}}_g)(T^{-(\nu_g+\delta)/2}\mathbf{F}_g^{0\prime}\widehat{\mathbf{F}}_g)^{-1} \\
&\times (N^{-1}\boldsymbol{\Gamma}_g^{0\prime}\boldsymbol{\Gamma}_g^0)^{-1}\boldsymbol{\gamma}_{g,i}^0\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0) \\
&= \frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_T\boldsymbol{\gamma}_{g,j}^{0\prime}(\boldsymbol{\Gamma}_g^{0\prime}\boldsymbol{\Gamma}_g^0)^{-1}\boldsymbol{\gamma}_{g,i}^0\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0) \\
&= \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_Ta_{ji}\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0). \qquad \text{(B.87)}
\end{aligned}$$

We will use this expression for $\mathbf{L}_2$ later.

Let us now move on to $\mathbf{L}_3$.

$$\mathbf{L}_3 = \frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\mathbf{F}_g^0\boldsymbol{\gamma}_{g,j}^0(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0$$

$$= \frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\sum_{g=1}^{G}(\mathbf{F}_g^0-\widehat{\mathbf{F}}_g\mathbf{H}_g^{-1})$$

$$\times\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\boldsymbol{\gamma}_{g,j}^0(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0, \tag{B.88}$$

where, by using arguments that are similar to those used in the proofs of Lemmas B.3 and B.5,

$$\begin{aligned}
&\frac{1}{NT}\sum_{i=1}^{N}\left\|\sum_{g=1}^{G}(\mathbf{F}_g^0-\widehat{\mathbf{F}}_g\mathbf{H}_g^{-1})\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\boldsymbol{\gamma}_{g,j}^0(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\right\|^2\\
&\le\frac{1}{NT}\sum_{i=1}^{N}\left(\sum_{g=1}^{G}\|\mathbf{F}_g^0-\widehat{\mathbf{F}}_g\mathbf{H}_g^{-1}\|\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\|\boldsymbol{\gamma}_{g,j}^0(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{D}_T^{-1}\mathbf{D}_T\mathbf{X}_j'\widehat{\mathbf{F}}_g\|\|\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\|\right)^2\\
&\le O_P(1)\sum_{g=1}^{G}T^{-\nu_g}\|\mathbf{F}_g^0-\widehat{\mathbf{F}}_g\mathbf{H}_g^{-1}\|^2\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|^2\\
&=o_P(\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|^2)
\end{aligned}\tag{B.89}$$

implying

$$\begin{aligned}
\|\mathbf{L}_3\|&\le\left(\frac{1}{NT}\sum_{i=1}^{N}\|\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\|^2\right)^{1/2}\\
&\times\left(\frac{1}{NT}\sum_{i=1}^{N}\left\|\sum_{g=1}^{G}(\mathbf{F}_g^0-\widehat{\mathbf{F}}_g\mathbf{H}_g^{-1})\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\boldsymbol{\gamma}_{g,j}^0(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)'\mathbf{X}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\right\|^2\right)^{1/2}\\
&=o_P(\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|).
\end{aligned}\tag{B.90}$$

The same steps can be used to show that $\mathbf{L}_4$ and $\mathbf{L}_5$ are of the same order.

For $\mathbf{L}_6$, write

$$\begin{aligned}
\mathbf{L}_6&=\frac{1}{NT}\sum_{g=1}^{G}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\mathbf{X}_j(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\boldsymbol{\varepsilon}_j'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\\
&\le\sum_{g=1}^{G}\frac{1}{N^2T^{(\nu_g+\delta)/2+1}}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\boldsymbol{\varepsilon}_j'\mathbf{F}_g^0\mathbf{H}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\\
&+\sum_{g=1}^{G}\frac{1}{N^2T^{(\nu_g+\delta)/2+1}}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\boldsymbol{\varepsilon}_j'(\widehat{\mathbf{F}}_g-\mathbf{F}_g^0\mathbf{H}_g)\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0,
\end{aligned}\tag{B.91}$$

where the first term on the right is bounded by

$$\sum_{g=1}^{G}\frac{1}{N^2T^{(\nu_g+\delta)/2+1}}\sum_{i=1}^{N}\sum_{j=1}^{N}\|\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_T\boldsymbol{\varepsilon}_j'\mathbf{F}_g^0\mathbf{H}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\|\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|$$

$$
\begin{aligned}
&\leq O_P(1)\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|\sum_{g=1}^{G}\frac{1}{N^2T^{(\nu_g+\delta)/2+1}}\left(\sum_{i=1}^{N}\|\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\|^2\right)^{1/2}\left(\sum_{i=1}^{N}\|\boldsymbol{\gamma}_{g,i}^0\|^2\right)^{1/2}\\
&\times\left(\sum_{j=1}^{N}\|\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_T\|^2\right)^{1/2}\left(\sum_{j=1}^{N}\|\boldsymbol{\varepsilon}_j'\mathbf{F}_g^0\|^2\right)^{1/2}\|\mathbf{H}_g\|\\
&\leq O_P(1)\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|\sum_{g=1}^{G}T^{-(\nu_g+\delta)/2-1}O_P(\sqrt{T})O_P(\sqrt{T})O_P(T^{\nu_g/2})O_P(T^{-(\nu_g-\delta)/2})\\
&=o_P(\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|). \qquad \text{(B.92)}
\end{aligned}
$$

The second term is of the same order. Thus, $\|\mathbf{L}_6\|=o_P(\|\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0-\widehat{\boldsymbol{\beta}}_0)\|)$. The same is true for $\|\mathbf{L}_7\|$.

We now examine $\mathbf{L}_8$. Let us define $\boldsymbol{\Sigma}_\varepsilon=N^{-1}\sum_{i=1}^N\boldsymbol{\Sigma}_{\varepsilon,i}$, in which $\boldsymbol{\Sigma}_{\varepsilon,i}$ has been defined in Assumption 3. $\mathbf{L}_8$ can then be written as

$$
\begin{aligned}
\mathbf{L}_8=&\sum_{g=1}^{G}\frac{1}{NT^{(\nu_g+\delta)/2+1}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\Sigma}_\varepsilon\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\\
&+\sum_{g=1}^{G}\frac{1}{N^2T^{(\nu_g+\delta)/2+1}}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'(\boldsymbol{\varepsilon}_j\boldsymbol{\varepsilon}_j'-\boldsymbol{\Sigma}_{\varepsilon,j})\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\\
&-\sum_{g=1}^{G}\frac{1}{N^2T^{(\nu_g+\delta)/2+1}}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{P}_{\widehat{F}}(\boldsymbol{\varepsilon}_j\boldsymbol{\varepsilon}_j'-\boldsymbol{\Sigma}_{\varepsilon,j})\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0. \qquad \text{(B.93)}
\end{aligned}
$$

For the first term on the right,

$$
\begin{aligned}
&\sqrt{NT}\left\|\sum_{g=1}^{G}\frac{1}{NT^{(\nu_g+\delta)/2+1}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\Sigma}_\varepsilon\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\right\|\\
&\leq\sum_{g=1}^{G}\frac{1}{\sqrt{N}T^{(\nu_g+\delta+1)/2}}\sum_{i=1}^{N}\|\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\Sigma}_{\varepsilon,j}\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\|\\
&\leq O_P(1)\sum_{g=1}^{G}N^{-1/2}T^{-(\nu_g+\delta+1)/2}O_P(NT)O_P(T^{(\delta-1)/2})\\
&\leq O_P(1)\sum_{g=1}^{G}N^{1/2}T^{-\nu_g/2}=O_P(\sqrt{N}T^{-\nu_G/2})=O_P(1), \qquad \text{(B.94)}
\end{aligned}
$$

where the last equality holds by $NT^{-\nu_G}<\infty$. Let

$$
\overline{\mathbf{A}}_1=\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{F_0}\boldsymbol{\Sigma}_\varepsilon T^{-(\nu_G-1)/2}\mathbf{F}_G^0(T^{-\nu_G}\mathbf{F}_G^{0\prime}\mathbf{F}_G^0)^{-1}(N^{-1}\boldsymbol{\Gamma}_G^{0\prime}\boldsymbol{\Gamma}_G^0)^{-1}\boldsymbol{\gamma}_{G,i}^0. \qquad \text{(B.95)}
$$

Note how $\mathbf{A}_1=\text{plim}_{N,T\to\infty}\mathbb{E}(\overline{\mathbf{A}}_1|\mathcal{C})$. In this notation,

$$
\sqrt{NT}\sum_{g=1}^{G}\frac{1}{NT^{(\nu_g+\delta)/2+1}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\Sigma}_\varepsilon\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0
$$

$$= \sqrt{NT} \sum_{g=1}^{G} \frac{1}{NT^{(\nu_g+\delta)/2+1}} \sum_{i=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{F_0} \boldsymbol{\Sigma}_\varepsilon \widehat{\mathbf{F}}_g (T^{-(\nu_g+\delta)/2} \mathbf{F}_g^{0\prime} \widehat{\mathbf{F}}_g)^{-1} (N^{-1} \boldsymbol{\Gamma}_g^{0\prime} \boldsymbol{\Gamma}_g^0)^{-1} \boldsymbol{\gamma}_{g,i}^0 (1 + o_P(1))$$
$$= \sum_{g=1}^{G} \sqrt{\frac{N}{T^{\nu_g}}} \frac{1}{NT} \sum_{i=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{F_0} \boldsymbol{\Sigma}_\varepsilon T^{-(\nu_g-1)/2} \mathbf{F}_g^0 (T^{-\nu_g} \mathbf{F}_g^{0\prime} \mathbf{F}_g^0)^{-1} (N^{-1} \boldsymbol{\Gamma}_g^{0\prime} \boldsymbol{\Gamma}_g^0)^{-1} \boldsymbol{\gamma}_{g,i}^0 (1 + o_P(1))$$
$$= \sqrt{\frac{N}{T^{\nu_G}}} \frac{1}{NT} \sum_{i=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{F_0} \boldsymbol{\Sigma}_\varepsilon T^{-(\nu_G-1)/2} \mathbf{F}_G^0 (T^{-\nu_G} \mathbf{F}_G^{0\prime} \mathbf{F}_G^0)^{-1} (N^{-1} \boldsymbol{\Gamma}_G^{0\prime} \boldsymbol{\Gamma}_G^0)^{-1} \boldsymbol{\gamma}_{G,i}^0 (1 + o_P(1))$$
$$= \sqrt{N} T^{-\nu_G/2} \overline{\mathbf{A}}_1 (1 + o_P(1)), \tag{B.96}$$

where the second equality follows from arguments similar to those used in (B.30). Note also that $\lim \sqrt{N} T^{-\nu_G/2} < \infty$. Further use of the same steps used by Jiang et al. (2021) establishes that the second and third terms of $\mathbf{L}_8$ are $o_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\|) + o_P((NT)^{-1/2})$. Hence,

$$\mathbf{L}_8 = T^{-(\nu_G+1)/2} \overline{\mathbf{A}}_1 + o_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\|) + o_P((NT)^{-1/2}). \tag{B.97}$$

$\mathbf{L}_9$ can be written as

$$\mathbf{L}_9 = \sum_{g=1}^{G} \frac{1}{N^2 T^{(\nu_g+\delta)/2+1}} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{\widehat{F}} (\mathbf{F}_g^0 - \widehat{\mathbf{F}}_g \mathbf{H}_g^{-1}) \boldsymbol{\gamma}_{g,j}^0 \boldsymbol{\varepsilon}_j' \mathbf{F}_g^0 \mathbf{H}_g \mathbf{Q}_g^{-1} \boldsymbol{\gamma}_{g,i}^0$$
$$+ \sum_{g=1}^{G} \frac{1}{N^2 T^{(\nu_g+\delta)/2+1}} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{\widehat{F}} (\mathbf{F}_g^0 - \widehat{\mathbf{F}}_g \mathbf{H}_g^{-1}) \boldsymbol{\gamma}_{g,j}^0 \boldsymbol{\varepsilon}_j' (\widehat{\mathbf{F}}_g - \mathbf{F}_g^0 \mathbf{H}_g) \mathbf{Q}_g^{-1} \boldsymbol{\gamma}_{g,i}^0, \tag{B.98}$$

where the first term on the right-hand side is bounded by

$$\sum_{g=1}^{G} \frac{1}{N^2 T^{(\nu_g+\delta)/2+1}} \sum_{i=1}^{N} \|\mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{\widehat{F}}\| \|\mathbf{F}_g^0 - \widehat{\mathbf{F}}_g \mathbf{H}_g^{-1}\| \|T^{-(\nu_g-1)/2} \boldsymbol{\Gamma}_g^{0\prime} \boldsymbol{\varepsilon} \mathbf{F}_g^0\|$$
$$\times T^{(\nu_g-1)/2} \|\mathbf{H}_g\| \|\mathbf{Q}_g^{-1} \boldsymbol{\gamma}_{g,i}^0\|$$
$$\leq O_P(1) \sum_{g=1}^{G} N^{-1} T^{-(\nu_g+\delta)/2-1} O_P(\sqrt{T}) \|\mathbf{F}_g^0 - \widehat{\mathbf{F}}_g \mathbf{H}_g^{-1}\| O_P(\sqrt{NT}) O_P(T^{(\nu_g-1)/2}) O_P(T^{-(\nu_g-\delta)/2})$$
$$\leq O_P(1) \sum_{g=1}^{G} (NT)^{-1/2} T^{-\nu_g/2} \|\mathbf{F}_g^0 - \widehat{\mathbf{F}}_g \mathbf{H}_g^{-1}\|$$
$$= o_P((NT)^{-1/2}), \tag{B.99}$$

where the first inequality follows from $\|T^{-(\nu_g-1)/2} \boldsymbol{\Gamma}_g^{0\prime} \boldsymbol{\varepsilon} \mathbf{F}_g^0\| = O_P(\sqrt{NT})$. The second term on the right-hand side of $\mathbf{L}_9$ is also $o_P((NT)^{-1/2})$. We therefore conclude that $\|\mathbf{L}_9\| = o_P((NT)^{-1/2})$.

$\mathbf{L}_{10}$ can be written more compactly as

$$\mathbf{L}_{10} = \frac{1}{NT} \sum_{g=1}^{G} \sum_{i=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{\widehat{F}} \frac{1}{NT^{(\nu_g+\delta)/2}} \sum_{j=1}^{N} \boldsymbol{\varepsilon}_j \boldsymbol{\gamma}_{g,j}^{0\prime} \mathbf{F}_g^{0\prime} \widehat{\mathbf{F}}_g \mathbf{Q}_g^{-1} \boldsymbol{\gamma}_{g,i}^0$$

$$= \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_j a_{ij}, \tag{B.100}$$

which we will again make use of later.

Let us move on to $\mathbf{L}_{11}$. We begin by rewriting $\mathbf{e}_{g,j}$ as

$$\begin{aligned}\mathbf{e}_{g,j} &= \sum_{d=1}^{g-1}[\mathbf{F}_d^0\boldsymbol{\gamma}_{d,j}^0 - \widehat{\mathbf{F}}_d T^{-\delta}\widehat{\mathbf{F}}_d'(\mathbf{y}_j - \mathbf{X}_j\widehat{\boldsymbol{\beta}}_0 - \widehat{\mathbf{F}}_d\widehat{\boldsymbol{\gamma}}_{d,j})]\\ &= \sum_{d=1}^{g-1}[-\mathbf{P}_{\widehat{F}_d}\mathbf{X}_j(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0) + \mathbf{M}_{\widehat{F}_d}\mathbf{F}_d^0\boldsymbol{\gamma}_{d,j}^0 - \mathbf{P}_{\widehat{F}_d}\mathbf{F}_{+d}^0\boldsymbol{\gamma}_{+d,j}^0 - \mathbf{P}_{\widehat{F}_d}\mathbf{e}_{d,j} - \mathbf{P}_{\widehat{F}_d}\boldsymbol{\varepsilon}_j],\end{aligned} \tag{B.101}$$

where $\mathbf{F}_{+d}^0 = (\mathbf{F}_{d+1}^0, \ldots, \mathbf{F}_G^0)$ and $\boldsymbol{\gamma}_{+d,j}^0 = (\boldsymbol{\gamma}_{d+1,j}^{0\prime}, \ldots, \boldsymbol{\gamma}_{G,j}^{0\prime})'$, as in Appendix A. By inserting this into $\mathbf{L}_{11}$, we get

$$\begin{aligned}\mathbf{L}_{11} = \sum_{g=1}^{G}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\boldsymbol{\varepsilon}_j\sum_{d=1}^{g-1}[-\mathbf{P}_{\widehat{F}_d}\mathbf{X}_j(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\\ + \mathbf{M}_{\widehat{F}_d}\mathbf{F}_d^0\boldsymbol{\gamma}_{d,j}^0 - \mathbf{P}_{\widehat{F}_d}\mathbf{F}_{+d}^0\boldsymbol{\gamma}_{+d,j}^0 - \mathbf{P}_{\widehat{F}_d}\mathbf{e}_{d,j} - \mathbf{P}_{\widehat{F}_d}\boldsymbol{\varepsilon}_j]'\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0.\end{aligned} \tag{B.102}$$

Hence, $\mathbf{L}_{11}$ can be written as a sum of five terms. There is no need to study the fourth term, the one due to $\mathbf{P}_{\widehat{F}_d}\mathbf{e}_{d,j}$, as we can keep expanding $\mathbf{e}_{d,j}$ until we cannot. The first and fifth terms are $o_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\|)$ and $o_P((NT)^{-1/2})$, respectively, by the same arguments used for evaluating $\mathbf{L}_7$ and $\mathbf{L}_8$. Moreover, the steps used for evaluating $\mathbf{L}_9$ can be used to show that

$$\begin{aligned}&\left\|\sum_{g=1}^{G}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\boldsymbol{\varepsilon}_j\sum_{d=1}^{g-1}\boldsymbol{\gamma}_{d,j}^{0\prime}\boldsymbol{F}_d^{0\prime}\mathbf{M}_{\widehat{F}_d}\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\right\|\\ &= \left\|\sum_{g=1}^{G}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{d=1}^{g-1}\boldsymbol{\varepsilon}'\boldsymbol{\Gamma}_d^0\mathbf{F}_d^{0\prime}\mathbf{M}_{\widehat{F}_d}\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\right\|\\ &= \left\|\sum_{g=1}^{G}\frac{1}{N^2T^{(\nu_g+\delta)/2+1}}\sum_{i=1}^{N}\sum_{d=1}^{g-1}\mathbf{D}_T\mathbf{X}_i'\boldsymbol{\varepsilon}'\boldsymbol{\Gamma}_d^0(\mathbf{F}_d^0 - \widehat{\mathbf{F}}_d\mathbf{H}_d^{-1})'\mathbf{M}_{\widehat{F}_d}\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\right\|\\ &+ \left\|\sum_{g=1}^{G}\frac{1}{N^2T^{(\nu_g+\delta)/2+1}}\sum_{i=1}^{N}\sum_{d=1}^{g-1}\mathbf{D}_T\mathbf{X}_i'\mathbf{P}_{\widehat{F}}\boldsymbol{\varepsilon}'\boldsymbol{\Gamma}_d^0(\mathbf{F}_d^0 - \widehat{\mathbf{F}}_d\mathbf{H}_d^{-1})'\mathbf{M}_{\widehat{F}_d}\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\right\|\\ &= o_P((NT)^{-1/2}).\end{aligned} \tag{B.103}$$

The second term in $\mathbf{L}_{11}$ is therefore negligible. It remains to consider the third term, which is

$$\left\|\sum_{g=1}^{G}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\frac{1}{NT^{(\nu_g+\delta)/2}}\sum_{j=1}^{N}\boldsymbol{\varepsilon}_j\sum_{d=1}^{g-1}\boldsymbol{\gamma}_{+d,j}^{0\prime}\mathbf{F}_{+d}^{0\prime}\mathbf{P}_{\widehat{F}_d}\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1}\boldsymbol{\gamma}_{g,i}^0\right\|$$

$$
= \left\| \sum_{g=1}^{G} \frac{1}{NT} \sum_{i=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{\widehat{F}} \frac{1}{NT^{(\nu_g+\delta)/2}} \sum_{d=1}^{g-1} \boldsymbol{\varepsilon}' \boldsymbol{\Gamma}_{+d}^0 \mathbf{F}_{+d}^{0\prime} \mathbf{P}_{\widehat{F}_d} \widehat{\mathbf{F}}_g \mathbf{Q}_g^{-1} \boldsymbol{\gamma}_{g,i}^0 \right\|
$$

$$
= o_P(\|\mathbf{L}_{10}\|) = o_P((NT)^{-1/2}) \tag{B.104}
$$

where we have used Assumption 4. Therefore, $\|\mathbf{L}_{11}\| = o_P((NT)^{-1/2})$.

For $\mathbf{L}_{12}$,

$$
\begin{aligned}
\mathbf{L}_{12} &= \sum_{g=1}^{G} \frac{1}{NT} \sum_{i=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{\widehat{F}} \frac{1}{NT^{(\nu_g+\delta)/2}} \sum_{j=1}^{N} \mathbf{e}_{g,j} \boldsymbol{\varepsilon}_j' \widehat{\mathbf{F}}_g \mathbf{Q}_g^{-1} \boldsymbol{\gamma}_{g,i}^0 \\
&= \sum_{g=1}^{G} \frac{1}{NT} \sum_{i=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{\widehat{F}} \frac{1}{NT^{(\nu_g+\delta)/2}} \sum_{j=1}^{N} \mathbf{F}_g^0 \boldsymbol{\gamma}_i^0 \boldsymbol{\varepsilon}_j' \widehat{\mathbf{F}}_g \mathbf{Q}_g^{-1} \boldsymbol{\gamma}_{g,i}^0 = \mathbf{L}_9,
\end{aligned} \tag{B.105}
$$

where the second equality follows from the construction of $\mathbf{e}_{g,j}$. Hence, since $\mathbf{L}_9$ is negligible, $\mathbf{L}_{12}$ is also negligible.

Next up is $\mathbf{L}_{13}$, whose order can be worked out in the following way:

$$
\begin{aligned}
\|\mathbf{L}_{13}\| &\leq \sum_{g=1}^{G} \frac{1}{N^2T} \sum_{i=1}^{N} \left\| \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{\widehat{F}} \sum_{d=1}^{g-1} \mathbf{F}_d^0 \boldsymbol{\Gamma}_d^{0\prime} \boldsymbol{\Gamma}_g^0 (N^{-1} \boldsymbol{\Gamma}_g^{0\prime} \boldsymbol{\Gamma}_g^0)^{-1} \boldsymbol{\gamma}_{g,i}^0 \right\| \\
&\leq \sum_{g=1}^{G} \frac{1}{N^2T} \sum_{i=1}^{N} \left\| \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{\widehat{F}} \sum_{d=1}^{g-1} (\mathbf{F}_d^0 - \widehat{\mathbf{F}}_d \mathbf{H}_d^{-1}) \boldsymbol{\Gamma}_d^{0\prime} \boldsymbol{\Gamma}_g^0 (N^{-1} \boldsymbol{\Gamma}_g^{0\prime} \boldsymbol{\Gamma}_g^0)^{-1} \boldsymbol{\gamma}_{g,i}^0 \right\| \\
&\leq O_P(1) \sum_{g=1}^{G} \frac{1}{N^2T} \sum_{d=1}^{g-1} \sum_{i=1}^{N} \|\mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{\widehat{F}}\| \|\mathbf{F}_d^0 - \widehat{\mathbf{F}}_d \mathbf{H}_d^{-1}\| \|\boldsymbol{\Gamma}_d^{0\prime} \boldsymbol{\Gamma}_g^0\| \|\boldsymbol{\gamma}_{g,i}^0\| \\
&\leq O_P(1) \sum_{g=1}^{G} \frac{1}{NT} \sum_{d=1}^{g-1} O_P(\sqrt{T}) \|\mathbf{F}_d^0 - \widehat{\mathbf{F}}_d \mathbf{H}_d^{-1}\| O_P(N^p) \\
&\leq O_P(1) \sum_{g=1}^{G} \frac{1}{N^{1-p}\sqrt{T}} \sum_{d=1}^{g-1} \|\mathbf{F}_d^0 - \widehat{\mathbf{F}}_d \mathbf{H}_d^{-1}\| = o_P((NT)^{-1/2}),
\end{aligned} \tag{B.106}
$$

where the first inequality follows from the construction of $\mathbf{e}_{g,j}$, and the last equality follows by going through a development similar to the first result of Lemma B.5 and further using a development similar to (B.80). The same arguments show that $\|\mathbf{L}_{14}\|$ and $\|\mathbf{L}_{15}\|$ are negligible, too.

We now put everything together. This yields

$$
\begin{aligned}
\mathbf{L} &= -(\mathbf{L}_1 + \cdots + \mathbf{L}_{15}) \\
&= -\mathbf{L}_2 - \mathbf{L}_8 - \mathbf{L}_{10} + o_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\|) + o_P((NT)^{-1/2}) \\
&= \left(-\frac{1}{NT} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{\widehat{F}} \mathbf{X}_j \mathbf{D}_T a_{ji} + o_P(1) \right) \mathbf{D}_T^{-1} (\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0) - T^{(\nu_G+1)/2} \overline{\mathbf{A}}_1 \\
&\quad - \frac{1}{NT} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{M}_{\widehat{F}} a_{ji} \boldsymbol{\varepsilon}_j + o_P((NT)^{-1/2}),
\end{aligned} \tag{B.107}
$$

which in turn implies

$$\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}^0) = \mathbf{B}^{-1}\left(\mathbf{L} + \frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_i\right)$$
$$= (\mathbf{B}^{-1}\mathbf{N} + o_P(1))\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0) + \mathbf{B}^{-1}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i$$
$$- T^{-(\nu_G+1)/2}\mathbf{B}^{-1}\overline{\mathbf{A}}_1 + o_P((NT)^{-1/2}), \tag{B.108}$$

where $\mathbf{Z}_i(\mathbf{F})$ is defined in Assumption 2, and

$$\mathbf{N} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_T a_{ij}.$$

This expression for $\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}^0)$ can be inserted into $\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)$, giving

$$\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0) = \mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}^0) - \mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)$$
$$= (\mathbf{B}^{-1}\mathbf{N} + o_P(1))\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0) + \mathbf{B}^{-1}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i$$
$$- \mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0) - T^{-(\nu_G+1)/2}\mathbf{B}^{-1}\overline{\mathbf{A}}_1 + o_P((NT)^{-1/2}), \tag{B.109}$$

which can be solved for $\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)$

$$\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0) = \mathbf{B}(\widehat{\mathbf{F}})^{-1}\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i - \mathbf{B}\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0) - T^{-(\nu_G+1)/2}\overline{\mathbf{A}}_1\right)$$
$$+ o_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\|) + o_P((NT)^{-1/2}). \tag{B.110}$$

Also, making use of Assumption 3, it is not difficult to show that $(NT)^{-1}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i = O_P((NT)^{-1/2})$. Hence,

$$\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0) + \mathbf{B}(\widehat{\mathbf{F}})^{-1}\mathbf{B}\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)$$
$$= \mathbf{B}(\widehat{\mathbf{F}})^{-1}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i - T^{-(\nu_G+1)/2}\mathbf{B}(\widehat{\mathbf{F}})^{-1}\overline{\mathbf{A}}_1$$
$$+ o_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\|) + o_P((NT)^{-1/2}). \tag{B.111}$$

By using the fact that

$$\mathbf{B}(\widehat{\mathbf{F}})^{-1}\mathbf{B}\mathbf{D}_T^{-1} = \mathbf{D}_T^{-1}\left(\sum_{i=1}^{N}\mathbf{Z}_i(\widehat{\mathbf{F}})'\mathbf{Z}_i(\widehat{\mathbf{F}})\right)^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_i, \tag{B.112}$$

the left-hand side of this last equation can be written as

$$\begin{aligned}
&\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0) + \mathbf{B}(\widehat{\mathbf{F}})^{-1}\mathbf{B}\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0) \\
&= \mathbf{D}_T^{-1}\left[\widehat{\boldsymbol{\beta}}_0 + \left(\sum_{i=1}^N \mathbf{Z}_i(\widehat{\mathbf{F}})'\mathbf{Z}_i(\widehat{\mathbf{F}})\right)^{-1}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_i(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0) - \boldsymbol{\beta}^0\right]. \qquad \text{(B.113)}
\end{aligned}$$

It follows that

$$\begin{aligned}
&\sqrt{NT}\mathbf{D}_T^{-1}\left[\widehat{\boldsymbol{\beta}}_0 + \left(\sum_{i=1}^N \mathbf{Z}_i(\widehat{\mathbf{F}})'\mathbf{Z}_i(\widehat{\mathbf{F}})\right)^{-1}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_i(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0) - \boldsymbol{\beta}^0\right] \\
&= \mathbf{B}(\widehat{\mathbf{F}})^{-1}\frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i - \sqrt{N}T^{-\nu_G/2}\mathbf{B}(\widehat{\mathbf{F}})^{-1}\overline{\mathbf{A}}_1 \\
&+ o_P(\sqrt{NT}\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\|) + o_P(1). \qquad \text{(B.114)}
\end{aligned}$$

Consider the $o_P(\sqrt{NT}\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\|)$ reminder term. By the first result of this lemma, $\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\| = O_P((NT)^{-1/2} \vee \|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\|)$. This can be inserted into (B.111), giving

$$\begin{aligned}
\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0) &= \mathbf{B}(\widehat{\mathbf{F}})^{-1}\frac{1}{NT}\sum_{i=1}^N \mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i + o_P((NT)^{-1/2} \vee \|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\|) \\
&= O_P((NT)^{-1/2}) + o_P(\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\|), \qquad \text{(B.115)}
\end{aligned}$$

which in turn implies

$$\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\| = O_P((NT)^{-1/2}). \qquad \text{(B.116)}$$

The second result then follows. ■

B.4 Proofs of main results

Proof of Lemma 1.

Without loss of generality, we assume that $\boldsymbol{\beta}^0 = \mathbf{0}_{d_x\times 1}$. This implies

$$\begin{aligned}
&(NT)^{-1}[\mathrm{SSR}(\boldsymbol{\beta}, \mathbf{F}) - \mathrm{SSR}(\boldsymbol{\beta}^0, \mathbf{F}^0)] \\
&= \frac{1}{NT}\sum_{i=1}^N (\mathbf{X}_i\boldsymbol{\beta} + \mathbf{F}^0\boldsymbol{\gamma}_i^0)'\mathbf{M}_F(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{F}^0\boldsymbol{\gamma}_i^0) \\
&+ \boldsymbol{\beta}'\frac{2}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i + \frac{2}{NT}\sum_{i=1}^N \boldsymbol{\gamma}_i^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}_F\boldsymbol{\varepsilon}_i + \frac{1}{NT}\sum_{i=1}^N \boldsymbol{\varepsilon}_i'(\mathbf{P}_{F^0} - \mathbf{P}_F)\boldsymbol{\varepsilon}_i \\
&= \frac{1}{NT}\sum_{i=1}^N (\mathbf{X}_i\mathbf{D}_T\mathbf{D}_T^{-1}\boldsymbol{\beta} + \mathbf{F}^0\boldsymbol{\gamma}_i^0)'\mathbf{M}_F(\mathbf{X}_i\mathbf{D}_T\mathbf{D}_T^{-1}\boldsymbol{\beta} + \mathbf{F}^0\boldsymbol{\gamma}_i^0)
\end{aligned}$$

$$+\boldsymbol{\beta}'\mathbf{D}_T^{-1}\frac{2}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i+\frac{2}{NT}\sum_{i=1}^{N}\boldsymbol{\gamma}_i^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}_F\boldsymbol{\varepsilon}_i+O_P(N^{-1}\vee T^{-1})$$

$$=\boldsymbol{\beta}'\mathbf{D}_T^{-1}\mathbf{B}(\mathbf{F})\mathbf{D}_T^{-1}\boldsymbol{\beta}+\boldsymbol{\theta}'\mathbf{B}\boldsymbol{\theta}+\boldsymbol{\beta}'\mathbf{D}_T^{-1}\frac{2}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i$$

$$+\frac{2}{NT}\sum_{i=1}^{N}\boldsymbol{\gamma}_i^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}_F\boldsymbol{\varepsilon}_i+O_P(N^{-1}\vee T^{-1}), \tag{B.117}$$

where the second equality follows from Lemma B.2, and $\boldsymbol{\theta}=\boldsymbol{\eta}+\mathbf{B}^{-1}\mathbf{C}\boldsymbol{\beta}$ and $\mathbf{B}=N^{-1}\boldsymbol{\Gamma}^{0\prime}\boldsymbol{\Gamma}^0\otimes\mathbf{I}_T$ are defined as on page 1265 of Bai (2009). Note that for $\boldsymbol{\beta}\in\mathbb{R}^{d_x}$, we may have

$$\sup_{\boldsymbol{\beta}\in\mathbb{R}^{d_x},\mathbf{F}\in\mathbb{D}_F}\left|\boldsymbol{\beta}'\mathbf{D}_T^{-1}\frac{2}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right|\neq o_P(1). \tag{B.118}$$

In our proof of consistency, we consider two cases; (i) $\|\mathbf{D}_T^{-1}\boldsymbol{\beta}\|\leq C$ and (ii) $\|\mathbf{D}_T^{-1}\boldsymbol{\beta}\|>C$, where C is a large positive constant. Under (i),

$$\sup_{\|\mathbf{D}_T^{-1}\boldsymbol{\beta}\|\leq C,\mathbf{F}\in\mathbb{D}_F}\left|\boldsymbol{\beta}'\mathbf{D}_T^{-1}\frac{2}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right|=O_P(N^{-1/2}\vee T^{-1/2}) \tag{B.119}$$

by Lemma B.2. The expression given in (B.117) for $(NT)^{-1}[\mathrm{SSR}(\boldsymbol{\beta},F)-\mathrm{SSR}(\boldsymbol{\beta}^0,\mathbf{F}^0)]$ therefore reduces to

$$(NT)^{-1}[\mathrm{SSR}(\boldsymbol{\beta},\mathbf{F})-\mathrm{SSR}(\boldsymbol{\beta}^0,\mathbf{F}^0)]=\boldsymbol{\beta}'\mathbf{D}_T^{-1}\mathbf{B}(\mathbf{F})\mathbf{D}_T^{-1}\boldsymbol{\beta}+\boldsymbol{\theta}'\mathbf{B}\boldsymbol{\theta}+\frac{2}{NT}\sum_{i=1}^{N}\boldsymbol{\gamma}_i^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}_F\boldsymbol{\varepsilon}_i$$

$$+O_P(N^{-1/2}\vee T^{-1/2}), \tag{B.120}$$

where $\boldsymbol{\beta}'\mathbf{D}_T^{-1}\mathbf{B}(\mathbf{F})\mathbf{D}_T^{-1}\boldsymbol{\beta}$ does not involve $\mathbf{F}^0$ and $\sum_{i=1}^{N}\boldsymbol{\gamma}_i^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}_F\boldsymbol{\varepsilon}_i$ is independent of $\boldsymbol{\beta}$. Hence, provided $d_{max}\geq d_f$, the consistency of $\mathbf{D}_T^{-1}\widehat{\boldsymbol{\beta}}_0$ in case (i) follows from the same arguments as in Bai (2009).

Under (ii), (B.117) can be written as

$$(NT)^{-1}[\mathrm{SSR}(\boldsymbol{\beta},\mathbf{F})-\mathrm{SSR}(\boldsymbol{\beta}^0,\mathbf{F}^0)]$$

$$=\boldsymbol{\beta}'\mathbf{D}_T^{-1}\mathbf{B}(\mathbf{F})\mathbf{D}_T^{-1}\boldsymbol{\beta}+\boldsymbol{\theta}'\mathbf{B}\boldsymbol{\theta}+\boldsymbol{\beta}'\mathbf{D}_T^{-1}\frac{2}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i+\frac{2}{NT}\sum_{i=1}^{N}\boldsymbol{\gamma}_i^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}_F\boldsymbol{\varepsilon}_i$$

$$+O_P(N^{-1/2}\vee T^{-1/2})$$

$$\geq c_0\|\mathbf{D}_T^{-1}\boldsymbol{\beta}\|^2+\boldsymbol{\beta}'\mathbf{D}_T^{-1}\frac{2}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i+\boldsymbol{\theta}'\mathbf{B}\boldsymbol{\theta}+\frac{2}{NT}\sum_{i=1}^{N}\boldsymbol{\gamma}_i^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}_F\boldsymbol{\varepsilon}_i$$

$$+O_P(N^{-1/2}\vee T^{-1/2})$$

$$\geq \frac{c_0}{2}C^2 + \boldsymbol{\theta}'\mathbf{B}\boldsymbol{\theta} + \frac{2}{NT}\sum_{i=1}^{N}\boldsymbol{\gamma}_i^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}_F\boldsymbol{\varepsilon}_i + O_P(N^{-1/2}\vee T^{-1/2}), \tag{B.121}$$

where c_0 is defined in Assumption 2, and the second inequality follows from the fact that the quadratic term dominates the linear one for large values of C. Hence, $(NT)^{-1}[\mathrm{SSR}(\boldsymbol{\beta},\mathbf{F}) - \mathrm{SSR}(\boldsymbol{\beta}^0,\mathbf{F}^0)] > 0$, but from the definition of $\widehat{\boldsymbol{\beta}}_0$ we also know that $\mathrm{SSR}(\widehat{\boldsymbol{\beta}}_0,\widehat{\mathbf{F}}) - \mathrm{SSR}(\boldsymbol{\beta}^0,\mathbf{F}^0) \leq 0$, which means that $\mathbf{D}_T^{-1}\widehat{\boldsymbol{\beta}}_0$ cannot belong to (ii).

Note that since $\|(NT)^{-1}\sum_{i=1}^N \mathbf{D}_T\mathbf{X}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\| = O_P(N^{-1/2}\vee T^{-1/2})$ by Lemma B.2, all we need is $C = C^0(N^{-1/2}\vee T^{-1/2})$ for some large constant C^0 in order to ensure that the last inequality of (B.121) holds. This implies $\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0) = O_P(N^{-1/2}\vee T^{-1/2})$, so the proof is complete. ■

Proof of Lemma 2.

We start with part (a). Suppose first that $d_1 = 0$, such that $d_f = 0$. In this case,

$$\frac{1}{NT}\sum_{i=1}^{N}\|\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}_0\|^2 = \frac{1}{NT}\sum_{i=1}^{N}\|\mathbf{X}_i\mathbf{D}_T\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0) + \boldsymbol{\varepsilon}_i\|^2$$
$$= \frac{1}{NT}\sum_{i=1}^{N}\|\boldsymbol{\varepsilon}_i\|^2 + o_P(1). \tag{B.122}$$

This implies $\tau_N \asymp 1/\ln(T\vee N)$, which is much larger than $\widehat{\lambda}_{1,1}/\widehat{\lambda}_{1,0}$. The result then follows immediately.

Suppose now instead that $d_1 > 0$. Straightforward algebra reveals that

$$\frac{1}{NT^{\nu_1}}\sum_{i=1}^{N}\|\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}_0\|^2 = \frac{1}{NT^{\nu_1}}\sum_{i=1}^{N}\|\mathbf{X}_i\mathbf{D}_T\mathbf{D}_T^{-1}(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0) + \mathbf{F}^0\boldsymbol{\gamma}_i^0 + \boldsymbol{\varepsilon}_i\|^2$$
$$= \frac{1}{NT^{\nu_1}}\sum_{i=1}^{N}\|\mathbf{F}_1^0\boldsymbol{\gamma}_{1,i}^0\|^2 + o_P(1)$$
$$= (\mathrm{vec}\,\boldsymbol{\Sigma}_{\Gamma_1^0})'\mathrm{vec}\,\boldsymbol{\Sigma}_{F_1^0} + o_P(1), \tag{B.123}$$

which together with Assumption 1 implies $\tau_N \asymp 1/\ln(T\vee N)$. Note that for $d = 1,\ldots,d_1$,

$$\frac{T^{\nu_1}}{\lambda_{1,d}} = \frac{T^{2\nu_1}}{\mathbf{h}_{1,d}^{0\prime}\mathbf{F}_1^{0\prime}\boldsymbol{\Sigma}_1^0\mathbf{F}_1^0\mathbf{h}_{1,d}^0} = \frac{NT^{2\nu_1}}{\mathbf{h}_{1,d}^{0\prime}\mathbf{F}_1^{0\prime}\mathbf{F}_1^0\boldsymbol{\Gamma}_1^{0\prime}\boldsymbol{\Gamma}_1^0\mathbf{F}_1^{0\prime}\mathbf{F}_1^0\mathbf{h}_{1,d}^0}$$
$$= \frac{1}{(T^{-\nu_1}\mathbf{h}_{1,d}^{0\prime}\mathbf{F}_1^{0\prime}\mathbf{F}_1^0)(N^{-1}\boldsymbol{\Gamma}_1^{0\prime}\boldsymbol{\Gamma}_1^0)(T^{-\nu_1}\mathbf{F}_1^{0\prime}\mathbf{F}_1^0\mathbf{h}_{1,d}^0)}$$
$$= \frac{1}{(T^{-(\nu_1+\delta)/2}\widehat{\mathbf{F}}_{1,d}^{0\prime}\mathbf{F}_1^0)(N^{-1}\boldsymbol{\Gamma}_1^{0\prime}\boldsymbol{\Gamma}_1^0)(T^{-(\nu_1+\delta)/2}\mathbf{F}_1^{0\prime}\widehat{\mathbf{F}}_{1,d}^0)}(1+o_P(1)) \asymp 1, \tag{B.124}$$

where the fourth equality follows from (B.32) and the last step is due to (B.28) and Assumption 1.

Further use of Lemma B.3, (B.123) and (B.124), we obtain

$$\frac{\widehat{\lambda}_{1,1}}{\widehat{\lambda}_{1,0}} = \frac{T^{-\nu_1}\widehat{\lambda}_{1,1}}{(NT^{\nu_1})^{-1}\sum_{i=1}^N \|\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}_0\|^2} \asymp 1, \tag{B.125}$$

$$\frac{\widehat{\lambda}_{1,d+1}}{\widehat{\lambda}_{1,d}} = \frac{T^{-\nu_1}\widehat{\lambda}_{1,d+1}}{T^{-\nu_1}\widehat{\lambda}_{1,d}} \asymp 1 \tag{B.126}$$

for $d = 1, \ldots, d_1 - 1$ and

$$\frac{\widehat{\lambda}_{1,d_1+1}}{\widehat{\lambda}_{1,d_1}} = O_P(T^{-(\nu_1-\nu_2)}). \tag{B.127}$$

For $d = d_1 + 1, \ldots, d_{max}$, we get

$$\frac{T^{-\nu_1}\widehat{\lambda}_{1,d}}{T^{-\nu_1}\widehat{\lambda}_{1,0}} = O_P(T^{-(\nu_1-\nu_2)}), \tag{B.128}$$

which is less than τ_N by (B.123) and Lemma B.3. The required result follows from this and the definition of $\widehat{d}_1$.

Each step in the sequential procedure of Step 2 introduces additional remainder terms that all converge to zero under Assumptions 3 and 4 by Lemmas B.3 and B.5. This proves (a).

Part (b) follows from the (rotational) consistency of $\widehat{\mathbf{F}}_g$ established as a part of the proofs of Lemmas B.3 and B.5. ■

Proof of Theorem 1.

By Lemma B.6, the rate of convergence given in Lemma 1 is not the best one possible. However, under Assumptions 1–6, we have

$$\|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0)\| = O_P((NT)^{-1/2} \vee \|\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}^0)\|) = O_P((NT)^{-1/2}), \tag{B.129}$$

which can be inserted back into (B.114), leading to

$$\begin{aligned} &\sqrt{NT}\mathbf{D}_T^{-1}\left[\widehat{\boldsymbol{\beta}}_0 + \left(\sum_{i=1}^N \mathbf{Z}_i(\widehat{\mathbf{F}})'\mathbf{Z}_i(\widehat{\mathbf{F}})\right)^{-1}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_i(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0) - \boldsymbol{\beta}^0\right] \\ &= \mathbf{B}(\widehat{\mathbf{F}})^{-1}\frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i - \sqrt{N}T^{-\nu_G/2}\mathbf{B}(\widehat{\mathbf{F}})^{-1}\overline{\mathbf{A}}_1 + o_P(1). \end{aligned} \tag{B.130}$$

Note that $\mathbf{Z}_i(\widehat{\mathbf{F}})$ is $\widehat{\mathbf{Z}}_i$ with $\widehat{a}_{ij}$ replaced by a_{ij}. We now show that the effect of the estimation of a_{ij}

is negligible. We begin by noting how

$$\frac{1}{NT}\sum_{i=1}^{N}[\mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\mathbf{Z}_i(\widehat{\mathbf{F}})\mathbf{D}_T - \mathbf{D}_T\widehat{\mathbf{Z}}_i'\widehat{\mathbf{Z}}_i\mathbf{D}_T]$$
$$= \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_T\widehat{a}_{ij} - \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_T a_{ij}. \tag{B.131}$$

Here,

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_T\widehat{a}_{ij}$$
$$= \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_T\sum_{g=1}^{G}(\mathbf{y}_j - \mathbf{X}_j\widehat{\boldsymbol{\beta}}_0 - \widehat{\mathbf{F}}_{-g}\widehat{\boldsymbol{\gamma}}_{-g,j})'\widehat{\mathbf{F}}_g$$
$$\times (\widehat{\mathbf{F}}_g'\widehat{\boldsymbol{\Sigma}}_g\widehat{\mathbf{F}}_g)^{-1}\widehat{\mathbf{F}}_g'(\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}_0 - \widehat{\mathbf{F}}_{-g}\widehat{\boldsymbol{\gamma}}_{-g,i})$$
$$= \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_T\sum_{g=1}^{G}[\mathbf{X}_j(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0) + \mathbf{F}_g^0\boldsymbol{\gamma}_{g,j}^0 + \mathbf{e}_{g,j} + \boldsymbol{\varepsilon}_j]'\widehat{\mathbf{F}}_g$$
$$\times T^{-\delta}\mathbf{V}_g^{-1}\widehat{\mathbf{F}}_g'[\mathbf{X}_j(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0) + \mathbf{F}_g^0\boldsymbol{\gamma}_{g,i}^0 + \mathbf{e}_{g,i} + \boldsymbol{\varepsilon}_i]$$
$$= \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_T\sum_{g=1}^{G}[\mathbf{X}_j(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0) + \mathbf{F}_g^0\boldsymbol{\gamma}_{g,j}^0 + \mathbf{e}_{g,j} + \boldsymbol{\varepsilon}_j]'\widehat{\mathbf{F}}_g$$
$$\times T^{-(\nu_g+\delta)}(T^{-\nu_g}\mathbf{V}_g)^{-1}\widehat{\mathbf{F}}_g'[\mathbf{X}_j(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0) + \mathbf{F}_g^0\boldsymbol{\gamma}_{g,i}^0 + \mathbf{e}_{g,i} + \boldsymbol{\varepsilon}_i](1 + o_P(1))$$
$$= \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_j\mathbf{D}_T a_{ij}(1 + o_P(1)), \tag{B.132}$$

where the second equality here follows from the definition of $\mathbf{V}_g^0$, which is again based on taking d_g as known, while the last equality follows from direct calculation. The effect of the estimation of a_{ij} in $\widehat{\mathbf{Z}}_i$ is therefore negligible, which in turn implies that the right-hand side of (B.130) becomes

$$\sqrt{NT}\mathbf{D}_T^{-1}\left[\widehat{\boldsymbol{\beta}}_0 + \left(\sum_{i=1}^{N}\mathbf{Z}_i(\widehat{\mathbf{F}})'\mathbf{Z}_i(\widehat{\mathbf{F}})\right)^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_i(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0) - \boldsymbol{\beta}^0\right]$$
$$= \sqrt{NT}\mathbf{D}_T^{-1}\left[\widehat{\boldsymbol{\beta}}_0 + \left(\sum_{i=1}^{N}\widehat{\mathbf{Z}}_i'\widehat{\mathbf{Z}}_i\right)^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\mathbf{X}_i(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_0) - \boldsymbol{\beta}^0\right] + o_P(1)$$
$$= \sqrt{NT}\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0) + o_P(1). \tag{B.133}$$

It follows that

$$\sqrt{NT}\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0) = \mathbf{B}(\widehat{\mathbf{F}})^{-1}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i - \sqrt{N}T^{-\nu_G/2}\mathbf{B}(\widehat{\mathbf{F}})^{-1}\overline{\mathbf{A}}_1$$
$$+ o_P(1). \tag{B.134}$$

Consider $(NT)^{-1/2}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i$. From the definition of $\mathbf{Z}_i(\mathbf{F})$,

$$\begin{aligned}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{Z}_i(\mathbf{F}^0)'\boldsymbol{\varepsilon}_i + \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{D}_T[\mathbf{Z}_i(\widehat{\mathbf{F}}) - \mathbf{Z}_i(\mathbf{F}^0)]'\boldsymbol{\varepsilon}_i \\ &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{Z}_i(\mathbf{F}^0)'\boldsymbol{\varepsilon}_i - \sqrt{NT}(\mathbf{R}_1 - \mathbf{R}_2), \qquad \text{(B.135)}\end{aligned}$$

where

$$\mathbf{R}_1 = \frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'(\mathbf{P}_{\widehat{F}} - \mathbf{P}_{F^0})\boldsymbol{\varepsilon}_i, \qquad \text{(B.136)}$$

$$\mathbf{R}_2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}a_{ij}\mathbf{D}_T\mathbf{X}_j'(\mathbf{P}_{\widehat{F}} - \mathbf{P}_{F^0})\boldsymbol{\varepsilon}_i. \qquad \text{(B.137)}$$

Here,

$$\begin{aligned}\mathbf{R}_1 &= \frac{1}{NT^{1+\delta}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'(\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H})\mathbf{H}'\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i + \frac{1}{NT^{1+\delta}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'(\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H})(\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H})'\boldsymbol{\varepsilon}_i \\ &\quad + \frac{1}{NT^{1+\delta}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{F}^0\mathbf{H}(\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H})'\boldsymbol{\varepsilon}_i + \frac{1}{NT^{1+\delta}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{F}^0[\mathbf{H}\mathbf{H}'^{\delta}(\mathbf{F}^{0\prime}\mathbf{F}^0)^{-1}]\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i \\ &= \mathbf{R}_{11} + \mathbf{R}_{12} + \mathbf{R}_{13} + \mathbf{R}_{14}, \qquad \text{(B.138)}\end{aligned}$$

with obvious implicit definitions of $\mathbf{R}_{11}, \ldots, \mathbf{R}_{14}$ and $\mathbf{H} = \text{diag}(\mathbf{H}_1, \ldots, \mathbf{H}_G)$. Let $\mathbf{R}_{1m,j}$ be the j-th row of $\mathbf{R}_{1m}$ for $m \in \{1, \ldots, 4\}$. In this notation,

$$\begin{aligned}\|\mathbf{R}_{11,j}\| &\le \left\|\frac{1}{NT^{\delta/2+1}}\sum_{i=1}^{N}(\boldsymbol{\varepsilon}_i'\mathbf{F}^0\mathbf{H}) \otimes (T^{-\kappa_j/2}\mathbf{X}_{j,i})'\right\| \|T^{-\delta/2}\text{vec}(\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H})\| \\ &= O_P((NT)^{-1/2})T^{-\delta/2}\|\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H}\| = o_P((NT)^{-1/2}), \qquad \text{(B.139)}\end{aligned}$$

where the equality follows from Lemma B.2 and the fact that $T^{(\nu_g-\delta)/2}\mathbf{H}_g = O_P(1)$. Similarly, for $\mathbf{R}_{12}$,

$$\begin{aligned}\|\mathbf{R}_{12,j}\| &\le \left\|\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i^{\prime-\kappa_j/2}\mathbf{X}_{j,i})'\right\| \|T^{-\delta}\text{vec}[(\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H})(\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H})']\| \\ &= O_P(N^{-1/2})T^{-\delta}\|\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H}\|^2 = o_P((NT)^{-1/2}), \qquad \text{(B.140)}\end{aligned}$$

where the second equality is due to (B.54).

Consider $\mathbf{R}_{14}$. From $T^{-\delta/2}\|\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H}\| = o_P(1)$ and $\|T^{(\nu_g-\delta)/2}\mathbf{H}_g\| = O_P(1)$, we obtain

$$\|\mathbf{I}_{d_f} - T^{\delta}(\mathbf{H}'\mathbf{F}^{0\prime}\mathbf{F}^0\mathbf{H})^{-1}\| = o_P(1), \qquad \text{(B.141)}$$

$$\|T^{\nu_g-\delta}\mathbf{H}_g\mathbf{H}_g'^{-\nu_g}\mathbf{F}_g^{0\prime}\mathbf{F}_g^0)^{-1}\| = o_P(1). \tag{B.142}$$

Together with Lemma B.2 this implies

$$\begin{aligned}
\|\mathbf{R}_{14,j}\| &= \left\| \frac{1}{NT^{1+\delta}} \sum_{i=1}^{N} T^{-\kappa_j/2}\mathbf{X}_{j,i}'\mathbf{F}^0[\mathbf{H}\mathbf{H}'^{\delta}\mathbf{H}(\mathbf{H}'\mathbf{F}^{0\prime}\mathbf{F}^0\mathbf{H})^{-1}\mathbf{H}']\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i \right\| \\
&= \left\| \frac{1}{NT^{1+\delta}} \sum_{i=1}^{N} T^{-\kappa_j/2}\mathbf{X}_{j,i}'\mathbf{F}^0\mathbf{H}[\mathbf{I}_{d_f} - T^{\delta}(\mathbf{H}'\mathbf{F}^{0\prime}\mathbf{F}^0\mathbf{H})^{-1}]\mathbf{H}'\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i \right\| \\
&\le \left\| \frac{1}{NT^{1+\delta}} \sum_{i=1}^{N} (\boldsymbol{\varepsilon}_i'\mathbf{F}^0\mathbf{H}) \otimes (T^{-\kappa_j/2}\mathbf{X}_{j,i}'\mathbf{F}^0\mathbf{H}) \right\| \|\mathbf{I}_{d_f} - T^{\delta}(\mathbf{H}'\mathbf{F}^{0\prime}\mathbf{F}^0\mathbf{H})^{-1}\| \\
&= O_P((NT)^{-1/2})\|\mathbf{I}_{d_f} - T^{\delta}(\mathbf{H}'\mathbf{F}^{0\prime}\mathbf{F}^0\mathbf{H})^{-1}\| = o_P((NT)^{-1/2}).
\end{aligned} \tag{B.143}$$

Finally, let us consider $\mathbf{R}_{13}$.

$$\begin{aligned}
\mathbf{R}_{13} &= \frac{1}{NT^{1+\delta}} \sum_{i=1}^{N} \mathbf{D}_T\mathbf{X}_i'\mathbf{F}^0\mathbf{H}(\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H})'\boldsymbol{\varepsilon}_i \\
&= \sum_{g=1}^{G} \frac{1}{NT^{1+\delta}} \sum_{i=1}^{N} \mathbf{D}_T\mathbf{X}_i'\mathbf{F}_g^0\mathbf{H}_g(\widehat{\mathbf{F}}_g - \mathbf{F}_g^0\mathbf{H}_g)'\boldsymbol{\varepsilon}_i \\
&= \sum_{g=1}^{G} \frac{1}{NT^{1+\delta}} \sum_{i=1}^{N} \mathbf{D}_T\mathbf{X}_i'\mathbf{F}_g^0\mathbf{H}_g\mathbf{H}_g'(\widehat{\mathbf{F}}_g\mathbf{H}_g^{-1} - \mathbf{F}_g^0)'\boldsymbol{\varepsilon}_i
\end{aligned} \tag{B.144}$$

Expanding $\widehat{\mathbf{F}}_g\mathbf{H}_g^{-1} - \mathbf{F}_g^0$ as we did for $\mathbf{L}$ above, we then just need to focus on the leading terms equivalent to $\mathbf{L}_2$, $\mathbf{L}_8$, and $\mathbf{L}_{10}$.

$$\begin{aligned}
\mathbf{R}_{13} &= \sum_{g=1}^{G} \frac{1}{NT^{1+\delta}} \sum_{i=1}^{N} \mathbf{D}_T\mathbf{X}_i'\mathbf{F}_g^0\mathbf{H}_g\mathbf{H}_g' \left(\frac{1}{NT^{(\nu_g+\delta)/2}} \sum_{j=1}^{N} \mathbf{X}_j(\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0)\boldsymbol{\gamma}_{g,j}^{0\prime}\mathbf{F}_g^{0\prime}\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1} \right)' \boldsymbol{\varepsilon}_i \\
&+ \sum_{g=1}^{G} \frac{1}{NT^{1+\delta}} \sum_{i=1}^{N} \mathbf{D}_T\mathbf{X}_i'\mathbf{F}_g^0\mathbf{H}_g\mathbf{H}_g' \left(\frac{1}{NT^{(\nu_g+\delta)/2}} \sum_{j=1}^{N} \boldsymbol{\varepsilon}_j\boldsymbol{\gamma}_{g,j}^{0\prime}\mathbf{F}_g^{0\prime}\widehat{\mathbf{F}}_g\mathbf{Q}_g^{-1} \right)' \boldsymbol{\varepsilon}_i \\
&+ \sum_{g=1}^{G} \frac{1}{NT^{1+\delta}} \sum_{i=1}^{N} \mathbf{D}_T\mathbf{X}_i'\mathbf{F}_g^0\mathbf{H}_g\mathbf{H}_g' \left(\frac{1}{\sqrt{T^{\nu_g+1}}}\boldsymbol{\Sigma}_{\varepsilon}\frac{\mathbf{F}_g^0}{T^{(\nu_g-1)/2}} \left(\frac{\mathbf{F}_g^{0\prime}\mathbf{F}_g^0}{T^{\nu_g}}\right)^{-1} \left(\frac{\boldsymbol{\Gamma}_g^{0\prime}\boldsymbol{\Gamma}_g^0}{N}\right)^{-1} \right)' \boldsymbol{\varepsilon}_i \\
&+ o_P\left(\frac{1}{\sqrt{NT}}\right) \\
&= \mathbf{R}_{131} + \mathbf{R}_{132} + \mathbf{R}_{133} + o_P\left(\frac{1}{\sqrt{NT}}\right).
\end{aligned}$$

It is easy to see that $\|\mathbf{R}_{131}\| = o_P(\|\boldsymbol{\beta}^0 - \widehat{\boldsymbol{\beta}}_0\|)$, so negligible. We now examine $\mathbf{R}_{133}$. Then we can write

$$\mathbf{R}_{133} = \sum_{g=1}^{G} \frac{1}{NT^{1+\nu_g}} \sum_{i=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{F}_g^0 \left(\frac{\mathbf{F}_g^{0\prime} \mathbf{F}_g^0}{T^{\nu_g}} \right)^{-1} \left(\frac{\mathbf{\Sigma}_\varepsilon}{\sqrt{T^{\nu_g+1}}} \frac{\mathbf{F}_g^0}{T^{(\nu_g-1)/2}} \left(\frac{\mathbf{F}_g^{0\prime} \mathbf{F}_g^0}{T^{\nu_g}} \right)^{-1} \left(\frac{\mathbf{\Gamma}_g^{0\prime} \mathbf{\Gamma}_g^0}{N} \right)^{-1} \right)' \boldsymbol{\varepsilon}_i$$

$$= \sum_{g=1}^{G} \frac{1}{T^{\nu_g+1}} \frac{1}{NT} \sum_{i=1}^{N} \mathbf{D}_T \mathbf{X}_i' \frac{\mathbf{F}_g^0}{T^{(\nu_g-1)/2}} \left(\frac{\mathbf{F}_g^{0\prime} \mathbf{F}_g^0}{T^{\nu_g}} \right)^{-1} \left(\frac{\mathbf{\Gamma}_g^{0\prime} \mathbf{\Gamma}_g^0}{N} \right)^{-1} \left(\frac{\mathbf{F}_g^{0\prime} \mathbf{F}_g^0}{T^{\nu_g}} \right)^{-1} \frac{\mathbf{F}_g^{0\prime}}{T^{(\nu_g-1)/2}} \mathbf{\Sigma}_\varepsilon \boldsymbol{\varepsilon}_i$$

$$= o_P\left(\frac{1}{\sqrt{NT}} \right), \tag{B.145}$$

where the first equality follows from (B.142), and the last step follows from some routine analysis using Assumption 3. Below, we investigate $\mathbf{R}_{132}$, which is one source of the bias term.

$$\mathbf{R}_{132} = \sum_{g=1}^{G} \frac{1}{NT^{1+\delta}} \sum_{i=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{F}_g^0 \mathbf{H}_g \mathbf{H}_g' \left(\frac{1}{NT^{(\nu_g+\delta)/2}} \sum_{j=1}^{N} \boldsymbol{\varepsilon}_j \boldsymbol{\gamma}_{g,j}^{0\prime} \mathbf{F}_g^{0\prime} \widehat{\mathbf{F}}_g \mathbf{Q}_g^{-1} \right)' \boldsymbol{\varepsilon}_i$$

$$= \sum_{g=1}^{G} \frac{1}{NT^{1+\nu_g}} \sum_{i=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{F}_g^0 (T^{\nu_g-\delta} \mathbf{H}_g \mathbf{H}_g') \left(\frac{1}{N} \sum_{j=1}^{N} \boldsymbol{\varepsilon}_j \boldsymbol{\gamma}_{g,j}^{0\prime} \left(\frac{\mathbf{\Gamma}_g^{0\prime} \mathbf{\Gamma}_g^0}{N} \right)^{-1} \right)' \boldsymbol{\varepsilon}_i$$

$$= \sum_{g=1}^{G} \frac{1}{NT^{1+\nu_g}} \sum_{i=1}^{N} \mathbf{D}_T \mathbf{X}_i' \mathbf{F}_g^0 \left(\frac{\mathbf{F}_g^{0\prime} \mathbf{F}_g^0}{T^{\nu_g}} \right)^{-1} \left(\frac{1}{N} \sum_{j=1}^{N} \boldsymbol{\varepsilon}_j \boldsymbol{\gamma}_{g,j}^{0\prime} \left(\frac{\mathbf{\Gamma}_g^{0\prime} \mathbf{\Gamma}_g^0}{N} \right)^{-1} \right)' \boldsymbol{\varepsilon}_i (1 + o_P(1))$$

$$= \sum_{g=1}^{G} \frac{1}{NT^{(\nu_g-1)/2}} \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\mathbf{D}_T \mathbf{X}_i' \mathbf{F}_g^0}{T^{(\nu_g+1)/2}} \left(\frac{\mathbf{F}_g^{0\prime} \mathbf{F}_g^0}{T^{\nu_g}} \right)^{-1} \left(\frac{\mathbf{\Gamma}_g^{0\prime} \mathbf{\Gamma}_g^0}{N} \right)^{-1}$$

$$\times \boldsymbol{\gamma}_{g,j}^0 \frac{\boldsymbol{\varepsilon}_j' \boldsymbol{\varepsilon}_i}{T} (1 + o_P(1)), \tag{B.146}$$

where the second equality follows from the definition of $\mathbf{Q}_g$, and the third equality follows from (B.142). Then we are able to further write

$$\sqrt{NT} \mathbf{R}_{32} = \sum_{g=1}^{G} \frac{T}{\sqrt{N} T^{\nu_g/2}} \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\mathbf{D}_T \mathbf{X}_i' \mathbf{F}_g^0}{T^{(\nu_g+1)/2}} \left(\frac{\mathbf{F}_g^{0\prime} \mathbf{F}_g^0}{T^{\nu_g}} \right)^{-1} \left(\frac{\mathbf{\Gamma}_g^{0\prime} \mathbf{\Gamma}_g^0}{N} \right)^{-1} \boldsymbol{\gamma}_{g,j}^0 \frac{\boldsymbol{\varepsilon}_j' \boldsymbol{\varepsilon}_i}{T} (1 + o_P(1))$$

$$= \frac{T}{\sqrt{N} T^{\nu_G/2}} \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\mathbf{D}_T \mathbf{X}_i' \mathbf{F}_G^0}{T^{(\nu_G+1)/2}} \left(\frac{\mathbf{F}_G^{0\prime} \mathbf{F}_G^0}{T^{\nu_G}} \right)^{-1} \left(\frac{\mathbf{\Gamma}_G^{0\prime} \mathbf{\Gamma}_G^0}{N} \right)^{-1}$$

$$\times \boldsymbol{\gamma}_{G,j}^0 \frac{\boldsymbol{\varepsilon}_j' \boldsymbol{\varepsilon}_i}{T} (1 + o_P(1)). \tag{B.147}$$

Hence, by adding the results,

$$\sqrt{NT} \mathbf{R}_1 = \sqrt{NT} (\mathbf{R}_{11} + \mathbf{R}_{12} + \mathbf{R}_{13} + \mathbf{R}_{14})$$

$$= T^{(2-\nu_G)/2} N^{-1/2} \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} T^{-(\nu_G+1)/2} \mathbf{D}_T \mathbf{X}_i' \mathbf{F}_G^0 (T^{-\nu_G} \mathbf{F}_G^{0\prime} \mathbf{F}_G^0)^{-1}$$

$$\times (N^{-1} \mathbf{\Gamma}_G^{0\prime} \mathbf{\Gamma}_G^0)^{-1} \boldsymbol{\gamma}_{G,j}^0 T^{-1} \boldsymbol{\varepsilon}_j' \boldsymbol{\varepsilon}_i + o_P(1) \tag{B.148}$$

$\sqrt{NT}\mathbf{R}_2$ can be evaluated in exactly the same way, and the limiting representation is analogous to the one given above for $\sqrt{NT}\mathbf{R}_1$ with $\mathbf{X}_i$ replaced by $-\sum_{j=1}^N \mathbf{X}_j a_{ij}$. Moreover, $\|\mathbf{B}(\widehat{\mathbf{F}})-\mathbf{B}(\mathbf{F}^0)\| = o_P(1)$. It follows that if we define

$$\overline{\mathbf{A}}_2 = \frac{1}{N}\sum_{i=1}^N\sum_{j=1}^N T^{-(\nu_G+1)/2}\mathbf{D}_T\mathbf{Z}_i(0)'\mathbf{F}_G^0(T^{-\nu_G}\mathbf{F}_G^{0\prime}\mathbf{F}_G^0)^{-1}(N^{-1}\mathbf{\Gamma}_G^{0\prime}\mathbf{\Gamma}_G^0)^{-1}\boldsymbol{\gamma}_{G,j}^0 T^{-1}\boldsymbol{\varepsilon}_j'\boldsymbol{\varepsilon}_i, \tag{B.149}$$

where $\mathbf{Z}_i(0) = \mathbf{X}_i - \sum_{j=1}^N \mathbf{X}_j a_{ij}$, then (B.135) reduces to

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i = \frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{D}_T\mathbf{Z}_i(\mathbf{F}^0)'\boldsymbol{\varepsilon}_i - T^{(2-\nu_G)/2}N^{-1/2}\overline{\mathbf{A}}_2 + o_P(1), \tag{B.150}$$

which in turn implies that (B.134) becomes

$$\begin{aligned}\sqrt{NT}\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0) &= \mathbf{B}(\mathbf{F}^0)^{-1}\frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{D}_T\mathbf{Z}_i(\mathbf{F}^0)'\boldsymbol{\varepsilon}_i - \sqrt{N}T^{-\nu_G/2}\mathbf{B}(\mathbf{F}^0)^{-1}\overline{\mathbf{A}}_1 \\ &\quad - T^{(2-\nu_G)/2}N^{-1/2}\mathbf{B}(\mathbf{F}^0)^{-1}\overline{\mathbf{A}}_2 + o_P(1).\end{aligned} \tag{B.151}$$

The required result is now implied by Assumption 6. ■

C Conditions that ensure asymptotic unbiasedness

In this section, we provide a set of assumptions that ensure that the asymptotic distribution of $\sqrt{NT}\mathbf{D}_T(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0)$ given in Theorem 1 is free of bias without for that matter requiring that $\varepsilon_{i,t}$ is serially and cross-sectionally independent. One way to accomplish this is to assume that $\rho_1 = \rho_2 = 0$, as in Corollary 1. The assumptions considered here, which are stated in Assumption C.1, can be seen as alternatives to this last condition. In terms of the notation of Theorem 1, they ensure that $\mathbf{A}_1 = \mathbf{A}_2 = \mathbf{0}_{d_x\times 1}$.

Assumption C.1 (No asymptotic bias)**.** One of the following set of conditions is met:

(a) $T^{1-\nu_g}\mathbb{E}(\mathbf{f}_{g,t}^{0\prime}\mathbf{f}_{g,s}^0|\mathbf{x}_t,\mathbf{x}_s) = \phi_{ts}$ w.p.a.1 and $\sum_{t=1}^T\sum_{s=1}^T|\phi_{ts}| = O(T)$, where $\mathbf{x}_t = (\mathbf{x}_{1,t},\dots,\mathbf{x}_{N,t})'$. If $G\geq 2$, then $q < (\nu_G+\nu_{G-1})/2 - 1/4$ and $\nu_{g-1}-\nu_g > 1/2$ for $g = 2,\dots,G$.

(b) $T/N \to c_4 \in (0,\infty)$ and $\nu_G > 1$. If $G \geq 2$, then $q < (\nu_G+\nu_{G-1}-1)/2$ and $\nu_{g-1}-\nu_g > 1/2$.

(c) $T^{1-\nu_g-\kappa_j}\mathbb{E}(\sum_{t=1}^T\sum_{s=1}^T x_{j,i,t}x_{j,k,s}\mathbf{f}_{g,t}^{0\prime}\mathbf{f}_{g,s}^0) = O(T^{2-r})$, where $r<2$, $r+\nu_G-1>0$ and $x_{j,i,t}$ is the j-th row of $\mathbf{x}_{i,t}$. If $G\geq 2$, then $\nu_{g-1}-\nu_g > 1/2$.

Assumption C.1 ensures that the asymptotic distribution of $\sqrt{NT}\mathbf{D}_T(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0)$ is bias-free. It is, however, not necessary, and (a)–(b) should therefore be viewed as examples of conditions under which there is no asymptotic distribution bias. These conditions all have their strengths and weaknesses, and so their suitability will in general depend on the context. Take as an example condition (b), which has the advantage of not requiring any more moment conditions than those that are already in Assumption 1. It does, however, require that $\nu_G > 1$, which rules out both signal-weak and stationary factors ($\nu_G \leq 1$). Conditions (a) and (c) are more general in this regard, but then at the expense of requiring additional moment conditions. Condition (c) is a functional central limit theorem style moment condition.

The next corollary to Theorem 1 verifies that the asymptotic distribution of $\sqrt{NT}\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0)$ is indeed bias-free under Assumption C.1.

Corollary C.1 (Unbiased asymptotic distribution)**.** *Suppose that Assumptions 1–4, 6, and C.1 are met and that $N/T^{\nu_G} \to 0$. Then, as $N, T \to \infty$,*

$$\sqrt{NT}\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0) \to_D MN(\mathbf{0}_{d_x \times 1}, \mathbf{B}_0^{-1}\boldsymbol{\Omega}\mathbf{B}_0^{-1}).$$

Proof: Under $NT^{-\nu_G} \to 0$, (B.134) in the proof of Theorem 1 may be written as

$$\sqrt{NT}\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0) = \mathbf{B}(\mathbf{F}^0)^{-1}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{Z}_i(\widehat{\mathbf{F}})'\boldsymbol{\varepsilon}_i + o_P(1).$$

Consider $(NT)^{-1}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_i$, which we can write as

$$\begin{aligned}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{\widehat{F}}\boldsymbol{\varepsilon}_i &= \frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{F^0}\boldsymbol{\varepsilon}_i - \frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'(\mathbf{P}_{\widehat{F}} - \mathbf{P}_{F^0})\boldsymbol{\varepsilon}_i \\ &= \frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{M}_{F^0}\boldsymbol{\varepsilon}_i - \mathbf{R}_1. \end{aligned} \tag{C.1}$$

As in the proof of Theorem 1,

$$\begin{aligned}\mathbf{R}_1 &= \frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'^{-\delta}\widehat{\mathbf{F}}\widehat{\mathbf{F}}' - \mathbf{P}_{F^0})\boldsymbol{\varepsilon}_i \\ &= \frac{1}{NT^{1+\delta}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'(\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H})\mathbf{H}'\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i + \frac{1}{NT^{1+\delta}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'(\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H})(\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H})'\boldsymbol{\varepsilon}_i \\ &+ \frac{1}{NT^{1+\delta}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{F}^0\mathbf{H}(\widehat{\mathbf{F}} - \mathbf{F}^0\mathbf{H})'\boldsymbol{\varepsilon}_i + \frac{1}{NT^{1+\delta}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'\mathbf{F}^0[\mathbf{H}\mathbf{H}' - (\mathbf{F}^{0\prime}\mathbf{F}^0)^{-1}]\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i \\ &= \mathbf{R}_{11} + \mathbf{R}_{12} + \mathbf{R}_{13} + \mathbf{R}_{14}, \end{aligned} \tag{C.2}$$

where $\|\mathbf{R}_{11}\|$, $\|\mathbf{R}_{12}\|$ and $\|\mathbf{R}_{14}\|$ are all $o_P((NT)^{-1/2})$, just as before. Let us therefore consider $\mathbf{R}_{13}$, which is the source of the bias in the asymptotic distribution of $\sqrt{NT}\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0)$. The purpose of Assumption C.1 is to control this term. Suppose first that condition (b) holds. Let $x_{j,k,t}$ denote the

j-th row of $\mathbf{x}_{k,t}$. In this notation,

$$\begin{aligned}
&\mathbb{E}\left\|\frac{1}{NT^{\delta/2+1}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i'\otimes(T^{-\kappa_j/2}\mathbf{X}_{j,i}'\mathbf{F}_g^0T^{-(\nu_g-\delta)/2})\right\|^2\\
&=\frac{1}{N^2T^{\delta+2}}\sum_{s=1}^{T}\mathbb{E}\left\|\sum_{i=1}^{N}\varepsilon_{i,s}\sum_{t=1}^{T}T^{-\kappa_j/2}x_{j,i,t}T^{-(\nu_g-1)/2}\mathbf{f}_{g,t}^{0\prime}T^{(\nu_g-1)/2}T^{-(\nu_g-\delta)/2}\right\|^2\\
&=\frac{1}{(NT)^2}\sum_{i=1}^{N}\sum_{k=1}^{N}\mathbb{E}\left[\left(\sum_{t=1}^{T}T^{-\kappa_j/2}x_{j,i,t}T^{-(\nu_g-1)/2}\mathbf{f}_{g,t}^{0\prime}\right)\left(\sum_{t=1}^{T}T^{-(\nu_g-1)/2}\mathbf{f}_{g,t}^{0}T^{-\kappa_j/2}x_{j,k,t}\right)\right]\sigma_{\varepsilon,ik}\\
&=O(N^{-1}). \qquad \text{(C.3)}
\end{aligned}$$

Moreover, by applying Assumption 6 (b) to the expression given for $\|T^{-\delta/2}\widehat{\mathbf{F}}_{2,d}-T^{-\nu_2/2}\mathbf{F}_2^0\mathbf{h}_{2,d}^0\|$ in the proof of Lemma B.5, we can show that

$$\|T^{-\delta/2}\text{vec}(\widehat{\mathbf{F}}-\mathbf{F}^0\mathbf{H})'\|=o_P(T^{-1/2}). \qquad \text{(C.4)}$$

Making use of these results, we obtain

$$\begin{aligned}
\|\mathbf{R}_{3,j}\|&\leq\left\|\frac{1}{NT^{\delta/2+1}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i'^{-\kappa_j/2}\mathbf{X}_{j,i}'\mathbf{F}^0\mathbf{H})\right\|\|T^{-\delta/2}\text{vec}(\widehat{\mathbf{F}}-\mathbf{F}^0\mathbf{H})'\|\\
&=o_P((NT)^{-1/2}). \qquad \text{(C.5)}
\end{aligned}$$

Alternatively, we may invoke Assumption 6 (a) to arrive at the same result. In this case, $\|T^{-\delta/2}\text{vec}(\widehat{\mathbf{F}}-\mathbf{F}^0\mathbf{H})'\|=o_P(1)$, but we also have

$$\begin{aligned}
&\mathbb{E}\left\|\frac{1}{NT^{\delta/2+1}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i'\otimes(T^{-\kappa_j/2}\mathbf{X}_{j,i}'\mathbf{F}_g^0T^{-(\nu_g-\delta)/2})\right\|^2\\
&=\frac{1}{(NT)^2}\sum_{i=1}^{N}\sum_{k=1}^{N}\mathbb{E}\left[\left(\sum_{t=1}^{T}T^{-\kappa_j/2}x_{j,i,t}T^{-(\nu_g-1)/2}\mathbf{f}_{g,t}^{0\prime}\right)\left(\sum_{t=1}^{T}T^{-(\nu_g-1)/2}\mathbf{f}_{g,t}^{0}T^{-\kappa_j/2}x_{j,k,t}\right)\right]\sigma_{\varepsilon,ik}\\
&=\frac{1}{(NT)^2}\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{k=1}^{N}\mathbb{E}[T^{-\kappa_j}x_{j,i,t}x_{j,k,t}T^{-(\nu_g-1)}\mathbb{E}(\|\mathbf{f}_{g,t}^0\|^2|\boldsymbol{\varepsilon}_t)]\sigma_{\varepsilon,ik}\\
&+\frac{2}{(NT)^2}\sum_{t=2}^{T}\sum_{s<t}\sum_{i=1}^{N}\sum_{k=1}^{N}\mathbb{E}[T^{-\kappa_j}x_{j,i,t}x_{j,k,s}T^{-(\nu_g-1)}\mathbb{E}(\mathbf{f}_{g,t}^{0\prime}\mathbf{f}_{g,s}^{0}|\boldsymbol{\varepsilon}_t,\boldsymbol{\varepsilon}_s)]\sigma_{\varepsilon,ik}\\
&=O(1)\frac{1}{(NT)^2}\sum_{t=2}^{T}\sum_{s<t}\sum_{i=1}^{N}\sum_{k=1}^{N}|\phi_{ts}||\sigma_{\varepsilon,ik}|=O((NT)^{-1}), \qquad \text{(C.6)}
\end{aligned}$$

and so $\|\mathbf{R}_{3,j}\|$ is of the same order as before. The proof under Assumption 6 (c) is simpler and is therefore omitted.

Hence, by adding the results,

$$\sqrt{NT}\|\mathbf{R}_1\| \leq \sqrt{NT}(\|\mathbf{R}_{11}\| + \|\mathbf{R}_{12}\| + \|\mathbf{R}_{13}\| + \|\mathbf{R}_{14}\|) = o_P(1). \tag{C.7}$$

We have therefore shown that

$$\sqrt{NT}\left\|\frac{1}{NT}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{X}_i'(\mathbf{P}_{\widehat{F}} - \mathbf{P}_{F^0})\boldsymbol{\varepsilon}_i\right\| = o_P(1), \tag{C.8}$$

and we can similarly show that

$$\sqrt{NT}\left\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{D}_T\mathbf{X}_i'(\mathbf{P}_{\widehat{F}} - \mathbf{P}_{F^0})\mathbf{X}_j\mathbf{D}_T a_{ij}\right\| = o_P(1). \tag{C.9}$$

These results can be inserted into (B.134), giving

$$\sqrt{NT}\mathbf{D}_T^{-1}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0) = \mathbf{B}(\mathbf{F}^0)^{-1}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{D}_T\mathbf{Z}_i(\mathbf{F}^0)'\boldsymbol{\varepsilon}_i + o_P(1). \tag{C.10}$$

The sought result now follows from Assumption 6. ■

The asymptotic distribution in Corollary C.1 is the same as the one given in Corollary 1.

If ρ_1, ρ_2, $\mathbf{A}_1$ and $\mathbf{A}_2$ are all different from zero, one possibility is to use bias correction. Dhaene and Jochmans (2015) were first to bring attention to the relevance of the half-panel jacknife approach for bias correction in panel data. These authors focus on the fixed effects case, but Chen et al. (2021) have shown that the Jackknife can be used to correct for bias also in models with interactive effects. In our setting, the value of ν_G is unknown, and therefore the standard Jackknife is not directly applicable. We thus turn to the hybrid half-panel Jackknife correction proposed by Fernández-Val and Weidner (2018). The bias-corrected version of the IPC estimator is given by

$$\widehat{\boldsymbol{\beta}}_{\text{bc}} = 3\widehat{\boldsymbol{\beta}} - \frac{1}{2}(\widehat{\boldsymbol{\beta}}_{1,N} + \widehat{\boldsymbol{\beta}}_{2,N} + \widehat{\boldsymbol{\beta}}_{1,T} + \widehat{\boldsymbol{\beta}}_{2,T}), \tag{C.11}$$

where $\widehat{\boldsymbol{\beta}}$ is the standard IPC estimator defined in (9), $\widehat{\boldsymbol{\beta}}_{1,N}$ and $\widehat{\boldsymbol{\beta}}_{2,N}$ are defined in the same way but applied to cross-sectional units $\{1, \ldots, \lfloor N/2 \rfloor\}$ and $\{\lfloor N/2 \rfloor + 1, \ldots, N\}$, respectively, and $\widehat{\boldsymbol{\beta}}_{1,T}$ and $\widehat{\boldsymbol{\beta}}_{2,T}$ are the IPC estimators applied to odd and even numbered time periods, respectively. The splitting in even and odd numbered time periods is needed in order to account for the behaviour of the factors. See Fernández-Val and Weidner (2018) for a more detailed discussion of the hybrid Jackknife approach.

D Empirical illustration

Economists have become concerned that recently house prices have grown too quickly, and that prices are now too high relative to per capita incomes. If this is correct and there is any truth to the theory on the matter, prices should stagnate or fall until they are better aligned with income, which in statistical terms mean that house prices should be cointegrated with income. The validity of this assumption has important implications for policy, because a failure could be due to a housing bubble.

In this section, we revisit the real house price data set of Holly et al. (2010), which comprises data on log real house prices ($p_{i,t}$) and log real per capita income ($w_{i,t}$) for 49 US states across the 1975-2003 period. According to the economic theory, $p_{i,t}$ and $w_{i,t}$ should be cointegrated with cointegrating vector $(1, -1)'$. The previous empirical evidence of this prediction has, however, been mixed and far from convincing (see, for example, Gallin, 2006). Holly et al. (2010) argue that this lack of empirical support can be attributed in part to a failure to account for cross-sectional dependence, leading to deceptive conclusions. The authors therefore apply the "CIPS" panel unit root test of Pesaran (2007), which allows for cross-section dependence in the form of a common factor. The test is applied both to $p_{i,t}$ and $w_{i,t}$ separately, and to $p_{i,t} - w_{i,t}$. According to the results, while the variables are unit root non-stationary, their difference is not. Holly et al. (2010) also report CCE results suggesting that the estimated income elasticity is indeed close to one. They therefore conclude that $p_{i,t}$ and $w_{i,t}$ are cointegrated with cointegrating vector $(1, -1)'$, just as predicted by the theory.

Our interest in the work of Holly et al. (2010) stems from their preference to apply the CIPS test, which tests for a unit root in the defactored data. This means that if $p_{i,t}$ and $w_{i,t}$ are not cointegrated by themselves but only when conditioning on unit root common factors, because of the way that the data are defactored prior to the testing, the unit root null hypothesis is likely to be rejected by the CIPS test. That is, the test is likely to lead to the conclusion of cointegration when in fact there is none. In this section, we use IPC as a means to investigate this possibility.

Insert Figure D.1 about here

We begin by plotting the variables. This is done in Figure D.1. As expected, both variables are highly persistent and the ADF test provides no evidence against the unit root null. This corroborates the unit root test results reported by Holly et al. (2010). However, we also see that the trending behaviors of $p_{i,t}$ and $w_{i,t}$ are very different, suggesting that their stochastic trends are not the same, which they should be under cointegration. We also see that the trending behavior is very similar across states, which is suggestive of non-stationary common factors. Of course, the IPC procedure does not require cointegration and it does allow for very general types of factors. We therefore proceed with the estimation of the model.

Insert Table D.1 and Figure D.2 about here

In the usual notation, $y_{i,t} = p_{i,t}$ and $\mathbf{x}_{i,t} = w_{i,t}$ in this illustration. For comparison purposes, the IPC results are presented together with the results obtained by applying the PC estimator of Bai (2009), as well as the usual OLS estimator with time and state fixed effects. The results are reported in Table D.1. The first thing to note is that the estimated slopes vary substantially depending on the estimation methods used. Interestingly, the point estimates are increasing in the generality of the estimator with fixed effects OLS (IPC) leading to the lowest (highest) estimate. We therefore begin by considering the IPC results. The point estimate of 2.1024 is far from the theoretically predicted value of one, which is also not included in the reported 95% confidence interval.

As for the factors, we estimate $\widehat{d}_1 = \widehat{d}_2 = \widehat{d}_3 = 1$, implying that the estimated number of factors is equal to $\widehat{d}_f = \widehat{d}_1 + \widehat{d}_2 + \widehat{d}_3 = 3$ and that the estimated number of groups equals $\widehat{G} = 3$. The estimated factors based on setting $\delta = 1$, denoted $\widehat{f}_{1,t}$, $\widehat{f}_{2,t}$ and $\widehat{f}_{3,t}$, are plotted in Figure D.2. As is well known in the literature, and as we formally prove in Section B, $\widehat{\mathbf{F}}$ is not consistent for $\mathbf{F}^0$ itself but only for a certain rotation of $\mathbf{F}^0$. This means that while the level and sign cannot be interpreted, we can look at the behavior of $\widehat{f}_{1,t}$, $\widehat{f}_{2,t}$ and $\widehat{f}_{3,t}$ over time. As expected given Figure D.1, all three factors are highly persistent, and $\widehat{f}_{1,t}$ and $\widehat{f}_{2,t}$ are even trending, which we take as evidence against cointegration between $p_{i,t}$ and $w_{i,t}$, since under cointegration the factors should be stationary. We also note that $\widehat{f}_{1,t}$ and $\widehat{f}_{2,t}$ look rather similar even though we estimate $\widehat{G} = 3$ distinct groups, which could be due to over-specification of the number of groups. Of course, since in this paper $\nu_1, \cdots, \nu_G$ are not necessarily well separated integers, it may be that in this application ν_1 and ν_2 are similar yet different enough for the IPC procedure to put $\widehat{f}_{1,t}$ and $\widehat{f}_{2,t}$ in different groups. As a measure of the difference between ν_1 and ν_2, we look at $\sum_{i=1}^{N} \widehat{\gamma}_{1,i}^2 / \sum_{i=1}^{N} \widehat{\gamma}_{2,i}^2$. By using the results provided in Section B, we can show that this ratio should be $O_P(T^{\nu_1 - \nu_2})$, implying $\ln(\sum_{i=1}^{N} \widehat{\gamma}_{1,i}^2 / \sum_{i=1}^{N} \widehat{\gamma}_{2,i}^2)/\ln T = \nu_1 - \nu_2 + o_P(1)$. By plugging in the known values of T and $\sum_{i=1}^{N} \widehat{\gamma}_{1,i}^2 / \sum_{i=1}^{N} \widehat{\gamma}_{2,i}^2$, we get $\ln(\sum_{i=1}^{N} \widehat{\gamma}_{1,i}^2 / \sum_{i=1}^{N} \widehat{\gamma}_{2,i}^2)/\ln T \approx 0.41$, which is thus an estimate of $\nu_1 - \nu_2$. There is therefore a clear difference in the degree of trending of the two factors, which we take as evidence in support of the three-group IPC estimate. By repeating this exercise, we estimate $\nu_2 - \nu_3$ to $\ln(\sum_{i=1}^{N} \widehat{\gamma}_{2,i}^2 / \sum_{i=1}^{N} \widehat{\gamma}_{3,i}^2)/\ln T \approx 0.73$. Hence, as expected given Figure D.1, $\widehat{f}_{2,t}$ is more similar to $\widehat{f}_{2,t}$ than to $\widehat{f}_{3,t}$.

Because we estimate three time-varying factors, fixed effects OLS is invalid as it only allows for a common time effect. Moreover, since the factors come from three distinct groups, and are not all stationary, PC is invalid, too. This leaves us with the IPC estimator, which again provides strong evidence against the theoretically predicted one-to-one cointegrated relationship between $p_{i,t}$ and $w_{i,t}$. Holly et al. (2010, page 172) conclude that "[o]ur results support the hypothesis that real house prices have been rising in line with fundamentals (real incomes), and there seems little evidence of house price bubbles at the national level." The results reported here reveal a completely different picture with housing prices being long run disconnected with real income.

References

Bai, J. (2009), 'Panel data models with interactive fixed effects', *Econometrica* **77**, 1229–1279.

Chen, M., Fernández-Val, I. and Weidner, M. (2021), 'Nonlinear factor models for network and panel data', *Journal of Econometrics* **220**, 296–324.

Dhaene, G. and Jochmans, K. (2015), 'Split-panel Jackknife Estimation of Fixed-effect Models', *The Review of Economic Studies* **82**(3), 991–1030.

Fernández-Val, I. and Weidner, M. (2018), 'Fixed effects estimation of large-t panel data models', *Annual Review of Economics* **10**, 109–138.

Gallin, J. (2006), 'The long run relationship between house prices and income: Evidence from local housing markets', *Real Estate Economics* **34**, 417–438.

Holly, S., Pesaran, M. H. and Yamagata, T. (2010), 'A spatio-temporal model of house prices in the USA', *Journal of Econometrics* **158**, 160–173.

Jiang, B., Yang, Y., Gao, J. and Hsiao, C. (2021), 'Recursive estimation in large panel data models: Theory and practice', *Journal of Econometrics* **224**, 439–465.

Lam, C., Yao, Q. and Bathia, N. (2011), 'Estimation of latent factors for high-dimensional time series', *Biometrika* **98**, 901–918.

Magnus, J. R. and Neudecker, H. (2007), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, third edn, John Wiley & Sons Ltd.

Pesaran, M. H. (2007), 'A simple panel unit root test in the presence of cross section dependence', *Journal of Applied Econometrics* **22**, 265–312.

Figure D.1: Plots of the variables.

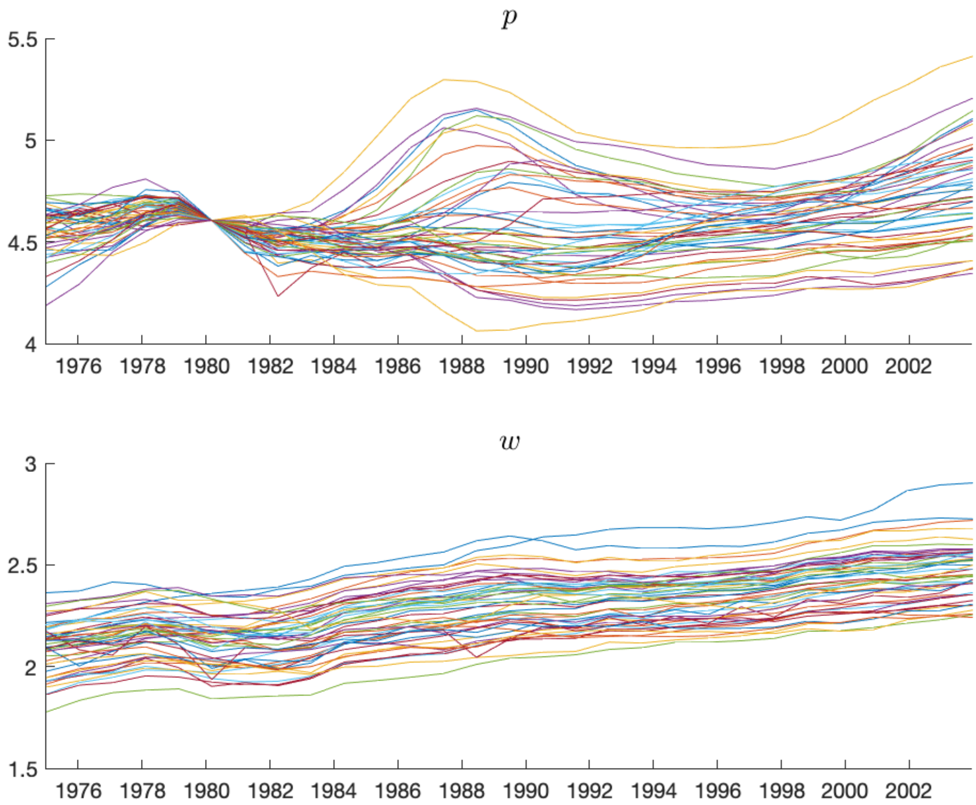

Table D.1: Empirical estimation results.

Estimator	Point estimate	Wald p-value
FE	0.3453	0.0000
PC	1.1602	0.0000
IPC	2.1024	0.0000

Notes: "FE", "PC" and "IPC" refer to the two-way fixed effects OLS estimator, the PC estimator of Bai (2009) and the proposed IPC estimator, respectively, in a regression of $p_{i,t}$ onto $w_{i,t}$. The reported Wald p-values test the null hypothesis that the relevant coefficient is zero.

Figure D.2: Plots of the estimated factors.