Heterogeneous Panel Data Models with Cross-Sectional Dependence

Article *in* SSRN Electronic Journal · September 2017

DOI: 10.2139/ssrn.3033430

2 authors, including:

Jiti Gao
Monash University (Australia)

Heterogeneous Panel Data Models with Cross–Sectional Dependence

JITI GAO[†] AND KAI XIA[‡] *

Monash University, Australia[†]

and Xiamen University, China[‡]

September 7, 2017

Abstract

This paper considers a semiparametric panel data model with heterogeneous coefficients and individual-specific trending functions, where the random errors are assumed to be serially correlated and cross-sectionally dependent. We propose mean group estimators for the coefficients and trending functions involved in the model. It can be shown that the proposed estimators can achieve an asymptotic consistency with rates of root$-NT$ and root$-NTh$, respectively as $(N, T) \to (\infty, \infty)$, where N is allowed to increase faster than T. Furthermore, a statistic for testing homogeneous coefficients is constructed based on the difference between the mean group estimator and a pooled estimator. Its asymptotic distributions are established under both the null and a sequence of local alternatives, even if the difference between these estimators vanishes considerably fast (can achieve root-NT^2 rate at most under the null) and consistent estimator available for the covariance matrix is not required explicitly. The finite sample performance of the proposed estimators together with the size and local power properties of the test are demonstrated by simulated data examples, and an empirical application with the OECD health care expenditure dataset is also provided.

Key words: Health care expenditure, Nonlinear trending function, Nonstationary time series

JEL Classification Numbers: C14; C22; G17.

---

*The authors would like to thank the Australian Research Council Discovery Grants Program for its support under Grant numbers: DP150101012 and DP170104421. *Corresponding author*: Jiti Gao, Department of Econometrics and Business Statistics, Monash University, Caulfield East, Victoria 3145, Australia. Email: jiti.gao@monash.edu.

1 Introduction

Unobserved heterogeneity is a pervasive feature among microeconomic individual responses as suggested by Heckman (2001), see also Durlauf et al. (2001) for an example of multi-country studies. There has been a rich literature introducing such heterogeneity into panel data models, with different parsimonious structures to achieve more accurate inference. For example, Pesaran (2006) considers panel data models with i.i.d. random coefficients across individuals, where the common population mean is estimated with root-N consistency. Lin and Ng (2012) and Su et al. (2014) consider group-specific slope coefficients with unknown group structure, where accuracy is improved by pooling within each group. Boneva et al. (2015) impose the heterogeneous covariate functions with a common component structure, where the basis functions can be estimated with a root-NTh rate, also see Vogt and Linton (2015) for a nonparametric panel data model with a group structure.

In addition to the individual response, the underlying time-varying feature, such as technology growth or climate change captured by trending functions (see Robinson, 2012; Chen et al., 2012), can also vary across individuals. Zhang et al. (2012) propose a nonparametric test for common trends assumption as in Chen et al. (2012) and find it rejected in both empirical examples. Recently, Bonhomme and Manresa (2015) extend the conventional fixed effects into group-specific time-varying effects. Su et al. (2015) consider a group-specific time-varying coefficient model where the residuals are assumed to be cross-sectionally independent.

In this paper, instead of imposing restrictions on the heterogeneity structure, we utilize the panel data information by using the weighted averaging or mean-group estimators based on individual regressions. In particular, we consider both the slope coefficients and the time-varying trends to be heterogeneous across sections. In detail, we consider the following semiparametric panel data model:

$$Y_{it} = X_{it}^{\top}\beta_i + f_{it} + \alpha_i + e_{it}, \quad \text{for} \quad 1 \le i \le N, \quad 1 \le t \le T \tag{1.1}$$

for $i \ge 1$, where β_i is a $d \times 1$ vector of unknown parameters, $f_{it} = f_i(\tau_t)$, with $\tau_t = t/T$ is a trending function, both β_i and $f_i(\tau)$ are allowed to be heterogeneous across sections to account for unobserved heterogeneity caused by socioeconomic and demographic difference, and both of them are assumed to be nonrandom[1], α_i is the fixed–effects allowed to be correlated with X_{it} with possibly unknown structure, X_{it} is a $d \times 1$ covariate vector satisfies $X_{it} = g_{it} + x_i + v_{it}$, in which g_{it} is a trending function like f_{it}, thus both X_{it} and Y_{it} can be non-stationary because of these unobservable trending functions, x_i is an individual specific effect, and X_{it} is allowed to be correlated with e_{it} through x_i, and e_{it} and v_{it} are assumed to be serially correlated and cross-sectionally dependent. In addition, N and T both can tend to infinity, where N is allowed

[1] Hsiao (2014) suggests that fixed and different β_i is more appropriate if it is correlated with explanatory variables or comes from heterogeneous population. Besides, we can show that the main results in this paper still holds for some random coefficients structures.

to grow faster than T.

The estimation of interest in model Eqs. (1.1) will be β_i, $f_i(\tau)$ for $i \geq 1$ and $0 \leq \tau \leq 1$, and their weighted averages $\overline{\beta} = \frac{1}{N}\sum_{i=1}^N \beta_i$, $\overline{f}(\tau) = \frac{1}{N}\sum_{i=1}^N f_i(\tau)$ for $0 \leq \tau \leq 1$. As Heckman (2001) concludes, while the "representative agent" paradigm is shown to lack empirical support, the "average" person becomes as a popular alternative. To estimate such average effects, we calculate mean group estimators based on separated semiparametric profile likelihood estimations. The resulting estimators $\widehat{\beta}_{mg}$ achieves a root-NT consistency to $\overline{\beta}$ and $\widehat{f}_{mg}(\tau)$ achieves a root-NTh consistency to $\overline{f}(\tau)$. In particular, if $\beta_i = (\beta_1^\top, \beta_{2,i}^\top)^\top$, we have $\widehat{\beta}_{1,mg} \to \beta_1$ and $\widehat{\beta}_{2,mg} \to \overline{\beta}_2$ with root-NT rate, which may not holds for the pooled estimators($\widehat{\beta}_p$). Similarly, such an approach could also be useful when practitioners are interested in average effects of particular groups or classes divided by gender, geographic location or observed socioeconomic indicators.

A further topic discussed in this paper is to test whether all the $\beta_i's$ in Eqs. (1.1) are homogeneous or not where f_{it} can be either homogeneous or heterogeneous. We propose an easily applicable test statistic based on $(\widehat{\beta}_{mg} - \widehat{\beta}_p)$. The test statistic proposed can be written as

$$l_{NT}^2(\widehat{\beta}_{mg} - \widehat{\beta}_p)^\top \Sigma_\beta^{-1}(\widehat{\beta}_{mg} - \widehat{\beta}_p) \tag{1.2}$$

where Σ_β is the covariance matrix involved in the asymptotic distribution of $l_{NT}(\widehat{\beta}_{mg} - \widehat{\beta}_p)$, in which $l_{NT} \to \infty$ is a sequence of real numbers.

One difficulty with such approach is that the difference $(\widehat{\beta}_{mg} - \widehat{\beta}_p)$ can be too small under both the null and the alternative. Under the null, for the homoscedasticity case, the asymptotic variances of $\widehat{\beta}_{mg}$ and $\widehat{\beta}_p$ are identical, l_{NT} must tend to infinity faster than root-NT rate. Any additional assumption imposed might be too restrictive (to rule out the homoscedasticity case) and can still be criticized for possible poor power performance under some alternatives(Pesaran and Yamagata, 2008). Secondly, there is an indeterminacy in l_{NT} since $(\widehat{\beta}_{mg} - \widehat{\beta}_p)$ converges with root-NT rate if there is enough heteroskedasticity in X_{it} and converges faster than root-NT rate, otherwise. Finally, the choice of an asymptotic covariance matrix for $\widehat{\beta}_p$ has already become a difficult task in the literature (see Kim and Sun, 2013, for a recent study). As a consequence, test statistics may not be feasible since Σ_β involved are too complicated.

To solve the above difficulties, we make some improvements to the conventional approach. On the one hand, there is no restriction imposed on the asymptotic variances of $\widehat{\beta}_{mg}$ and $\widehat{\beta}_p$ under the null. For example, if $\widehat{\beta}_{mg}$ and $\widehat{\beta}_p$ share the same asymptotic distribution, the test statistic is constructed using the asymptotic distribution of $\sqrt{NT^2}(\widehat{\beta}_{mg} - \widehat{\beta}_p)$. On the other hand, we propose using a data-dependent transformed statistic of Eqs. (1.2) of the form:

$$NT(\widehat{\beta}_{mg} - \widehat{\beta}_p)^\top \widehat{H}_\beta^{-1}(\widehat{\beta}_{mg} - \widehat{\beta}_p) \tag{1.3}$$

The transformation has been introduced to avoid estimating covariance matrices involved in both time series data(Kiefer et al., 2000) and cross-sectional data (Lee and Robinson, 2016). Here in Eqs. (1.3), $\widehat{H}_\beta$ is calculated based on the observed data, no consistent covariance matrix is involved explicitly (the resulting asymptotic distribution will be non-standard). In addition, $\widehat{H}_\beta \xrightarrow{P} 0$ if $\sqrt{NT}(\widehat{\beta}_{mg} - \widehat{\beta}_p) \xrightarrow{P} 0$. Further calculation shows in this case there will be no indeterminacy issues.

The structure of this paper is organized as follows. Section 2 proposes our estimation method. Section 3 establishes asymptotic distributions for the proposed estimators. Section 4 introduces a statistic for testing coefficient homogeneity. Section 5 evaluates the finite sample performance by simulated data examples. Section 6 provides a real data example using the health care expenditure dataset from OECD countries. Section 7 makes the conclusion. Appendix A provides some justifications and explanations for Assumptions 1 to 7. Appendix B and C prove the main theorems in this paper. In the supplemental material, Appendices D and E prove some propositions and lemmas that have been used in the proofs of the main results. Appendix F of the supplemental material provides more details on the empirical application.

Throughout the paper, $\iota_T = (1, \cdots, 1)^\top$ is a $T \times 1$ vector; I_d is the $d \times d$ identity matrix; $A^\top$ is the transpose of matrix or vector A; $\mathrm{tr}(A)$, $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ denote the trace, the maximum and the minimum eigenvalues of a square matrix A; $||A|| = \mathrm{tr}\left(AA^\top\right)$ denotes the Euclidean norm of a matrix or a vector A; $A \otimes B$ is the Kronecker product of matrices A and B; $\xrightarrow{P}$ and $\xrightarrow{D}$ stand for converging in probability and in distribution respectively, while $\Rightarrow$ denotes weak convergence; $[a]$ means the largest integer part of a; $\int g(w)dw$ represents $\int_{-\infty}^{\infty} g(w)dw$.

2 Estimation method

In this section, we propose a mean group approach based on a semiparametric profile likelihood estimation method of individual parameter vectors and trending functions. In addition, a pooled estimator is provided by treating parameter vectors as identical to each other across individuals.

2.1 Mean group or weighted averaging estimators

The mean group or weighted averaging estimators are calculated by the following steps:

Step(i): For $i \geq 1$, denote

$$Y_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} Y_{it}, \quad X_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} X_{it}, \quad f_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} f_i(\tau_t), \quad e_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} e_{it} \tag{2.1}$$

By taking average over time of model Eqs. (1.1), we have

$$Y_{i\cdot} = X_{i\cdot}^\top \beta_i + f_{i\cdot} + \alpha_i + e_{i\cdot}. \tag{2.2}$$

Then we subtract Eqs. (2.2) from Eqs. (1.1) to eliminate the fixed effects α_i,

$$Y_{it} - Y_{i\cdot} = (X_{it}^\top - X_{i\cdot}^\top)\beta_i + f_i(\tau_t) - f_{i\cdot} + e_{it} - e_{i\cdot}. \tag{2.3}$$

Denote

$$Y_i = \begin{pmatrix} Y_{i1} \\ \vdots \\ Y_{iT} \end{pmatrix}, \quad X_i = \begin{pmatrix} X_{i1}^\top \\ \vdots \\ X_{iT}^\top \end{pmatrix}, \quad f_i = \begin{pmatrix} f_i(\tau_1) \\ \vdots \\ f_i(\tau_T) \end{pmatrix}, \quad e_i = \begin{pmatrix} e_{i1} \\ \vdots \\ e_{iT} \end{pmatrix}. \tag{2.4}$$

Eqs. (2.3) can be rewritten as:

$$Y_i - Y_{i\cdot}\iota_T = (X_i - X_{i\cdot}^\top \otimes \iota_T)\beta_i + f_i - f_{i\cdot}\iota_T + e_i - e_{i\cdot}\iota_T. \tag{2.5}$$

Step(ii): For $i \geq 1$, treat β_i as given, let $K(\cdot)$ be a kernel function, h_i be a bandwidth, and denote

$$Z_i(\tau) = \begin{pmatrix} 1 & \frac{\tau_1-\tau}{h_i} \\ \vdots & \vdots \\ 1 & \frac{\tau_T-\tau}{h_i} \end{pmatrix}, \quad W_i(\tau) = \begin{pmatrix} K(\frac{\tau_1-\tau}{h_i}) & & \\ & \ddots & \\ & & K(\frac{\tau_T-\tau}{h_i}) \end{pmatrix},$$

then by Taylor expansion, at $\tau \in [0,1]$ we have

$$f_i - f_{i\cdot}\iota_T \approx Z_i(\tau)(f_i(\tau) - f_{i\cdot}, h_i f_i'(\tau))^\top. \tag{2.6}$$

By plugging it into Eqs. (2.5), given β_i, the local linear estimator of $f_i(\tau) - f_{i\cdot}$ can be calculated as

$$\widehat{f}_i(\tau, \beta_i) = s_i(\tau)(Y_i - Y_{i\cdot}i_T - (X_i - X_{i\cdot}^\top \otimes \iota_T)\beta_i), \tag{2.7}$$

where $s_i(\tau) = (1 \quad 0)[Z_i^\top(\tau)W_i(\tau)Z_i(\tau)]^{-1}Z_i^\top(\tau)W_i(\tau)$.

Step(iii): Letting $S_i = \left(s_i^\top(\tau_1), \cdots, s_i^\top(\tau_T)\right)^\top$, then we have

$$\widetilde{Y}_i \approx \widetilde{X}_i\beta_i + e_i - e_{i\cdot}\iota_T, \tag{2.8}$$

where $\widetilde{Y}_i = (I_T - S_i)(Y_i - Y_{i\cdot}\iota_T)$, $\widetilde{X}_i = (I_T - S_i)(X_i - X_{i\cdot}^\top \otimes \iota_T)$. Then β_i can be estimated by

$$\widehat{\beta}_i = (\widetilde{X}_i^\top \widetilde{X}_i)^{-1}\widetilde{X}_i^\top \widetilde{Y}_i \tag{2.9}$$

and $f_i(\tau)$ is estimated by

$$\widehat{f}_i(\tau) = s_i(\tau)(Y_i - Y_{i\cdot}\iota_T - (X_i - X_{i\cdot}^\top \otimes \iota_T)\widehat{\beta}_i). \tag{2.10}$$

Step(iv): Finally, the mean group estimators are calculated as

$$\widehat{\beta}_{mg} = \frac{1}{N}\sum_{i=1}^N \widehat{\beta}_i, \quad \widehat{f}_{mg}(\tau) = \frac{1}{N}\sum_{i=1}^N \widehat{f}_i(\tau) \tag{2.11}$$

More generally, given constants $w_1, \cdots, w_N$, we consider the weighted averaging estimators by

$$\widehat{\beta}_w = \sum_{i=1}^N w_i\widehat{\beta}_i, \quad \text{and} \quad \widehat{f}_w = \sum_{i=1}^N w_i\widehat{f}_i(\tau). \tag{2.12}$$

2.2 Pooled estimators

Pooled estimators for β are considered here for two purposes. On the one hand, pooled estimators are usually accurate in homogeneous models, and they may be asymptotically unbiased or consistent for some heterogeneous scenarios. On the other hand, they are essentially useful in constructing test statistics for homogeneity hypotheses testing.

Chen et al. (2012) propose a pooled profile likelihood estimation procedure when $\beta_i = \beta$ and $f_{it} = f_t$ hold for all i. However, the pooled estimator of β might be severely biased even if $\beta_i = \beta$ holds for all i. For example, by assuming $\beta_i = \beta$, Eqs. (1.1) can be rearranged as:

$$Y_{it} = X_{it}^{\top}\beta + f_t + \alpha_i + u_{it}, \tag{2.13}$$

where $u_{it} = f_{it} - f_t + e_{it}$. As a consequence, $\mathrm{E}\,(u_{it}) \neq 0$ if $f_{it} \neq f_t$.

Since unobserved trends, such as technology processes are likely to be diversified, we propose another pooled estimator robust to heterogeneous trending functions. The procedure is to smooth out $f_i(\tau)$ from Eqs. (2.2) to Eqs. (2.7), and then letting $\beta_i = \beta$ in Eqs. (2.8) such that

$$\widetilde{Y}_i \approx \widetilde{X}_i\beta + e_i - e_{i\cdot}\iota_T. \tag{2.14}$$

Then we pool Eqs. (2.14) and estimate β by the least–square method:

$$\widehat{\beta}_p = \left(\sum_{i=1}^{N} \widetilde{X}_i^{\top}\widetilde{X}_i\right)^{-1}\sum_{i=1}^{N} \widetilde{X}_i^{\top}\widetilde{Y}_i. \tag{2.15}$$

Note that the definitions of $\widetilde{X}_i$ and $\widetilde{Y}_i$ are different from those in Chen et al. (2012), as we use the individual rather than the pooled information to eliminate f_{it} and g_{it}. Also, in this paper, we assume for all $i \geq 1$, the same h_i is employed in Eqs. (2.8) and Eqs. (2.15), thus the definitions of $\widetilde{X}_i$ and $\widetilde{Y}_i$ within $\widehat{\beta}_{mg}$ and $\widehat{\beta}_p$ are identical to each other. Such an assumption may not be necessary, but will facilitate the calculations, especially in the hypotheses testing involving both $\widehat{\beta}_{mg}$ and $\widehat{\beta}_p$. Mathematically, we may discuss how to deal with the case where h_i can be different according to i, but $h_i = \rho_i\, h_0$, in which h_0 is a kind of leading order, and ρ_i is to allowed to depend on i. In large sample theory, certain additional conditions are needed on ρ_i, while the choice of ρ_i is quite straightforward. We wish to leave this type of technical extensions for future research.

3 The asymptotic theory of estimators

In this section we introduce some regularity conditions, and then establish asymptotic distributions for $\widehat{\beta}_w$, $\widehat{\beta}_{mg}$ and $\widehat{f}_{mg}$.

3.1 Assumptions

Denote that $v_t = (v_{1t}, \cdots, v_{Nt})^\top$, $e_t = (e_{1t}, \cdots, e_{Nt})^\top$, $\eta_{it} = \Omega_{v,i}^{-1} v_{it} v_{it}^\top - I_d$, and $\xi_{it} = \Omega_{v,i}^{-1} v_{it} e_{it}$, where $\Omega_{v,i} = \mathrm{E}\left(v_{it} v_{it}^\top\right)$. We impose the following assumptions.

Assumption 1 (i) Suppose that $\{v_t, e_t : t \geq 1\}$ is a vector of strictly stationary α-mixing time series with zero mean and $E(\xi_{it}) = 0$. The mixing coefficient $\alpha(\cdot)$ satisfies $\sum_{k=1}^{\infty} k^{\phi} (\alpha_v(k))^{\frac{\delta}{4+\delta}} < \infty$ for some $\phi > 1$, and $\delta > 0$ as involved in (iii) below.
(ii) $\min_{1 \leq i \leq N} \lambda_{\min}\left(\Omega_{v,i}\right) > c_v$ for some constant $c_v > 0$.
(iii) $\mathrm{E}\left|\left|v_{it}\right|\right|^{4+\delta} < \infty$, $\mathrm{E}\left|\left|e_{it}\right|\right|^{4+\delta} < \infty$ for the same constant $\delta > 0$ as in (i) above.

Assumption 2 As $(N, T) \to (\infty, \infty)$,
(i) $\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{T} |\mathrm{E}(e_{i,1} e_{j,1+k})| = O(N)$, $\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \left|\left|\mathrm{E}(\xi_{it_1} \xi_{jt_2}^\top)\right|\right| = O(NT)$.
(ii) $\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t_1=1}^{T_1} \sum_{t_2=1}^{T_1} \sum_{s_1=1}^{T_2} \sum_{s_2=1}^{T_2} \left|\left|\mathrm{E}\left(\eta_{it_1} \xi_{is_1} \xi_{js_2}^\top \eta_{jt_2}^\top\right)\right|\right| = O(NT_1T_2)$.
(iii) $\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t_1=1}^{T_1} \sum_{t_2=1}^{T_1} \sum_{s_1=1}^{T_2} \sum_{s_2=1}^{T_2} \left|\left|\mathrm{E}\left(v_{it_1} v_{jt_2}^\top e_{is_1} e_{js_2}\right)\right|\right| = O\left(NT_1T_2\right)$.
(iv) $\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t_1=1}^{T} \cdots \sum_{t_4=1}^{T} \left|\left|\mathrm{E}\left(v_{it_1}^\top \xi_{it_2} v_{jt_3}^\top \xi_{jt_4}\right)\right|\right| = O(NT^2)$.

Assumption 3
(i) Denote $\gamma_{N,w} = \left(\sum_{i=1}^{N} w_i^2\right)^{-1}$, and let $\gamma_{N,w} \sum_{i=1}^{N} \sum_{j=1}^{N} w_i \Omega_{v,i}^{-1} \Omega_{v,e}(i,j) \Omega_{v,i}^{-1} w_j \to \Sigma_{ve,w}$ for some $d \times d$ positive definite matrix $\Sigma_{ve,w}$, where $\Omega_{v,e}(i,j) = \mathrm{E}(\xi_{i1} \xi_{j1}^\top) + 2\sum_{k=2}^{\infty} \mathrm{E}(\xi_{i1} \xi_{jk}^\top)$.
(ii) Assume that $\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} C_K(i,j) \sigma_e(i,j) \to \sigma_e^2$ for some $0 < \sigma_e^2 < \infty$, where $\sigma_e(i,j) = \mathrm{E}(e_{i1} e_{j1}) + 2\sum_{k=2}^{\infty} \mathrm{E}(e_{i1} e_{jk})$, and $C_K(i,j) = \int \frac{1}{\rho_i \rho_j} K\left(\frac{v}{\rho_i}\right) K\left(\frac{v}{\rho_j}\right) dv$, with $h = \min_{1 \leq i \leq N} h_i$ and $\rho_i = h_i / h$.
(iii) $\mathrm{E}\left|\left|\gamma_{N,w}^{1/2} \sum_{i=1}^{N} w_i \xi_{it}\right|\right|^{2+\delta} < \infty$ and $\sum_{i_1=1}^{N} \cdots \sum_{i_4=1}^{N} |\mathrm{E}(e_{i1,t} e_{i2,t} e_{i3,t} e_{i4,t})| = O(N^2)$.

Assumption 4
(i) $K(\cdot)$ is a continuous and symmetric probability density function with a compact support.
(ii) $f_i(\cdot)$ and $g_i(\cdot)$ have continuous derivatives of up to the second order, and the identification condition $\int_0^1 f_i(\tau) d\tau = 0$ holds for $i \geq 1$.
(iii) As $(N, T) \to (\infty, \infty)$, (a) $NTh^8 \to 0$; (b) $Nh^4 \to 0$; (c) $\frac{N}{T^2 h^2} \to 0$; (d) $\frac{N}{T^{1+\frac{\delta^*}{2+\delta^*/2}}} \to 0$; (e) $Th^{2+r} \to \infty$; (f) $\frac{Nh}{T} \to 0$, where $0 < \delta^* < \delta$, and $r > 0$.

Justifications for Assumptions 1 to 4 are given by the Appendix A below.

3.2 Asymptotic distributions of the given estimators

The following two theorems give the asymptotic distributions of $\widehat{\beta}_i$ and $\widehat{f}_i(\tau)$ for any fixed $i \geq 1$.

Theorem 3.1: Under Assumptions 1 and 4, for any fixed $i \geq 1$, as $T \to \infty$

$$\sqrt{T}\left(\widehat{\beta}_i - \beta_i\right) \xrightarrow{D} N\left(0_d, \Omega_{v,i}^{-1} \Omega_{v,e}(i,i) \Omega_{v,i}^{-1}\right). \tag{3.1}$$

Denote $\mu_j = \int v^j K(v) dv$, and $\nu_j = \int v^j K^2(v) dv$, we have

Theorem 3.2: Under Assumptions 1 and 4, for any fixed $i \geq 1$, and $0 < \tau < 1$, as $T \to \infty$

$$\sqrt{Th_i}\left(\widehat{f}_i(\tau) - f_i(\tau) - b_i(\tau)h_i^2 + o(h_i^2)\right) \xrightarrow{D} N\left(0, \nu_0\sigma_e(i,i)\right), \tag{3.2}$$

where $b_i(\tau) = \frac{1}{2}\mu_2 f_i''(\tau)$.

Theorems 3.1 and 3.2 are standard results in semiparametric profile likelihood estimation. The next theorem gives the asymptotic distribution of the weighted average estimator $\widehat{\beta}_w = \sum w_i \widehat{\beta}_i$.

Theorem 3.3: Under Assumptions 1 to 4, as $(N,T) \to (\infty,\infty)$, we have

$$\sqrt{\gamma_{N,w}T}\left(\widehat{\beta}_w - \overline{\beta}_w\right) \xrightarrow{D} N\left(0_d, \Sigma_{ve,w}\right), \tag{3.3}$$

where $\overline{\beta}_w = \sum_i w_i \beta_i$.

Note $\widehat{\beta}_{mg}$ is a special case of $\widehat{\beta}_w$ when $w_i = \frac{1}{N}$ and $\gamma_{N,w} = N$. In this case, $\Sigma_{ve,w}$ is simplified into $\frac{1}{N}\sum_{i=1}^N \sum_{j=1}^N \Omega_{v,i}^{-1}\Omega_{v,e}(i,j)\Omega_{v,j}^{-1} \to \Sigma_{ve}$. Then the asymptotic distribution of $\widehat{\beta}_{mg}$ is given by the following corollary.

Corollary 3.1: Under Assumptions 1 to 4, as $(N,T) \to (\infty,\infty)$, we have

$$\sqrt{NT}\left(\widehat{\beta}_{mg} - \overline{\beta}\right) \xrightarrow{D} N\left(0_d, \Sigma_{ve}\right). \tag{3.4}$$

Remark 1. (i) In Corollary 3.1, $\widehat{\beta}_{mg}$ achieves an asymptotic normality with root NT rate. In addition, since $\overline{\beta} = \beta$ if $\beta_i = \beta$ for all i, $\widehat{\beta}_{mg}$ is also root-NT consistency to β under the homogeneity setting.

(ii) Consider that only some entries of the parameter vector β_i are heterogeneous, such as $\beta_i = (\beta_{i,1}^\top, \beta_2^\top)^\top$. If we divide $\widehat{\beta}_{mg}$ into $\widehat{\beta}_{mg} = (\widehat{\beta}_{mg,1}^\top, \widehat{\beta}_{mg,2}^\top)^\top$ accordingly, $\widehat{\beta}_{mg,2}$ will be root-NT consistency to the true parameter β_2. However, if we divide the pooled estimator $\widehat{\beta}_p$ into $\widehat{\beta}_p = (\widehat{\beta}_{p,1}^\top, \widehat{\beta}_{p,2}^\top)^\top$, by our calculations (i.e., the proof of Theorem 4.2 below, $\widehat{\beta}_{p,2}$ will be a biased estimator of β_2 unless $\Omega_{v,i}$ satisfies some block-diagonal restrictions.

For the trending functions, we only establish the asymptotic distribution of $\widehat{f}_{mg}(\tau)$, since the asymptotic distribution of $\widehat{f}_w(\tau)$ will involve more specific assumptions on the weighting coefficients and the individual bandwidth.

Theorem 3.4: Under Assumptions 1 to 4, as $(N,T) \to (\infty,\infty)$, for $0 < \tau < 1$, we have

$$\sqrt{NTh}\left(\widehat{f}_{mg}(\tau) - \overline{f}(\tau) - b_f(\tau)h^2 + o(h^2)\right) \xrightarrow{D} N\left(0, \sigma_{eK}^2\right), \tag{3.5}$$

where $b_f(\tau) = \frac{1}{2N}\sum_{i=1}^N \mu_2 \rho_i^2 f_i''(\tau)$.

Note that for the given group-division of individuals, our estimators within groups will be consistent to the corresponding group averages, with rates depending on the group sizes and T. This will be useful when the practitioner is interested in group effects, where the grouping division are given, and the individual coefficient is heterogeneous within given groups. In Section 6, we illustrate such applications by dividing OECD countries into 6 groups according to their geographic locations.

4 Homogeneity test

In this section, we propose a hypotheses testing procedure towards checking coefficient homogeneity when the residual e_{it} is both serially correlated and cross–sectionally dependent. It will establish a criterion to understand whether the coefficients are heterogeneous or not, and help the practitioner to decide which estimator to use.

4.1 Hypotheses structure

We aim to construct a test statistic for the following null hypothesis:

$$H_0 : \beta_i = \beta_0 \quad \text{for } 1 \le i \le N. \tag{4.1}$$

A natural idea is to construct a test statistic based on the difference between $\widehat{\beta}_{mg}$ and $\widehat{\beta}_p$. The ideal situation is that $\widehat{\beta}_p - \widehat{\beta}_{mg}$ results in an asymptotic normality under H_0 with a root NT rate, and deviates from this distribution significantly under an alternative hypothesis.

However, under H_0, both $\widehat{\beta}_{mg}$ and $\widehat{\beta}_p$ result in an asymptotic normality with the root NT rate. In certain cases, their asymptotic covariance matrces may be close or even identical such that

$$\sqrt{NT}\left(\widehat{\beta}_{mg} - \widehat{\beta}_p\right) \xrightarrow{p} 0. \tag{4.2}$$

This brings some difficulties to construct a test statistic and establish its asymptotic distribution under H_0. Furthermore, the pooled estimator, $\widehat{\beta}_p$ may be unbiased under the alternative (see an example in Pesaran and Yamagata, 2008). Thus the two estimators may still be close to each other under certain alternatives. In this case, there may be some issues related to the power of this type of test statistics.

In this section, we will propose a test statistic that avoids dealing with such issues. This statistic will allow $\widehat{\beta}_p$ and $\widehat{\beta}_{mg}$ to have exactly the same asymptotic distribution with a root NT rate under H_0. Furthermore, its power can be satisfactory even when $\widehat{\beta}_p$ is unbiased or consistent to $\overline{\beta}$ under a suitable alternative.

In constructing the statistic, we focus on the asymptotic distribution of $\sqrt{\gamma_{NT} NT}\left(\widehat{\beta}_{mg} - \widehat{\beta}_p\right)$ rather than $\sqrt{NT}\left(\widehat{\beta}_{mg} - \widehat{\beta}_p\right)$, where γ_{NT} is defined as

$$\gamma_{NT} = \left(\max\left(\frac{1}{T}, \gamma_N\right)\right)^{-1}, \quad \gamma_N = \frac{1}{N}\sum_{i=1}^{N}\left|\left|\widetilde{\Omega}_{v,i}\right|\right|^2, \quad \widetilde{\Omega}_{v,i} = \Omega_{v,i} - \overline{\Omega}_N, \tag{4.3}$$

with $\overline{\Omega}_N = \frac{1}{N}\sum_{i=1}^{N}\Omega_{v,i}$. Note γ_{NT} is a slower rate between T and γ_N^{-1} (if $\gamma_N \neq 0$). Thus if Eqs. (4.2) is true, we automatically have $\gamma_{NT} \to \infty$ such that $\widehat{\beta}_{mg} - \widehat{\beta}_p$ achieves asymptotic normality with a super $\sqrt{\gamma_{NT} NT}$ rate. On the other hand, under H_1, the difference between $\widehat{\beta}_{mg}$ and $\widehat{\beta}_p$ will be amplified by this $\sqrt{\gamma_{NT} NT}$ rate. Thus, the test statistic has good power performance even if $\widehat{\beta}_p$ is consistent to $\overline{\beta}$.

To establish such asymptotic results, some extra conditions are required.

Assumption 5. Denote $\zeta_{N,t} = \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \sum_{i=1}^{N} \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i} \Delta_{Ti,1} \right) \xi_{it}$, where $\Delta_{Ti,1} = \frac{1}{T} \sum_{t=1}^{T} \eta_{it}$.
(i) $\overline{\Omega}_N \to \Omega_v$ for some $d \times d$ positive definite Ω_v.
(ii) $\mathrm{Var}\left(\sum_{t=1}^{T} \zeta_{N,t} \right) \to \Sigma_{ve,1}$ for some $d \times d$ positive definite matrix $\Sigma_{ve,1}$ and $\mathrm{E}\left|\left| \sqrt{T} \zeta_{N,t} \right|\right|^{2+\delta} < \infty$.
(iii) $\sum_{t_1=1}^{[r_1 T]} \sum_{t_2=[r_2 T]+1}^{T} \mathrm{E}\left(\zeta_{N,t_1} \zeta_{N,t_2}^\top \right) = o(1)$ for any $0 \leq r_1 \leq r_2 \leq 1$, and $\mathrm{E}\left|\left| \sum_{t=1}^{T} \zeta_{N,t} \right|\right|^{2+\delta} < \infty$.
(iv) $\frac{N \log(2T)}{T^2 h^2} \to 0$.

If $\gamma_{NT} = T$, the following extra conditions are required.

Assumption 6.
(i) $\mathrm{E}\,||v_{it}||^{8+\delta} < \infty$, and $\mathrm{E}\,||\xi_{it}||^{4+\delta} < \infty$.
(ii) $\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t_1=1}^{T_1} \sum_{t_2=1}^{T_1} \sum_{t_3=1}^{T_2} \cdots \sum_{t_6=1}^{T_2} \left|\left| \mathrm{E}\left(\eta_{it_3}^\top v_{it_4} e_{it_1} e_{jt_2} v_{jt_5}^\top \eta_{jt_6} \right) \right|\right| = O\left(NT_2^2 T_1 \right)$.
(iii) $\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t_1=1}^{T_1} \sum_{t_2=1}^{T_1} \sum_{t_3=1}^{T_2} \cdots \sum_{t_6=1}^{T_2} \left| \mathrm{E}\left(\xi_{it_1}^\top v_{it_3} v_{jt_4}^\top v_{it_5} v_{jt_6}^\top \xi_{jt_2} \right) \right| = O\left(NT_2^2 T_1 \right)$.
(iv) $\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t_1=1}^{T_1} \sum_{t_2=1}^{T_1} \sum_{t_3=1}^{T_2} \cdots \sum_{t_6=1}^{T_2} \left| \mathrm{E}\left(\eta_{it_1} \eta_{it_3} \xi_{jt_5} \xi_{jt_6}^\top \eta_{it_4}^\top \eta_{jt_2}^\top \right) \right| = O\left(NT_2^2 T_1 \right)$.
(v) $\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t_1=1}^{T} \cdots \sum_{t_4=1}^{T} \left|\left| \mathrm{E}\left(v_{it_1} v_{it_2}^\top v_{jt_3} v_{jt_4}^\top \right) \right|\right| = O\left(NT^2 \right)$.
(vi) $\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \left|\left| \mathrm{E}\left(\eta_{it_1} \eta_{jt_2}^\top \right) \right|\right| = O(NT)$.

Assumptions 5 and 6 are analogous versions of Assumptions 3 and 4. Especially, the role γ_{NT} here is similar to the $\gamma_{N,w}$, and both are used to control the scale of the asymptotic variance. The key difference is that $\widetilde{\Omega}_{v,i}$ is allowed to equal zero for all i simultaneously, which is the reason why another random components are included in $\zeta_{N,t}$ and the extra assumption are imposed in Assumption 6. More specific discussions are included in Appendix A.

The following lemma describes the asymptotic distribution of $\widehat{\beta}_p - \widehat{\beta}_{mg}$.

Lemma 4.1: Under H_0 and by Assumptions 1 to 6 , as $(N, T) \to (\infty, \infty)$, we have

$$(\gamma_{NT} NT)^{1/2} \, (\widehat{\beta}_p - \widehat{\beta}_{mg}) \xrightarrow{D} N(0_d, \Sigma_{ve,2}), \tag{4.4}$$

where $\Sigma_{ve,2} = \Omega_v^{-1} \Sigma_{ve,1} \Omega_v^{-1}$.

Remark 2. To illustrate the result in Lemma 4.1, we provide two examples as follows.

- Consider $d = 1$, $\Omega_{v,i} = 1$ for $i = 1, \cdots, [N/2]$, and $\Omega_{v,i} = 2$ otherwise, then $\gamma_N = \frac{1}{4}$ is a constant. Lemma 4.1 indicates $\sqrt{NT}\left(\widehat{\beta}_{mg} - \widehat{\beta}_p \right)$ converges to a non-degenerate random variable, and $\widehat{\beta}_{mg}$ and $\widehat{\beta}_p$ have different asymptotic distributions under H_0.

- Consider $\Omega_{v,i} = \Omega_v$ for all i, thus $\gamma_N = 0$ and $\gamma_{NT} = T$. In this case, $\widehat{\beta}_p$ and $\widehat{\beta}_{mg}$ will have the same asymptotic distribution, thus $\sqrt{NT}\left(\widehat{\beta}_p - \widehat{\beta}_{mg} \right) \xrightarrow{P} 0$. Nevertheless, Eqs. (4.4) reveals that in this case $\widehat{\beta}_{mg} - \widehat{\beta}_p$ converges to zero with a $\sqrt{NT^2}$ rate.

If both γ_{NT} and consistent estimate of $\Sigma_{ve,2}$ were available, the standard practice to construct statistic based on Lemma 4.1 would be:

$$\gamma_{NT} NT \left(\widehat{\beta}_{mg} - \widehat{\beta}_p \right)^\top \Sigma_{ve,2}^{-1} \left(\widehat{\beta}_{mg} - \widehat{\beta}_p \right) \xrightarrow{D} \chi_d^2. \tag{4.5}$$

However, Eqs. (4.5) is generally infeasible, since γ_{NT} is unknown, and it is difficult to provide any consistent estimator for $\Sigma_{ve,2}$ with weak dependence in both cross-sectional and serial dimensions. To overcome such difficulties, our statistic is constructed in the following way:

$$\widehat{J}_{NT} = NT\left(\widehat{\beta}_p - \widehat{\beta}_{mg}\right)^{\top} \widehat{H}_{NT}^{-1}\left(\widehat{\beta}_p - \widehat{\beta}_{mg}\right). \tag{4.6}$$

where $\widehat{H}_{NT} = \frac{1}{T}\sum_{m=1}^{T} \widehat{L}_{NT,m}\widehat{L}_{NT,m}^{\top}$, in which

$$\widehat{L}_{NT,m} = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N}\sum_{t=1}^{m}\left(\widehat{\Omega}_{x,Ti}^{-1} - \widehat{\Omega}_{x,NT}^{-1}\right)\widetilde{X}_{it}\widehat{e}_{it}, \tag{4.7}$$

and $\widehat{\Omega}_{x,Ti} = \frac{1}{T}\widetilde{X}_i^{\top}\widetilde{X}_i$, $\widehat{\Omega}_{x,NT} = \frac{1}{N}\sum_{i=1}^{N}\widehat{\Omega}_{x,Ti}$, with $\widetilde{Y}_{it}$ and $\widetilde{X}_{it}$ being the t-th elements of $\widetilde{Y}_i$ and $\widetilde{X}_i$, respectively, and $\widehat{e}_{it} = \widetilde{Y}_{it} - \widetilde{X}_{it}^{\top}\widehat{\beta}_i$.

$\widehat{J}_{NT}$ becomes a feasible version of Eqs. (4.5) for two reasons. Firstly, $\widehat{L}_{NT,m}$ has a similar structure with $\widehat{\beta}_p - \widehat{\beta}_{mg}$. Thus the $\sqrt{\gamma_{NT}}$ from $\left(\widehat{\beta}_p - \widehat{\beta}_{mg}\right)$ and $\widehat{L}_{NT,m}$ will be canceled out with each other within $\widehat{J}_{NT}$. Secondly, $\Sigma_{ve,2}$ can be replaced by a stochastic transformation $\widehat{H}_{NT}$, which can be calculated from the observed data. Note that such transformed version $\widehat{H}_{NT}$ will not converge to the population covariance matrix (see Kiefer et al., 2000; Vogelsang, 2012; Lee and Robinson, 2016). For example, $\gamma_{NT}\widehat{H}_{NT}$ will converge in distribution to another random variable constructed by the integration of a Brownian bridge.

Theorem 4.1: Under H_0 and Assumptions 1 to 6, as $(N,T) \to (\infty,\infty)$, we have

$$\widehat{J}_{NT} \xrightarrow{D} W_d^{\top}(1)\Phi_d^{-1}W_d(1), \tag{4.8}$$

where $\Phi_d = \int_0^1 (W_d(r) - rW_d(1))(W_d(r) - rW_d(1))^{\top} dr$ and $W_d(\cdot)$ is a $d\times 1$ vector of independent Brownian motion processes.

To evaluate the asymptotic power properties of $\widehat{J}_{NT}$, we consider the following alternative hypothesis:

$$H_1 : \beta_i = \beta_0 + \delta_{NT}\epsilon_i \quad \text{for } 1 \le i \le N, \tag{4.9}$$

with δ_{NT} and a $d \times 1$ vector of real numbers ϵ_i satisfying the following assumption.

Assumption 7. Denote $\widetilde{\epsilon}_i = \epsilon_i - \frac{1}{N}\sum_{i=1}^{N}\epsilon_i$. (i) $\delta_{NT}\sqrt{\gamma_{NT}NT} \to \infty$; (ii) $\delta_{NT}\sqrt{\gamma_{NT}} \to 0$; (iii) $\frac{1}{N}\sum_{i=1}^{N}\Omega_{v,i}\widetilde{\epsilon}_i \to \psi$, where ψ is a $d\times 1$ vector of constant with $||\psi|| > 0$.

Assumption 7(i) specifies the minimum scale required on the departures from β_0, Assumption 7(ii) is imposed to simplify the calculations, and $\frac{1}{N}\sum_{i=1}^{N}\Omega_{v,i}\widetilde{\epsilon}_i$ determines the bias introduced by using $\widehat{\beta}_p$ under H_1. Some other versions of (ii) and (iii) are discussed in Appendix A.

The following theorem states the asymptotic distribution of $\widehat{J}_{NT}$ under H_1.

Theorem 4.2: Under H_1 and Assumptions 1 to 7, as $(N,T) \to (\infty,\infty)$, we have

$$\left(\delta_{NT}^2\gamma_{NT}NT\right)^{-1}\widehat{J}_{NT} \xrightarrow{D} \psi^{\top}\Omega_v^{-1}\Phi_d^{-1}\Omega_v^{-1}\psi. \tag{4.10}$$

Theorems 4.1 and 4.2 imply the size and the power of the proposed test statistics, which can be concluded as the following corollary.

Corollary 4.1: Under Assumptions 1 to 7, we have as $(N, T) \to (\infty, \infty)$,

$$P\left(\widehat{J}_{NT} > c_\alpha | H_0\right) \to \alpha, \tag{4.11}$$

$$P\left(\widehat{J}_{NT} > c_\alpha | H_1\right) \to 1, \tag{4.12}$$

where c_α is the critical value at the significance level α of the distribution of the random variable $W_d^\top(1)\Phi_d^{-1}W_d(1)$.

4.2 Bootstrap based critical value

In implementation, the test statistic $\widehat{J}_{NT}$ can be calculated based on the observed data, and the critical values of the asymptotic distribution in Eqs. (4.8) are available in a table by (Kiefer et al., 2000). One disadvantage arises since the weak convergence of $L_{NT,m}$ to the Brownian bridge requires large enough T. Then $\widehat{J}_{NT}$ may converge to its asymptotic distribution very slowly since the contribution by increasing N seems small. Thus the size calculated by using the critique values from the table may not be satisfactory unless T is large enough. To improve the finite sample performance for the case where T is relatively small, we propose to use a bootstrap based critical value for $\widehat{J}_{NT}$ in each case. The bootstrap procedure includes the following steps.

S.1 Estimate the heterogeneous model in Eqs. (1.1), and obtain the estimated residual $\widehat{e}_{it} = \widetilde{Y}_{it} - \widetilde{X}_{it}^\top\widehat{\beta}_i$, the mean group estimator $\widehat{\beta}_{mg}$, together with $\widehat{\Omega}_{x,Ti}$ and $\widehat{\Omega}_{x,NT}$. And estimate the pooled estimator $\widehat{\beta}_p$ in Eqs. (2.15) and then calculate the test statistic $\widehat{J}_{NT}$ based on $\{\widehat{e}_{it}, \widehat{\beta}_{mg}, \widehat{\beta}_p, \widehat{\Omega}_{x,Ti}, \widehat{\Omega}_{x,NT}\}$.

S.2 Obtain the corresponding residual in the pooled estimation $\widehat{u}_{it} = \widetilde{Y}_{it} - \widetilde{X}_{it}^\top\widehat{\beta}_p$. Then use it to calculate the bootstrap error $u_{it}^* = \widehat{u}_{it}\zeta_{it}$, where $\{\zeta_{it}, 1 \le i \le N, 1 \le t \le T\}$ is i.i.d. random variables drawn from

$$\mathrm{P}\left(\zeta_{it} = -\frac{\sqrt{5}-1}{2}\right) = \frac{\sqrt{5}+1}{2\sqrt{5}}, \quad \mathrm{P}\left(\zeta_{it} = \frac{\sqrt{5}+1}{2}\right) = \frac{\sqrt{5}-1}{2\sqrt{5}} \tag{4.13}$$

and then obtain $Y_{it}^* = X_{it}^\top\widehat{\beta}_p + u_{it}^*$.

S.3 Use the bootstrap sample $\{Y_{it}^*, X_{it}\}$ to re-estimate $\widehat{\beta}_{mg}^*$, $\widehat{\beta}_p^*$ and $\widehat{e}_{it}^*$. And calculate the $\widehat{J}_{NT}^*$ based on $\{\widehat{e}_{it}^*, \widehat{\beta}_{mg}^*, \widehat{\beta}_p^*, \widehat{\Omega}_{x,Ti}, \widehat{\Omega}_{x,NT}\}$.

S.4 Repeat steps S.2 to S.3 by $B = 250$ times and produce the sequence $\{\widehat{J}_{NT,b}^*\}$ for $b = 1, \cdots, B$. Then calculate the α quantile c_α^* of the empirical distribution based on $\{\widehat{J}_{NT,b}^*\}$, and use c_α^* as the bootstrap critical value for size and power calculation.

5 Simulated studies

In this section, we carry out simulation studies to compare the finite sample performance of the proposed estimators, together with the asymptotic size and local power of the homogeneity test.

5.1 Estimation biases and standard deviations

Consider the following data generating processes (DGPs) based on Eqs. (1.1):

$$Y_{it} = X_{it}^{\top}\beta_i + f_i(\tau_t) + \alpha_i + e_{it},$$

where $f_i(u) = \frac{i}{N}\sqrt{u} - \frac{2}{3}\frac{i}{N}$, $\alpha_i = \max\{\frac{1}{T}\sum_{t=1}^{T} X_{it,1}, \frac{1}{T}\sum_{t=1}^{T} X_{it,2}\}$. X_{it} follows $X_{it} = g_i(\tau_t) + x_i + v_{it}$, with $g_i(u) = ((\frac{i}{N})^{1/2}u, \quad 2\frac{i}{N}\cos(\pi u))^{\top}$ and $x_i = \frac{1}{T}\sum_{t=1}^{T} e_{it}$.

For the cross-sectional and serial dependent structure, we consider the case where e_{it} follows $e_{it} = 0.3e_{i,t-1} + \varepsilon_{it}$. For $t \geq 1$, $\varepsilon_t = (\varepsilon_{1t}, \cdots, \varepsilon_{Nt})^{\top}$ is drawn from $N(0_N, \Omega_e)$, with the (i,j)-th element of Ω_e being $\Omega_{e,ij} = \frac{1}{1+(i-j)^2}$. Similarly, $v_{it} = 0.3v_{i,t-1} + \eta_{it}$. Denote $\eta_t = (\eta_{1t}^{\top}, \cdots, \eta_{Nt}^{\top})^{\top}$, and for $t \geq 1$, $vec(\eta_t)$ is drawn independently from $N(0_{Nd}, \Omega_{v,d})$ with $\Omega_{v,d} = I_d \otimes \Omega_v$, where $\Omega_{v,ij} = \frac{1}{1+(i-j)^2}$ for $i \neq j$, and $\Omega_{v,ii} = \sigma_{v,i}^2$.

Furthermore, the specifications on the coefficient heterogeneity (β_i) and heteroskedasticity $(\sigma_{v,i}^2)$ are different across DGPs as follows:

- DGP 1: $\beta_i = \left(\frac{2i-1}{2N} - \frac{1}{2}, 4\cos(\pi\frac{i}{N}) + \frac{4}{N}\right)^{\top}$; $\sigma_{v,i}^2 = 1$ for all i.
- DGP 2: $\beta_i = \left(\frac{2i-1}{2N} - \frac{1}{2}, 4\cos(\pi\frac{i}{N}) + \frac{4}{N}\right)^{\top}$; $\sigma_{v,i}^2 = 1$ for $i = 1, \cdots, [N/2]$, and $\sigma_{v,i}^2 = 2$, otherwise.

Data simulated by DGP1 and DGP2 will be used to evaluate the finite sample performance of the mean group estimators (MGEs) in comparison to the pooled estimators, which include the pooled estimator (PE1) in Chen et al. (2012) and the pooled estimator (PE2) in Eqs. (2.15).

In our implementation, Epanechnikove kernel is used throughout this section, the bandwidth h_i is chosen by a leave-one-out cross-validation metho:

$$h_{i,CV} = \arg\min_{h} \sum_{t=1}^{T} \left(Y_{it} - X_{it}^{\top}\widehat{\beta}_{i,-t}(h) - \widehat{f}_{i,-t}(\tau_t, h)\right)^2, \tag{5.1}$$

where $\widehat{\beta}_{i,-t}(h)$ and $\widehat{f}_{i,-t}(\tau_t, h)$ are the individual estimators with the sample point (X_{it}, Y_{it}) being removed. For each simulation indexed as $m = 1, \cdots, M$, we calculate $\frac{1}{\sqrt{d}}||\widehat{\beta}_m - \overline{\beta}||$ for all the estimators. The procedure is repeated by $M = 500$ times, eventually for each estimator, we obtain an absolute error sequence $\{\frac{1}{\sqrt{d}}||\widehat{\beta}_m - \overline{\beta}||\}_{m=1}^{M}$.

Table 1: Means and SDs of the absolute errors for estimating $\overline{\beta}$ in DGP1 and DGP2

		DGP1				**DGP2**			
	$N \backslash T$	10	20	50	100	10	20	50	100
MGE	10	0.149	0.094	0.056	0.039	0.126	0.079	0.045	0.032
		(0.078)	(0.052)	(0.028)	(0.020)	(0.069)	(0.043)	(0.023)	(0.017)
	20	0.108	0.066	0.041	0.028	0.087	0.056	0.034	0.023
		(0.057)	(0.037)	(0.022)	(0.015)	(0.049)	(0.030)	(0.019)	(0.012)
	50	0.069	0.041	0.025	0.018	0.058	0.035	0.021	0.014
		(0.035)	(0.021)	(0.013)	(0.009)	(0.030)	(0.018)	(0.011)	(0.007)
	100	0.048	0.030	0.018	0.012	0.041	0.024	0.015	0.011
		(0.025)	(0.016)	(0.009)	(0.006)	(0.021)	(0.013)	(0.008)	(0.005)
PE1	10	0.281	0.242	0.222	0.218	0.172	0.127	0.088	0.068
		(0.125)	(0.097)	(0.065)	(0.048)	(0.087)	(0.065)	(0.045)	(0.034)
	20	0.313	0.290	0.269	0.263	0.166	0.135	0.115	0.108
		(0.104)	(0.076)	(0.050)	(0.035)	(0.083)	(0.060)	(0.043)	(0.032)
	50	0.398	0.374	0.358	0.352	0.247	0.221	0.209	0.210
		(0.073)	(0.050)	(0.035)	(0.023)	(0.062)	(0.049)	(0.033)	(0.024)
	100	0.466	0.446	0.429	0.424	0.328	0.313	0.305	0.305
		(0.055)	(0.040)	(0.023)	(0.016)	(0.043)	(0.034)	(0.024)	(0.017)
PE2	10	0.172	0.120	0.073	0.051	0.199	0.175	0.164	0.158
		(0.092)	(0.065)	(0.038)	(0.028)	(0.096)	(0.075)	(0.048)	(0.035)
	20	0.127	0.084	0.054	0.037	0.172	0.172	0.161	0.158
		(0.066)	(0.044)	(0.029)	(0.019)	(0.075)	(0.055)	(0.037)	(0.027)
	50	0.081	0.055	0.034	0.024	0.160	0.160	0.160	0.159
		(0.045)	(0.029)	(0.018)	(0.013)	(0.053)	(0.037)	(0.022)	(0.017)
	100	0.056	0.039	0.025	0.017	0.159	0.159	0.159	0.160
		(0.030)	(0.021)	(0.012)	(0.009)	(0.040)	(0.026)	(0.018)	(0.011)

In Table 1, we report the mean and standard error (SD, in parentheses) results in each case for the three estimators. First of all, it shows that MGE is more accurate and more efficient in estimating $\overline{\beta}$ than PE1 and PE2 in both DGP1 and DGP2. And an increase in either N or T will result in a substantial decrease in MGE. Secondly, the bias of PE1 is increasing as N becomes large, and this bias seems to expand as N and T increase by the same rate. Finally, the pooled estimator PE2 is consistent to $\overline{\beta}$ under homoscedasticity (DGP1), and the reason is that the bias under H_1 will be zero if $\Omega_{v,i} = \Omega_v$ for all i. But PE2 is still less accurate and less efficient than MGE since it induces other random errors. Under homoscedasticity (DGP2), the bias term is nonzero, and PE2 becomes a biased estimator for $\overline{\beta}$ even for large N or T or both.

Similarly, Table 2 reports the means and SDs of $\{MSE_m(f)\}_{m=1}^{M}$ for MEG and PE1[2], where $MSE_m(f) = \frac{1}{T}\sum_{t=1}^{T}(\widehat{f}_m(\tau_t) - \overline{f}(\tau_t))^2$. It shows that MGE is more accurate and efficient than PE1. While the bias of MGE decreases as either N or T increases, the bias of PE1 increases significantly as N increases.

[2]PE2 is not included in this comparison since it does not estimate the trending functions.

Table 2: Means and SDs of the absolute errors for estimating $\overline{f}$ in DGP1 and DGP2

		DGP1				DGP2			
	$N\backslash T$	10	20	50	100	10	20	50	100
MGE	10	0.219	0.086	0.032	0.016	0.198	0.082	0.030	0.014
		(0.297)	(0.094)	(0.034)	(0.015)	(0.241)	(0.096)	(0.031)	(0.013)
	20	0.142	0.055	0.022	0.010	0.130	0.053	0.021	0.010
		(0.180)	(0.055)	(0.023)	(0.010)	(0.156)	(0.059)	(0.022)	(0.010)
	50	0.097	0.033	0.013	0.007	0.088	0.032	0.012	0.006
		(0.119)	(0.034)	(0.014)	(0.007)	(0.117)	(0.039)	(0.013)	(0.006)
	100	0.078	0.030	0.010	0.005	0.069	0.023	0.009	0.005
		(0.101)	(0.036)	(0.012)	(0.006)	(0.083)	(0.032)	(0.010)	(0.006)
PE1	10	0.383	0.205	0.092	0.050	0.394	0.241	0.168	0.132
		(0.402)	(0.205)	(0.085)	(0.050)	(0.415)	(0.200)	(0.117)	(0.070)
	20	0.418	0.245	0.151	0.118	0.288	0.179	0.119	0.091
		(0.412)	(0.227)	(0.126)	(0.090)	(0.278)	(0.162)	(0.080)	(0.048)
	50	1.231	0.920	0.810	0.761	0.443	0.303	0.248	0.244
		(0.810)	(0.524)	(0.331)	(0.226)	(0.355)	(0.232)	(0.149)	(0.114)
	100	3.433	2.971	2.804	2.811	1.649	1.520	1.405	1.412
		(1.330)	(0.950)	(0.559)	(0.382)	(0.867)	(0.610)	(0.379)	(0.253)

Table 3: Empirical rejection ratios of homogeneity test statistic

	acv				bcv			
$N\backslash T$	10	20	50	100	10	20	50	100
DGP3 (H_0): $\beta_i = \beta_0$, homoscedasticity								
10	0.204	0.110	0.080	0.054	0.057	0.054	0.055	0.044
20	0.190	0.102	0.088	0.044	0.054	0.054	0.055	0.045
50	0.188	0.106	0.086	0.057	0.052	0.046	0.054	0.042
100	0.019	0.092	0.076	0.052	0.050	0.066	0.054	0.046
DGP4 (H_0): $\beta_i = \beta_0$, heteroscedasticity								
10	0.196	0.126	0.084	0.058	0.058	0.064	0.060	0.062
20	0.218	0.099	0.072	0.058	0.070	0.058	0.064	0.051
50	0.182	0.100	0.066	0.048	0.044	0.040	0.052	0.050
100	0.224	0.104	0.066	0.044	0.066	0.050	0.052	0.032
DGP5 (H_1): $\beta_i = \beta_0 + \frac{\log(N)\log(T)}{T^{3/4}}\epsilon_i$, homoscedasticity								
10	0.372	0.366	0.446	0.518	0.154	0.282	0.414	0.506
20	0.450	0.454	0.572	0.624	0.194	0.348	0.526	0.618
50	0.492	0.608	0.708	0.768	0.240	0.478	0.674	0.760
100	0.566	0.662	0.714	0.808	0.304	0.584	0.688	0.788
DGP6 (H_1) : $\beta_i = \beta_0 + \frac{\log(N)\log(T)}{2\sqrt{NT}}\epsilon_i$, heteroscedasticity								
10	0.292	0.298	0.446	0.576	0.106	0.218	0.416	0.606
20	0.354	0.342	0.630	0.746	0.146	0.266	0.590	0.762
50	0.352	0.508	0.756	0.894	0.168	0.388	0.746	0.890
100	0.374	0.596	0.864	0.944	0.198	0.492	0.850	0.938

5.2 Homogeneity test results

In this part, we consider four new DGPs with $\beta_i = \beta_0 + \delta_{NT}\epsilon_i$, where $\beta_0 = (1, 2)^\top$, and

- DGP 3: $\epsilon_i = 0_d$; $\sigma^2_{v,i} = 1$ for all i.
- DGP 4: $\epsilon_i = 0_d$; $\sigma^2_{v,i} = 1$ for $i = 1, \cdots, [N/2]$, and $\sigma^2_{v,i} = 2$ otherwise.
- DGP 5: $\delta_{NT} = \frac{\log(N)\log(T)}{T^{3/4}}$; $\epsilon_i \sim N(0, I_d)$ for all i; $\sigma^2_{v,i} = 1$ for all i.
- DGP 6: $\delta_{NT} = \frac{\log(N)\log(T)}{2\sqrt{NT}}$; $\epsilon_i = (1, -1)^\top$, $\sigma^2_{v,i} = 1$ for $i = 1, \cdots, [N/2]$; otherwise $\epsilon_i = (-1, 1)^\top$ and $\sigma^2_{v,i} = 2$.

Data simulated by DGP3 to DGP6 will be used to test the null hypothesis:

$$H_0 : \beta_i = \beta_0, \tag{5.2}$$

for all $i \geq 1$. Note H_0 is true under DGP3 (homoscedasticity) and DGP4 (heteroscedasticity). In comparison, DGP5 considers a random departure sequence from β_0 with the scale $\delta_{NT} = \frac{\log(N)\log(T)}{T^{3/4}}$, and DGP6 introduces departure sequence with $\delta_{NT} = \frac{\log(N)\log(T)}{2\sqrt{NT}}$. To decide whether to accept or reject the null hypothesis, we first use the 5% critical value of the non-standard asymptotic distribution in Eqs. (4.8) from Kiefer et al. (2000). Since the test statistic may converge relatively slow to its asymptotic distribution under H_0, we also employ the bootstrap procedure proposed in last section. All simulations are repeated by 1000 times, and the bootstrap steps are repeated by $B = 250$ times in each simulation.

The upper panels of Table 3 compare the rejection ratios by using an asymptotic critical value (acv) and the bootstrap critical value (bcv) for DGP3 and DGP4. It shows the rejections using the given critical values approach to the right size as T increases, but the contribution by increasing N seems small. Thus there exists an oversize problem by this approach when T is small. In comparison, the size by using the bootstrap critical value in each case is much more accurate when T is small. For example, the rejection ratio is close to 0.05 for the sample size of $N = T = 10$. Eventually, the rejections based on both the acv and bcv are around 0.05 when $T = 100$.

On the other hand, the rejection ratios increase substantially as either N or T increases under DGP5 and DGP6, which indicates the test statistic has good local power performance. In DGP5, $\widehat{\beta}_{mg}$ and $\widehat{\beta}_p$ have identical asymptotic distributions, and the variance of β_i converges to 0 as T increases. In DGP6, the departure of β_i to β_0 is of the order $O\left(\frac{\log(N)\log(T)}{2\sqrt{NT}}\right)$. The fact that the rejection ratios approach 1 in such cases indicates that the power of the test statistic does not require a large departure under H_1.

6 Empirical application

In the related literature, the heterogeneous features of health care expenditure across OECD countries have been commonly attributed to factors like income, population structure, physician supply, public finance and technology growth (see Baltagi and Moscone, 2010, for a review). While income growth plays an important role in explaining health care spending diversities, its coefficients, the income elasticity of health care spending, are unlikely to be identical across countries due to different economic or demographic conditions. Similarly, the underlying technology progresses or efficiencies of health care system are impossible to be identical across all OECD countries. Such unobserved heterogeneities are considered in this application to illustrate the proposed model.

6.1 The dataset and econometric model

In this empirical example, we study the determinants of health care expenditure in OECD countries where those elasticities and underlying progresses are allowed to be heterogeneous. We collect the annual data on 32 OECD countries[3] from 1990 to 2012 for five variables from OECD Health Statistic 2015 database including health care expenditure (HE), GDP, the dependent population (under 15 or above 65) over working population (dependency rate,DR), physicians per 1000 persons (PH), and the share of government health care expenditure in total health care spending (GHE). Both HE and GDP are per capita data translated into US Dollars by purchasing power parity and expressed in natural logarithm. All the variables have been extracted by the time average to eliminate the fixed effects. Data plots for these variables are presented in Figure 2 and Figure 1.

The model we use can be rewritten as follows:

$$\mathrm{HE}_{it} = \beta_{1,i}\mathrm{GDP}_{it} + \beta_{2,i}\mathrm{PO}_{it} + \beta_{3,i}\mathrm{PH}_{it} + \beta_{4,i}\mathrm{GHE}_{it} + f_{it} + \alpha_i + e_{it}, \tag{6.1}$$

where the covariates are assumed to follow $X_{it} = g_{it} + x_i + v_{it}$. Thus they can be deterministically nonstationary because g_{it} is driven by e_{it} through x_i. Both f_{it} and β_i may vary across i to account for the heterogeneous technology process and covariates effects.

6.2 The estimated coefficients

Similar to the simulted data examples, we provide the results of MGE, PE1 and PE2 for comparison.

Figure 3 plots the estimation results for the individual $\widehat{\beta}_i$ (scatter points) in comparison to their averages $\widehat{\beta}_{mg}$ (blue dash line). First of all, the income elasticities of health care spending in Figure 3 for most countries are distributed between 0 and 1, indicating that health care is a necessity good and governmental intervention is necessary at least for the majority of the OECD countries. Secondly, the estimates of DR and GHE in OECD countries are distributed closely around zero, indicating they are relatively trivial in determining health care expenditure. Thirdly, the estimates of PH are quite diversified: the supply effects seem to be quite large in

[3] Chile and Slovak Republic are not include due to inadequate data

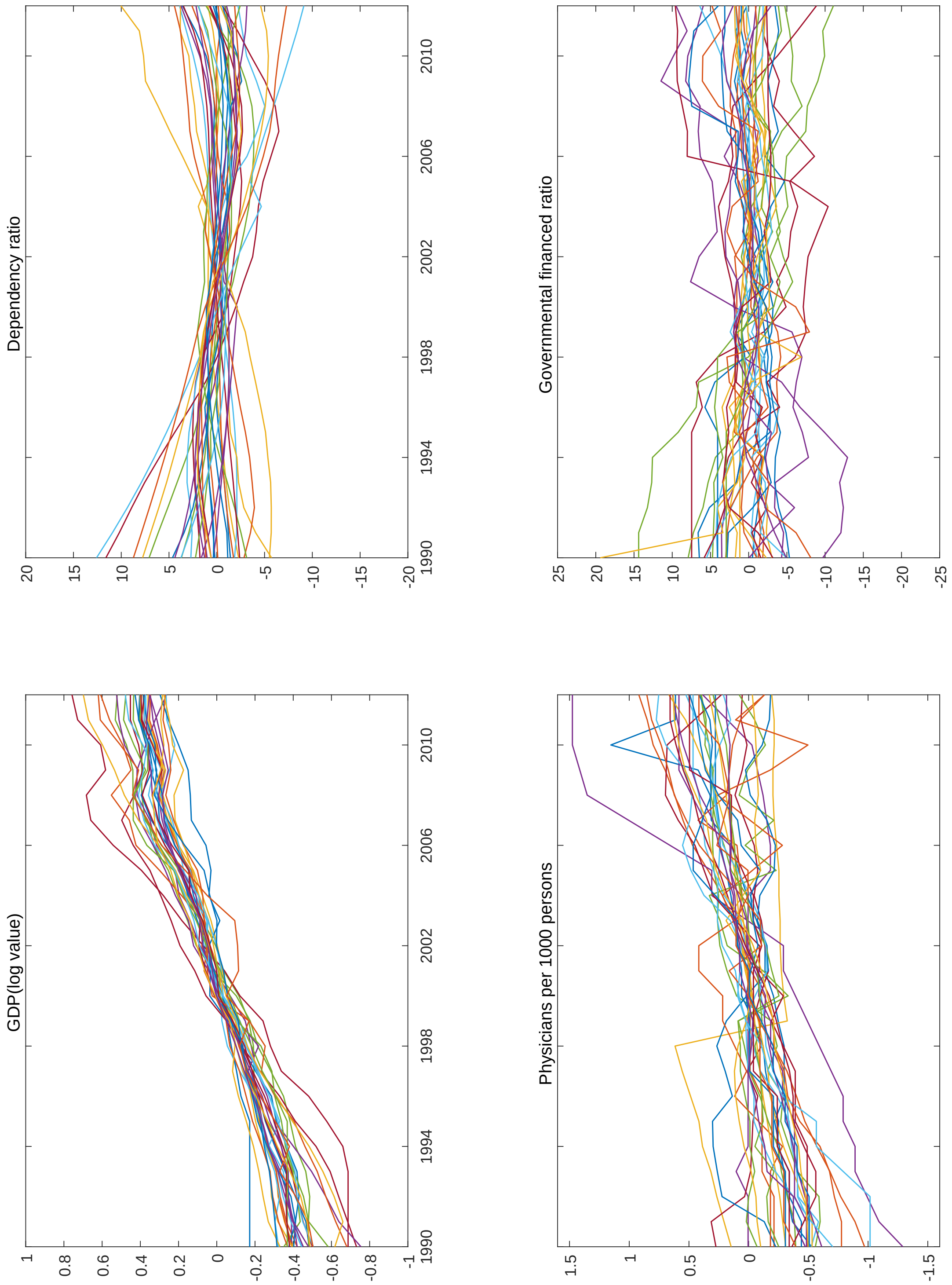

Figure 1: Plots of explanatory variables in OECD countries

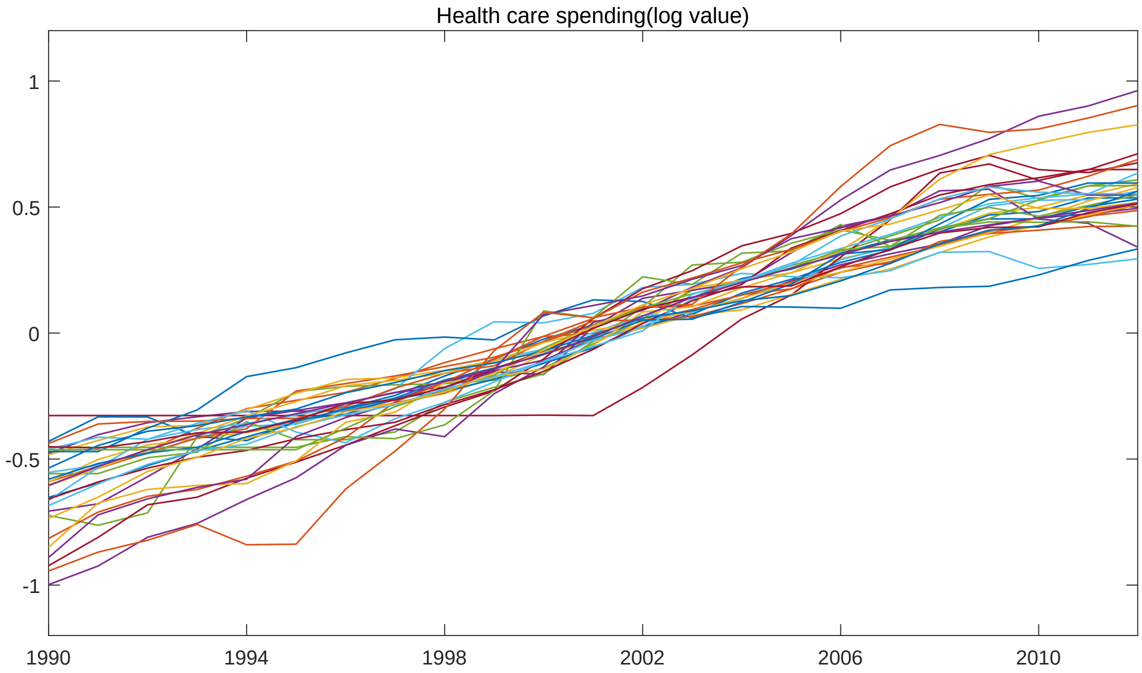

Figure 2: Plots of health care spending in OECD countries

some Asian countries, while they are close to zero for the others. And overall, the MGE shows the average effect of PH in OECD countries is close to zero[4].

In Table 4, we provide the MGE, PE1, and PE2 estimates for the average effects, where the standard errors (in parentheses) are calculated using a wild bootstrap procedure, and "*" indicating a significance under the 5% level (two-tailed) by the bootstrap procedure. On the one hand, the MGE estimates of income and public financed ratio are significantly positive (0.324 and 0.006 respectively), while the estimated effects for dependent ratio and physicians supply are small and insignificant, which are consistent with the findings obtained in the related literature(see Baltagi and Moscone, 2010). On the other hand, the PE1 estimate of the income elasticity is about two times of the MGE estimate. Based on our earlier analysis in Section 4, the results of PE1 are not reliable due to the unobservable heterogeneity in f_{it}. The estimates of MGE and PE2 are more convincible especially when they appear to be similar.

Theorem 3.4 also indicates that we can use the group version of MGE to analyse the average effects for the given groups, which will be more accurate in comparison to the individual estimates. Thus, we divide the 32 OECD countries into six groups by their geographic locations as follows:

- Asia and Australasia (G1): Korea, Japan, Israel, Turkey, Australia, New Zealand;
- Former Eastern Europe (G2): Czech Republic, Estonia, Hungary, Poland, Slovenia, Aus-

[4]More detail of individual coefficients estimation is provided in the supplementary file, together with their bootstrap version standard errors.

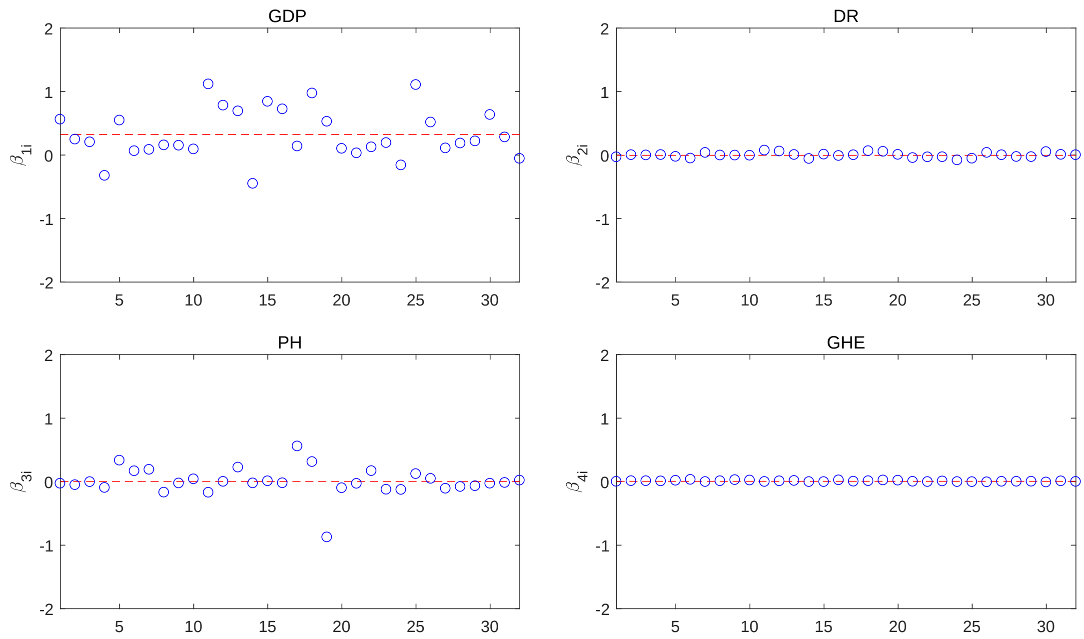

Figure 3: Plots of $\widehat{\beta}_i$ for OECD countries

Table 4: Mean group and pooled estimators

	MGE	PE1	PE2		MGE	PE1	PE2
GDP	0.324*	0.663*	0.369*	PH	0.000	0.021	−0.020
	(0.029)	(0.062)	(0.041)		(0.023)	(0.015)	(0.011)
DR	−0.004	−0.008*	−0.004	GHE	0.006*	0.010*	−0.001
	(0.003)	(0.002)	(0.004)		(0.001)	(0.001)	(0.001)

tria;

- Former Western Europe (G3): Belgium, France, Germany, Luxembourg, Netherlands, Switzerland;
- Former Southern Europe (G4): Greece, Italy, Portugal, Spain;
- Former Northern Europe (G5): Denmark, Finland, Iceland, Norway, Sweden;
- North America, Ireland and United Kingdom (G6): Canada, Ireland, Mexico, United Kingdom, United States.

Table 5: Mean group estimators in each group

	G1	G2	G3	G4	G5	G6
GDP	0.544*	0.352*	0.209*	0.762*	0.256*	−0.091
	(0.064)	(0.101)	(0.049)	(0.098)	(0.058)	(0.068)
DR	0.009	0.004	−0.006	0.001	−0.023*	−0.010*
	(0.006)	(0.010)	(0.007)	(0.012)	(0.005)	(0.004)
PH	0.164	0.088*	−0.146*	−0.046	0.002	−0.044
	(0.088)	(0.038)	(0.023)	(0.040)	(0.059)	(0.051)
GHE	−0.001	0.002	0.013*	0.004	0.012*	0.005*
	(0.001)	(0.003)	(0.001)	(0.002)	(0.003)	(0.002)

The MGE results in each group together with their standard errors are reported in Table 5, where "*" indicating a significance of the 5% level similar to Table 4. First of all, the income elasticities are positive and significant in all groups but G6. One percent of income increase will bring about 0.76 and 0.54 percent of health care spending increase in the Former Southern Europe and Asia, respectively, which are much higher than other groups. Secondly, both the population structure and the public finance are significant factors in some groups (the Former Northern Europe and the Group of North America, Ireland and United Kingdom), but the effects are quite small. Finally, health care supply factor plays a positive role in Asia and the Former Eastern Europe, and significantly negative in the Former Western Europe, which may be related to the characteristics of the local health care systems.

6.3 Coefficient homogeneity test on OECD data set

Although Figure 3 gives a rough picture on $\widehat{\beta}_i$ across different OECD countries, a more explicit criterion is necessary to decide whether the underlying coefficients are heterogeneous or not. Similarly, in Table 4 we have no clue about which one of MGE and PE2 is more preferable. And as we discussed in Section 3, even only one parameter in β_i is heterogeneous, PE2 is likely to be biased in other homogeneous parameters.

We therefore consider testing whether the coefficients are identical across OECD countries:

$$H_0 : \beta_i = \beta_0, \quad \text{for all } i \geq 1. \tag{6.2}$$

Furthermore, we would like to investigate whether the coefficients are homogeneous within each geographic group, that is, for any $k = 1, \cdots, 6$,

$$H_0^{(k)} : \beta_i = \beta_0, \quad \text{for } i \in G_k. \tag{6.3}$$

The test statistics for H_0 and $H_0^{(1)}$ to $H_0^{(6)}$ are calculated based on Eqs. (4.6). According to the simulated data examples, using an asymptotic critical value is likely to cause a size distortion problem when T is relatively small. Thus we utilize the procedure proposed in Section 4 to calculate the bootstrap critical value and the p-value in each case. These results are reported in Table 7, together with the asymptotic critical value from Kiefer et al. (2000) in the last line, where "*" indicates a rejection of the null hypothesis at the 5% significance level.

From Table 7, H_0 in Eqs. (6.2) has been rejected based on the bcv, with a $p-$value being 0.01. Further testing results on Eqs. (6.3) help us understand the heterogeneity within each groups. On the one hand, although the null has been rejected at the OECD level, the coefficients are not necessarily heterogeneous within every group. For example, the null hypothesis in Eqs. (6.3) has been accepted in the Asian group and the Former Northern Europe group, with p-values being 0.57 and 0.99, respectively. On the other hand, using the acv in the last row, the conclusions are not changed by using either the bcv or the acv in each case. In addition, more details and discussions are provided in Appendix F of the supplementary material.

Table 6: Mean group estimators in each group

	OECD	G1	G2	G3	G4	G5	G6
Test statistic	588.60*	66.22	254.72*	759.17*	1678.63*	4.53	539.49*
bcv	313.20	328.73	291.73	425.96	481.74	389.95	479.83
Bootstrap p-value	0.01	0.57	0.07	0.00	0.00	0.99	0.03
acv				261.32			

6.4 The estimated trending functions

The estimated trending functions for each country are presented in Figure 4 by groups. Also, the estimated trends together with their bootstrap standard errors are provided in Appendix F of the supplementary material. In Figure 4, first of all, we find a common rising tendency among all the estimated trends, this indicates that technology advances will commonly generate growth in cost and expanded longevity (Chandra and Skinner, 2012) in all countries. Secondly, it is also natural to see a substantial diversity among these trends since there is no unique pattern for technology progress among OECD countries. Thirdly, the estimated trends are almost identical

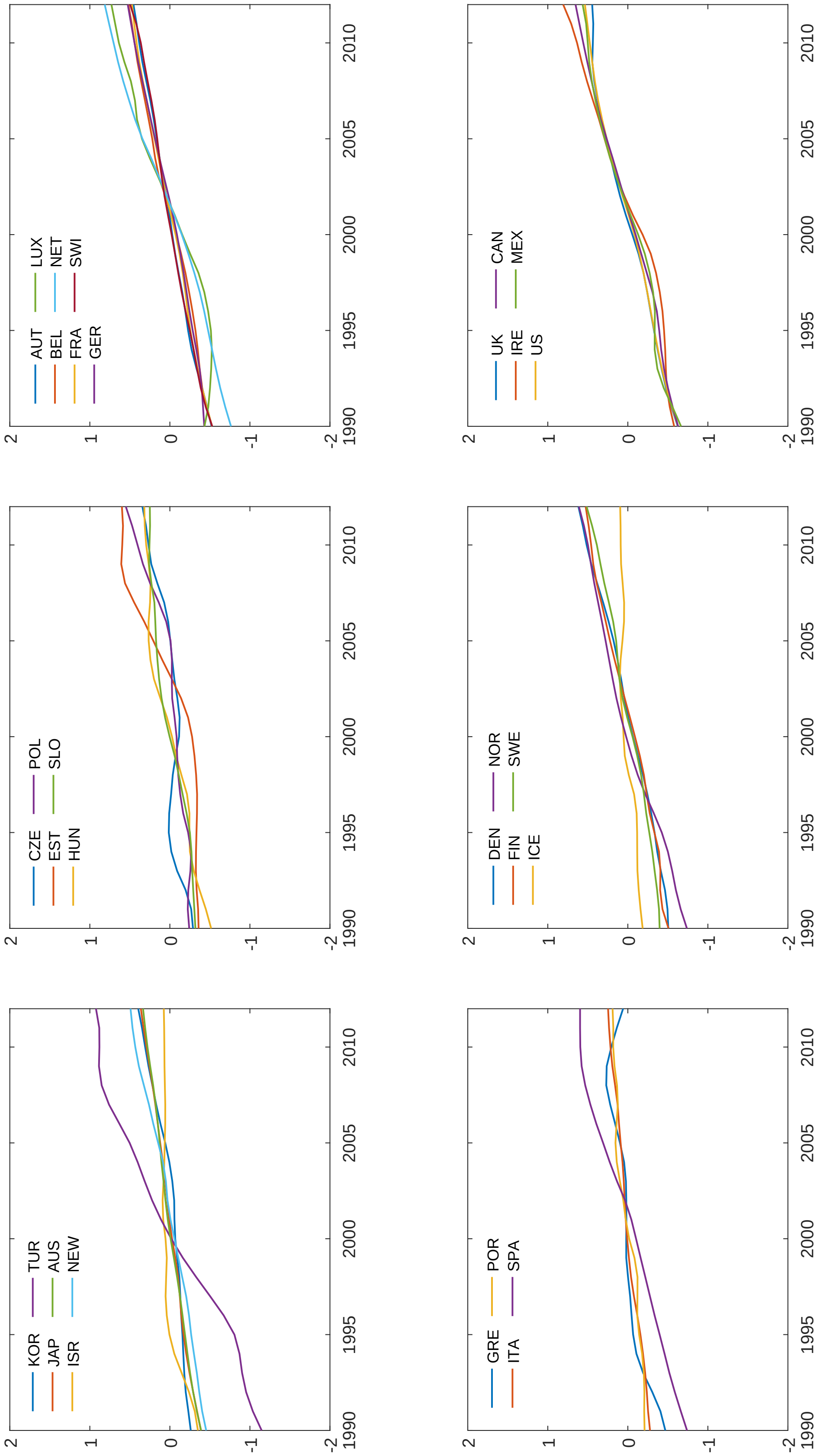

Figure 4: The estimated trends in OECD countries by groups

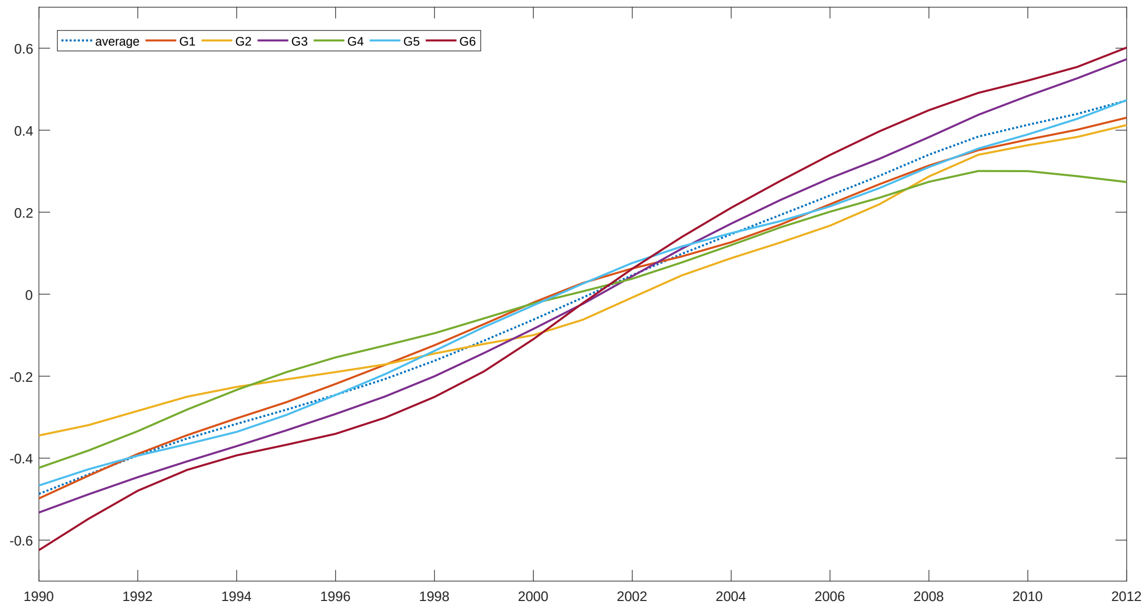

Figure 5: The mean group estimated trends for each group

in the Former Western European countries except in the Netherlands and Luxembourg. This indicates the similarities among the countries in some groups are much stronger than the average level.

Finally, in Figure 5, we compare the MGE estimaed trends in the six groups in comparison to the OECD average. It is obvious that the estimated trends in G5 and G6 are steeper than the OECD average, which may support that technology and R&D sector should be more advanced in U.S. and the Former Western European countries. Besides, although several groups experienced a slowdown around 2008, the Former Southern Europe group is the only one with a substantial dropping ever since the Subprime Crisis.

7 Conclusions and discussions

In this paper, we have introduced a heterogeneous panel data model with cross-sectional dependence. The mean group estimators, the weighted averaging estimators together with a pooled estimator have all been proposed and then studied. Asymptotic results have been derived and we have shown that each of these estimators achieve the possible optimal rate of convergence as $(N, T) \to (\infty, \infty)$, while N is allowed to increase faster than T. A new statistic has been proposed to test the homogeneity hypotheses. In our test, there is no need to discriminate the pooled estimator from the mean group estimator under H_0, and the pooled estimator is allowed to be consistent under H_1. In addition, there is also no need to require any consistent estimator for the covariance matrix involved in the test statistic.

Simulated examples have been given to demonstrate that the mean group estimators are more accurate and efficient than the pooled estimators in the finite sample performance. Additional simulated examples are used to illustrate the size and power of the proposed test. In particular, the bootstrap critical value is helpful in achieving a desirable size especially when the sample size is small. In addition, a real data example using an OECD health case expenditure dataset is studied to show the applicability of the proposed model.

One main limitation in this paper is X_{it} can be correlated with e_{it} only through x_i. A more general setting up is to further introduce endogeneity by allowing v_{it} to be correlated with e_{it}.

Appendix A Related discussion of the main results

Justification of Assumption 1: Assumption 1(i) requires v_t and e_t to be a joint α-mixing sequence where the mixing coefficient satisfies a moderate rate such as $o\left(k^{-(4+\delta)(1+\phi)/\delta}\right)$. Assumption 1(ii) ensures $\Omega_{v,i}$ to be nonsingular such that $\Omega_{v,i}^{-1}$ exists. And Assumption 1(iii) includes some regular moment conditions on v_{it} and e_{it}.

Justification of Assumption 2: Assumption 2 imposes some mild conditions on e_{it} and v_{it}. Take Assumption 2(i) for example, it is satisfied if either of the following cases holds,

- If e_{it} is cross-sectionally independent, and be serially independent or weakly dependent (such as α-mixing), then Assumption 2(i) holds.
- Consider a spatial autoregressive (SAR) process such as $e_t = \lambda W_N e_t + \varepsilon_t = S_N^{-1}\varepsilon_t$. If we assume S_N^{-1} is uniformly bounded in both row and column sums in absolute value as in Lee and Yu (2015), and ε_t is i.i.d. or α-mixing sequence over time, it is easy to check such a SAR process satisfies Assumption 2(i).
- For e_{it} to be the random errors with a linear structure, Assumption 2(i) is satisfied by imposing appropriate restrictions on the linear coefficients (see Kim and Sun, 2013; Chudik and Pesaran, 2015; Lee and Robinson, 2016).
- In addition, see Dong et al. (2015a,b) for an illustration that weak factor models are included by Assumption 2(i).

Assumption 2(ii)(iii)(iv) is included to handle the errors introduced by the semiparametric estimation procedure. It is clear that if v_{it} and e_{js} are independent for all i, j, t, s, then these conditions are ensured by Assumption 2(i). In addition, the superscript T_1 and T_2 in (ii) and (iii) will be used for some uniform convergence results for establishing the asymptotic distribution of the proposed homogeneity test statistic.

Justification of Assumption 3: Assumption 3 imposes some conditions necessary for establishing the central limit theorems (CLTs) in Theorems 3.3 and 3.4. In particular, (i) and (ii) require the positive definiteness of the asymptotic covariance matrices. (iii) will be used in the Lindeberg condition for proving the CLTs. There are ρ_i and ρ_j contained in σ^2_{eK} because different bandwidths can be allowed across different individuals. If the same bandwidth is chosen, $\sigma^2_{eK} = \nu_0 \frac{1}{N}\sum_{i=1}^N \sum_{j=1}^N \sigma_e(i,j)$, which is the same as the pooled estimators.

Justification of Assumption 4: Assumption 4(i) specifies the condition on the kernel function used in estimation, which is quite common in the literature. Assumption 4(ii) gives a regular smooth condition in the local linear fitting. Condition $\int_0^1 f_i(\tau) = 0$ helps in identifying $f_i(\tau)$ from α_i. Assumption 4(iii) includes some restrictions on h and expansion rate between N and T. If we assume $h \propto (NT)^{-1/5}$ and $N = [T^c]$, Assumption 4 will require $c < 1 + \delta^*/(2+\delta^*/2)$ for $\delta^* < \delta$.

Justification of Assumption 5 : Assumption 5 is imposed to derive the asymptotic distribution of $\widehat{\beta}_{mg} - \widehat{\beta}_p$. It actually includes two different scenarios. For the homoscedasticity ($\Omega_{v,i} = \Omega_v$ for all i), $\widehat{\beta}_p$ and $\widehat{\beta}_{mg}$ has the same $\sqrt{NT}$ asymptotic distribution, thus the leading term of $\widehat{\beta}_p - \widehat{\beta}_{mg}$ will be $\frac{1}{NT}\sum_{i=1}^N \sum_{t=1}^T \Omega_{v,i}\Delta_{Ti,1}\xi_{it}$, which is of order $O_P\left(\frac{1}{\sqrt{NT^2}}\right)$. In this case, $\zeta_{N,t} = \frac{1}{\sqrt{NT}}\sum_{i=1}^N \sum_{s=1}^T \Omega_{v,i}\eta_{is}\xi_{it}$. Otherwise, $\gamma_{NT} = \gamma_N^{-1}$, $\zeta_{N,t} = \frac{1}{\sqrt{\gamma_N NT}}\sum_{i=1}^N \widetilde{\Omega}_{v,i}\xi_{it}$, which is similar to the situation in Assumption 3.

The condition (ii) will be used to establish the CLT results for $\widehat{\beta}_p - \widehat{\beta}_{mg}$. Conditions (iii) and (iv) are imposed for the establishment of the asymptotic distribution of the test statistic. In particular, the two relationships in (iii) will be used to verify the finite distribution and tightness conditions in the proof of the functional central limit theorem for $L_{NT}(r)$.

Justification of Assumption 6: The major part of Assumption 6 is roughly an analogous version of Assumption 2 under $\gamma_{NT} = T$. In standard practice, such as in Theorem 3.3, we normally ignore small items of order $o_P\left(\frac{1}{\sqrt{NT}}\right)$. However, if $\gamma_{NT} = T$, the asymptotic distribution in Lemma 4.1 will be established with a rate of $\sqrt{NT^2}$. Thus some technical calculations and assumptions are then required to verify those items that are normally ignored.

Justification of Assumption 7: Assumption 7(i) specifies the rate of δ_{NT} such that the departure under H_1 can not be too small (no slower than $\frac{1}{\sqrt{\gamma_{NT}NT}}$). Assumption 7(ii) requires an upper bound for δ_{NT} to simplify the discussion. Assumption 7(iii), together with Assumption 7(i), ensures that $\widehat{\beta}_p$ has certain bias under H_1, thus the test has a good power performance.

If we drop condition (ii), thus δ_{NT} is required to expand at a rate faster than $\gamma_{NT}^{-1/2}$. In this case, the leading terms of $\widehat{\beta}_p - \widehat{\beta}_{mg}$ under H_1 will include an extra random error dominated by $\frac{1}{\sqrt{NT}}\sum_{i=1}^N \sum_{t=1}^T \eta_{it}\Omega_{v,i}\widetilde{\epsilon}_i$. Based on our calculation, the power of the test should not be a problem since large δ_{NT} implies a more significant deviation from the null. But if we want to derive the asymptotic distribution as in Theorem 4.2, some detailed assumptions will be needed to ensure the joint asymptotic distribution of $\frac{1}{\sqrt{NT}}\sum_{i=1}^N \sum_{t=1}^T \eta_{it}\Omega_{v,i}\widetilde{\epsilon}_i$ with $\widehat{H}_{NT}$.

In addition, if ϵ_i is random, for example, if $\epsilon_i \sim i.i.d(0,\sigma^2)$, and is exogenous with v_{it} and e_{it}, it is clear that $\frac{1}{\sqrt{N}}\left(\frac{1}{N}\sum_{i=1}^N \Omega_{v,i}\sigma^2\Omega_{v,i}\right)^{-1/2}\sum_{i=1}^N \Omega_{v,i}\epsilon_i \xrightarrow{d} N(0_d, I_d)$. Then it is straightforward that $\widehat{J}_{NT}$ will have a similar asymptotic distribution under H_1 as long as δ_{NT} grows to zero at an order faster than $\frac{1}{\gamma_{NT}T}$.

Some corresponding examples for Assumption 7 have been given in the simulated data studies.

Appendix B Summaries of the main proofs

Throughout the rest of this paper, let C be a generic positive constant whose value may vary from place to place but not dependent on both i and t. Since Theorems 3.1 and 3.2 are relatively standard results in semiparametric time series regression, we omit their proofs.

Proof of Theorem 3.3.

Recall ι_T is a $T \times 1$ vector of ones, and by definition of $s_i(\tau)$ we have $s_i(\tau)\iota_T = 1$ for $1 \le i \le N$, $0 \le \tau \le 1$. Note that $S_i = \left(s_i^\top(1/T), \cdots, s_i^\top(t/T)\right)^\top$, then $(I_T - S_i)\iota_T = 0$. Thus we have

$$\widetilde{Y}_i = (I_T - S_i)\left(Y_i - \iota_T Y_{i\cdot}\right) = (I_T - S_i)Y_i, \tag{B.1}$$

$$\widetilde{X}_i = (I_T - S_i)\left(X_i - X_{i\cdot}^\top \otimes \iota_T\right) = (I_T - S_i)X_i. \tag{B.2}$$

Recall $\widehat{\beta}_w = \sum_{i=1}^N w_i \left(\widetilde{X}_i^\top \widetilde{X}_i\right)^{-1} \widetilde{X}_i \widetilde{Y}_i$, and $\widehat{\Omega}_{x,Ti} = \frac{1}{T}\widetilde{X}_i^\top \widetilde{X}_i$. Then together with Eqs. (B.1) and Eqs. (B.2), we have

$$\begin{aligned}
\widehat{\beta}_w - \overline{\beta}_w &= \frac{1}{T}\sum_{i=1}^N w_i \widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_i^\top (I_T - S_i)Y_i - \sum_{i=1}^N w_i\beta_i \\
&= \frac{1}{T}\sum_{i=1}^N w_i \widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_i^\top (I_T - S_i)(X_i\beta_i + \alpha_i\iota_T + f_i + e_i) - \sum_{i=1}^N w_i\beta_i \\
&= \sum_{i=1}^N w_i \left(\widehat{\Omega}_{x,Ti}^{-1}\left(\frac{1}{T}\widetilde{X}_i^\top \widetilde{X}_i\right)\beta_i - \beta_i\right) + \frac{1}{T}\sum_{i=1}^N w_i \widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_i^\top (I_T - S_i)\iota_T\alpha_i \\
&\quad + \frac{1}{T}\sum_{i=1}^N w_i \widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_i^\top \left(\widetilde{e}_i + \widetilde{f}_i\right) \\
&\equiv A_{NT,1} + A_{NT,2} + A_{NT,3},
\end{aligned} \tag{B.3}$$

where $\widetilde{f}_i = (I_T - S_i)f_i$, $\widetilde{e}_i = (I_T - S_i)e_i$. Note that $\widehat{\Omega}_{x,Ti} = \frac{1}{T}\widetilde{X}_i^\top \widetilde{X}_i$ and $(I_T - S_i)\iota_T = 0$, we have

$$A_{NT,1} = 0, \tag{B.4}$$

$$A_{NT,2} = 0. \tag{B.5}$$

For $A_{NT,3}$, denote

$$Q_{Ti} = \left(\widehat{\Omega}_{x,Ti}^{-1} - \Omega_{v,i}^{-1}\right)\Omega_{v,i}, \quad \Delta_{Ti} = \Omega_{v,i}^{-1}\left(\widehat{\Omega}_{x,Ti} - \Omega_{v,i}\right), \tag{B.6}$$

$$\widetilde{Q}_{Ti} = Q_{Ti} I(||Q_{Ti}|| \le \mathcal{M}), \quad \widetilde{Q}_{Ti}^c = Q_{Ti} - \widetilde{Q}_{Ti}, \tag{B.7}$$

where $\mathcal{M}$ is a bounded constant such that $\mathcal{M} > \sqrt{d} + c_v^{-1} d^{3/2} \max_i ||\Omega_{v,i}||$, where c_v is defined in Assumption 1(ii). Then we can decompose $\widehat{\Omega}_{x,Ti}$ as follows: calculations.

$$\begin{aligned}
\widehat{\Omega}_{x,Ti}^{-1} &= \Omega_{v,i}^{-1} + \left(\widehat{\Omega}_{x,Ti}^{-1} - \Omega_{v,i}^{-1}\right) = \Omega_{v,i}^{-1} - \widehat{\Omega}_{x,Ti}^{-1}\left(\widehat{\Omega}_{x,Ti} - \Omega_{v,i}\right)\Omega_{v,i}^{-1} \\
&= \Omega_{v,i}^{-1} - \widehat{\Omega}_{x,Ti}^{-1}\Omega_{v,i}\Delta_{Ti}\Omega_{v,i}^{-1} = \Omega_{v,i}^{-1} - \left(I_d + \widetilde{Q}_{Ti} + \widetilde{Q}_{Ti}^c\right)\Delta_{Ti}\Omega_{v,i}^{-1} \\
&= \left(I_d - \left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti} - \widetilde{Q}_{Ti}^c\Delta_{Ti}\right)\Omega_{v,i}^{-1}.
\end{aligned} \tag{B.8}$$

In addition, let $\widetilde{e}_{it}$ and $\widetilde{f}_{it}$ be the t-th entry of $\widetilde{e}_i$ and $\widetilde{f}_i$, and denote,

$$\widetilde{\xi}_{it} = \Omega_{v,i}^{-1}\widetilde{X}_{it}\left(\widetilde{e}_{it} + \widetilde{f}_{it}\right), \quad \xi_{it} = \Omega_{v,i}^{-1} v_{it} e_{it}, \quad \xi_{it}^c = \widetilde{\xi}_{it} - \xi_{it}. \tag{B.9}$$

Then we can rewrite $A_{NT,3}$ as:

$$\begin{aligned} A_{NT,3} &= \frac{1}{T}\sum_{i=1}^N w_i \left(I_d - \left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti} - \widetilde{Q}_{Ti}^c \Delta_{Ti}\right)\sum_{t=1}^T \widetilde{\xi}_{it} \\ &= \frac{1}{T}\sum_{i=1}^N\sum_{t=1}^T w_i\xi_{it} + \frac{1}{T}\sum_{i=1}^N\sum_{t=1}^T w_i \xi_{it}^c - \frac{1}{T}\sum_{i=1}^N\sum_{t=1}^T w_i\left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it} \\ &-\frac{1}{T}\sum_{i=1}^N\sum_{t=1}^T w_i \widetilde{Q}_{Ti}^c \Delta_{Ti}\widetilde{\xi}_{it} \equiv A_{NT,3,1} + \cdots + A_{NT,3,4}. \end{aligned} \tag{B.10}$$

In Propositions (ii) and (iii) in Appendix C below, we have shown that

$$\mathrm{E}\,||A_{NT,3,2}||^2 \leq \frac{C\sqrt{N}}{\gamma_{N,w}T^2 h}, \quad \mathrm{E}\,||A_{NT,3,3}|| \leq C\frac{N^{1/2}}{\gamma_{N,w}^{1/2} T^{1+\frac{\delta^*}{4+\delta^*}}}. \tag{B.11}$$

By $\frac{N}{T^2h^2} \to 0$ and $\frac{N^{1/2}}{T} \to 0$ in Assumption 4, we have

$$A_{NT,3,2} = o_P\left(\frac{1}{\sqrt{\gamma_{N,w}T}}\right), \quad A_{NT,3,3} = o_P\left(\frac{1}{\sqrt{\gamma_{N,w}T}}\right). \tag{B.12}$$

For $A_{NT,3,4}$, by the definition of $\widetilde{Q}_{Ti}^c$ in Eqs. (B.7), we can rewrite it as $Q_{Ti} I\left(\left|\left|\widetilde{Q}_{Ti}\right|\right| > \mathcal{M}\right)$. Then $A_{NT,3,4}$ is zero unless $||Q_{T,i}|| > \mathcal{M}$ holds for at least one $i \geq 1$. Thus for any $\epsilon > 0$, we have

$$\begin{aligned} &P\left(||A_{NT,3,4}|| > \epsilon\right) \leq P\left(||A_{NT,3,4}|| \neq 0\right) \\ &\leq P\left(\max_{1\leq i\leq N} ||Q_{T,i}|| > \mathcal{M}\}\right) \leq \sum_{i=1}^N P\left(||Q_{T,i}|| > \mathcal{M}\right) = o(1), \end{aligned} \tag{B.13}$$

where the last step is proved in Proposition C.1(i). Then we have

$$A_{NT,3,4} = o_P\left(\frac{1}{\sqrt{\gamma_{N,w}T}}\right). \tag{B.14}$$

Denote $V_{N,t} = \gamma_{N,w}^{-1/2}\sum_{i=1}^N w_i\xi_{it}$. To finish the proof of Theorem 3.3, we only need to show

$$\frac{1}{\sqrt{T}}\sum_{t=1}^T V_{N,t} \xrightarrow{D} N\left(0, \Sigma_{ve,w}\right). \tag{B.15}$$

By Assumption 1, $\{V_{N,t}, t \geq 1\}$ is an α-mixing sequence with $\mathrm{E}\left(V_{N,t}\right) = 0$ with zero mean. By $\mathrm{E}\,||V_{N,t}||^{2+\delta} \leq \infty$ and $\mathrm{Var}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T V_{N,t}\right) \to \Sigma_{ve,w}$ in Assumption 3, Eqs. (B.15) can be verified by a central limit theorem for α-mixing, such as Theorem 2.21 in Fan and Yao (2003).

Proof of Theorem 3.4.

By similar calculations to Eqs. (B.3), we have

$$\widehat{f}_{mg}(\tau) - \overline{f}(\tau) = \frac{1}{N}\sum_{i=1}^N s_i(\tau)\left(Y_i - Y_{i\cdot} i_T - \left(X_i - X_{i\cdot}^\top \otimes \iota_T\right)\widehat{\beta}_i\right) - \overline{f}(\tau)$$

$$
\begin{aligned}
&= \frac{1}{N}\sum_{i=1}^{N}\left(s_i(\tau)\left(Y_i - X_i\widehat{\beta}_i\right) - \left(Y_{i\cdot} - X_{i\cdot}^{\top}\widehat{\beta}_i\right)\right) - \overline{f}(\tau) \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(s_i(\tau)f_i - f_i(\tau)\right) - \frac{1}{N}\sum_{i=1}^{N} e_{i\cdot} - \frac{1}{N}\sum_{i=1}^{N} f_{i\cdot} \\
&\quad + \frac{1}{N}\sum_{i=1}^{N}\left(s_i(\tau)X_i - X_{i\cdot}^{\top}\right)\left(\beta_i - \widehat{\beta}_i\right) + \frac{1}{N}\sum_{i=1}^{N} s_i(\tau)e_i \\
&\equiv A_{NT,4} + \cdots + A_{NT,8}.
\end{aligned} \tag{B.16}
$$

By standard results of local linear fitting, we have

$$
A_{NT,4} = b_f(\tau)h^2 + o(h^2). \tag{B.17}
$$

By the moment condition in Assumption 2(i), $\mathrm{E}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} e_{it}\right)^2 = O(NT)$, then we have

$$
A_{NT,5} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} e_{it} = O_p\left(\frac{1}{\sqrt{NT}}\right) = o_P\left(\frac{1}{\sqrt{NTh}}\right). \tag{B.18}
$$

By the definition of Riemann integral and $\int_0^1 f_i(\tau)d\tau = 0$ in Assumption 4(ii), we have

$$
A_{NT,6} = O\left(\frac{1}{T}\right). \tag{B.19}
$$

By Proposition C.2 we have

$$
A_{NT,7} = o_P\left(\frac{1}{\sqrt{NTh}}\right), \tag{B.20}
$$

$$
\sqrt{NTh}A_{NT,8} \xrightarrow{D} N(0, \sigma_{eK}^2), \tag{B.21}
$$

which complete the proof of Theorem 3.4.

Proof of Lemma 4.1

Denote $\widehat{\Omega}_{x,NT} = \frac{1}{NT}\sum_{i=1}^{N}\widetilde{X}_i^{\top}\widetilde{X}_i$. Under $\beta_i = \beta$ for all $i \geq 1$, by the same calculations as in Eqs. (B.3), it is straightforward to see that

$$
\widehat{\beta}_p - \beta = \frac{1}{NT}\widehat{\Omega}_{x,NT}^{-1}\sum_{i=1}^{N}\widetilde{X}_i^{\top}\left(\widetilde{e}_i + \widetilde{f}_i\right), \quad \widehat{\beta}_{mg} - \beta = \frac{1}{NT}\sum_{i=1}^{N}\widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_i\left(\widetilde{e}_i + \widetilde{f}_i\right). \tag{B.22}
$$

Since $\widehat{\Omega}_{x,NT}^{-1} - \widehat{\Omega}_{x,Ti}^{-1} = \widehat{\Omega}_{x,NT}^{-1}\left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right)\widehat{\Omega}_{x,Ti}^{-1}$, then we have

$$
\begin{aligned}
\widehat{\beta}_p - \widehat{\beta}_{mg} &= \frac{1}{NT}\widehat{\Omega}_{x,NT}^{-1}\sum_{i=1}^{N}\left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right)\widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_i^{\top}(\widetilde{e}_i + \widetilde{f}_i) \\
&= \frac{1}{NT}\Omega_v^{-1}\sum_{i=1}^{N}\left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right)\widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_i^{\top}(\widetilde{e}_i + \widetilde{f}_i) \\
&\quad + \frac{1}{NT}Q_{NT}\sum_{i=1}^{N}\left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right)\widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_i^{\top}(\widetilde{e}_i + \widetilde{f}_i) \\
&\equiv A_{NT,9} + A_{NT,10},
\end{aligned} \tag{B.23}
$$

where $Q_{NT} = \widehat{\Omega}_{x,NT}^{-1} - \Omega_v^{-1}$.

Denote $A_{NT,11} = \frac{1}{NT}\sum_{i=1}^N \left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right)\widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_i^\top(\widetilde{e}_i + \widetilde{f}_i)$. Note that $Q_{NT} \xrightarrow{P} 0$, then for Lemma 4.1, we only need to prove

$$A_{NT,11} = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti,1}\right)\xi_{it} + o_P\left(\frac{1}{\sqrt{\gamma_{NT}NT}}\right) \tag{B.24}$$

$$\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}}\sum_{i=1}^N\sum_{t=1}^T \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti,1}\right)\xi_{it} \xrightarrow{D} N\left(0_d, \Sigma_{ve,1}\right). \tag{B.25}$$

Step(i): Proof of Eqs. (B.24).

Recall that $\widetilde{\Omega}_{v,i} = \Omega_{v,i} - \overline{\Omega}_N$, and $\Delta_{NT} = \widehat{\Omega}_{x,NT} - \overline{\Omega}_N$, $\Delta_{Ti} = \Omega_{v,i}^{-1}\left(\widehat{\Omega}_{x,Ti} - \Omega_{v,i}\right)$. Thus

$$\begin{aligned}\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT} &= \left(\Omega_{v,i} - \overline{\Omega}_N\right) + \left(\widehat{\Omega}_{x,Ti} - \Omega_{v,i}\right) - \left(\widehat{\Omega}_{x,NT} - \overline{\Omega}_N\right)\\ &= \widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti} - \Delta_{NT}.\end{aligned} \tag{B.26}$$

Then we have

$$\begin{aligned}A_{NT,11} &= \frac{1}{NT}\sum_{i=1}^N \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti}\right)\widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_i^\top(\widetilde{e}_i + \widetilde{f}_i) - \frac{1}{NT}\sum_{i=1}^N \Delta_{NT}\widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_i^\top(\widetilde{e}_i + \widetilde{f}_i)\\ &\equiv A_{NT,11,1} + A_{NT,11,2}.\end{aligned} \tag{B.27}$$

Recall in Eqs. (B.7) we have $\widehat{\Omega}_{x,Ti}^{-1} = \Omega_{v,i}^{-1} - \left(I_d + \widetilde{Q}_{Ti} + \widetilde{Q}_{Ti}^c\right)\Delta_{Ti}\Omega_{v,i}^{-1}$ and $\widetilde{\xi}_{it} = \Omega_{v,i}^{-1}\widetilde{X}_{it}\left(\widetilde{e}_{it} + \widetilde{f}_{it}\right)$, we have

$$\begin{aligned}A_{NT,11,1} = &\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti}\right)\widetilde{\xi}_{it} - \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti}\right)\left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it}\\ &- \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti}\right)\widetilde{Q}_{Ti}^c\Delta_{Ti}\widetilde{\xi}_{it} \equiv A_{NT,11,1,1} + \cdots + A_{NT,11,1,3}.\end{aligned} \tag{B.28}$$

Recall also that $\Omega_{v,i}^{-1}\widetilde{X}_{it}\widetilde{e}_{it} = \xi_{it} + \xi_{it}^c$, thus,

$$\begin{aligned}A_{NT,11,1,1} = &\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti,1}\right)\xi_{it} + \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \Omega_{v,i}\Delta_{Ti,2}\xi_{it}\\ &+ \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti}\right)\xi_{it}^c \equiv A_{NT,11,1,1,1} + \cdots + A_{NT,11,1,1,3}.\end{aligned} \tag{B.29}$$

In Proposition C.1(ii), Proposition C.3(ii) and (iii), we have shown that

$$A_{NT,11,1,1,2} = o_P\left(\frac{1}{\sqrt{\gamma_{NT}NT}}\right), \quad A_{NT,11,1,1,3} = o_P\left(\frac{1}{\sqrt{\gamma_{NT}NT}}\right) \tag{B.30}$$

Thus we have

$$A_{NT,11,1,1} = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti,1}\right)\xi_{it} + o_P\left(\frac{1}{\sqrt{\gamma_{NT}NT}}\right) \tag{B.31}$$

In Propositions C.1(iii) and C.3(iv) below, we have shown that

$$A_{NT,11,1,2} = o_P\left(\frac{1}{\sqrt{\gamma_{NT}NT}}\right). \tag{B.32}$$

Recall the definition of $\widetilde{Q}_{Ti}^c$, by the argument in Eqs. (B.13), we have

$$A_{NT,11,1,3} = o_P\left(\frac{1}{\sqrt{\gamma_{NT}NT}}\right). \tag{B.33}$$

Meanwhile, we have $\Delta_{NT} = o_P\left(\frac{1}{\sqrt{\gamma_{NT}}}\right)$ in Proposition C.3(i), together with the proof of Theorem 3.3, we have

$$A_{NT,11,2} = o_P\left(\frac{1}{\sqrt{\gamma_{NT}NT}}\right), \tag{B.34}$$

which completes the proof of Eqs. (B.24).

Step(ii): Proof of Eqs. (B.25).

Recall that $\zeta_{N,t} = \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}}\sum_{i=1}^N\left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti,1}\right)\xi_{it}$, we need to show

$$\sum_{t=1}^T \zeta_{N,t} \xrightarrow{D} N(0_d, \Sigma_{ve,1}). \tag{B.35}$$

Let $\mathscr{F}_{t-m}^{t+m} = \sigma\left\{(v_{t-m}, e_{t-m}), \cdots, (v_{t+m}, e_{t+m})\right\}$. In Proposition C.3(v), we have shown that,

$$\mathrm{E}\left|\left|\zeta_{N,t} - \mathrm{E}\left(\zeta_{N,t}|\mathscr{F}_{t-m}^{t+m}\right)\right|\right|^2 < v_m d_t \tag{B.36}$$

with $v_m = o\left(m^{-1}\right)$, and d_t irrelevant with m. This means that $\zeta_{N,t}$ is a L_2-NED process defined on α mixing sequence $\{(e_t, v_t), t \geq 1\}$. By Eqs. (B.36), Assumptions 5 and 6, $\{\zeta_{N,t}, t \geq 1\}$ satisfies the following conditions with $c_t = \frac{1}{\sqrt{T}}$: (a)$\mathrm{E}\left(\zeta_{N,t}\right) = 0$ and $\mathrm{Var}\left(\sum_{t=1}^T \zeta_{N,t}\right) \to \Sigma_{ve,1}$; (b)$\sup_T \mathrm{E}\left|\left|\zeta_{N,t}/c_t\right|\right|^{2+\delta} < \infty$; (c) $\zeta_{N,t}$ is of size -1 on $\{(v_t, e_t), t \geq 1\}$, which is α-mixing of size $-\phi$ such that $\phi > \frac{2+\delta}{\delta}$; and (d) $\sup_T TC_T^2 < \infty$. Thus by CLT for NED processes, such as Corollary 24.7 in Davidson (1994), we have

$$\sum_{t=1}^T \zeta_t \xrightarrow{D} N(0_d, \Sigma_{ve}), \tag{B.37}$$

Proof of Theorem 4.1

Denote $\widehat{L}_{NT}(r) = \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{t=1}^{[rT]}\left(\widehat{\Omega}_{x,NT}^{-1} - \widehat{\Omega}_{x,Ti}^{-1}\right)\widetilde{X}_{it}\widehat{e}_{it}$. If we can prove

$$\begin{pmatrix} \sqrt{\gamma_{NT}}\widehat{L}_{NT}(r) \\ \sqrt{\gamma_{NT}NT}(\widehat{\beta}_{mg} - \widehat{\beta}_p) \end{pmatrix} \Rightarrow \begin{pmatrix} \Sigma_{ve,2}^{1/2}W_d(r) - r\Sigma_{ve,2}^{1/2}W_d(1) \\ \Sigma_{ve,2}^{1/2}W_d(1) \end{pmatrix}. \tag{B.38}$$

Then, by using the definition of Riemann integral and continuous mapping theorem we have

$$\gamma_{NT}\widehat{H}_{NT} = \frac{1}{T}\sum_{m=1}^T \sqrt{\gamma_{NT}}\widehat{L}_{NT}\left(\frac{m}{T}\right)\sqrt{\gamma_{NT}}\widehat{L}_{NT}^\top\left(\frac{m}{T}\right) \xrightarrow{D} \Sigma_{ve,2}^{1/2}\Phi_d\Sigma_{ve,2}^{1/2\top}. \tag{B.39}$$

By Eqs. (B.38) and Eqs. (B.39), and continuous mapping theorem again, it is direct to see that Theorem 4.1 holds. Thus we only need to show Eqs. (B.38).

To prove Eqs. (B.38), denote that $L_{NT}(r) = \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}}\sum_{i=1}^N\sum_{t=1}^{[rT]}\left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti,1}\right)\xi_{it}$, note that by Lemma 4.1 we have

$$\sup_{0\leq r\leq 1}||\sqrt{\gamma_{NT}NT}\left(\widehat{\beta}_p - \widehat{\beta}_{mg}\right) - \Omega_v^{-1}L_{NT}(1)|| = o_P(1). \tag{B.40}$$

And in Proposition C.4 we have

$$\sup_{0\le r\le 1}\left|\left|\sqrt{\gamma_{NT}}\widehat{L}_{NT}(r)-\Omega_v^{-1}\left(L_{NT}(r)-rL_{NT}(1)\right)\right|\right|=o_P(1). \tag{B.41}$$

Then we need only to show

$$\begin{pmatrix} \Omega_v^{-1}\left(L_{NT}(r)-rL_{NT}(1)\right) \\ \Omega_v^{-1}L_{NT}(1) \end{pmatrix} \Rightarrow \begin{pmatrix} \Sigma_{ve,2}^{1/2}W_d(r)-r\Sigma_{ve,2}^{1/2}W_d(1) \\ \Sigma_{ve,2}^{1/2}W_d(1) \end{pmatrix}, \tag{B.42}$$

which requires to prove $L_{NT}(r)\Rightarrow L(r)$, where $L(r)=\Sigma_{ve,1}^{1/2}W_d(r)$ and $\Sigma_{ve,2}=\Omega_v^{-1}\Sigma_{ve,1}\Omega_v^{-1}$.

To prove the functional central limit theorem for the $d\times 1$ vector of stochastic processes $L_{NT}(r)$ defined on $D[0,1]^d$, by Theorem 29.16 in Davidson (1994), we need to show that for any fixed λ with $\lambda^\top\lambda=1$, we have

$$\lambda^\top L_{NT}(r)\Rightarrow\lambda^\top L(r) \tag{B.43}$$

and

$$P\left(L(r)\in C[0,1]^d\right)=1. \tag{B.44}$$

Condition Eqs. (B.44) is satisfied directly by the definition of $L(r)$. To verify the weak convergence for the univariate function $\lambda^\top L_{NT}(r)$, by Theorem 13.5 in Billingsley (1968)), we go through two steps to prove the finite dimensional distribution results: for arbitrary integer k and distinct $r_1,\cdots,r_k$,

$$(\lambda^\top L_{NT}(r_1),\cdots,\lambda^\top L_{NT}(r_k))\xrightarrow{D}(\lambda^\top L(r_1),\cdots,\lambda^\top L(r_k)) \tag{B.45}$$

and the tightness condition: for any $0\le r_1\le r_2\le r_3\le 1$, and $\beta>0$, $\alpha>\frac{1}{2}$, $C>0$,

$$\mathrm{E}\left((\lambda^\top L_{NT}(r_2)-\lambda^\top L_{NT}(r_1))^{2\beta}(\lambda^\top L_{NT}(r_3)-\lambda^\top L_{NT}(r_2))^{2\beta}\right)\le C|r_3-r_1|^{2\alpha}. \tag{B.46}$$

Step(i): Finite dimensional distribution. To show Eqs. (B.45), by Cramér-Wold device, it suffices to show for any $c=(c_1,\cdots,c_k)^\top$,

$$\sum_{l=1}^k c_l\lambda^\top L_{NT}(r_l)\xrightarrow{D}\sum_{l=1}^k c_l\lambda^\top L(r_l). \tag{B.47}$$

On the one hand, by Assumption 5(ii) and (iii), for any $0<r_1\le r_2\le 1$, we have

$$\mathrm{E}\left(L_{NT}(r_1)L_{NT}^\top(r_1)\right)\to\Sigma_{ve,1}, \tag{B.48}$$

$$\mathrm{E}\left(L_{NT}(r_1)L_{NT}^\top(r_2-r_1)\right)\to 0. \tag{B.49}$$

Thus by standard calculation, we have

$$\mathrm{Var}\left(\sum_{l=1}^k c_l\lambda^\top L_{NT}(r_l)\right)\to\mathrm{Var}\left(\sum_{l=1}^k c_l\lambda^\top L(r_l)\right). \tag{B.50}$$

On the other hand, recall that in Assumption 5, $\zeta_{N,t}=\frac{\sqrt{\gamma_{NT}}}{\sqrt{N}}\sum_{i=1}^N\left(\widetilde{\Omega}_{v,i}+\Omega_{v,i}\Delta_{Ti,1}\right)\xi_{it}$. Thus

$$\sum_{l=1}^k c_l\lambda^\top L_{NT}(r_l)=\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}}\sum_{l=1}^k c_l\sum_{i=1}^N\sum_{t=1}^{[r_lT]}\lambda^\top\left(\widetilde{\Omega}_{v,i}+\Omega_{v,i}\Delta_{Ti,1}\right)\xi_{it}$$

$$= \sum_{t=1}^{T} c_{T,t} \lambda^\top \zeta_{N,t}, \tag{B.51}$$

where $c_{T,t} = \frac{1}{\sqrt{T}} \sum_{l=1}^{k} (c_l I(t \leq [r_l T]))$. Note that for any given $0 \leq r_1, \cdots, \leq r_k \leq 1$, $\{c_{T,t}, t \geq 1\}$ will be a deterministic sequence of vectors with $T \left(\sup_{1 \leq t \leq T} ||c_{T,t}||^2 \right) < \infty$. Thus, $\{c_{T,t} \lambda^\top \zeta_{N,t}, t \geq 1\}$ is also a L_2-NED process satisfying all the requirements (a) to (d) in the proof of Lemma 4.1, which suffices to prove Eqs. (B.47).

Step(ii) Tightness condition. For the tightness condition in Eqs. (B.46), recall that $\zeta_{N,t} = \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \sum_{i=1}^{N} \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i} \Delta_{Ti,1} \right) \xi_{it}$, by Assumption 5, for any $0 \leq r_1 \leq r_2 \leq 1$, we have

$$\begin{aligned}
\mathrm{E} \, ||L_{NT}(r_2) - L_{NT}(r_1)||^{2+\delta} &\leq \mathrm{E} \left|\left| \sum_{t=[r_1 T]+1}^{[r_2 T]} \zeta_{N,t} \right|\right|^{2+\delta} \\
&\leq C_1 \left(\frac{[r_2 T] - [r_1 T]}{T} \right)^{\frac{2+\delta^*}{2}} \leq C (r_2 - r_1)^{\frac{2+\delta^*}{2}}.
\end{aligned} \tag{B.52}$$

Then by Cauchy-Schwarz inequality and $0 \leq r_1 \leq r_2 \leq r_3 \leq 1$

$$\begin{aligned}
&\mathrm{E} \left(\left(\lambda^\top L_{NT}(r_2) - \lambda^\top L_{NT}(r_1) \right)^{\frac{2+\delta^*}{2}} \left(\lambda^\top L_{NT}(r_3) - \lambda^\top L_{NT}(r_2) \right)^{\frac{2+\delta^*}{2}} \right) \\
&\leq \left(\mathrm{E} \, ||L_{NT}(r_2) - L_{NT}(r_1)||^{2+\delta^*} \right)^{1/2} \left(\mathrm{E} \, ||L_{NT}(r_3) - L_{NT}(r_2)||^{2+\delta^*} \right)^{1/2} \\
&\leq C (r_2 - r_1)^{\frac{2+\delta^*}{4}} (r_3 - r_2)^{\frac{2+\delta^*}{4}} \leq C |r_3 - r_1|^{\frac{2+\delta^*}{2}},
\end{aligned} \tag{B.53}$$

which verifies the tightness condition in Eqs. (B.46) with $\alpha = \beta = 1 + \delta^*/4$.

Proof of Theorem 4.2. By similar calculations in the proof of Lemma 4.1, we have

$$\begin{aligned}
\widehat{\beta}_p - \widehat{\beta}_{mg} &= \left(\widehat{\beta}_p - \beta_0 \right) - \left(\widehat{\beta}_{mg} - \beta_0 \right) \\
&= \frac{1}{NT} \sum_{i=1}^{N} \left(\widehat{\Omega}_{x,NT}^{-1} - \widehat{\Omega}_{x,Ti}^{-1} \right) \widetilde{X}_i^\top \left(\widetilde{e}_i + \widetilde{f}_i \right) + \delta_{NT} \left(\widehat{\Omega}_{x,NT}^{-1} \frac{1}{N} \sum_{i=1}^{N} \widehat{\Omega}_{x,Ti} \epsilon_i - \overline{\epsilon} \right) \\
&\equiv \theta_{NT,1} + \theta_{NT,2}.
\end{aligned} \tag{B.54}$$

Note that $\widehat{\Omega}_{x,NT} = \frac{1}{N} \sum_{i=1}^{N} \widehat{\Omega}_{x,Ti}$, thus $\overline{\epsilon} = \widehat{\Omega}_{x,NT}^{-1} \frac{1}{N} \sum_{i=1}^{N} \widehat{\Omega}_{x,Ti} \overline{\epsilon}$, recall $\widetilde{\epsilon}_i = \epsilon_i - \overline{\epsilon}$, we have

$$\begin{aligned}
\theta_{NT,2} &= \delta_{NT} \widehat{\Omega}_{x,NT}^{-1} \frac{1}{N} \sum_{i=1}^{N} \widehat{\Omega}_{x,Ti} \widetilde{\epsilon}_i \\
&= \delta_{NT} \widehat{\Omega}_{x,NT}^{-1} \frac{1}{N} \sum_{i=1}^{N} \Omega_{v,i} \widetilde{\epsilon}_i + \delta_{NT} \widehat{\Omega}_{x,NT}^{-1} \frac{1}{N} \sum_{i=1}^{N} \Omega_{v,i} \Delta_{Ti} \widetilde{\epsilon}_i \\
&\equiv \theta_{NT,2,1} + \theta_{NT,2,2}.
\end{aligned} \tag{B.55}$$

By Assumption 7(ii) and (iii) and results in Proposition C.3(i) below, we have

$$\theta_{NT,2,2} = O_P \left(\sqrt{\frac{1}{\gamma_{NT} NT}} \right). \tag{B.56}$$

Along with Eqs. (B.56), under Assumption 7(iii) in H_1, and by $Q_{NT} = o_P(1)$ in Proposition C.3(i), we have

$$(\delta_{NT})^{-1} \theta_{NT,2} \to \Omega_v^{-1}\psi. \tag{B.57}$$

Next, the statistics can be rewritten as

$$\begin{aligned}
&\widehat{J}_{NT} = NT \left(\theta_{NT,1} + \theta_{NT,2}\right)^\top \widehat{H}_{NT}^{-1} \left(\theta_{NT,1} + \theta_{NT,2}\right) \\
&= NT\theta_{NT,1}^\top \widehat{H}_{NT}^{-1}\theta_{NT,1} + NT\theta_{NT,2}^\top \widehat{H}_{NT}^{-1}\theta_{NT,1} + NT\theta_{NT,1}^\top \widehat{H}_{NT}^{-1}\theta_{NT,2} + NT\theta_{NT,2}^\top \widehat{H}_{NT}^{-1}\theta_{NT,2} \\
&\equiv A_{NT,12} + \cdots + A_{NT,15},
\end{aligned} \tag{B.58}$$

where $\theta_{NT,1}$ is the difference between the two estimators under H_0. Then by the proof of Theorem 4.1,

$$A_{NT,12} \xrightarrow{D} W_d^\top(1)\Phi_d W_d^\top(1). \tag{B.59}$$

By Lemma 4.1, Eqs. (B.39) and Eqs. (B.57), we have

$$\begin{aligned}
&\left(\delta_{NT}\sqrt{\gamma_{NT}NT}\right)^{-1} A_{NT,13} = \left(\delta_{NT}^{-1}\theta_{NT,2}^\top\right)\left(\gamma_{NT}\widehat{H}_{NT}\right)^{-1}\left(\sqrt{\gamma_{NT}NT}\theta_{NT,1}\right) \\
&\xrightarrow{D} \psi^\top\Omega_v^{-1}\Phi_d W_d^\top(1).
\end{aligned} \tag{B.60}$$

Similarly, we have

$$\left(\delta_{NT}\sqrt{\gamma_{NT}NT}\right)^{-1} A_{NT,14} \xrightarrow{D} W_d^\top(1)\Phi_d\Omega_v^{-1}\psi, \tag{B.61}$$

$$\left(\delta_{NT}^2\gamma_{NT}NT\right)^{-1} A_{NT,15} \xrightarrow{D} \psi^\top\Omega_v^{-1}\Phi_d\Omega_v^{-1}\psi. \tag{B.62}$$

Since $\delta_{NT}\sqrt{\gamma_{NT}NT} \to \infty$ in Assumption 7(i), we have

$$\left(\delta_{NT}^2\gamma_{NT}NT\right)^{-1} \widehat{J}_{NT} = \left(\delta_{NT}^2\gamma_{NT}NT\right)^{-1} A_{NT,15} + O_P\left(\left(\delta_{NT}\sqrt{\gamma_{NT}NT}\right)^{-\frac{1}{2}}\right), \tag{B.63}$$

which completes the proof of Theorem 4.2.

Appendix C Propositions

In this part, we give some necessary propositions that have been used in the proofs of the main theorems. The proofs of these propositions are provided in Appendix D in the supplementary material.

Proposition C.1. Under Assumptions 1 to 4, we have
(i) $\sum_{i=1}^N P\left(||Q_{Ti}|| > \mathcal{M}\right) = o(1)$ for constant $\mathcal{M}$ large enough;
(ii) $\mathrm{E}\left|\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \widetilde{\Omega}_{v,i}\xi_{it}^c\right|\right| \le \frac{C\sqrt{\gamma_N}}{N^{1/4}\sqrt{T^2h}}$, and $\mathrm{E}\left|\left|\sum_{i=1}^N\sum_{t=1}^T w_i\xi_{it}^c\right|\right|^2 \le \frac{CN^{1/4}}{\sqrt{\gamma_{N,w}h}}$;
(iii) $\mathrm{E}\left|\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \widetilde{\Omega}_{v,i}\left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it}\right|\right| \le \frac{C\sqrt{\gamma_N}}{T^{1+\frac{\delta^*}{4+\delta^*}}}$, and
$\mathrm{E}\left|\left|\sum_{i=1}^N\sum_{t=1}^T w_i\left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it}\right|\right| \le \frac{CN^{1/2}}{\gamma_{N,w}^{1/2}T^{\frac{\delta^*}{4+\delta^*}}}$.

Proposition C.2. Under Assumptions 1 to 4, we have
(i) $\frac{1}{N}\sum_{i=1}^N\left(s_i(\tau)X_i - X_{i\cdot}^\top\right)\left(\beta_i - \widehat{\beta}_i\right) = o_P\left(\frac{1}{\sqrt{NTh}}\right)$;
(ii) $\sqrt{\frac{Th}{N}}\sum_{i=1}^N s_i(\tau)e_i \xrightarrow{D} N(0, \sigma_e^2)$.

Proposition C.3. Under Assumptions 1 to 6, we have
(i) $\Delta_{NT} = o\left(\frac{1}{\sqrt{\gamma_{NT}}}\right)$, $Q_{NT} = o_P(1)$, and $\frac{1}{N}\sum_{i=1}^N \Omega_{v,i}\Delta_{Ti}\bar{\epsilon}_i = O_P\left(\sqrt{\frac{1}{NT}}\right)$;
(ii) $\mathrm{E}\left|\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \Omega_{v,i}\Delta_{Ti,2}\xi_{it}\right|\right| \leq \frac{C}{\sqrt{NT^3h^2}}$;
(iii) $\mathrm{E}\left|\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \Omega_{v,i}\Delta_{Ti}\xi_{it}^c\right|\right| \leq \frac{C_1}{\sqrt{NT^3h^2}} + \frac{C_2}{\sqrt{T^4h^2}}$;
(iv) $\mathrm{E}\left|\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \Omega_{v,i}\Delta_{Ti}\left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it}\right|\right| \leq \frac{C}{\sqrt{T^4h^2}}$;
(v) $\mathrm{E}\left|\left|\zeta_{N,t} - \mathrm{E}\left(\zeta_{N,t}|\mathscr{F}_{t-m}^{t+m}\right)\right|\right|^2 \leq v_m d_t$ for some $v_m = O\left(m^{-\phi}\right)$ with $\phi > 1$ and d_t irrelevant with m.

Proposition C.4. Under Assumptions 1 to 6, $\sup_{0\leq r\leq 1}\left|\left|\sqrt{\gamma_{NT}}\widehat{L}_{NT}(r) - L_{NT}(r) - rL_{NT}(1)\right|\right| = o_P(1)$.

Appendix D Proof of Propositions

In Appendix D, we provide the proofs of the propositions in Appendix C. Several useful lemmas and their proofs are given in Appendix E. In Appendix F, we include some additional results for the OECD empirical application.

Proof of Proposition C.1(i). Recall that $Q_{Ti} = \widehat{\Omega}_{x,Ti}^{-1}\Omega_{v,i} - I_d$, then $||Q_{Ti}|| \leq \left|\left|\widehat{\Omega}_{x,Ti}^{-1}\right|\right| \, ||\Omega_{v,i}|| + ||I_d||$, and denote $\mathcal{M}^* = \frac{\mathcal{M}-\sqrt{d}}{\max_i||\Omega_{v,i}||}$. Then we have the following relationship between these two sets:

$$\{||Q_{Ti}|| > \mathcal{M}\} \subset \left\{\left|\left|\widehat{\Omega}_{x,Ti}^{-1}\right|\right| > \mathcal{M}^*\right\}. \tag{D.1}$$

Note $\widehat{\Omega}_{x,Ti}$ is a symmetric positive definite matrix, and consider its eigenvalue decomposition $\widehat{\Omega}_{x,Ti} = D_i\Lambda_i D_i^\top$, where Λ_i is a diagonal matrix of its eigenvalues, and D_i as the unitary matrix of eigenvectors, such that $D_iD_i^\top = I_d$. Thus for any $i \geq 1$

$$\begin{aligned} &||\widehat{\Omega}_{x,Ti}^{-1}|| = ||D_i\Lambda_i^{-1}D_i^\top|| \leq ||D_i|| \times ||\Lambda_i^{-1}|| \times ||D_i^\top|| \\ =&\mathrm{tr}\left(D_iD_i^\top\right)\sqrt{\sum_{l=1}^d \lambda_l^{-2}\left(\widehat{\Omega}_{x,Ti}\right)}. \leq d^{3/2}\lambda_{\min}^{-1}\left(\widehat{\Omega}_{x,Ti}\right) \end{aligned} \tag{D.2}$$

Then by Eqs. (D.2), we have

$$\left\{\left|\left|\widehat{\Omega}_{x,Ti}^{-1}\right|\right| > \mathcal{M}^*\right\} \subset \left\{d^{3/2}\lambda_{\min}^{-1}\left(\widehat{\Omega}_{x,Ti}\right) > \mathcal{M}^*\right\}. \tag{D.3}$$

Recall that $\min_{1\leq i\leq N}\lambda_{\min}(\Omega_{v,i}) > c_v$ for some constant $c_v > 0$. Then if $\mathcal{M}^*$ is chosen large enough such that $c_v - \frac{d^{3/2}}{\mathcal{M}^*} > \varepsilon$ for some constant $\varepsilon > 0$, then Eqs. (D.3) further indicates

$$\left\{d^{3/2}\lambda_{\min}^{-1}\left(\widehat{\Omega}_{x,Ti}\right) > \mathcal{M}^*\right\} \subset \left\{c_v - \lambda_{\min}\left(\widehat{\Omega}_{x,Ti}\right) > \varepsilon\right\}. \tag{D.4}$$

Recall in the proof of Theorem 3.3, we have defined $\widehat{\Omega}_{x,Ti} = \Omega_{v,i} + \Omega_{v,i}\Delta_{Ti}$. And by a standard matrix theory, for square matrices A and B, we have $\lambda_{\min}(A+B) \geq \lambda_{\min}(A) - ||B||$, thus we have

$$\begin{aligned} &c_v - \lambda_{\min}\left(\widehat{\Omega}_{x,Ti}\right) \leq c_v - \left(\lambda_{\min}\left(\Omega_{v,i}\right) - ||\Omega_{v,i}\Delta_{Ti}||\right) \\ &= ||\Omega_{v,i}\Delta_{Ti}|| - \left(\lambda_{\min}\left(\Omega_{v,i}\right) - c_v\right) \leq ||\Omega_{v,i}\Delta_{Ti}||, \end{aligned} \tag{D.5}$$

where in the last step we have used $\lambda_{\min}(\Omega_{v,i}) > c_v$ for all i in Assumption 1. Then, combining Eqs. (D.4) and Eqs. (D.5), we have

$$\left\{c_v - \lambda_{\min}\left(\widehat{\Omega}_{x,Ti}\right) > \varepsilon\right\} \subset \{||\Omega_{v,i}\Delta_{Ti}|| > \varepsilon\}. \tag{D.6}$$

Denote $\Delta_{Ti,1} = \frac{1}{N}\sum_{t=1}^{T}\eta_{it}$ and $\Delta_{Ti,2} = \Delta_{Ti} - \Delta_{Ti,1}$, note that

$$\{||\Omega_{v,i}\Delta_{Ti}|| > \varepsilon\} \subset \{||\Omega_{v,i}\Delta_{Ti,1}|| > \frac{\varepsilon}{2}\} \cup \{||\Omega_{v,i}\Delta_{Ti,2}|| > \frac{\varepsilon}{2}\}. \tag{D.7}$$

Combining Eqs. (D.1) to Eqs. (D.7), we have

$$\sum_{i=1}^{N} P\left(||Q_{Ti}|| > \mathcal{M}\right) \leq \sum_{i=1}^{N} P\left(||\Omega_{v,i}\Delta_{Ti,1}|| > \frac{\varepsilon}{2}\right) + \sum_{i=1}^{N} P\left(||\Omega_{v,i}\Delta_{Ti,2}|| > \frac{\varepsilon}{2}\right). \tag{D.8}$$

By Chebyshev's inequality and results in Lemma E.2(ii) and (iii) below, we have

$$\sum_{i=1}^{N} P\left(||\Omega_{v,i}\Delta_{Ti,1}|| > \frac{\varepsilon}{2}\right) \leq N\frac{\mathrm{E}\,||\Omega_{v,i}\Delta_{Ti,1}||^{2+\delta^*/2}}{(\varepsilon/2)^{2+\delta^*/2}} = O\left(\frac{N}{T^{1+\delta^*/4}}\right), \tag{D.9}$$

$$\sum_{i=1}^{N} P\left(||\Omega_{v,i}\Delta_{Ti,2}|| > \frac{\varepsilon}{2}\right) \leq N\frac{\mathrm{E}\,||\Omega_{v,i}\Delta_{Ti,2}||^{2}}{(\varepsilon/2)^{2}} = O\left(\frac{N}{T^{2}h}\right), \tag{D.10}$$

which complete the proof of Proposition C.1(i).

Proof of Proposition C.1(ii).

Denote that $\widetilde{f}_{it} = f_i(\tau_t) - s(\tau_t)f_i$, and $\widetilde{g}_{it} = g_i(\tau_t) - s(\tau_t)g_i$. To prove Proposition C.1, we decompose ξ^c_{it} into $\xi^c_{it,1}$ and $\xi^c_{it,2}$ in the following way:

$$\xi^c_{it,1} = -\Omega^{-1}_{v,i}v_{it}s_i(\tau_t)e_i - \Omega^{-1}_{v,i}v_i^\top s_i^\top(\tau_t)e_{it} + \Omega^{-1}_{v,i}v_i^\top s_i^\top(\tau_t)s_i(\tau_t)e_i, \tag{D.11}$$

$$\xi^c_{it,2} = \Omega^{-1}_{v,i}\widetilde{g}_{it}e_{it} - \Omega^{-1}_{v,i}\widetilde{g}_{it}s_i(\tau_t)e_i + \Omega^{-1}_{v,i}v_{it}\widetilde{f}_{it} + \Omega^{-1}_{v,i}v_i^\top s_i^\top(\tau_t)\widetilde{f}_{it} + \Omega^{-1}_{v,i}\widetilde{g}_{it}\widetilde{f}_{it}, \tag{D.12}$$

where the different orders of $\xi^c_{it,1}$ and $\xi^c_{it,2}$ are specified in Lemma E.2(iii) and (iv) below. Then we have

$$\begin{aligned}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{\Omega}_{v,i}\xi^c_{it} &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{\Omega}_{v,i}\xi^c_{it,1} + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{\Omega}_{v,i}\xi^c_{it,2}\\ &\equiv A^*_{NT,1} + A^*_{NT,2}.\end{aligned} \tag{D.13}$$

Note that

$$\begin{aligned}\mathrm{E}\left|\left|A^*_{NT,1}\right|\right|^2 &= \frac{1}{N^2T^2}\mathrm{tr}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\widetilde{\Omega}_{v,i}\mathrm{E}\left(\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\xi^c_{it_1,1}{\xi^c_{jt_2,1}}^\top\right)\widetilde{\Omega}_{v,j}\right)\\ &\leq \frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\left|\widetilde{\Omega}_{v,i}\right|\right|\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(\xi^c_{it_1,1}{\xi^c_{jt_2,1}}^\top\right)\right|\right|\left|\left|\widetilde{\Omega}_{v,j}\right|\right|.\end{aligned} \tag{D.14}$$

Meanwhile, consider a $N \times N$ matrix $\mathbf{\Xi}_{NT,\xi^c}$ where its (i,j)-element is

$$\mathbf{\Xi}_{NT,\xi^c}(i,j) = \sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(\xi^c_{it_1,1}\left(\xi^c_{jt_2,1}\right)^\top\right)\right|\right|. \tag{D.15}$$

In Lemma E.4(iii) and Lemma E.2(vi), we have shown that

$$\sum_{i=1}^{N}\sum_{j=1}^{N}||\mathbf{\Xi}_{NT,\xi^c}(i,j)|| < \frac{CN}{h}, \quad \max_{1\leq i,j\leq N}||\mathbf{\Xi}_{NT,\xi^c}(i,j)|| < \frac{C}{h}. \tag{D.16}$$

Then by a standard matrix theory, we have

$$\lambda_{\max}\left(\mathbf{\Xi}_{NT,\xi^c}\right) \leq ||\mathbf{\Xi}_{NT,\xi^c}|| = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}||\mathbf{\Xi}_{NT,\xi^c}(i,j)||^2} \leq C\frac{\sqrt{N}}{h}. \tag{D.17}$$

Recall that $\widetilde{\mathbf{\Omega}}_v = \left(\left|\left|\widetilde{\Omega}_{v,1}\right|\right|,\cdots,\left|\left|\widetilde{\Omega}_{v,N}\right|\right|\right)^\top$, and $\gamma_N = \frac{1}{N}\sum_{i=1}^{N}||\widetilde{\Omega}_{v,i}||^2 = \frac{1}{N}\widetilde{\mathbf{\Omega}}_v^\top\widetilde{\mathbf{\Omega}}_v$, then by a standard matrix theory, such as Theorem 4 in Chapter 11 of Magnus et al. (1995), we have

$$\begin{aligned}
\mathrm{E}\left|\left|A^*_{NT,1}\right|\right|^2 &\leq \frac{1}{N^2T^2}\widetilde{\mathbf{\Omega}}_v^\top\mathbf{\Xi}_{NT,\xi^c}\widetilde{\mathbf{\Omega}}_v \leq \frac{1}{N^2T^2}\lambda_{\max}\left(\mathbf{\Xi}_{NT,\xi^c}\right)\widetilde{\mathbf{\Omega}}_v^\top\widetilde{\mathbf{\Omega}}_v \\
&\leq \frac{\gamma_N}{NT^2}\lambda_{\max}\left(\mathbf{\Xi}_{NT,\xi^c}\right) \leq C\frac{\gamma_N\sqrt{N}}{NT^2h}.
\end{aligned} \tag{D.18}$$

For $A^*_{NT,2}$, let R_N be a $N\times N$ matrix of ones, note that $\lambda_{\max}(R_N) \leq ||R_N|| = \sqrt{N^2}$. Thus

$$\left(\frac{1}{N}\sum_{i=1}^{N}\left|\left|\widetilde{\Omega}_{v,i}\right|\right|\right)^2 = \frac{1}{N^2}\widetilde{\mathbf{\Omega}}_{v,i}^\top R_N\widetilde{\mathbf{\Omega}}_{v,i} \leq \frac{1}{N^2}\sum_{i=1}^{N}\left|\left|\widetilde{\Omega}_{v,i}\right|\right|^2\lambda_{\max}\left(R_N\right) = \frac{1}{N}\sum_{i=1}^{N}\left|\left|\widetilde{\Omega}_{v,i}\right|\right|^2 = \gamma_N. \tag{D.19}$$

In Lemma E.2(iv) we have shown that $\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^{T}\xi^c_{it,2}\right|\right|^2 \leq \frac{Ch^4}{T}$. Then by standard calculation,

$$\begin{aligned}
\mathrm{E}\left|\left|A^*_{NT,2}\right|\right| &\leq \frac{1}{N}\sum_{i=1}^{N}\left|\left|\widetilde{\Omega}_{v,i}\right|\right|\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^{T}\xi^c_{it,2}\right|\right| \\
&\leq \frac{Ch^2}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\left|\left|\widetilde{\Omega}_{v,i}\right|\right| = \frac{Ch^2\sqrt{\gamma_N}}{\sqrt{T}},
\end{aligned}$$

which suffices to prove the first part of Proposition C.1(ii). Similarly, the second part of Proposition C.1(ii) can be verified by replacing $\widetilde{\Omega}_{v,i}$ with w_i.

Proof of Proposition C.1(iii).

Recall that $\widetilde{\eta}_{it} = \eta_{it} + \eta^c_{it}$, $\widetilde{\xi}_{it} = \xi_{it} + \xi^c_{it}$, and $\Delta_{Ti} = \frac{1}{N}\sum_{t=1}^{T}\widetilde{\eta}_{it}$ we have

$$\begin{aligned}
&\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{\Omega}_{v,i}\left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it} = \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\widetilde{\Omega}_{v,i}\left(I_d + \widetilde{Q}_{Ti}\right)\widetilde{\eta}_{it_1}\widetilde{\xi}_{it_2} \\
&= \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\widetilde{\Omega}_{v,i}\eta_{it_1}\xi_{it_2} + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\widetilde{\Omega}_{v,i}\widetilde{Q}_{Ti}\eta_{it_1}\xi_{it_2} \\
&\quad + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\widetilde{\Omega}_{v,i}\left(I_d + \widetilde{Q}_{Ti}\right)\eta_{it_1}\xi^c_{it_2} + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\widetilde{\Omega}_{v,i}\left(I_d + \widetilde{Q}_{Ti}\right)\eta^c_{it_1}\widetilde{\xi}_{it_2} \\
&\equiv A^*_{NT,3} + \cdots + A^*_{NT,6}.
\end{aligned} \tag{D.20}$$

For $A_{NT,3}$, similar to the calculation with $A^*_{NT,1}$, consider the $N\times N$ matrix $\mathbf{\Xi}_{NT,\eta,\xi}$ with its (i,j) element defined as

$$\mathbf{\Xi}_{NT,\eta,\xi}(i,j) = \frac{1}{T^2}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T}\left|\left|\mathrm{E}\left(\eta_{it_1}\xi_{it_2}\xi^\top_{it_4}\eta^\top_{it_3}\right)\right|\right|. \tag{D.21}$$

By Assumption 2, Lemma E.4(iii) and follow the same argument in the proof of Proposition C.1(ii), we have

$$\lambda_{\max}\left(\mathbf{\Xi}_{NT,\eta,\xi}\right) \leq C\sqrt{N} \tag{D.22}$$

Thus we have

$$\begin{aligned}
\mathrm{E}\left|\left|A^*_{NT,3}\right|\right|^2 &\leq \frac{1}{N^2T^2}\mathrm{tr}\left(\widetilde{\mathbf{\Omega}}_v^\top \mathbf{\Xi}_{NT,\eta,\xi}\widetilde{\mathbf{\Omega}}_v\right) \\
&\leq \frac{1}{NT^2}\frac{1}{N}\sum_{i=1}^N \left|\left|\widetilde{\Omega}_{v,i}\right|\right|^2 \lambda_{\min}\left(\mathbf{\Xi}_{NT,\eta,\xi}\right) \leq \frac{C\gamma_N\sqrt{N}}{NT^2}. \qquad \text{(D.23)}
\end{aligned}$$

For $A^*_{NT,4}$, by Eqs. (D.19), and results in Lemma E.2(ii), (iv), (v), together with Cauchy-Swarchz inequality, we have

$$\begin{aligned}
\mathrm{E}\left|\left|A^*_{NT,4}\right|\right| &\leq \frac{C_1}{T^2}\left(\frac{1}{N}\sum_{i=1}^N\left|\left|\widetilde{\Omega}_{v,i}\right|\right|\right)\left(\mathrm{E}\left|\left|\sum_{t_2=1}^T \xi_{it}\right|\right|^{2+\delta^*/2}\right)^{\frac{1}{2+\delta^*/2}} \\
&\times\left(\mathrm{E}\left(\left|\left|\sum_{t_1=1}^T \eta_{it_1}\right|\right|^2\right)^{\frac{2+\delta^*/2}{2}}\right)^{\frac{1}{2+\delta^*/2}}\left(C_2\mathrm{E}\left|\left|\widetilde{Q}_{Ti}\right|\right|^2\right)^{\frac{\delta^*}{4+\delta^*}} \leq \sqrt{\frac{C_1\gamma_N}{T^{2+\frac{\delta^*}{2+\delta^*/2}}}}. \qquad \text{(D.24)}
\end{aligned}$$

Since $\left|\left|I_d+\widetilde{Q}_{Ti}\right|\right|$ is bounded, by results in Eqs. (D.19), Lemma E.2(ii), (v) and Cauchy-Swarchz inequality, we have

$$\begin{aligned}
\mathrm{E}\left|\left|A^*_{NT,5}\right|\right| &\leq \frac{1}{T}\left(\frac{1}{N}\sum_{i=1}^N\left|\left|\widetilde{\Omega}_{v,i}\right|\right|\right)\left(\mathrm{E}\left|\left|\sum_{t_1=1}^T \eta_{it_1}\right|\right|^2\right)^{\frac{1}{2}}\left(\mathrm{E}\left|\left|\frac{1}{T}\sum_{t_2=1}^T \xi^c_{it_2}\right|\right|^2\right)^{\frac{1}{2}} \\
&\leq \frac{C_1\sqrt{\gamma_N}}{T\sqrt{Th}}. \qquad \text{(D.25)}
\end{aligned}$$

And similarly,

$$\begin{aligned}
\mathrm{E}\left|\left|A^*_{NT,6}\right|\right| &\leq \frac{1}{T}\left(\frac{1}{N}\sum_{i=1}^N\left|\left|\widetilde{\Omega}_{v,i}\right|\right|\right)\left(\mathrm{E}\left|\left|\sum_{t_1=1}^T \eta^c_{it_1}\right|\right|^2\right)^{\frac{1}{2}}\left(\mathrm{E}\left|\left|\frac{1}{T}\sum_{t_2=1}^T \widetilde{\xi}_{it_2}\right|\right|^2\right)^{\frac{1}{2}} \\
&\leq \frac{C_1\sqrt{\gamma_N}}{T\sqrt{Th}}. \qquad \text{(D.26)}
\end{aligned}$$

Together with the rates in Assumption 4, we have prove Proposition C.1(iii).

Proof of Proposition C.2(i):

Use the same calculation in the proof of Theorem 3.3, and $\widehat{\Omega}^{-1}_{x,Ti} = \Omega^{-1}_{v,i} + \widetilde{Q}_{Ti}\Omega^{-1}_{v,i} + \widetilde{Q}^c_{Ti}\Omega^{-1}_{v,i}$, we have

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^N\left(s_i(\tau)X_i - X_{i\cdot}^\top\right)(\beta_i - \widehat{\beta}_i) = \frac{1}{NT}\sum_{i=1}^N\left(s_i(\tau)X_i - X_{i\cdot}^\top\right)\widehat{\Omega}^{-1}_{x,Ti}\widetilde{X}_i^\top\left(\widetilde{e}_i + \widetilde{f}_i\right) \\
&= \frac{1}{NT}\sum_{i=1}^N\left(s_i(\tau)X_i - X_{i\cdot}^\top\right)\left(I_d + \widetilde{Q}_{Ti} + \widetilde{Q}^c_{Ti}\right)\sum_{t=1}^T\widetilde{\xi}_{it} \\
&= \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\left(s_i(\tau)X_i - X_{i\cdot}^\top\right)\widetilde{\xi}_{it} + \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\left(s_i(\tau)X_i - X_{i\cdot}^\top\right)\widetilde{Q}_{Ti}\widetilde{\xi}_{it} \\
&+\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\left(s_i(\tau)X_i - X_{i\cdot}^\top\right)\widetilde{Q}^c_{Ti}\widetilde{\xi}_{it} \equiv A^*_{NT,7} + \cdots + A^*_{NT,9}. \qquad \text{(D.27)}
\end{aligned}$$

Since $s_i(\tau)\iota_T = 1$, by standard calculation we have $s_i(\tau)X_i - X_{i\cdot}^\top = s_i(\tau)v_i + s_i(\tau)g_i - v_{i\cdot}^\top - g_{i\cdot}^\top$. Thus

$$A^*_{NT,7} = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\left(s_i(\tau)v_i - v_{i\cdot}^\top\right)\xi_{it} + \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\left(s_i(\tau)g_i - g_{i\cdot}^\top\right)\xi_{it}$$

$$
\begin{aligned}
&+\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(s_i(\tau)v_i - v_{i\cdot}^\top\right)\xi_{it}^c + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(s_i(\tau)g_i - g_{i\cdot}^\top\right)\xi_{it}^c \\
&\equiv A_{NT,7,1}^* + \cdots + A_{NT,7,4}.
\end{aligned} \tag{D.28}
$$

In Lemma E.4(i), we have shown

$$
A_{NT,7,1}^* = o_P\left(\frac{1}{\sqrt{NTh}}\right) \tag{D.29}
$$

For $A_{NT,7,2}^*$, note that $s_i(\tau)g_i$ is the local linear estimator of $g_i(\tau)$, then $s_i(\tau)g_i \to g_i(\tau)$. Thus $g_i(\tau)$ is bounded. In addition, by Assumption 4, $g_{i\cdot}$ will also be bounded. Therefore we have

$$
\begin{aligned}
\mathrm{E}\left|\left|A_{NT,7,2}\right|\right|^2 &= \frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left(s_i(\tau)g_i - g_{i\cdot}^\top\right)\mathrm{E}\left(\xi_{it}\xi_{it}^\top\right)\left(s_i(\tau)g_i - g_{i\cdot}^\top\right)^\top \\
&\le \frac{C}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(\xi_{it}\xi_{it}^\top\right)\right|\right| = O\left(\frac{1}{NT}.\right)
\end{aligned} \tag{D.30}
$$

Thus

$$
A_{NT,7,2}^* = O_P\left(\frac{1}{\sqrt{NT}}\right) = o_P\left(\frac{1}{\sqrt{NTh}}\right) \tag{D.31}
$$

By Lemma E.2(i), $\mathrm{E}\left|\left|s_i(\tau)v_i\right|\right|^4 \le \frac{C}{T^2h^2}$, and by mixing condition in Assumption 1, $\mathrm{E}\left|\left|v_{i\cdot}\right|\right|^4 \le \frac{C}{T^2}$. In addition, by Lemma E.2, $\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^{T}\xi_{it}^c\right|\right|^2 \le \frac{C}{\sqrt{T^2h}}$, together with Cauchy-Schwarz inequality,

$$
\mathrm{E}\left|\left|A_{NT,7,3}^*\right|\right| \le \frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{E}\left|\left|s_i(\tau)v_i - v_{i\cdot}^\top\right|\right|^2\right)^{1/2}\left(\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^{T}\xi_{it}^c\right|\right|^2\right)^{1/2} = O\left(\frac{1}{\sqrt{NTh}}\sqrt{\frac{N}{T^2h}}\right). \tag{D.32}
$$

Thus

$$
A_{NT,7,3}^* = o_P\left(\frac{1}{\sqrt{NTh}}\right). \tag{D.33}
$$

In Lemma E.4(ii), we have

$$
A_{NT,7,4}^* = o_P\left(\frac{1}{\sqrt{NTh}}\right). \tag{D.34}
$$

From Eqs. (D.29) to Eqs. (D.34), we have

$$
A_{NT,7}^* = o_P\left(\frac{1}{\sqrt{NTh}}\right). \tag{D.35}
$$

Next, note that

$$
\begin{aligned}
A_{NT,8}^* &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(s_i(\tau)v_i - v_{i\cdot}^\top\right)\widetilde{Q}_{Ti}\widetilde{\xi}_{it} + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(s_i(\tau)g_i - g_{i\cdot}^\top\right)\widetilde{Q}_{Ti}\widetilde{\xi}_{it} \\
&\equiv A_{NT,8,1}^* + A_{NT,8,2}^*.
\end{aligned} \tag{D.36}
$$

Similar to the proof of $A_{NT,7,3}^*$, and $\mathrm{E}\left|\left|\sum_{t=1}^{T}\widetilde{\xi}_{it}\right|\right|^2 \le C\sqrt{T}$, and $\mathrm{E}\left|\left|\widetilde{Q}_{Ti}\right|\right|^4 \le \frac{C}{T}$, then together with Cauchy-Schwarz inequality,

$$
\mathrm{E}\left|\left|A_{NT,8,1}^*\right|\right| \le \frac{1}{NT}\sum_{i=1}^{N}\left(\mathrm{E}\left|\left|s_i(\tau)v_i - v_{i\cdot}^\top\right|\right|^4\right)^{1/4}\left(\mathrm{E}\left|\left|\widetilde{Q}_{Ti}\right|\right|^4\right)^{1/4}\left(\mathrm{E}\left|\left|\sum_{t=1}^{T}\widetilde{\xi}_{it}\right|\right|^2\right)^{1/2}
$$

$$= O\left(\frac{1}{\sqrt{Th}}\frac{1}{T^{1/4}}\frac{1}{\sqrt{T}}\right) = O\left(\frac{1}{\sqrt{NTh}}\sqrt{\frac{N}{T^{3/2}}}\right). \tag{D.37}$$

Hence

$$A^*_{NT,8,1} = o_P\left(\frac{1}{\sqrt{NTh}}\right). \tag{D.38}$$

Similarly, since $\left|\left|s_i(\tau)g_i - g_{i\cdot}^\top\right|\right|$ is bounded, we have

$$\begin{aligned}
\mathrm{E}\left|\left|A^*_{NT,8,2}\right|\right| &\leq \frac{1}{NT}\sum_{i=1}^{N}\left(\mathrm{E}\left|\left|\widetilde{Q}_{Ti}\right|\right|^2\right)^{1/4}\left(\mathrm{E}\left|\left|\sum_{t=1}^{T}\widetilde{\xi}_{it}\right|\right|^2\right)^{1/2} \\
&= O\left(\frac{1}{T}\right) = O\left(\frac{1}{\sqrt{NTh}}\sqrt{\frac{Nh}{T}}\right).
\end{aligned} \tag{D.39}$$

By Assumption 4, $Nh/T \to 0$, then

$$A^*_{NT,8,2} = o_P\left(\frac{1}{\sqrt{NTh}}\right). \tag{D.40}$$

By Eqs. (D.38) and Eqs. (D.40), we have

$$A^*_{NT,8} = o_P\left(\frac{1}{\sqrt{NTh}}\right). \tag{D.41}$$

Recall the definition of $\widetilde{Q}_i^c$ and by the same argument in Eqs.(B.13), we have

$$A^*_{NT,9} = o_P\left(\frac{1}{\sqrt{NTh}}\right). \tag{D.42}$$

which completes the proof of Proposition C.2(i).

Proof of Proposition C.2(ii).

Denote $K_{it,j}(\tau) = \left(\frac{t/T-\tau}{h_i}\right)^j K\left(\frac{t/T-\tau}{h_i}\right)$ and $w_{T,it} = (1 \quad 0)\left(Z_i^\top(\tau)W(\tau)Z_i(\tau)\right)^{-1}(K_{it,0}(\tau), K_{it,1}(\tau))^\top$, then $s_i(\tau)e_i = \sum_{t=1}^T w_{T,it}e_{it}$. Denote $V_{N,t} = \frac{\sqrt{Th}}{\sqrt{N}}\sum_{i=1}^N w_{T,it}(\tau)e_{it}$, and we need to show

$$\sum_{t=1}^{T} V_{N,t} \xrightarrow{D} N(0, \sigma^2_{eK}). \tag{D.43}$$

Here we employ the large-block and small-block arguments to prove asymptotic normality in Eqs. (D.43). First we partition the set $\{1, \cdots, T\}$ into $2\kappa_T + 1$ subsets with large blocks of size l_T and small block of size s_T and the remaining set with size $T - \kappa_T(l_T + s_T)$ where $l_T = O(T^{(r-1)/r})$, $s_T = O(T^{1/r})$, and $\kappa_T = O(T^{1/r})$ for any $r > 2$. Then for $p = 1, \cdots, \kappa_T$, define

$$\widetilde{V}_{N,p} = \sum_{t=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T} V_{N,t}, \quad \overline{V}_{N,p} = \sum_{t=pl_T+(p-1)s_T+1}^{p(l_T+s_T)} V_{N,t}, \quad \widehat{V}_N = \sum_{t=k_T(l_T+s_T)+1}^{T} V_{N,t}. \tag{D.44}$$

For the small blocks, by Assumption 2 and the results in Lemma E.1(ii) and Lemma E.3(i), we have

$$\mathrm{E}\left(\sum_{p=1}^{\kappa_T}\overline{V}_{N,p}\right)^2 = \frac{Th}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{p_1=1}^{\kappa_T}\sum_{t_1=p_1l_T+(p_1-1)s_T+1}^{p_1(l_T+s_T)}\sum_{p_2=1}^{\kappa_T}\sum_{t_2=p_2l_T+(p_2-1)s_T+1}^{p_2(l_T+s_T)} w_{T,it_1}(\tau)$$

$$
\begin{aligned}
&\times w_{T,jt_2}(\tau)\mathrm{E}(e_{it_1}e_{jt_2})\\
&\leq \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\sum_{p_1=1}^{\kappa_T}\sum_{t_1=p_1 l_T+(p_1-1)s_T+1}^{p_1(l_T+s_T)}|w_{T,it_1}(\tau)|\right)\sum_{p_2=1}^{\kappa_T}\sum_{k=(p_2-1)(l_T+s_T)+1}^{(p_2-1)(l_T+s_T)+s_T}|\mathrm{E}(e_{i,1}e_{j,1+k})|\\
&\leq \max_i\left(\sum_{p_1=1}^{\kappa_T}\sum_{t_1=p_1 l_T+(p_1-1)s_T+1}^{p_1(l_T+s_T)}|w_{T,it_1}(\tau)|\right)\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{p_2=1}^{\kappa_T}\sum_{k=(p_2-1)(l_T+s_T)+1}^{(p_2-1)(l_T+s_T)+s_T}|\mathrm{E}(e_{i,1}e_{j,1+k})|\right)\\
&= O\left(\frac{\kappa_T s_T}{T}\right)=o\left(1\right).
\end{aligned}
\tag{D.45}
$$

The last step uses $\kappa_T s_T = T^{2/r}$ where $r>2$.

For the remaining set, note that $T-\kappa_T(l_T+s_T)=O(l_T)$, then

$$
\begin{aligned}
\mathrm{E}\left(\widehat{V}_N\right)^2 &= \frac{Th}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=\kappa_T(l_T+s_T)+1}^{T}\sum_{t_2=\kappa_T(l_T+s_T)+1}^{T} w_{T,it_1}(\tau)w_{T,jt_2}(\tau)\mathrm{E}(e_{it_1}e_{jt_2})\\
&\leq \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\sum_{t_1=\kappa_T(l_T+s_T)+1}^{T}|w_{T,it_1}(\tau)|\right)\sum_{k=1}^{T-\kappa_T(l_T+s_T)}|\mathrm{E}(e_{i,1}e_{j,1+k})|\\
&\leq \max_i\left(\sum_{t_1=\kappa_T(l_T+s_T)+1}^{T}|w_{T,it_1}(\tau)|\right)\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{T-\kappa_T(l_T+s_T)}|\mathrm{E}(e_{i,1}e_{j,1+k})|\\
&= O\left(\frac{l_T}{Th}\right)=o\left(1\right),
\end{aligned}
\tag{D.46}
$$

where in the last step we require that $\frac{l_T}{Th}\to 0$, and this is guaranteed by $Th^{2+\delta}\to\infty$ for some $\delta>0$.

By Eqs. (D.45) and Eqs. (D.46), we have

$$
\sum_{t=1}^{T}V_{N,t}=\sum_{p=1}^{\kappa_T}\widetilde{V}_p+o_p(1) \tag{D.47}
$$

By Lemma E.3(ii) and (iii) we have shown the Lindeberg condition and the Feller condition,

$$
\sum_{p=1}^{\kappa_T}\mathrm{E}\left(\widetilde{V}_p^2 I\left(|\widetilde{V}_p|\geq\varepsilon\sigma_{eK}^2\right)\right)\leq\frac{Cl_T}{Th},\quad \sum_{p=1}^{\kappa_T}\mathrm{E}\left(\widetilde{V}_p^2\right)\to\sigma_{eK}^2 \tag{D.48}
$$

for some constant C. By inequalities for α-mixing, we have

$$
\left|\mathrm{E}\left(\exp\left(i\lambda\sum_{p=1}^{\kappa_T}\widetilde{V}_p\right)\right)-\prod_{p=1}^{\kappa_T}\mathrm{E}\left(\exp\left(i\lambda\widetilde{V}_p\right)\right)\right|\leq C(\kappa_T-1)\alpha(s_T)\to 0. \tag{D.49}
$$

Following the standard procedure to prove Lindeberg theorem (see Theorem 23.6 Davidson, 1994), the Lindeberg condition and the Feller condition suffice to show

$$
\left|\prod_{p=1}^{\kappa_T}\mathrm{E}\left(\exp\left(i\lambda\widetilde{V}_p\right)\right)-\exp\left(-\lambda^2\sigma_{eK}^2/2\right)\right|\to 0 \tag{D.50}
$$

which completes the proof of Proposition C.2(ii).

Proof of Proposition C.3(i).

Note that

$$
\Delta_{NT}=\frac{1}{N}\sum_{i=1}^{N}\Delta_{Ti,1}+\frac{1}{N}\sum_{i=1}^{N}\Delta_{Ti,2}\equiv A_{NT,10}^*+A_{NT,11}^*. \tag{D.51}
$$

By Assumption 6(vi), we have

$$\mathrm{E}\left|\left|A^{*}_{NT,10}\right|\right|^{2}=\frac{1}{N^{2}T^{2}}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\mathrm{E}\left(\eta_{it_1}\eta_{jt_2}^{\top}\right)=O\left(\frac{1}{NT}\right). \tag{D.52}$$

By Lemma E.4(ii), we have

$$A^{*}_{NT,11}=o\left(\frac{1}{\sqrt{NT}}\right), \tag{D.53}$$

which proves the first part of Proposition C.3(i).

Thus we have shown $\Delta_{NT}=o_P(1)$, which along with Assumption 5(i) indicates $\widehat{\Omega}_{x,NT}\xrightarrow{P}\Omega_v$. Thus it is obvious to see that $Q_{NT}=o_P(1)$.

Next, similar to the decomposition of ξ^c_{it}, we can decompose η^c_{it} by

$$\begin{aligned}
\eta^c_{it,1}&=-v_i^{\top}s^{\top}(\tau_t)v_{it}^{\top}-v_{it}s(\tau_t)v_i+v_i^{\top}s^{\top}(\tau_t)s(\tau_t)v_i,\\
\eta^c_{it,2}&=\widetilde{g}_{it}\widetilde{g}_{it}^{\top}-\widetilde{g}_{it}s(\tau_t)v_i-v_i^{\top}s^{\top}(\tau_t)\widetilde{g}_{it}^{\top}+\widetilde{g}_{it}v_{it}^{\top}+v_{it}\widetilde{g}_{it}^{\top}.
\end{aligned} \tag{D.54}$$

Thus we can rewrite that

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\Omega_{v,i}\Delta_{Ti}\widetilde{\epsilon}_i=\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\Omega_{v,i}\eta_{it}\widetilde{\epsilon}_i+\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\Omega_{v,i}\eta^c_{it,1}\widetilde{\epsilon}_i+\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\Omega_{v,i}\eta^c_{it,2}\widetilde{\epsilon}_i\\
&\equiv A^{*}_{NT,12}+\cdots+A^{*}_{NT,14}.
\end{aligned} \tag{D.55}$$

Similar to $B_{NT,3}$, by Assumption 6(vi), we have

$$\mathrm{E}\left|\left|A^{*}_{NT,12}\right|\right|^{2}\leq\frac{C}{N^{2}T^{2}}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(\eta_{it_1}\eta_{jt_2}^{\top}\right)\right|\right|=O\left(\frac{1}{NT}\right). \tag{D.56}$$

For $B_{NT,6}$, consider a $N\times N$ matrix Ξ_{NT,η^c} with its (i,j)-th element defined as

$$\Xi_{NT,\eta^c}(i,j)=\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\mathrm{E}\left(\eta^c_{it_1,1}\left(\eta^c_{jt_2,1}\right)^{\top}\right). \tag{D.57}$$

By Lemma E.4(iii) and Lemma E.2(vi), we have

$$\sum_{i=1}^{N}\sum_{j=1}^{N}||\Xi_{NT,\eta^c}(i,j)||<\frac{C_1N}{h},\quad\max_{1\leq i,j\leq N}||\Xi_{NT,\xi^c}(i,j)||<\frac{C_2}{h}. \tag{D.58}$$

Then

$$\lambda_{\max}\left(\Xi_{NT,\eta^c}\right)\leq C\frac{\sqrt{N}}{h}. \tag{D.59}$$

Thus, we have

$$\begin{aligned}
\mathrm{E}\left|\left|A^{*}_{NT,13}\right|\right|^{2}&=\frac{1}{N^{2}T^{2}}\sum_{i=1}^{N}\sum_{j=1}^{N}\widetilde{\epsilon}_i^{\top}\Omega_{v,i}\Xi_{NT,\eta^c}(i,j)\Omega_{v,j}\widetilde{\epsilon}_i\\
&\leq\frac{C}{N^{2}T^{2}}\sum_{i=1}^{N}\left|\left|\epsilon_i^{\top}\Omega_{v,i}\right|\right|^{2}\lambda_{\min}\left(\Xi_{NT,\eta^c}\right)=O\left(\frac{1}{NT}\frac{\sqrt{N}}{Th}\right).
\end{aligned} \tag{D.60}$$

For $B_{NT,7}$, by Lemma E.2(iv) we have $\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^{T}\eta_{it,2}^{c}\right|\right| \leq \frac{Ch^2}{\sqrt{T}}$. Thus

$$\mathrm{E}\left|\left|A_{NT,14}^{*}\right|\right| \leq \frac{1}{N}\sum_{i=1}^{N}||\Omega_{v,i}\widetilde{\epsilon}_{v}||\,\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^{T}\eta_{it,2}^{c}\right|\right| \leq \frac{C\sqrt{N}h^2}{\sqrt{NT}}, \tag{D.61}$$

which completes the proof of Proposition C.3(i).

Proof of Proposition C.3(ii).

Recall that $\Delta_{Ti,2} = \frac{1}{T}\sum_{t=1}^{T}\eta_{it}^{c}$ and $\eta_{it}^{c} = \eta_{it,1}^{c} + \eta_{it,2}^{c}$, then we have

$$\begin{aligned}
&\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\Omega_{v,i}\Delta_{Ti,2}\xi_{it} = \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\eta_{it_2,1}^{c}\xi_{it_1} + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\eta_{it_2,2}^{c}\xi_{it_1} \\
&\equiv A_{NT,15}^{*} + A_{NT,16}^{*}.
\end{aligned} \tag{D.62}$$

For $A_{NT,15}^{*}$, by definition of $\eta_{it,1}$, we have

$$\begin{aligned}
A_{NT,15}^{*} &= \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}v_i^{\top}s^{\top}(\tau_{t_2})s(\tau_{t_2})v_i\xi_{it_1} - \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}v_i^{\top}s^{\top}(\tau_{t_2})v_{it_2}^{\top}\xi_{it_1} \\
&\quad -\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}v_{it_2}s(\tau_{t_2})v_i\xi_{it_1} \\
&= A_{NT,15,1}^{*} + \cdots + A_{NT,15,3}^{*}.
\end{aligned} \tag{D.63}$$

Recall that $s_i(\tau)v_i = \sum_{t=1}^{T}w_{it}(\tau)v_{it}^{\top}$, then by using Assumption 6 and results in Lemma E.1(ii) and (iii), we have

$$\begin{aligned}
\mathrm{E}\left|\left|A_{NT,15,1}^{*}\right|\right|^2 &= \frac{1}{N^2T^4}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\cdots\sum_{t_8=1}^{T}w_{T,it_3}(\tau_{t_7})w_{T,it_4}(\tau_{t_7}) \\
&\quad\times w_{T,jt_5}(\tau_{t_8})w_{T,jt_6}(\tau_{t_8})\left(\mathrm{E}\left(\xi_{it_1}^{\top}v_{it_3}v_{it_4}^{\top}v_{jt_5}v_{it_6}^{\top}\xi_{it_2}\right)\right) \\
&= \frac{1}{N^2T^4}\max_{i,t_3}\left|\sum_{t_7=1}^{T}w_{T,it_3}(\tau_{t_7})\right|\max_{j,t_8}\left|\sum_{t_8=1}^{T}w_{T,jt_5}(\tau_{t_8})\right|\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\cdots\sum_{t_6=1}^{T} \\
&\quad\times|w_{T,it_4}(\tau_{t_7})|\,|w_{T,jt_6}(\tau_{t_8})|\left|\mathrm{E}\left(\xi_{it_1}^{\top}v_{it_3}v_{it_4}^{\top}v_{jt_5}v_{it_6}^{\top}\xi_{it_2}\right)\right| \\
&= \frac{1}{N^2T^6h^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\cdots\sum_{t_6=1}^{T}\left|\mathrm{E}\left(\xi_{it_1}^{\top}v_{it_3}v_{it_4}^{\top}v_{jt_5}v_{it_6}^{\top}\xi_{it_2}\right)\right| \\
&\leq \frac{C}{NT^3h^2}
\end{aligned} \tag{D.64}$$

Similarly, for $A_{NT,15,2}^{*}$ and $A_{NT,15,3}^{*}$, we have

$$\begin{aligned}
&\mathrm{E}\left|\left|A_{NT,15,2}^{*}\right|\right|^2 = \mathrm{E}\left|\left|A_{NT,15,3}^{*}\right|\right|^2 = \frac{1}{N^2T^4}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\cdots\sum_{t_6=1}^{T} \\
&\times w_{T,it_3}(\tau_{t_4})w_{T,it_6}(\tau_{t_5})\left(\mathrm{E}\left(\xi_{it_1}^{\top}v_{it_3}v_{it_4}^{\top}v_{jt_5}v_{it_6}^{\top}\xi_{it_2}\right)\right) \leq \frac{C}{NT^3h^2}.
\end{aligned} \tag{D.65}$$

For $A_{NT,16}^{*}$, note that $\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^{T}\eta_{it,2}^{c}\right|\right|^2 < \frac{h^4}{T}$, thus

$$\mathrm{E}\left|\left|A_{NT,16}^{*}\right|\right| \leq \frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{E}\left|\left|\frac{1}{T}\sum_{t_2=1}^{T}\eta_{it_2,2}^{c}\right|\right|^2\right)^{1/2}\left(\mathrm{E}\left|\left|\frac{1}{T}\sum_{t_1=1}^{T}\xi_{it_1}\right|\right|^2\right)^{1/2} \leq \frac{C\sqrt{N}h^2}{\sqrt{NT}}, \tag{D.66}$$

which proves Proposition C.3(ii).

Proof of Proposition C.3(iii).

Note that $\xi_{it}^c = \xi_{it,1}^c + \xi_{it,2}^c$, then

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\Omega_{v,i}\Delta_{Ti}\xi_{it}^c = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\eta_{it_1}\xi_{it_2,1}^c + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\eta_{it_1}\xi_{it_2,2}^c$$
$$+\frac{1}{NT}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\eta_{it_1}^c\xi_{it}^c \equiv A_{NT,17}^* + \cdots + A_{NT,19}^*. \tag{D.67}$$

For $A_{NT,17}^*$, recall that $\eta_{it} = \Omega_{v,i}^{-1}v_{it}v_{it}^\top - I_d$, then

$$\Omega_{v,i}\Delta_{Ti,1}\Omega_{v,i}^{-1} = \Omega_{v,i}\frac{1}{T}\sum_{t=1}^{T}\left(\Omega_{v,i}^{-1}v_{it}v_{it}^\top - I_d\right)\Omega_{v,i}^{-1}$$
$$= \frac{1}{T}\sum_{t=1}^{T}\left(v_{it}v_{it}^\top\Omega_{v,i}^{-1} - I_d\right) = \frac{1}{T}\sum_{t=1}^{T}\eta_{it}^\top. \tag{D.68}$$

By the definition of $\xi_{it,1}^c$, we then have

$$A_{NT,17}^* = -\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\eta_{it_2}^\top v_{it_1}s_i(\tau_{t_1})e_i - \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\eta_{it_2}^\top v_i^\top s_i^\top(\tau_{t_1})e_{it_1}$$
$$+\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\eta_{it_2}^\top v_i^\top s_i^\top(\tau_{t_1})s_i(\tau_{t_1})e_i$$
$$\equiv A_{NT,17,1}^* + \cdots + A_{NT,17,3}^*. \tag{D.69}$$

Using Assumption 6, Lemma E.1(ii), we have,

$$\mathrm{E}\left|\left|A_{NT,17,1}^*\right|\right|^2 = \frac{1}{N^2T^4}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\cdots\sum_{t_6=1}^{T}w_{T,it_5}(\tau_{t_1})w_{T,jt_6}(\tau_{t_2})$$
$$\times\mathrm{tr}\left(\mathrm{E}\left(\eta_{it_3}^\top v_{it_1}e_{it_5}e_{jt_6}v_{it_2}^\top\eta_{it_4}\right)\right)$$
$$\leq \frac{C}{N^2T^6h^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\cdots\sum_{t_6=1}^{T}\left|\left|\mathrm{E}\left(\eta_{it_3}^\top v_{it_1}e_{it_5}e_{jt_6}v_{it_2}^\top\eta_{it_4}\right)\right|\right|.$$
$$\leq \frac{C}{NT^3h^2} \tag{D.70}$$

By the same calculations, we can show that

$$\mathrm{E}\left|\left|A_{NT,17,2}^*\right|\right|^2 \leq \frac{C}{NT^3h^2}, \quad \mathrm{E}\left|\left|A_{NT,17,3}^*\right|\right|^2 \leq \frac{C}{NT^3h^2}. \tag{D.71}$$

For $A_{NT,18}^*$, by Lemma E.4(ii) and (iv) we have

$$\mathrm{E}\left|\left|A_{NT,18}^*\right|\right| \leq \frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{E}\left|\left|\frac{1}{T}\sum_{t_1=1}^{T}\eta_{it_1}\right|\right|^2\right)^{1/2}\left(\mathrm{E}\left|\left|\frac{1}{T}\sum_{t_2=1}^{T}\xi_{it_2,2}^c\right|\right|^2\right)^{1/2} \leq \frac{C\sqrt{N}h^2}{\sqrt{NT}}. \tag{D.72}$$

Similarly,

$$\mathrm{E}\left|\left|A_{NT,19}^*\right|\right| \leq \frac{1}{N}\sum_{i=1}^{N}\sqrt{\mathrm{E}\left|\left|\frac{1}{T}\sum_{t_2=1}^{T}\eta_{it_2}^c\right|\right|^2}\sqrt{\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^{T}\xi_{it}^c\right|\right|^2} \leq \frac{C}{\sqrt{NT^2}}\frac{\sqrt{N}}{Th}, \tag{D.73}$$

which completes the proof of Proposition C.3(iii).

Proof of Proposition C.3(iv) .

Recall that $\Delta_{Ti} = \frac{1}{T}\sum_{t=1}^{T}\widetilde{\eta}_{it}$ and $\Delta_{Ti} = \Delta_{Ti,1} + \Delta_{Ti,2}$, then we have

$$\begin{aligned}
&\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\Omega_{v,i}\Delta_{Ti}\left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it} \\
&= \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\Delta_{Ti,1}\left(I_d + \widetilde{Q}_{Ti}\right)\widetilde{\eta}_{it_1}\widetilde{\xi}_{it_2} \\
&\quad + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\Delta_{Ti,2}\left(I_d + \widetilde{Q}_{Ti}\right)\widetilde{\eta}_{it_1}\widetilde{\xi}_{it_2} \equiv A^*_{NT,20} + A^*_{NT,21}.
\end{aligned} \tag{D.74}$$

For $A^*_{NT,20}$, note that

$$\begin{aligned}
A^*_{NT,20} &= \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\Delta_{Ti,1}\eta_{it_1}\xi_{it_2} \\
&\quad + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\Delta_{Ti,1}\widetilde{Q}_{Ti}\eta_{it_1}\xi_{it_2} \\
&\quad + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\Delta_{Ti,1}\left(I_d + \widetilde{Q}_{Ti}\right)\eta_{it_1}\xi^c_{it_2} \\
&\quad + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\Delta_{Ti,1}\left(I_d + \widetilde{Q}_{Ti}\right)\eta^c_{it_1}\widetilde{\xi}_{it_2} \\
&\equiv A^*_{NT,20,1} + \cdots + A^*_{NT,20,4}.
\end{aligned} \tag{D.75}$$

Recall the definition of $\Delta_{Ti,1}$ and using Assumption 6, we have

$$\begin{aligned}
\mathrm{E}\left|\left|A^*_{NT,20,1}\right|\right|^2 &= \frac{1}{N^2T^4}\mathrm{E}\left|\left|\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\left(\frac{1}{T}\sum_{t_3=1}^{T}\eta_{it_3}\right)\eta_{it_1}\xi_{it_2}\right|\right|^2 \\
&\leq \frac{C}{N^2T^6}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\cdots\sum_{t_6=1}^{T}\left|\left|\mathrm{E}\left(\eta_{it_5}\eta_{it_1}\xi_{it_3}\xi^\top_{jt_4}\eta^\top_{jt_2}\eta^\top_{jt_6}\right)\right|\right|^2 \\
&\leq \frac{C_1}{NT^3} \leq \frac{C_2}{T^4h^2}.
\end{aligned} \tag{D.76}$$

By the moment condition in Assumption 6(i), using Lemma E.2 and Cauchy-Swarchz inequality, we have

$$\begin{aligned}
\mathrm{E}\left|\left|A^*_{NT,20,2}\right|\right| &\leq \frac{1}{NT^2}\sum_{i=1}^{N}\left(\mathrm{E}\left|\left|\Delta_{Ti,1}\right|\right|^{4+\delta^*}\right)^{\frac{1}{4+\delta^*}}\left(\mathrm{E}\left|\left|\widetilde{Q}_{Ti}\right|\right|^{\frac{4+\delta^*}{1+\delta^*}}\right)^{\frac{1+\delta^*}{4+\delta^*}} \\
&\quad\times\left(\mathrm{E}\left|\left|\sum_{t_1=1}^{T}\eta_{it_1}\right|\right|^{4+\delta^*}\right)^{\frac{1}{4+\delta^*}}\left(\mathrm{E}\left|\left|\sum_{t_2=1}^{T}\xi_{it_2}\right|\right|^{4+\delta^*}\right)^{\frac{1}{4+\delta^*}} \\
&\leq \frac{C_1}{T^{\frac{3}{2}+\frac{1+\delta^*}{4+\delta^*}}} \leq \frac{C_2}{T^2h}.
\end{aligned} \tag{D.77}$$

Similarly, together with the fact that $\left|\left|I_d + \widetilde{Q}_{Ti}\right|\right|$ is bounded, we have we have

$$\begin{aligned}
\mathrm{E}\left|\left|A^*_{NT,20,3}\right|\right| &\leq \frac{1}{NT^2}\sum_{i=1}^N \left(\mathrm{E}\left|\left|\Delta_{Ti,1}\right|\right|^4\right)^{\frac{1}{4}} \left(\mathrm{E}\left|\left|\sum_{t_1=1}^T \eta_{it_1}\right|\right|^4\right)^{\frac{1}{4}} \left(\mathrm{E}\left|\left|\sum_{t_2=1}^T \xi^c_{it_2}\right|\right|^2\right)^{\frac{1}{2}} \\
&\leq \frac{C_1}{\sqrt{T^4 h}} \leq \frac{C_2}{T^2 h}. \qquad \text{(D.78)}
\end{aligned}$$

Using the additional results in Lemma E.2(vi), by Cauchy-Swarchz inequality, we have,

$$\begin{aligned}
\mathrm{E}\left|\left|A^*_{NT,20,4}\right|\right| &\leq \frac{1}{NT}\sum_{i=1}^N \left(\mathrm{E}\left|\left|\Delta_{Ti,1}\right|\right|^4\right)^{\frac{1}{4}} \left(\mathrm{E}\left|\left|\sum_{t_1=1}^T \eta^c_{it_1}\right|\right|^4\right)^{\frac{1}{4}} \left(\mathrm{E}\left|\left|\frac{1}{T}\sum_{t_2=1}^T \widetilde{\xi}_{it_2}\right|\right|^2\right)^{\frac{1}{2}}. \\
&\leq \frac{C}{T^2 h} \qquad \text{(D.79)}
\end{aligned}$$

We also have

$$\begin{aligned}
\mathrm{E}\left|\left|A^*_{NT,21}\right|\right| &\leq \frac{1}{NT}\sum_{i=1}^N \left(\mathrm{E}\left|\left|\Delta_{Ti,2}\right|\right|^4\right)^{\frac{1}{4}} \left(\mathrm{E}\left|\left|\sum_{t_1=1}^T \widetilde{\eta}_{it_1}\right|\right|^4\right)^{\frac{1}{4}} \left(\mathrm{E}\left|\left|\frac{1}{T}\sum_{t_2=1}^T \widetilde{\xi}_{it_2}\right|\right|^2\right)^{\frac{1}{2}} \\
&\leq \frac{C}{T^2 h}, \qquad \text{(D.80)}
\end{aligned}$$

which completes the proof of Proposition C.3(iv).

Proof of Proposition C.3(v). By definition of $\zeta_{N,t}$, we have

$$\begin{aligned}
&\mathrm{E}\left|\left|\zeta_{N,t} - \mathrm{E}\left(\zeta_{N,t}|\mathscr{F}^{t+m}_{t-m}\right)\right|\right|^2 = \frac{\gamma_{NT}}{NT}\mathrm{E}\left|\left|\sum_{i=1}^N \Omega_{v,i}\Delta_{Ti,1}\xi_{it} - \mathrm{E}\left(\sum_{i=1}^N \Omega_{v,i}\Delta_{Ti,1}\xi_{it}|\mathscr{F}^{t+m}_{t-m}\right)\right|\right|^2 \\
&= \frac{\gamma_{NT}}{NT^3}\mathrm{E}\left|\left|\sum_{i=1}^N\sum_{s=1}^{t-m-1} \Omega_{v,i}\left(\eta_{is} - \mathrm{E}\left(\eta_{is}|\mathscr{F}^{t+m}_{t-m}\right)\right)\xi_{it} + \sum_{i=1}^N\sum_{s=t+m+1}^{T} \Omega_{v,i}\left(\eta_{is} - \mathrm{E}\left(\eta_{is}|\mathscr{F}^{t+m}_{t-m}\right)\right)\xi_{it}\right|\right|^2.
\end{aligned}$$

When $m \to \infty$, the first term in the right hand side will be zero, thus we only focus on the second term

$$\begin{aligned}
&\frac{\gamma_{NT}}{NT^3}\mathrm{E}\left|\left|\sum_{i=1}^N\sum_{s=t+m+1}^T \left(\eta_{is} - \mathrm{E}\left(\eta_{is}|\mathscr{F}^{t+m}_{t-m}\right)\right)\xi_{it}\right|\right|^2 \leq \frac{\gamma_{NT}}{NT^3}\mathrm{E}\left|\left|\sum_{i=1}^N\sum_{s=t+m+1}^T \eta_{is}\xi_{it}\right|\right|^2 \\
&\quad + \frac{2\gamma_{NT}}{NT^3}\mathrm{E}\left(\left|\left|\sum_{i=1}^N\sum_{s=t+m+1}^T \eta_{is}\xi_{it}\right|\right|\left|\left|\sum_{i=1}^N\sum_{s=t+m+1}^T \mathrm{E}\left(\eta_{is}|\mathscr{F}^{t+m}_{t-m}\right)\xi_{it}\right|\right|\right) \\
&\quad + \frac{\gamma_{NT}}{NT^3}\mathrm{E}\left|\left|\sum_{i=1}^N\sum_{s=t+m+1}^T \mathrm{E}\left(\eta_{is}|\mathscr{F}^{t+m}_{t-m}\right)\xi_{it}\right|\right|^2 \\
&\equiv A^*_{NT,22} + \cdots + A^*_{NT,24}. \qquad \text{(D.81)}
\end{aligned}$$

First note taht for $B_{NT,8}$, by Jensen's inequality for conditional expectation, we have

$$\begin{aligned}
A^*_{NT,24} &= \frac{\gamma_{NT}}{NT^3}\mathrm{E}\left|\left|\mathrm{E}\left(\sum_{i=1}^N\sum_{s=t+m+1}^T \eta_{is}\xi_{it}|\mathscr{F}^{t+m}_{t-m}\right)\right|\right|^2 \leq \frac{\gamma_{NT}}{NT^3}\mathrm{E}\left(\mathrm{E}\left|\left|\sum_{i=1}^N\sum_{s=t+m+1}^T \eta_{is}\xi_{it}\right|\right|^2|\mathscr{F}^{t+m}_{t-m}\right) \\
&\leq \frac{\gamma_{NT}}{NT^3}\mathrm{E}\left|\left|\sum_{i=1}^N\sum_{s=t+m+1}^T \eta_{is}\xi_{it}\right|\right|^2 = A^*_{NT,22}. \qquad \text{(D.82)}
\end{aligned}$$

Based on Eqs. (D.82), we have

$$A_{NT,23}^{*} \leq \frac{\gamma_{NT}}{NT^{3}} \sqrt{\mathrm{E}\left|\left|\sum_{i=1}^{N} \sum_{s=t+m+1}^{T} \eta_{is}\xi_{it}\right|\right|^{2}} \sqrt{\mathrm{E}\left|\left|\sum_{i=1}^{N} \sum_{s=t+m+1}^{T} \mathrm{E}\left(\eta_{is}|\mathscr{F}_{t-m}^{t+m}\right)\xi_{it}\right|\right|^{2}} \leq A_{NT,22}^{*}, \tag{D.83}$$

which implies

$$\mathrm{E}\left|\left|\zeta_{N,t} - \mathrm{E}\left(\zeta_{N,t}|\mathscr{F}_{t-m}^{t+m}\right)\right|\right|^{2} \leq CA_{NT,22}^{*}. \tag{D.84}$$

Next, by some properties for the α–mixing sequence, and for some constant $\phi > 1$,

$$\begin{aligned}
A_{NT,22}^{*} &\leq \frac{C\gamma_{NT}}{NT^{3}} \sum_{i=1}^{N}\sum_{j=1}^{N} \sum_{s_1=t+m+1}^{T} \sum_{s_2\geq s_1}^{T} \alpha^{\frac{\delta^*/2}{2+\delta^*/2}}\left(s_1 - t\right) \left|\left|\mathrm{E}\left(\eta_{is_1}^{\top}\right)\right|\right|^{4+\delta^*/2} \left|\left|\mathrm{E}\left(\eta_{js_2}\xi_{jt}\xi_{it}^{\top}\right)\right|\right|^{\frac{4+\delta^*/2}{3}} \\
&\leq m^{-\phi}\frac{C\gamma_{NT}N(T-m)m^{\phi}}{T^{3}} \sum_{k=m+1}^{T} \alpha^{\frac{\delta^*/2}{2+\delta^*/2}}(k) \\
&\leq m^{-\phi}\frac{C\gamma_{NT}N(T-m)}{T^{3}} \sum_{k=m+1}^{T} k^{\phi}\alpha^{\frac{\delta^*/2}{2+\delta^*/2}}(k) \leq m^{-\phi}\frac{C\gamma_{NT}N}{T^{2}}.
\end{aligned} \tag{D.85}$$

Letting $v_m = m^{-\phi}$ and $d_t = \frac{C\gamma_{NT}N}{T^2}$, thus we have completed the proof of Proposition C.3(v).

Proof of Proposition C.4.

Recall that $\widehat{e}_{it} = \widetilde{Y}_{it} - \widetilde{X}_{it}^{\top}\widehat{\beta}_i = \widetilde{X}_{it}^{\top}(\beta_i - \widehat{\beta}_i) + \widetilde{f}_{it} + \widetilde{e}_{it}$, $\widehat{\beta}_i - \beta_i = \frac{1}{T}\widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_i^{\top}\left(\widetilde{e}_i + \widetilde{f}_i\right)$, $\widehat{L}_{NT}(r) = \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{[rT]}\left(\widehat{\Omega}_{x,NT}^{-1} - \widehat{\Omega}_{x,Ti}^{-1}\right)\widetilde{X}_{it}\widehat{e}_{it}$, and $\widetilde{\xi}_{it} = \Omega_{v,i}^{-1}\widetilde{X}_{it}\left(\widetilde{e}_{it} + \widetilde{f}_{it}\right)$. Then by standard calculation,

$$\begin{aligned}
&\sqrt{\gamma_{NT}}\widehat{L}_{NT}(r) - \Omega_{v,i}^{-1}\left(L_{NT}(r) - rL_{NT}(1)\right) \\
&= \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \sum_{i=1}^{N}\sum_{t=1}^{[rT]} \left(\widehat{\Omega}_{x,NT}^{-1} - \widehat{\Omega}_{x,Ti}^{-1}\right)\widetilde{X}_{it}\left(\widetilde{X}_{it}^{\top}(\beta_i - \widehat{\beta}_i) + \widetilde{f}_{it} + \widetilde{e}_{it}\right) - \Omega_{v}^{-1}\left(L_{NT}(r) - rL_{NT}(1)\right) \\
&= -\left\{\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}T} \sum_{i=1}^{N}\sum_{t=1}^{[rT]}\sum_{s=1}^{T} \left(\widehat{\Omega}_{x,NT}^{-1} - \widehat{\Omega}_{x,Ti}^{-1}\right)\widetilde{X}_{it}\widetilde{X}_{it}^{\top}\widehat{\Omega}_{x,Ti}^{-1}\Omega_{v,i}\widetilde{\xi}_{is} - r\Omega_{v}^{-1}L_{NT}(1)\right\} \\
&\quad + \left\{\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \sum_{i=1}^{N}\sum_{t=1}^{[rT]} \left(\widehat{\Omega}_{x,NT}^{-1} - \widehat{\Omega}_{x,Ti}^{-1}\right)\Omega_{v,i}\widetilde{\xi}_{it} - \Omega_{v}^{-1}L_{NT}(r)\right\} \\
&\equiv A_{NT,1}^{*}(r) + A_{NT,2}^{*}(r).
\end{aligned} \tag{D.86}$$

Similar to the proof of Lemma 4.1, we have shown

$$\begin{aligned}
\widehat{\Omega}_{x,NT}^{-1} - \widehat{\Omega}_{x,Ti}^{-1} &= \widehat{\Omega}_{x,NT}^{-1}\left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right)\widehat{\Omega}_{x,Ti}^{-1} \\
&= \left(\Omega_{v}^{-1} + Q_{NT}\right)\left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right)\widehat{\Omega}_{x,Ti}^{-1}.
\end{aligned} \tag{D.87}$$

Thus

$$A_{NT,1}^{*}(r) = -\left\{\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}T}\Omega_{v}^{-1} \sum_{i=1}^{N}\sum_{t=1}^{[rT]}\sum_{s=1}^{T} \left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right)\widehat{\Omega}_{x,Ti}^{-1}\widetilde{X}_{it}\widetilde{X}_{it}^{\top}\widehat{\Omega}_{x,Ti}^{-1}\Omega_{v,i}\widetilde{\xi}_{is} - r\Omega_{v}^{-1}L_{NT}(1)\right\}$$

$$
-\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}T} Q_{NT} \sum_{i=1}^{N}\sum_{t=1}^{[rT]}\sum_{s=1}^{T} \left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right) \widehat{\Omega}_{x,Ti}^{-1} \widetilde{X}_{it} \widetilde{X}_{it}^{\top} \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} \widetilde{\xi}_{is}
$$

$$
\equiv A^*_{NT,1,1}(r) + A^*_{NT,1,2}(r). \tag{D.88}
$$

Recall $\widetilde{\eta}_{it} = \Omega_{v,i}^{-1} \widetilde{X}_{it} \widetilde{X}_{it}^{\top} - I_d$, $\widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} = I_d - \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} \Delta_{Ti}$, and $\widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} = I_d + \widetilde{Q}_{Ti} + \widetilde{Q}_{Ti}^{c}$, then

$$
\begin{aligned}
&\widehat{\Omega}_{x,Ti}^{-1} \widetilde{X}_{it} \widetilde{X}_{it}^{\top} \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} = \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} \left(I_d + \widetilde{\eta}_{it}\right) \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} \\
&= \left(I_d - \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} \Delta_{Ti}\right) \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} + \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} \widetilde{\eta}_{it} \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} \\
&= \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} + \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} \left(\widetilde{\eta}_{it} - \Delta_{Ti}\right) \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} \\
&= \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} + \left(I_d + \widetilde{Q}_{Ti} + \widetilde{Q}_{Ti}^{c}\right) \left(\widetilde{\eta}_{it} - \Delta_{Ti}\right) \left(I_d + \widetilde{Q}_{Ti} + \widetilde{Q}_{Ti}^{c}\right) \\
&= \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} + \left(I_d + \widetilde{Q}_{Ti}\right) \widetilde{\eta}_{it}^{*} \left(I_d + \widetilde{Q}_{Ti}\right) + \left(\widetilde{Q}_{Ti}^{c} \widetilde{\eta}_{it}^{*} \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} + \left(I_d + \widetilde{Q}_{Ti}\right) \widetilde{\eta}_{it}^{*} \widetilde{Q}_{Ti}^{c}\right),
\end{aligned}
$$

where $\widetilde{\eta}_{it}^{*} = \widetilde{\eta}_{it} - \Delta_{Ti}$. Thus

$$
\begin{aligned}
A^*_{NT,1,1}(r) = &-\left\{\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}T} \Omega_v^{-1} \sum_{i=1}^{N}\sum_{t=1}^{[rT]}\sum_{s=1}^{T} \left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right) \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} \widetilde{\xi}_{is} - r\Omega_v^{-1} L_{NT}(1)\right\} \\
&-\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}T} \Omega_v^{-1} \sum_{i=1}^{N}\sum_{t=1}^{[rT]}\sum_{s=1}^{T} \left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right) \left(I_d + \widetilde{Q}_{Ti}\right) \widetilde{\eta}_{it}^{*} \left(I_d + \widetilde{Q}_{Ti}\right) \widetilde{\xi}_{is} \\
&-\left\{\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}T} \Omega_v^{-1} \sum_{i=1}^{N}\sum_{t=1}^{[rT]}\sum_{s=1}^{T} \left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right) \left(\widetilde{Q}_{Ti}^{c} \widetilde{\eta}_{it}^{*} \widehat{\Omega}_{x,Ti}^{-1} \Omega_{v,i} + \left(I_d + \widetilde{Q}_{Ti}\right) \widetilde{\eta}_{it}^{*} \widetilde{Q}_{Ti}^{c}\right) \widetilde{\xi}_{is}\right\} \\
\equiv{}& A^*_{NT,1,1,1}(r) + \cdots + A^*_{NT,1,1,3}(r).
\end{aligned} \tag{D.89}
$$

In the proof of Lemma 4.1, we have already shown

$$
\sup_{0 \le r \le 1} ||A^*_{NT,1,1,1}(r)|| = o_P(1). \tag{D.90}
$$

In Lemma E.5(i), we have shown

$$
\sup_{0 \le r \le 1} ||A^*_{NT,1,1,2}(r)|| = o_P(1) \tag{D.91}
$$

and by the same calculation of Eqs.(B.13) as in the proof of Theorem 3.3, we have

$$
\sup_{0 \le r \le 1} ||A^*_{NT,1,1,3}(r)|| = o_P(1). \tag{D.92}
$$

Thus

$$
\sup_{0 \le r \le 1} ||A^*_{NT,1,1}(r)|| = o_P(1). \tag{D.93}
$$

By the results for $A_{NT,1,1}(r)$ and $Q_{NT} \to 0$, it is clear that

$$
\sup_{0 \le r \le 1} ||A^*_{NT,1,2}(r)|| = o_P(1). \tag{D.94}
$$

Hence

$$
\sup_{0 \le r \le 1} ||A^*_{NT,1}(r)|| = o_P(1). \tag{D.95}
$$

Similar to the calculations with $A_{NT,1}(r)^*$, we have

$$\begin{aligned}
A^*_{NT,2}(r) &= \left\{ \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \Omega_v^{-1} \sum_{i=1}^{N} \sum_{t=1}^{[rT]} \left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT} \right) \widehat{\Omega}^{-1}_{x,Ti} \Omega_{v,i} \widetilde{\xi}_{is} - \Omega_v^{-1} L_{NT}(r) \right\} \\
&\quad + \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} Q_{NT} \sum_{i=1}^{N} \sum_{t=1}^{[rT]} \left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT} \right) \widehat{\Omega}^{-1}_{x,Ti} \Omega_{v,i} \widetilde{\xi}_{is} \\
&\equiv A^*_{NT,2,1}(r) + A^*_{NT,2,2}(r).
\end{aligned} \tag{D.96}$$

Recall that $\widehat{\Omega}^{-1}_{x,Ti}\Omega_{v,i} = I_d - \left(I_d + \widetilde{Q}_{Ti} \right) \Delta_{Ti} - \widetilde{Q}^c_{Ti} \Delta_{Ti}$. Thus

$$\begin{aligned}
A^*_{NT,2,1}(r) &= \left\{ \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \Omega_v^{-1} \sum_{i=1}^{N} \sum_{t=1}^{[rT]} \left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT} \right) \left(I_d - \left(I_d + \widetilde{Q}_{Ti} \right) \Delta_{Ti} \right) \widetilde{\xi}_{is} - \Omega_v^{-1} L_{NT}(r) \right\} \\
&\quad + \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \Omega_v^{-1} \sum_{i=1}^{N} \sum_{t=1}^{[rT]} \left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT} \right) \widetilde{Q}^c_{Ti} \Delta_{Ti} \widetilde{\xi}_{is} \\
&= A^*_{NT,2,1,1}(r) + A^*_{NT,2,1,2}(r).
\end{aligned} \tag{D.97}$$

In Lemma E.5(ii), we have shown that

$$\sup_{0 \le r \le 1} ||A^*_{NT,2,1,1}(r)|| = o_P(1) \tag{D.98}$$

and by the same calculation of Eqs.(B.13) in the proof of Theorem 3.3, we have

$$\sup_{0 \le r \le 1} ||A^*_{NT,2,1,2}(r)|| = o_P(1). \tag{D.99}$$

Hence

$$\sup_{0 \le r \le 1} ||A^*_{NT,2,1}(r)|| = o_P(1). \tag{D.100}$$

Since $Q_{NT} \to 0$, we have

$$\sup_{0 \le r \le 1} ||A^*_{NT,2,2}(r)|| = o_P(1). \tag{D.101}$$

Finally, we have

$$\sup_{0 \le r \le 1} ||A^*_{NT,2}(r)|| = o_P(1), \tag{D.102}$$

which completes the proof of Proposition C.4.

Appendix E Lemmas

In this part, we prove some necessary lemmas which have been used in the proofs of the propositions.

Lemma E.1. Under Assumption 4, we have:
(i)$s_i(\tau)\iota_T = 1$, $\max_{it} |\widetilde{f}_{it}| = O(h^2)$, and $\max_{it} ||\widetilde{g}_{it}|| = O(h^2)$;
(ii)$\max_{it} |w_{T,it}(\tau)| = O(\frac{1}{Th})$ for $0 \le \tau \le 1$;
(iii) $\max_{it} \sum_{s=1}^{T} |w_{T,is}(\tau_t)| < \infty$ and $\max_{is} \sum_{t=1}^{T} |w_{T,is}(\tau_t)| < \infty$.

Lemma E.2. Under Assumptions 1 to 4,

(i) $\mathrm{E}\left|\left|s_i(\tau)v_i\right|\right|^4 \leq \frac{C}{T^2h^2}$;

(ii) $\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^T \eta_{it}\right|\right|^{2+\delta^*/2} \leq \frac{C}{T^{1+\delta^*/4}}$, and $\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^T \xi_{it}\right|\right|^{2+\delta^*/2} \leq \frac{C}{T^{1+\delta^*/4}}$;

(iii)$\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^T \eta_{it,1}^c\right|\right|^2 \leq \frac{C}{T^2h}$, and $\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^T \xi_{it,1}^c\right|\right|^2 \leq \frac{C}{T^2h}$;

(iv)$\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^T \eta_{it,2}^c\right|\right|^2 \leq \frac{Ch^4}{T}$, and $\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^T \xi_{it,2}^c\right|\right|^2 \leq \frac{Ch^4}{T}$;

(v)$\mathrm{E}\left|\left|\Delta_{Ti}\right|\right|^2 \leq \frac{C}{T}$, $\mathrm{E}\left|\left|\Delta_{Ti,2}\right|\right|^2 \leq \frac{C}{T^2h}$, $\mathrm{E}\left|\left|\widetilde{Q}_{Ti}\right|\right|^2 \leq \frac{C}{T}$, $\mathrm{E}\left|\left|\widetilde{Q}_{Ti}\right|\right|^4 \leq \frac{C}{T}$, and $\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^T \widetilde{\xi}_{it}\right|\right|^2 \leq \frac{C}{T}$;

(vi)$\sum_{t_1=1}^T\sum_{t_2=1}^T\left|\left|\mathrm{E}\left(\eta_{it_1,1}^c{\eta_{jt_2,1}^c}^\top\right)\right|\right| \leq \frac{C}{h}$, and $\sum_{t_1=1}^T\sum_{t_2=1}^T\left|\left|\mathrm{E}\left(\xi_{it_1,1}^c{\xi_{jt_2,1}^c}^\top\right)\right|\right| \leq \frac{C}{h}$;

Furthermore, if Assumption 6(i) holds in addition, we have

(vii)$\mathrm{E}\left|\left|\Delta_{Ti,2}\right|\right|^4 \leq \frac{C}{T^4h^4}$.

The notation of κ_T, l_T, and s_T in the next lemma are defined in the proof of Proposition C.2.

Lemma E.3. Under Assumption 1 to 4, we have

(i)$\sum_{p=1}^{\kappa_T}\sum_{t_1=pl_T+(p-1)s_T+1}^{p(l_T+s_T)} w_{T,it_1}(\tau) = O(\frac{\kappa_T s_T}{T})$;

(ii)$\sum_{p=1}^{\kappa_T}\mathrm{E}\left(\widetilde{V}_p^2 I\left(\left|\widetilde{V}_{N,p}\right| \geq \varepsilon\right)\right) \leq \frac{Cl_T}{Th_i}$ for some constant C;

(iii)$\sum_{p=1}^{\kappa_T}\mathrm{E}\left(\widetilde{V}_p^2\right) \to \sigma_{eK}^2$.

Lemma E.4. Under Assumptions 1 to 4, we have

(i) $\mathrm{E}\left|\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T s_i(\tau)v_i\xi_{it}\right|\right|^2 = O\left(\frac{1}{NT^2h^2}\right)$, and $\mathrm{E}\left|\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T v_{i\cdot}^\top\xi_{it}\right|\right|^2 = O\left(\frac{1}{NT^2h^2}\right)$;

(ii) $\mathrm{E}\left|\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\left(s_i(\tau)g_i - g_{i\cdot}^\top\right)\xi_{it}^c\right|\right|^2 \leq \frac{C}{NT^2h}$, $\mathrm{E}\left|\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \eta_{it}^c\right|\right|^2 \leq \frac{C}{NT^2h}$, ;

(iii)$\sum_{i=1}^N\sum_{j=1}^N\sum_{t_1=1}^T\sum_{t_2=1}^T\left|\left|\mathrm{E}\left(\eta_{it_1,1}^c{\eta_{jt_2,1}^c}^\top\right)\right|\right| \leq \frac{CN}{h}$, and $\sum_{i=1}^N\sum_{j=1}^N\sum_{t_1=1}^T\sum_{t_2=1}^T\left|\left|\mathrm{E}\left(\xi_{it_1,1}^c{\xi_{jt_2,1}^c}^\top\right)\right|\right|$
$\leq \frac{CN}{h}$.

Lemma E.5. Under Assumptions 1 to 8, we have

(i) $\sup_{0\leq r\leq 1}\left|\left|\frac{\sqrt{\gamma_{NT}}}{\sqrt{NTT}}\sum_{i=1}^N\sum_{t_1=1}^{[rT]}\sum_{t_2=1}^T\left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right)\left(I_d + \widetilde{Q}_{Ti}\right)\widetilde{\eta}_{it_1}^*\left(I_d + \widetilde{Q}_{Ti}\right)\widetilde{\xi}_{it_2}\right|\right| = o_P(1)$;

(ii) $\sup_{0\leq r\leq 1}\left|\left|\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}}\sum_{i=1}^N\sum_{t=1}^{[rT]}\left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right)\left(I_d - \left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\right)\widetilde{\xi}_{it} - L_{NT}(r)\right|\right| = o_P(1)$.

The following lemma given by Corollary 3.1 in Móricz et al. (1982) will be useful in proving Lemma E.5.

Lemma E.6. Let $X_1, \cdots, X_n$ be arbitrary random variables. The only restrictions on the join distribution of the $X_k's$ will be those imposed by the assumed bounds on $\mathrm{E}|S(i,j)|^\gamma$, $1 \leq i \leq j \leq n$, where $S(i,j) = \sum_{k=i}^j X_k$. Define $M(i,j) = \max\{|S(i,i)|, |S(i,i+1)|, \cdots, |S(i,j)|\}$, and some function $g(i,j)$ satisfying

$$
\begin{aligned}
&g(i,j) \geq 0, \quad \text{all} \quad 1 \leq i \leq j \leq n, \\
&g(i,j) \leq g(i,j+1), \quad \text{all} \quad 1 \leq i \leq j \leq n \\
&g(i,j) + g(j+1,k) \leq Qg(i,k), \quad 1 \leq i \leq j < k \leq n
\end{aligned} \tag{E.1}
$$

where $Q \geq 1$ Let $\gamma \geq 1$ be a given real. Suppose that

$$
\mathrm{E}\left|S(i,j)\right|^\gamma \leq g(i,j), \quad 1 \leq i \leq j \leq n
$$

Then

$$\mathrm{E}M^{\gamma}(1,n)\leq g(1,n)\left(\sum_{k=0}^{[\log n]}Q^{k/\gamma}\right)^{\gamma}$$

where $\left(\sum_{k=0}^{[\log n]}Q^{k/\gamma}\right)^{\gamma}$ is of order of $(\log(2T))^{\gamma}$ for $Q=1$ and $T^{\log Q}$ for $Q>1$.

Lemma E.1(i) are standard results in the local linear fitting, more details are available at Chapter 2 in Li and Racine (2007) and Appendix B in Chen et al. (2012).

Proof of Lemma E.1(ii).

Recall that

$$w_{T,it}(\tau)=\frac{1}{Th_i}(1,0)\left(\frac{1}{Th_i}Z_i^{\top}(\tau)W_i(\tau)Z_i(\tau)\right)^{-1}\left(K_{it,0}(\tau),K_{it,1}(\tau)\right)^{\top}, \tag{E.2}$$

where $K_{it,l}(\tau)=\left(\frac{t/T-\tau}{h_i}\right)^{l}K\left(\frac{t/T-\tau}{h_i}\right)$. Denote $S_{i,l}(\tau)=\frac{1}{Th_i}\sum_{t=1}^{T}K_{it,l}(\tau)$, by standard calculation we have

$$\left(\frac{1}{Th_i}Z_i^{\top}(\tau)W_i(\tau)Z_i(\tau)\right)^{-1}=\begin{pmatrix}\frac{S_{i,2}(\tau)}{S_{i,0}(\tau)S_{i,2}(\tau)-(S_{i,1}(\tau))^2} & -\frac{S_{i,1}(\tau)}{S_{i,0}(\tau)S_{i,2}(\tau)-(S_{i,1}(\tau))^2}\\ -\frac{S_{i,1}(\tau)}{S_{i,0}(\tau)S_{i,2}(\tau)-(S_{i,1}(\tau))^2} & \frac{S_{i,0}(\tau)}{S_{i,0}(\tau)S_{i,2}(\tau)-(S_{i,1}(\tau))^2}\end{pmatrix}. \tag{E.3}$$

Thus

$$w_{T,it}(\tau)=\frac{1}{Th_i}\left(S_{i,0}(\tau)S_{i,2}(\tau)-(S_{i,1}(\tau))^2\right)^{-1}\left(S_{i,2}(\tau)K_{it,0}(\tau)-S_{i,1}(\tau)K_{it,1}(\tau)\right). \tag{E.4}$$

To prove Lemma E.1(ii), by Assumption 4 $K_{it,j}(\tau)$ will be bounded uniformly for i and t, then we need to show that

$$|S_{i,l}(\tau)|\leq\mathcal{M}_1\quad\text{and} \tag{E.5}$$

$$\left|\left(S_{i,0}(\tau)S_{i,2}(\tau)-(S_{i,1}(\tau))^2\right)^{-1}\right|\leq\mathcal{M}_2 \tag{E.6}$$

hold for all i and τ uniformly for finite constant $\mathcal{M}_1$, $\mathcal{M}_2$.

To prove Eqs. (E.5), we have shown the uniform convergence of $S_{i,l}(\tau)$ by the definition of Riemann integral. Consider $(\tau_0,\tau_1,\cdots,\tau_T)$ as a partition of $[0,1]$, and denote

$$M_t(\tau)=\sup_{\tau_{t-1}\leq\tau_s\leq\tau_t}\left(\frac{\tau_s-\tau}{h_i}\right)^{l}K\left(\frac{\tau_s-\tau}{h_i}\right),\quad m_t(\tau)=\inf_{\tau_{t-1}\leq\tau_s\leq\tau_t}\left(\frac{\tau_s-\tau}{h_i}\right)^{l}K\left(\frac{\tau_s-\tau}{h_i}\right). \tag{E.7}$$

Since $K(\cdot)$ is a continuous density function defined on a compact support, we have

$$\sup_{0\leq\tau\leq1}\max_{1\leq t\leq T}\left(M_t(\tau)-m_t(\tau)\right)\leq\frac{C}{Th} \tag{E.8}$$

for some constant $0<C<\infty$. By the definition of Riemann integral,

$$\frac{1}{h_i}\frac{1}{T}\sum_{t=1}^{T}m_t(\tau)\leq\mu_l(\tau)\leq\frac{1}{h_i}\frac{1}{T}\sum_{t=1}^{T}M_t(\tau), \tag{E.9}$$

where $\mu_l(\tau) = \int_{-1}^{1} u^l K(u)du$ if $h \le \tau \le 1-h$, $\mu_l(\tau) = \int_{-c}^{1} u^l K(u)du$ if $\tau = ch$ with $0 \le c \le 1$, and $\mu_l(\tau) = \int_{-1}^{c} u^l K(u)du$ if $\tau = 1 - ch$ with $0 \le c \le 1$. And by the definitions of $M_t(\tau)$ and $m_t(\tau)$, it is clear that

$$\frac{1}{Th_i}\sum_{t=1}^{T} m_t(\tau) \le S_{i,l}(\tau) \le \frac{1}{Th_i}\sum_{t=1}^{T} M_t(\tau). \tag{E.10}$$

Thus by Eqs. (E.8) to Eqs. (E.10), we have

$$|S_{i,l}(\tau) - \mu_l(\tau)| \le \frac{1}{Th_i}\sum_{t=1}^{T}(M_t(\tau) - m_t(\tau)) \le \frac{C}{Th^2}, \tag{E.11}$$

which implies $S_{i,l}(\tau)$ converge to $\mu_l(\tau)$ uniformly for i and τ. Thus $S_{i,l}(\tau)$ is uniformly bounded.

To prove Eqs. (E.6), we need to show

$$\left| S_{i,2}(\tau)S_{i,0}(\tau) - (S_{i,1}(\tau))^2 \right| > c \tag{E.12}$$

for some constant $c > 0$ which does not depend on i and τ. By standard calculation, we have

$$\begin{aligned}
&S_{i,2}(\tau)S_{i,0}(\tau) - (S_{i,1}(\tau))^2 = \frac{1}{T^2h_i^2}\sum_{t_1=1}^{T} K_{it_1,0}(\tau)\sum_{t_2=1}^{T} K_{it_2,2}(\tau) - \frac{1}{T^2h_i^2}\left(\sum_{t=1}^{T} K_{it,1}(\tau)\right)^2 \\
&= \frac{1}{T^2h_i^2}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left(\frac{\tau_{t_1}-\tau}{h_i}\right)K\left(\frac{\tau_{t_1}-\tau}{h_i}\right)\left(\frac{\tau_{t_1}-\tau_{t_2}}{h_i}\right)K\left(\frac{\tau_{t_2}-\tau}{h_i}\right) \\
&= \frac{1}{T^2h_i^2}\sum_{t_1<t_2}^{T}\left(\frac{\tau_{t_1}-\tau}{h_i}\right)K\left(\frac{\tau_{t_1}-\tau}{h_i}\right)\left(\frac{\tau_{t_1}-\tau_{t_2}}{h_i}\right)K\left(\frac{\tau_{t_2}-\tau}{h_i}\right) \\
&\quad + \frac{1}{T^2h_i^2}\sum_{t_1>t_2}^{T}\left(\frac{\tau_{t_1}-\tau}{h_i}\right)K\left(\frac{\tau_{t_1}-\tau}{h_i}\right)\left(\frac{\tau_{t_1}-\tau_{t_2}}{h_i}\right)K\left(\frac{\tau_{t_2}-\tau}{h_i}\right) \\
&\equiv B_{Ti,1} + B_{Ti,2}.
\end{aligned} \tag{E.13}$$

Next, note the symmetric features of $B_{Ti,2}$, by replacing t_1 with s_2 and t_2 with s_1 for $1 \le s_1 < s_2 \le T$, we have

$$\begin{aligned}
B_{Ti,2} &= -\frac{1}{T^2h_i^2}\sum_{t_1>t_2}^{T}\left(\frac{\tau_{t_1}-\tau}{h_i}\right)K\left(\frac{\tau_{t_1}-\tau}{h_i}\right)\left(\frac{\tau_{t_2}-\tau_{t_1}}{h_i}\right)K\left(\frac{\tau_{t_2}-\tau}{h_i}\right) \\
&= -\frac{1}{T^2h_i^2}\sum_{s_1<s_2}^{T}\left(\frac{\tau_{s_2}-\tau}{h_i}\right)K\left(\frac{\tau_{s_2}-\tau}{h_i}\right)\left(\frac{\tau_{s_1}-\tau_{s_2}}{h_i}\right)K\left(\frac{\tau_{s_1}-\tau}{h_i}\right).
\end{aligned} \tag{E.14}$$

Comparing $B_{Ti,2}$ in Eqs. (E.14) with $B_{Ti,1}$ in Eqs. (E.13) again, we have

$$S_{i,2}(\tau)S_{i,0}(\tau) - (S_{i,1}(\tau))^2 = \frac{1}{T^2h_i^2}\sum_{s_1=1}^{T}\sum_{s_1<s_2} K\left(\frac{\tau_{s_1}-\tau}{h_i}\right)\left(\frac{\tau_{s_1}-\tau_{s_2}}{h_i}\right)^2 K\left(\frac{\tau_{s_2}-\tau}{h_i}\right). \tag{E.15}$$

Thus $S_{i,2}(\tau)S_{i,0}(\tau) - (S_{i,1}(\tau))^2$ will be positive for all i and τ.

Next, note $K(\cdot)$ is a continuous and symmetric density function, we can find an interval $\{\nu : a \le \nu \le b\}$ such that $\min_{a\le\nu\le b} K(\nu) > c_1$ for some constant $c_1 > 0$, and $b - a > c_2$ for some constant $c_2 > 0$. Then consider the following two subsets:

$$B_{Thi,1} = \{s : a \le \frac{s/T-\tau}{h_i} \le a + \frac{c_2}{3}\}, \quad B_{Thi,2} = \{s : b - \frac{c_2}{3} \le \frac{s/T-\tau}{h_i} \le b\}. \tag{E.16}$$

Then for any $s_1 \in B_{Th,1}$ and $s_2 \in B_{Th,2}$,

$$K\left(\frac{\tau_{s_1}-\tau}{h_i}\right) > c_1, \quad K\left(\frac{\tau_{s_2}-\tau}{h_i}\right) > c_1, \quad \left(\frac{\tau_{s_2}-\tau_{s_1}}{h_i}\right)^2 > \left(\frac{b-a}{3}\right)^2 = \frac{c_2^2}{9}. \tag{E.17}$$

Note all the terms within Eqs. (E.15) are positive. Thus

$$S_{i,2}(\tau)S_{i,0}(\tau) - (S_{i,1}(\tau))^2 \geq \frac{1}{T^2h_i^2}\sum_{s_1\in B_{Th,1}}\sum_{s_2\in B_{Th,2}} K\left(\frac{\tau_{s_2}-\tau}{h_i}\right)\left(\frac{\tau_{s_2}-\tau_{s_1}}{h_i}\right)^2 K\left(\frac{\tau_{s_1}-\tau}{h_i}\right)$$
$$\geq \frac{1}{T^2h_i^2}\left(\frac{[(b-a)Th_i]}{3}\right)^2\frac{c_1^2c_2^2}{9} \geq C\frac{c_1^2c_2^4}{81}, \tag{E.18}$$

which suffices to prove Eqs. (E.6).

Proof of Lemma E.2(iii)

By the same argument to prove Eqs. (E.6), it is direct that there exists some constant $C < \infty$ such that

$$\frac{1}{Th_i}\sum_{s=1}^{T}|K_{is,0}(\tau_t)| < C \tag{E.19}$$

uniformly for all i and t. Using Eqs. (E.19) along with Eqs. (E.4), Eqs. (E.5) and Eqs. (E.6), by standard calculations, we have

$$\sum_{s=1}^{T}|w_{T,is}(\tau_t)| \leq \frac{1}{Th_i}\sum_{s=1}^{T}\left|\left(S_{i,0}(\tau)S_{i,2}(\tau)-(S_{i,1}(\tau))^2\right)^{-1}\right||S_{i,2}(\tau_t)K_{is,0}(\tau)-S_{i,1}(\tau_t)K_{is,1}(\tau_t)|$$
$$\leq \frac{C_1}{Th_i}\sum_{s=1}^{T}|K_{is,0}(\tau_t)| + \frac{C_2}{Th_i}\sum_{s=1}^{T}|K_{is,1}(\tau_t)| < C_3 \tag{E.20}$$

holds for some constant $C_3 < \infty$ irrelevant with all i and t. And the remaining part of Lemma E.2(iii) can be verified by repeating the same calculations.

Proof of Lemma E.2(i).

Recall $s_i(\tau) = \sum_{t=1}^{T} w_{T,it}(\tau)v_{it}$, for the first part of Lemma E.2(iv), we need to show that

$$\sum_{t_1=1}^{T}\cdots\sum_{t_4=1}^{T}\prod_{k=1}^{4}|w_{T,it_k}(\tau)|\left\|\mathrm{E}\left(v_{it_1}^\top v_{it_2}v_{it_3}^\top v_{it_4}\right)\right\| \leq \frac{C}{T^2h^2}. \tag{E.21}$$

First consider the case where t_1 to t_4 are distinct. Assume $1 \leq s_1 < \cdots < s_4 \leq T$ is the permutation of $t_1, \cdots, t_4$ in ascending order with d_c being the c-th largest different between $s_{j+1}-s_j$ for $j = 1, \cdots, 3$. Without loss of generality, we assume $s_3 - s_2 < d_2$. By properties of α-mixing such as Lemma A.1 in Gao (2007), and Lemma E.1 (ii) and (iii),

$$\sum_{t_1=1}^{T}\cdots\sum_{t_4=1}^{T}\prod_{l=1}^{4}|w_{T,is_l}(\tau)|\left\|\mathrm{E}\left(v_{it_1}^\top v_{it_2}v_{it_3}^\top v_{it_4}\right)\right\|$$

$$\leq \sum_{s_1=1}^{T} \sum_{s_2=s_1+d_2} \sum_{0<s_3-s_2\leq d_2} \sum_{s_4=1}^{T} \prod_{l=1}^{4} |w_{T,is_l}(\tau)| \left|\left| \mathrm{E}\left(v_{is_1}^\top v_{is_2} v_{is_3}^\top v_{is_4}\right)\right|\right|$$
$$\leq \frac{C}{T^2h^2} \sum_{s_1=1}^{T} \sum_{d_2=1}^{T} \sum_{s_4=1}^{T} |w_{T,is_1}(\tau)|\,|w_{T,is_4}(\tau)|\, d_2 \left(\alpha(d_2)\right)^{\frac{\delta}{4+\delta}} \left(\mathrm{E}\,||v_{i,1}||^{4+\delta}\right)^{\frac{4}{4+\delta}}$$
$$\leq \frac{C}{T^2h^2} \left(\sum_{s_1=1}^{T} \sum_{s_4=1}^{T} |w_{T,is_1}(\tau)|\,|w_{T,is_4}(\tau)|\right) \left(\sum_{d_2=1}^{T} d_2 \left(\alpha(d_2)\right)^{\frac{\delta}{4+\delta}}\right) \leq \frac{C}{T^2h^2}. \tag{E.22}$$

The case of $s_2 - s_1 = d_1$ can be verified by the same calculation. In addition, if $s_2 - s_1 = d_3$, then $s_3 - s_2 = d_1$ or $s_3 - s_2 = d_2$ holds, which can be verified by the same arguments as in Eqs. (E.22). Thus we have completed the case where t_1 to t_4 are distinct. If there are any ties among t_1 to t_4, these results can be verified by similar calculations. Thus we have proved Eqs. (E.2)(i).

Proof of Lemma E.2(ii).

Recall that $\eta_{it} = \Omega_{v,i}^{-1} v_{it} v_{it}^\top - I_d$. Then by Assumption 1, $\{\eta_{it}, t \geq 1\}$ is an α–mixing sequence sharing the same mixing coefficient as for $\{v_t, t \geq 1\}$ for $i \geq 1$, with $\mathrm{E}\,(\eta_{it}) = 0$ and $\mathrm{E}\,||\eta_{it}||^{2+\delta/2} < \infty$. Then by Theorem 4.1 in Shao and Yu (1996), we have

$$\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^{T} \eta_{it}\right|\right|^{2+\delta^*/2} \leq \frac{C}{T^{1+\delta^*/4}} \mathrm{E}\,||\eta_{it}||^{2+\delta^*/2} \tag{E.23}$$

The second part of Lemma E.2(ii) can be verified by the same argument.

Proof of Lemma E.2(iii). Recall that

$$\Omega_{v,i} \frac{1}{T} \sum_{t=1}^{T} \eta_{it,1}^c = \frac{1}{T} \sum_{t=1}^{T} v_i^\top s^\top(\tau_t) s(\tau_t) v_i - \frac{1}{T} \sum_{t=1}^{T} v_i^\top s^\top(\tau_t) v_{it}^\top - \frac{1}{T} \sum_{t=1}^{T} v_{it} s(\tau_t) v_i$$
$$\equiv B_{Ti,3} + \cdots + B_{Ti,5}. \tag{E.24}$$

Note that

$$\mathrm{E}\,||B_{Ti,3}||^2 = \frac{C}{T^2} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \sum_{t_3=1}^{T} \cdots \sum_{t_6=1}^{T} C_w(t_1, \cdots, t_6) \mathrm{tr}\left(\mathrm{E}\left(v_{it_3} v_{it_4}^\top v_{it_5} v_{it_6}^\top\right)\right) \tag{E.25}$$

where $C_w(t_1, \cdots, t_6) = w_{T,it_3}(\tau_{t_1}) w_{T,it_4}(\tau_{t_1}) w_{T,it_5}(\tau_{t_2}) w_{T,it_6}(\tau_{t_2})$. Similar to the proof of Lemma E.2(i), we only consider the case where t_3 to t_6 are distinct. Assume $1 \leq s_1 < \cdots < s_4 \leq T$ is the permutation of $t_3, \cdots, t_6$ in ascending order with d_c be the c-th largest different between $s_{j+1} - s_j$ for $j = 1, \cdots, 3$.

First consider $s_2 - s_1 \geq d_2$, assume $s_3 - s_2 < d_2$,

$$\frac{C}{T^2} \sum_{t_1=1}^{T} \cdots \sum_{t_6=1}^{T} C_w(t_1, \cdots, t_6) \left|\left|\mathrm{E}\left(v_{it_3} v_{it_4}^\top v_{it_5} v_{it_6}^\top\right)\right|\right|$$
$$\leq \frac{C}{T^2} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \sum_{s_1=1}^{T} \sum_{s_2=s_1+1}^{T} \sum_{0<s_3-s_2\leq d_2} \sum_{s_4=s_3+1}^{T} C_w(t_1, \cdots, t_6) \left|\left|\mathrm{E}\left(v_{is_1} v_{is_2}^\top v_{is_3} v_{is_4}^\top\right)\right|\right|$$
$$\leq \frac{C}{T^3h_i} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \sum_{s_1=1}^{T} \sum_{s_2=s_1+1}^{T} \sum_{s_4=1}^{T} |C_{w,-s_3}(t_1, \cdots, t_6)|\, (s_2 - s_1) \left(\alpha(s_2 - s_1)\right)^{\frac{\delta}{4+\delta}} \left(\mathrm{E}\,||v_{i,1}||^{4+\delta}\right)^{\frac{4}{4+\delta}}$$

$$\leq \frac{C}{T^3 h_i} \max_{1\leq s_2 \leq T} \left(\sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \sum_{s_1=1}^{T} \sum_{s_4=1}^{T} |C_{w,-s_3}(t_1,\cdots,t_6)| \right) \sum_{k=1}^{T} k\left(\alpha(k)\right)^{\frac{\delta}{4+\delta}}, \tag{E.26}$$

where $C_{w,-s_3}(t_1,\cdots,t_6)$ removes $w_{T,is_3}(\cdot)$ within $C_w(t_1,\cdots,t_6)$. By Lemma E.1(ii) and (iii), we have

$$\begin{aligned}
&\max_{1\leq s_2\leq T} \left(\sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \sum_{s_1=1}^{T} \sum_{s_4=1}^{T} |C_{w,-s_3}(t_1,\cdots,t_6)| \right) \\
&= \begin{cases}
T\max_{t_4}\left(\sum_{t_1}|w_{T,it_4}(\tau_{t_1})|\right)\sum_{t_5}|w_{T,it_5}(\tau_{t_2})|\sum_{t_6}|w_{T,it_6}(\tau_{t_2})|, & \text{if } s_3=t_3, \quad s_2=t_4\\
T\max_{t_5}\left(\sum_{t_2}|w_{T,it_5}(\tau_{t_2})|\right)\sum_{t_4}|w_{T,it_4}(\tau_{t_1})|\sum_{t_6}|w_{T,it_6}(\tau_{t_2})|, & \text{if } s_3=t_3, \quad s_2=t_5\\
T\max_{t_6}\left(\sum_{t_2}|w_{T,it_6}(\tau_{t_2})|\right)\sum_{t_4}|w_{T,it_4}(\tau_{t_1})|\sum_{t_5}|w_{T,it_5}(\tau_{t_2})|, & \text{if } s_3=t_3, \quad s_2=t_6\\
T\max_{t_5}\left(\sum_{t_2}|w_{T,it_5}(\tau_{t_2})|\right)\sum_{t_3}|w_{T,it_3}(\tau_{t_1})|\sum_{t_6}|w_{T,it_6}(\tau_{t_2})|, & \text{if } s_3=t_4, \quad s_2=t_5\\
T\max_{t_6}\left(\sum_{t_2}|w_{T,it_6}(\tau_{t_2})|\right)\sum_{t_3}|w_{T,it_3}(\tau_{t_1})|\sum_{t_5}|w_{T,it_5}(\tau_{t_2})|, & \text{if } s_3=t_4, \quad s_2=t_6\\
T\max_{t_6}\left(\sum_{t_2}|w_{T,it_6}(\tau_{t_2})|\right)\sum_{t_3}|w_{T,it_3}(\tau_{t_1})|\sum_{t_4}|w_{T,it_4}(\tau_{t_1})|, & \text{if } s_3=t_5, \quad s_2=t_6
\end{cases}\\
&\leq CT.
\end{aligned} \tag{E.27}$$

By symmetry, the remaining cases, such as $s_3 = t_4$ with $s_2 = t_3$, can easily be verified. Together with the conditions assumed on the mixing coefficients in Assumption 1, we have

$$\frac{C}{T^3 h_i} \max_{1\leq s_2 \leq T} \left(\sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \sum_{s_1=1}^{T} \sum_{s_4=1}^{T} |C_{w,-s_3}(t_1,\cdots,t_6)| \right) \sum_{k=1}^{T} k\left(\alpha(k)\right)^{\frac{\delta}{4+\delta}} \leq \frac{C}{T^2 h_i}.$$

On the other hand, if $s_2 - s_1 = d_3$. In this case we can apply the above procedure regarding $s_3 - s_2 \geq d_2$, which gives the same results. Thus we have

$$\mathrm{E}\,||B_{Ti,3}||^2 \leq \frac{C}{T^2 h_i} \leq \frac{C}{T^2 h}. \tag{E.28}$$

Note that

$$\mathrm{E}\,||B_{Ti,4}||^2 \leq \frac{C}{T^2} \sum_{t_1=1}^{T} \cdots \sum_{t_4=1}^{T} C_w(t_1,\cdots,t_4) \mathrm{tr}\left(\mathrm{E}\left(v_{it_1}v_{it_2}^\top v_{it_3}v_{it_4}^\top\right)\right), \tag{E.29}$$

where $C_w(t_1,\cdots,t_4) = w_{T,it_3}(\tau_{t_1})w_{T,it_4}(\tau_{t_2})$. Here we only consider the case that t_1 to t_4 are distinct. Assume $1 \leq s_1 < \cdots < s_4 \leq T$ is the permutation of $t_1,\cdots,t_4$ in ascending order with d_c be the c-th largest different between $s_{j+1} - s_j$ for $j = 1,\cdots,3$. Similar to $B_{Ti,1}$, we first consider the case of $s_2 - s_1 \geq d_2$, assuming $s_3 - s_2 < d_2$,

$$\begin{aligned}
\mathrm{E}\,||B_{Ti,4}||^2 &\leq \frac{C}{T^2} \sum_{s_1=1}^{T} \sum_{s_2=s_1+1}^{T} \sum_{s_3-s_2\leq d_2} \sum_{s_4=s_3+1}^{T} C_w(t_1,\cdots,t_4) \left|\left|\mathrm{E}\left(v_{is_1}v_{is_2}^\top v_{is_3}v_{is_4}^\top\right)\right|\right| \\
&\leq \frac{C}{T^2} \sum_{s_1=1}^{T} \sum_{s_2=s_1+1}^{T} \sum_{s_3-s_2\leq d_2} \sum_{s_4=s_3+1}^{T} |C_w(t_1,\cdots,t_4)| \left(\alpha(s_2-s_1)\right)^{\frac{\delta}{4+\delta}} \left(\mathrm{E}\,||v_{i,1}||^{4+\delta}\right)^{\frac{4}{4+\delta}} \\
&\leq \frac{C}{T^2} \max_{1\leq s_2\leq s_3\leq T} \left(\sum_{s_1} \sum_{s_4} |C_w(t_1,\cdots,t_4)| \right) \sum_{k=1}^{T} k\left(\alpha(k)\right)^{\frac{\delta}{4+\delta}}.
\end{aligned} \tag{E.30}$$

In this case, by Lemma E.1(ii) and (iii), we have

$$\max_{1\leq s_2\leq s_3\leq T} \left(\sum_{s_1} \sum_{s_4} |w_{T,it_3}(\tau_{t_1})w_{T,it_4}(\tau_{t_2})| \right)$$

$$
\leq \begin{cases} \max_{t_3}\left(\sum_{t_1}|w_{T,it_3}(\tau_{t_1})|\right)\max_{t_4}\left(\sum_{t_2}|w_{T,it_4}(\tau_{t_2})|\right), & \text{if } s_1=t_1, s_4=t_2 \\ \frac{C}{h}\max_{t_3}\left(\sum_{t_1}|w_{T,it_3}(\tau_{t_1})|\right), \quad \text{if } s_1=t_1, s_4=t_3 & \\ \max_{t_3}\left(\sum_{t_1}|w_{T,it_3}(\tau_{t_1})|\right)\max_{t_2}\left(\sum_{t_3}|w_{T,it_4}(\tau_{t_2})|\right), & \text{if } s_1=t_1, s_4=t_4 \\ \max_{t_1}\left(\sum_{t_3}|w_{T,it_3}(\tau_{t_1})|\right)\max_{t_4}\left(\sum_{t_2}|w_{T,it_4}(\tau_{t_2})|\right), & \text{if } s_1=t_2, s_4=t_3 \\ \frac{C}{h}\max_{t_4}\left(\sum_{t_2}|w_{T,it_4}(\tau_{t_2})|\right), \quad \text{if } s_1=t_2, s_4=t_4 & \\ \max_{t_1}\left(\sum_{t_3}|w_{T,it_3}(\tau_{t_1})|\right)\max_{t_2}\left(\sum_{t_4}|w_{T,it_4}(\tau_{t_2})|\right), & \text{if } s_1=t_3, s_4=t_4 \end{cases}
$$

$$
\leq \frac{C}{h}. \tag{E.31}
$$

Then by symmetry again, other cases, such as $s_1 = t_2$ and $s_4 = t_1$, are included in the results above. Together with conditions on the mixing coefficients in Assumption 1, we have

$$
\frac{C}{T^2}\max_{1\leq s_2\leq s_3\leq T}\left(\sum_{s_1}\sum_{s_4}|C_w(t_1,\cdots,t_4)|\right)\sum_{k=1}^{T}k\left(\alpha(k)\right)^{\frac{\delta}{4+\delta}} \leq \frac{C}{T^2h}. \tag{E.32}
$$

In addition, if $s_2 - s_1 = d_3$, the same result can be verified by using $s_3 - s_2 \geq d_2$. Thus

$$
\mathrm{E}\,||B_{Ti,4}||^2 \leq \frac{C}{T^2h}, \tag{E.33}
$$

$$
\mathrm{E}\,||B_{Ti,5}||^2 \leq \frac{C}{T^2h}. \tag{E.34}
$$

Then by Eqs. (E.68) to Eqs. (E.34) and Cauchy-Schwarz inequality, we have proved

$$
\mathrm{E}\left|\left|\frac{1}{T}\sum_{t=1}^{T}\eta_{it,1}^{c}\right|\right|^2 \leq \frac{C}{T^2h}. \tag{E.35}
$$

The second part of Lemma E.2(iii) can be verified in the same way.

Proof of Lemma E.2(iv).

Note that

$$
\begin{aligned}
\Omega_{v,i}\frac{1}{T}\sum_{t=1}^{T}\eta_{it,2}^{c} &= \frac{1}{T}\sum_{t=1}^{T}\widetilde{g}_{it}\widetilde{g}_{it}^{\top} - \frac{1}{T}\sum_{t=1}^{T}\widetilde{g}_{it}s(\tau_t)v_i \\
&\quad -\frac{1}{T}\sum_{t=1}^{T}v_i^{\top}s^{\top}(\tau_t)\widetilde{g}_{it}^{\top} + \frac{1}{T}\sum_{t=1}^{T}\widetilde{g}_{it}v_{it}^{\top} + \frac{1}{T}\sum_{t=1}^{T}v_{it}\widetilde{g}_{it}^{\top} \\
&\equiv B_{Ti,6} + \cdots + B_{Ti,10}.
\end{aligned} \tag{E.36}
$$

Using the results as in Lemma E.1(i), we immediately have

$$
B_{Ti,6} \leq Ch^4. \tag{E.37}
$$

For $B_{Ti,5}$, recall that $s_i(\tau)v_i = \sum_{t=1}^{T} w_{T,it}(\tau)v_{it}^{\top}$, then

$$
\mathrm{E}\,||B_{Ti,7}||^2 = \frac{1}{T^2}\sum_{t_1=1}^{T}\cdots\sum_{t_4=1}^{T}w_{T,it_2}(\tau_{t_1})w_{T,it_4}(\tau_{t_3})\mathrm{tr}\left(\widetilde{g}_{it_1}\mathrm{E}\left(v_{it_2}^{\top}v_{it_4}\right)\widetilde{g}_{it_3}^{\top}\right)
$$

$$= \frac{1}{T^2}\sum_{t_2=1}^{T}\sum_{t_4=1}^{T}\left(\max_{t_2}\left(\sum_{t_1=1}^{T}|w_{T,it_2}(\tau_{t_1})|\right)\right)^2 \max_{1\le t\le T}||\widetilde{g}_{it}||^2\left|\mathrm{E}\left(v_{it_2}^\top v_{it_4}\right)\right|$$
$$\le \frac{Ch^4}{T^2}\sum_{t_2=1}^{T}\sum_{t_4=1}^{T}(\alpha(|t_2-t_4|))^{\frac{\delta}{2+\delta}}\left(\mathrm{E}\,||v_{i1}||^{2+\delta}\right)^{\frac{2}{2+\delta}} \le \frac{Ch^4}{T}. \tag{E.38}$$

By the same calculation, we have

$$\mathrm{E}\,||B_{Ti,8}||^2 \le \frac{C_1h^4}{T}, \tag{E.39}$$
$$\mathrm{E}\,||B_{Ti,9}||^2 \le \frac{C_1h^4}{T}, \tag{E.40}$$
$$\mathrm{E}\,||B_{Ti,10}||^2 \le \frac{C_1h^4}{T}. \tag{E.41}$$

Similarly, the second half of Lemma E.2(iv) can be verified in the same way.

Proof of Lemma E.2(v).

By Lemma E.2(ii) to Lemma E.2(iv) and the definition of Δ_{Ti} and $\Delta_{Ti,2}$, together with Cauchy-Schwarz inequality, it is straightforward to see

$$\mathrm{E}\,||\Delta_{Ti}||^2 \le \frac{C}{T}, \mathrm{E}\,||\Delta_{Ti,2}||^2 \le \frac{C}{T^2h}. \tag{E.42}$$

Next, recall $\widetilde{Q}_{Ti} = \left(\widehat{\Omega}_{x,Ti}^{-1}\Omega_{v,i} - I_d\right) I\left(||Q_{Ti}|| \le \mathcal{M}\right)$, then by standard calculation

$$\begin{aligned}
||\widetilde{Q}_{Ti}|| &= \left|\left|\left(\widehat{\Omega}_{x,Ti}^{-1}\Omega_{v,i} - I_d\right) I\left(||Q_{Ti}|| \le \mathcal{M}\right)\right|\right| \\
&= \left|\left|\widehat{\Omega}_{x,Ti}^{-1}\left(\widehat{\Omega}_{x,Ti} - \Omega_{v,i}\right) I\left(||Q_{Ti}|| \le \mathcal{M}\right)\right|\right| \\
&= \left|\left|\left(\Omega_{v,i}^{-1} + Q_{Ti}\right)\Omega_{v,i}\Delta_{T,i} I\left(||Q_{Ti}|| \le \mathcal{M}\right)\right|\right| \\
&\le ||\Delta_{T,i}|| + \left|\left|\widetilde{Q}_{Ti}\Omega_{v,i}\Delta_{T,i}\right|\right| \le C\,||\Delta_{T,i}||\,.
\end{aligned} \tag{E.43}$$

Thus we have

$$\mathrm{E}\left|\left|\widetilde{Q}_{Ti}\right|\right|^2 \le \frac{C}{T}. \tag{E.44}$$

Since $\left|\left|\widetilde{Q}_{Ti}\right|\right|$ is bounded, it is straightforward to see that

$$\mathrm{E}\left|\left|\widetilde{Q}_{Ti}\right|\right|^4 \le \frac{C}{T}. \tag{E.45}$$

The last part of Lemma E.2(v) can be verified by the same way of Δ_{Ti}.

Proof of Lemma E.2(vi)

$$\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(\eta_{it_1,1}^c\eta_{jt_2,1}^{c}{}^\top\right)\right|\right| \le \sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_i^\top s_i^\top(\tau_{t_1})s_i(\tau_{t_1})v_i v_j^\top s_j^\top(\tau_{t_2})s_j(\tau_{t_2})v_j\right)\right|\right|$$

$$
\begin{aligned}
&+\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_i^\top s_i^\top(\tau_{t_1})v_{it_1}^\top v_j^\top s_j^\top(\tau_{t_2})s_j(\tau_{t_2})v_j\right)\right|\right|+\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_{it_1}s_i(\tau_{t_1})v_iv_j^\top s_j^\top(\tau_{t_2})s_j(\tau_{t_2})v_j\right)\right|\right|\\
&+\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_i^\top s_i^\top(\tau_{t_1})s_i(\tau_{t_1})v_iv_{jt_2}s_j(\tau_{t_2})v_j\right)\right|\right|+\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_i^\top s_i^\top(\tau_{t_1})v_{it_1}^\top v_{jt_2}s_j(\tau_{t_2})v_j\right)\right|\right|\\
&+\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}||\mathrm{E}\left(v_{it_1}s_i(\tau_{t_1})v_iv_{jt_2}s_j(\tau_{t_2})v_j\right)||+\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_i^\top s_i^\top(\tau_{t_1})s_i(\tau_{t_1})v_iv_j^\top s_j^\top(\tau_{t_2})v_{jt_2}^\top\right)\right|\right|\\
&+\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_i^\top s_i^\top(\tau_{t_1})v_{it_1}^\top v_i^\top s_i^\top(\tau_{t_2})v_{it_2}^\top\right)\right|\right|+\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_{it_1}s_i(\tau_{t_1})v_iv_i^\top s_i^\top(\tau_{t_2})v_{it_2}^\top\right)\right|\right|\\
&\equiv B_{Ti,11}+\cdots+B_{Ti,19}.
\end{aligned}\tag{E.46}
$$

By similar calculations to the proof of Lemma E.2(iii), we have

$$
B_{Ti,11}=\frac{C}{T^2}\sum_{t_1=1}^{T}\cdots\sum_{t_6=1}^{T}C_w(t_1,\cdots,t_6)\left|\left|\mathrm{E}\left(v_{it_3}v_{it_4}^\top v_{it_5}v_{it_6}^\top\right)\right|\right|,\tag{E.47}
$$

where $C_w(t_1,\cdots,t_6)=w_{T,it_3}(\tau_{t_1})w_{T,it_4}(\tau_{t_1})w_{T,it_5}(\tau_{t_2})w_{T,it_6}(\tau_{t_2})$. Since the structure in Eqs. (E.47) is identical to the formula in the proof for $B_{Ti,1}$ (except for the multiplier $\frac{1}{T}$ at its beginning), then by the same argument, we directly have

$$
B_{Ti,11}\leq\frac{C}{h}.\tag{E.48}
$$

Thus, using the same arguments as in proving $B_{Ti,3}$ to $B_{Ti,5}$, it is straightforward to verify the results for $B_{Ti,12}$ to $B_{Ti,19}$, which suffice to prove the first part of Eqs. (E.2)(vi). Note that $\xi_{it,1}^c$ shares the same structure as with $\eta_{it,1}^c$, then the second part of Lemma E.2(vi) can be verified by the same calculations.

Proof of Lemma E.2(vii)

To verify the results in Lemma E.2(vii), we need to check the forth moments of $B_{Ti,1}$ to $B_{Ti,8}$. Such calculations will be easier since the results in Lemma E.2(vii) are not as tight as Lemma E.2(iii) and Lemma E.2(v), and also some new assumptions on the moment and mixing coefficient are imposed. Take $B_{Ti,1}$ for an example, we have

$$
\begin{aligned}
\mathrm{E}\,||B_{Ti,3}||^4&\leq\frac{1}{T^4}\sum_{t_1=1}^{T}\cdots\sum_{t_{10}=1}^{T}w_{T,it_3}(\tau_{t_1})\cdots w_{T,it_6}(\tau_{t_1})\\
&\quad\times w_{T,it_7}(\tau_{t_2})\cdots w_{T,it_{10}}(\tau_{t_2})\left|\left|\mathrm{E}\left(v_{it_3}v_{it_4}^\top\cdots v_{it_9}v_{it_{10}}^\top\right)\right|\right|\\
&=\frac{1}{T^{10}h^6}\sum_{t_3=1}^{T}\cdots\sum_{t_{10}=1}^{T}\left|\sum_{t_1=1}^{T}w_{T,it_3}(\tau_{t_1})\right|\left|\sum_{t_1=1}^{T}w_{T,it_7}(\tau_{t_2})\right|\left|\left|\mathrm{E}\left(v_{it_3}v_{it_4}^\top\cdots v_{it_9}v_{it_{10}}^\top\right)\right|\right|\\
&=\frac{1}{T^{10}h^6}\sum_{t_3=1}^{T}\cdots\sum_{t_{10}=1}^{T}\left|\left|\mathrm{E}\left(v_{it_3}v_{it_4}^\top\cdots v_{it_9}v_{it_{10}}^\top\right)\right|\right|\leq\frac{C}{T^6h^6},
\end{aligned}\tag{E.49}
$$

where the last step follows from the procedure of Lemma A.2 in Gao (2007) using the restrictions on the mixing coefficients in Assumption 6. The fourth moment results for the rest of $B_{Ti,2}$ to $B_{Ti,8}$ can be verified in a similar way, which suffice to prove Lemma E.2(vii).

Proof of Lemma E.3(i).

By the same calculations as in Eqs. (E.4), we have

$$\sum_{p=1}^{\kappa_T}\sum_{t_1=pl_T+(p-1)s_T+1}^{p(l_T+s_T)} w_{T,it_1}(\tau)=\left(S_{i,0}(\tau)S_{i,2}(\tau)-(S_{i,1}(\tau))^2\right)^{-1}$$
$$\times\frac{1}{Th_i}\sum_{p=1}^{\kappa_T}\sum_{t_1=pl_T+(p-1)s_T+1}^{p(l_T+s_T)}\left(S_{i,2}(\tau)K_{it,0}(\tau)-S_{i,1}(\tau)K_{it,1}(\tau)\right). \tag{E.50}$$

In the proof of Lemma E.1(ii), we have shown that $S_{i,l}(\tau)$ and $\left(S_{i,0}(\tau)S_{i,2}(\tau)-(S_{i,1}(\tau))^2\right)^{-1}$ are bounded uniformly, then we only need to show

$$\frac{1}{Th_i}\sum_{p=1}^{\kappa_T}\sum_{t=pl_T+(p-1)s_T+1}^{p(l_T+s_T)}|K_{ti,j}(\tau)|\leq\frac{C\kappa_T s_T}{T}. \tag{E.51}$$

Note that $K(\cdot)$ is a density function with compact support, which indicates that $|K_{ti,j}(\cdot)|$ is bounded. Thus

$$\sum_{t=pl_T+(p-1)s_T+1}^{p(l_T+s_T)}|K_{ti,j}(\tau)|\leq Cs_T \tag{E.52}$$

for all p. Note that the support of $K(\cdot)$ is bounded, thus $K(\cdot)$ is nonzero only for $t\in[\tau-CTh,\tau+CTh]$, for some $C<\infty$ depending on the type of kernel function. Meanwhile, the κ_T small block sets distributed uniformly on $\{1,\cdots,T\}$, then $K(\cdot)$ is nonzero in at most $\frac{\kappa_T}{T}\times 2CTh+1$ small block sets, therefore,

$$\frac{1}{Th_i}\sum_{p=1}^{\kappa_T}\sum_{t=pl_T+(p-1)s_T+1}^{p(l_T+s_T)}|K_{t,i,j}(\tau)|\leq\frac{1}{Th_i}\left(\frac{\kappa_T}{T}\times 2CTh+1\right)\sum_{t=pl_T+(p-1)s_T+1}^{p(l_T+s_T)}|K_{t,i,j}(\tau)|$$
$$\leq\frac{C\kappa_T s_T}{T}, \tag{E.53}$$

which proves Eqs. (E.51).

Proof of Lemma E.3(ii).

Recall that $V_{N,t}=\frac{\sqrt{Th}}{\sqrt{N}}\sum_{i=1}^{N}w_{T,it}(\tau)e_{it}$, and $\widetilde{V}_{N,p}=\sum_{t=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T}V_{N,t}$, then by standard calculation, we have

$$\sum_{p=1}^{\kappa_T}\mathrm{E}\left(\widetilde{V}_{N,p}^2 I\left(|\widetilde{\zeta}_p|\geq\varepsilon\right)\right)\leq\sum_{p=1}^{\kappa_T}\left(\mathrm{E}\left(\widetilde{V}_{N,p}^4\right)\right)^{1/2}\left(\mathrm{E}\left(I\left(|\widetilde{\zeta}_p|\geq\varepsilon\right)\right)^2\right)^{1/2}$$
$$\leq\sum_{p=1}^{\kappa_T}\left(\mathrm{E}\left(\widetilde{V}_{N,p}^4\right)\right)^{1/2}\left(\frac{\mathrm{E}\left(\widetilde{V}_{N,p}^4\right)}{\varepsilon^4}\right)^{1/2}$$
$$=\frac{\varepsilon^{-4}T^2h^2}{N^2}\sum_{p=1}^{\kappa_T}\mathrm{E}\left(\sum_{t=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T}\sum_{i=1}^{N}w_{Ti,t}(\tau)e_{it}\right)^4.$$

By the same argument as used in the proof Lemma E.3(i), there are at most $\left(2CTh\frac{\kappa_T}{T}+1\right)$ large block sets where $K(\cdot)$ is nonzero, together with Lemma E.1(ii) and Assumption 3(iii), we have

$$\sum_{p=1}^{\kappa_T} \mathrm{E}\left(\widetilde{V}_{N,p}^2 I\left(|\widetilde{\zeta}_p| \geq \varepsilon\right)\right) \leq \frac{\varepsilon^{-4}}{N^2T^2h^2}\left(2CTh\frac{\kappa_T}{T}+1\right)\sum_{t_1=1}^{l_T}\cdots\sum_{t_4=1}^{l_T}\sum_{i_1=1}^{N}\cdots\sum_{i_4=1}^{N}\left|\mathrm{E}\left(e_{i_1t_1}e_{i_2t_2}e_{i_3t_3}e_{i_4t_4}\right)\right|$$
$$\leq \frac{C\kappa_T}{N^2T^2h}N^2l_T^2 \leq \frac{Cl_T}{Th}\frac{\kappa_T l_T}{T}. \tag{E.54}$$

By the choice of $\frac{\kappa_T l_T}{T} < \infty$, thus we complete the proof of Lemma E.3(ii).

Proof of Lemma E.3(iii).

$$\sum_{p=1}^{\kappa_T}\mathrm{E}\left(\widetilde{V}_{N,p}^2\right) = \frac{Th}{N}\sum_{p=1}^{\kappa_T}\sum_{t_1=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T}\sum_{t_2=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T}\sum_{i=1}^{N}\sum_{j=1}^{N}w_{T,it_1}(\tau)w_{T,jt_2}(\tau)\mathrm{E}(e_{it_1}e_{jt_2})$$
$$= \frac{Th}{N}\sum_{p=1}^{\kappa_T}\sum_{t_1=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T}\sum_{t_2=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T}\sum_{i=1}^{N}\sum_{j=1}^{N}w_{T,it_1}(\tau)w_{T,jt_1}(\tau)\mathrm{E}(e_{it_1}e_{jt_2})$$
$$+\frac{Th}{N}\sum_{p=1}^{\kappa_T}\sum_{t_1=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T}\sum_{t_2=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T}\sum_{i=1}^{N}\sum_{j=1}^{N}w_{T,it_1}(\tau)\left(w_{T,jt_2}(\tau)-w_{T,jt_1}(\tau)\right)\mathrm{E}(e_{it_1}e_{jt_2})$$
$$\equiv B_{NT,1}+B_{NT,2}. \tag{E.55}$$

By Assumption 2, we have

$$B_{NT,1} = \frac{Th}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\sum_{p=1}^{\kappa_T}\sum_{t_1=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T}w_{T,it_1}(\tau)w_{T,jt_1}(\tau)\right)\left(\sum_{k=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T}\mathrm{E}(e_{i,1}e_{j,1+k})\right)$$
$$\to \sigma_{eK}^2. \tag{E.56}$$

For $B_{NT,2}$, note that within every large block set, $|t_1-t_2| \leq l_T$, and we have $l_T = o(Th)$, then by similar calculations in the proof of Lemma E.1, it is easy to see that

$$|w_{T,jt_2}(\tau)-w_{T,jt_1}(\tau)| \leq \frac{Cl_T}{T^2h^2} \tag{E.57}$$

for some finite C. Thus by the same argument in the proof of Lemma E.3(ii), we have

$$B_{NT,2} \leq \frac{Cl_T}{NTh}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\sum_{p=1}^{\kappa_T}\sum_{t_1=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T}|w_{T,it_1}(\tau)|\right)\left(\sum_{k=(p-1)(l_T+s_T)+1}^{pl_T+(p-1)s_T}\mathrm{E}(e_{i,1}e_{j,1+k})\right)$$
$$\leq O\left(\frac{Cl_T}{Th}\right), \tag{E.58}$$

which completes the proof of Lemma E.3(iii).

Proof of Lemma E.4(i).
Recall that $s_i(\tau)v_i = \sum_{t=1}^{T}w_{T,it}(\tau)v_{it}^\top$, using the moment condition in Assumption 2 and the results in Lemma E.1, we have

$$\mathrm{E}\left\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}s_i(\tau)v_i\xi_{it}\right\|^2 = \frac{1}{N^2T^2}\mathrm{E}\left|\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}w_{Ti,t_2}(\tau)v_{it_2}^\top\xi_{it_1}\right|^2$$

$$
\begin{aligned}
&= \frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\cdots\sum_{t_4=1}^{T} w_{T,it_1}(\tau)w_{T,it_3}(\tau)\mathrm{E}\left(v_{it_1}^\top \xi_{it_2} v_{jt_3}^\top \xi_{jt_4}\right) \\
&= \frac{1}{N^2T^4h^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\cdots\sum_{t_4=1}^{T} \left|\mathrm{E}\left(v_{it_1}^\top \xi_{it_2} v_{jt_3}^\top \xi_{jt_4}\right)\right| \\
&= O\left(\frac{1}{NT^2h^2}\right).
\end{aligned} \tag{E.59}
$$

Similarly,

$$
\begin{aligned}
\mathrm{E}\left|\left|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} v_{i\cdot}^\top \xi_{it}\right|\right|^2 &= \frac{1}{N^2T^4}\mathrm{E}\left|\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T} v_{it_2}^\top \xi_{it_1}\right|^2 \\
= \frac{1}{N^2T^4}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\cdots\sum_{t_4=1}^{T} &\left|\mathrm{E}\left(v_{it_1}^\top \xi_{it_2} v_{jt_3}^\top \xi_{jt_4}\right)\right| = O\left(\frac{1}{NT^2}\right),
\end{aligned} \tag{E.60}
$$

which completes the proof of Lemma E.4(ii).

Proof of Lemma E.4(ii).

To simplify the notation, we let $C_g = s_i(\tau)g_i - g_{i\cdot}^\top$. Note that C_g is deterministic and bounded in absolute value. Thus

$$
\begin{aligned}
\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} C_g \xi_{it}^c &= -\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} C_g \Omega_{v,i}^{-1} v_{it} s_i(\tau_t) e_i - \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} C_g \Omega_{v,i}^{-1} v_{i}^\top s_i^\top(\tau_t) e_{it} \\
&+ \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} C_g \Omega_{v,i}^{-1} v_{i}^\top s_i^\top(\tau_t) s_i(\tau_t) e_i + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} C_g \Omega_{v,i}^{-1} \widetilde{g}_{it} e_{it} \\
&- \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} C_g \Omega_{v,i}^{-1} \widetilde{g}_{it} s_i(\tau_t) e_i + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} C_g \Omega_{v,i}^{-1} v_{it} \widetilde{f}_{it} \\
&- \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} C_g \Omega_{v,i}^{-1} v_{i}^\top s_i^\top(\tau_t) \widetilde{f}_{it} + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} C_g \Omega_{v,i}^{-1} \widetilde{g}_{it} \widetilde{f}_{it} \\
&\equiv B_{NT,3} + \cdots + B_{NT,10}.
\end{aligned} \tag{E.61}
$$

Using the fact that $K(\cdot)$ is a density function with compact support, thus $K\left(\frac{\tau_t-\tau_s}{h}\right)$ will be zero for $|\tau_t - \tau_s| > ch$ for some $c > 0$, then by Assumption 6, together with Lemma E.1(ii) we have,

$$
\begin{aligned}
\mathrm{E}\left|\left|B_{NT,3}\right|\right|^2 &\leq \frac{C}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T} |w_{T,it_3}(\tau_{t_1})w_{T,jt_4}(\tau_{t_2})| \left|\left|\mathrm{E}\left(v_{it_1} v_{jt_2}^\top e_{it_3} e_{jt_4}\right)\right|\right| \\
&\leq \frac{C}{N^2T^4h^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{|t_3-t_1|\leq cTh}\sum_{|t_4-t_2|\leq cTh} \left|\left|\mathrm{E}\left(v_{it_1} v_{jt_2}^\top e_{it_3} e_{jt_4}\right)\right|\right| \\
&\leq \frac{C}{NT^2h}.
\end{aligned} \tag{E.62}
$$

By similar calculation, we can show that

$$
\mathrm{E}\left|\left|B_{NT,4}\right|\right|^2 \leq \frac{C}{NT^2h}. \tag{E.63}
$$

For $B_{NT,5}$, note $K\left(\frac{\tau_t-\tau_s}{h}\right)$ is zero for $|\tau_t-\tau_s|>ch$ for some $c>0$, then note that $w_{T,it_3}(\tau_{t_1})w_{T,it_4}(\tau_{t_1})$ will also be zero if $|\tau_{t_3}-\tau_{t_4}|>2ch$, because when t_3 moves too far away from t_4, one of the weights will be zero. Thus we have,

$$\begin{aligned}
\mathrm{E}\,||B_{NT,5}||^2 &\leq \frac{C}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\cdots\sum_{t_6=1}^{T}|w_{T,it_3}(\tau_{t_1})w_{T,it_5}(\tau_{t_2})| \\
&\quad\times |w_{T,it_4}(\tau_{t_1})w_{T,jt_6}(\tau_{t_2})|\left|\left|\mathrm{E}\left(e_{it_4}e_{jt_6}v_{it_3}v_{jt_5}^\top\right)\right|\right| \\
&\leq \frac{C}{N^2T^4h^2}\max_{it_3}\left(\sum_{t_1=1}^{T}|w_{T,it_3}(\tau_{t_1})|\right)\max_{jt_5}\left(\sum_{t_2=1}^{T}|w_{T,it_5}(\tau_{t_2})|\right)\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_3=1}^{T}\sum_{t_5=1}^{T} \\
&\quad\times\sum_{|t_4-t_3|\leq 2cTh}\ \sum_{|t_6-t_5|\leq 2cTh}\left|\left|\mathrm{E}\left(e_{it_4}e_{jt_6}v_{it_3}v_{jt_5}^\top\right)\right|\right| \leq \frac{C}{NT^2h}
\end{aligned}\tag{E.64}$$

On the other hand, we have

$$\begin{aligned}
\mathrm{E}\,||B_{NT,6}||^2 &\leq \frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\widetilde{g}_{it_1}^\top\widetilde{g}_{jt_2}\right|\left|\mathrm{E}\left(e_{it_1}e_{jt_2}\right)\right| \\
&\leq \frac{1}{N^2T^2}\max_{it}||\widetilde{g}_{it}||^2\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\mathrm{E}\left(e_{it_1}e_{jt_2}\right)\right| \leq \frac{C_1h^4}{NT}\leq\frac{C_2}{NT^2h},
\end{aligned}\tag{E.65}$$

where we have used $Th^4\to 0$. Similarly,

$$\begin{aligned}
\mathrm{E}\,||B_{NT,7}||^2 &= \frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T}\left|\widetilde{g}_{it_1}^\top\widetilde{g}_{jt_2}\right||w_{T,it_3}(\tau_{t_1})w_{T,it_4}(\tau_{t_2})|\left|\mathrm{E}\left(e_{it_3}e_{jt_4}\right)\right| \\
&\leq \frac{1}{N^2T^2}\max_{it}||\widetilde{g}_{it}||^2\max_{it_4}\left(\sum_{t_2=1}^{T}|w_{it_4}(\tau_{t_2})|\right)^2\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T}\left|\mathrm{E}\left(e_{it_3}e_{jt_4}\right)\right| \\
&\leq \frac{C_1h^4}{NT}\leq\frac{C_2}{NT^2h}.
\end{aligned}\tag{E.66}$$

Note that $B_{NT,8}$ and $B_{NT,9}$ are similar to $B_{NT,6}$ and $B_{NT,7}$ by replacing $\widetilde{g}_{it}$ with $\widetilde{f}_{it}$, and by Lemma E.1, we have

$$\mathrm{E}\,||B_{NT,8}||^2\leq\frac{C}{NT^2h},\quad \mathrm{E}\,||B_{NT,9}||^2\leq\frac{C}{NT^2h},\quad ||B_{NT,10}||^2\leq\frac{C}{NT^2h},\tag{E.67}$$

which prove the first part of Lemma E.4(iii). Due to the similar structure between η_{it}^c and ξ_{it}^c, the second part of Lemma E.4(iii) can be verified by the same calculation.

Proof of Lemma E.4(iii) Observe that

$$\begin{aligned}
&\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(\eta_{it_1,1}^c\eta_{jt_2,1}^{c\top}\right)\right|\right| \leq \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_i^\top s_i^\top(\tau_{t_1})s_i(\tau_{t_1})v_iv_j^\top s_j^\top(\tau_{t_2})s_j(\tau_{t_2})v_j\right)\right|\right| \\
&+\sum_{i=1}^{N}\cdots\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_i^\top s_i^\top(\tau_{t_1})v_{it_1}^\top v_j^\top s_j^\top(\tau_{t_2})s_j(\tau_{t_2})v_j\right)\right|\right| + \sum_{i=1}^{N}\cdots\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_{it_1}s_i(\tau_{t_1})v_iv_j^\top s_j^\top(\tau_{t_2})s_j(\tau_{t_2})v_j\right)\right|\right| \\
&+\sum_{i=1}^{N}\cdots\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_i^\top s_i^\top(\tau_{t_1})s_i(\tau_{t_1})v_iv_{jt_2}s_j(\tau_{t_2})v_j\right)\right|\right| + \sum_{i=1}^{N}\cdots\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_i^\top s_i^\top(\tau_{t_1})v_{it_1}^\top v_{jt_2}s_j(\tau_{t_2})v_j\right)\right|\right|
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=1}^{N}\cdots\sum_{t_2=1}^{T}||\mathrm{E}\left(v_{it_1}s_i(\tau_{t_1})v_iv_{jt_2}s_j(\tau_{t_2})v_j\right)|| + \sum_{i=1}^{N}\cdots\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_i^\top s_i^\top(\tau_{t_1})s_i(\tau_{t_1})v_iv_j^\top s_j^\top(\tau_{t_2})v_{jt_2}^\top\right)\right|\right| \\
&+\sum_{i=1}^{N}\cdots\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_i^\top s_i^\top(\tau_{t_1})v_{it_1}^\top v_i^\top s_i^\top(\tau_{t_2})v_{it_2}^\top\right)\right|\right| + \sum_{i=1}^{N}\cdots\sum_{t_2=1}^{T}\left|\left|\mathrm{E}\left(v_{it_1}s_i(\tau_{t_1})v_iv_i^\top s_i^\top(\tau_{t_2})v_{it_2}^\top\right)\right|\right| \\
&\equiv B_{NT,11}+\cdots+B_{NT,19}. \qquad \text{(E.68)}
\end{aligned}
$$

Since the components from $B_{NT,11}$ to $B_{NT,19}$ are similar, we give the calculation for $B_{NT,11}$ as an example. Note that with Assumption 6,

$$
\begin{aligned}
&\mathrm{E}\,||B_{NT,11}||^2 \le \frac{C}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\cdots\sum_{t_6=1}^{T}|w_{T,it_3}(\tau_{t_1})w_{T,it_4}(\tau_{t_1})| \\
&\times|w_{T,jt_5}(\tau_{t_2})w_{T,jt_6}(\tau_{t_2})|\left|\left|\mathrm{E}\left(v_{it_3}v_{it_4}^\top v_{jt_5}v_{jt_6}^\top\right)\right|\right| \\
&\le \frac{C}{N^2T^4h^2}\max_{it_3}\left(\sum_{t_1=1}^{T}|w_{T,it_3}(\tau_{t_1})|\right)\max_{jt_5}\left(\sum_{t_2=1}^{T}|w_{T,it_5}(\tau_{t_2})|\right)\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_3=1}^{T}\sum_{t_5=1}^{T} \\
&\times\sum_{|t_4-t_3|\le 2cTh}\ \sum_{|t_6-t_5|\le 2cTh}\left|\left|\mathrm{E}\left(v_{it_3}v_{it_4}^\top v_{jt_5}v_{jt_6}^\top\right)\right|\right| \le \frac{C}{NT^2h} \qquad \text{(E.69)}
\end{aligned}
$$

Similarly, the results for $B_{NT,11}$ to $B_{NT,19}$ can be verified by the same calculations. Then it suffices to prove the first part of Lemma E.4(iii). For the second part of Lemma E.4(iii), noting the similar structure of $\xi_{it,1}^c$ and $\eta_{it,1}^c$, it can be verified by the same calculations together with Assumption 6.

Proof of Lemma E.5(i).

Note that $\widetilde{\eta}_{it}^* = \widetilde{\eta}_{it} - \Delta_{Ti}$, we have

$$
\begin{aligned}
&\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}T}\sum_{i=1}^{N}\sum_{t_1=1}^{[rT]}\sum_{t_2=1}^{T}\left(\widehat{\Omega}_{x,Ti}-\widehat{\Omega}_{x,NT}\right)\left(I_d+\widetilde{Q}_{Ti}\right)\widetilde{\eta}_{it_1}^*\left(I_d+\widetilde{Q}_{Ti}\right)\widetilde{\xi}_{it_2} \\
&=-\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}T}\sum_{i=1}^{N}\sum_{t_1=1}^{[rT]}\sum_{t_2=1}^{T}\left(\widehat{\Omega}_{x,Ti}-\widehat{\Omega}_{x,NT}\right)\left(I_d+\widetilde{Q}_{Ti}\right)\Delta_{Ti}\left(I_d+\widetilde{Q}_{Ti}\right)\widetilde{\xi}_{it_2} \\
&\quad+\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}T}\sum_{i=1}^{N}\sum_{t_1=1}^{[rT]}\sum_{t_2=1}^{T}\left(\widehat{\Omega}_{x,Ti}-\widehat{\Omega}_{x,NT}\right)\left(I_d+\widetilde{Q}_{Ti}\right)\widetilde{\eta}_{it_1}\left(I_d+\widetilde{Q}_{Ti}\right)\widetilde{\xi}_{it_2} \\
&\equiv B_{NT,1}(r)+B_{NT,2}(r). \qquad \text{(E.70)}
\end{aligned}
$$

For $B_{NT,1}(r)$, recall that $\widehat{\Omega}_{x,Ti}-\widehat{\Omega}_{x,NT}=\widetilde{\Omega}_{v,i}+\Omega_{v,i}\Delta_{Ti}-\Delta_{NT}$, thus

$$
\begin{aligned}
B_{NT,1}(r) &= -\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T}\sum_{i=1}^{N}\sum_{t_2=1}^{T}\widetilde{\Omega}_{v,i}\left(I_d+\widetilde{Q}_{Ti}\right)\Delta_{Ti}\left(I_d+\widetilde{Q}_{Ti}\right)\widetilde{\xi}_{it_2} \\
&\quad-\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T}\sum_{i=1}^{N}\sum_{t_2=1}^{T}\Omega_{v,i}\Delta_{Ti}\left(I_d+\widetilde{Q}_{Ti}\right)\Delta_{Ti}\left(I_d+\widetilde{Q}_{Ti}\right)\widetilde{\xi}_{it_2} \\
&\quad+\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T}\Delta_{NT}\sum_{i=1}^{N}\sum_{t_2=1}^{T}\left(I_d+\widetilde{Q}_{Ti}\right)\Delta_{Ti}\left(I_d+\widetilde{Q}_{Ti}\right)\widetilde{\xi}_{it_2} \\
&= B_{NT,1,1}(r)+\cdots+B_{NT,1,3}(r). \qquad \text{(E.71)}
\end{aligned}
$$

Note that

$$
\begin{aligned}
B_{NT,1,1}(r) &= -\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T}\sum_{i=1}^{N}\sum_{t_2=1}^{T}\widetilde{\Omega}_{v,i}\left(I_d+\widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it_2} \\
&\quad -\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T}\sum_{i=1}^{N}\sum_{t_2=1}^{T}\widetilde{\Omega}_{v,i}\left(I_d+\widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{Q}_{Ti}\widetilde{\xi}_{it_2} \\
&\equiv B_{NT,1,1,1}(r)+B_{NT,1,1,2}(r). \qquad \text{(E.72)}
\end{aligned}
$$

In the proof of Proposition C.1(iii), we have already shown that

$$
\sup_{0\le r\le 1}||B_{NT,1,1,1}(r)|| = O_P\left(\sqrt{\frac{N}{T^{1+\frac{\delta^*}{2+\delta^*/2}}}}\right). \qquad \text{(E.73)}
$$

Recall that $\Delta_{Ti}=\frac{1}{T}\sum_{t=1}^{T}\widetilde{\eta}_{it}$. Thus

$$
\begin{aligned}
B_{NT,1,1,2}(r) &= -\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\widetilde{\Omega}_{v,i}\left(I_d+\widetilde{Q}_{Ti}\right)\eta_{it_1}\widetilde{Q}_{Ti}\xi_{it_2} \\
&\quad -\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\widetilde{\Omega}_{v,i}\left(I_d+\widetilde{Q}_{Ti}\right)\eta_{it_1}\widetilde{Q}_{Ti}\xi^{c}_{it_2} \\
&\quad -\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\widetilde{\Omega}_{v,i}\left(I_d+\widetilde{Q}_{Ti}\right)\eta^{c}_{it_1}\widetilde{Q}_{Ti}\widetilde{\xi}_{it_2} \\
&\equiv B_{NT,1,1,2,1}(r)+\cdots+B_{NT,1,1,2,2}(r). \qquad \text{(E.74)}
\end{aligned}
$$

Using Lemma E.2, Eqs. (D.19), and Cauchy Schwarz inequality, we have

$$
\begin{aligned}
&\mathrm{E}\left(\sup_{0\le r\le 1}||B_{NT,1,1,2,1}(r)||\right) \le \frac{C\sqrt{\gamma_{NT}}}{\sqrt{NT}T}\left(\sum_{i=1}^{N}\left|\left|\widetilde{\Omega}_{v,i}\right|\right|\right)\left(\mathrm{E}\left|\left|\sum_{t_2=1}^{T}\xi_{it}\right|\right|^{2+\delta^*/2}\right)^{\frac{1}{2+\delta^*/2}} \\
&\times\left(\mathrm{E}\left(\left|\left|\sum_{t_1=1}^{T}\eta_{it_1}\right|\right|^{2}\right)^{\frac{2+\delta^*/2}{2}}\right)^{\frac{1}{2+\delta^*/2}}\left(C_2\mathrm{E}\left|\left|\widetilde{Q}_{Ti}\right|\right|^{2}\right)^{\frac{\delta^*}{4+\delta^*}} \le C\sqrt{\frac{N}{T^{1+\frac{\delta^*}{2+\delta^*/2}}}}. \qquad \text{(E.75)}
\end{aligned}
$$

Since $\left|\left|I_d+\widetilde{Q}_{Ti}\right|\right|$ is bounded, by Eqs. (D.19), Lemma E.2(ii), (v) and Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
&\mathrm{E}\left(\sup_{0\le r\le 1}||B_{NT,1,1,2,2}(r)||\right) \le \frac{C\sqrt{\gamma_{NT}}}{\sqrt{NT}T}\left(\sum_{i=1}^{N}\left|\left|\widetilde{\Omega}_{v,i}\right|\right|\right)\left(\mathrm{E}\left|\left|\sum_{t_1=1}^{T}\eta_{it_1}\right|\right|^{2}\right)^{\frac{1}{2}}\left(\mathrm{E}\left|\left|\sum_{t_2=1}^{T}\xi^{c}_{it_2}\right|\right|^{2}\right)^{\frac{1}{2}} \\
&\le \frac{C\sqrt{N}}{\sqrt{T^2h}}. \qquad \text{(E.76)}
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
&\mathrm{E}\left(\sup_{0\le r\le 1}||B_{NT,1,1,2,3}(r)||\right) \le \frac{C\sqrt{\gamma_{NT}}}{\sqrt{NT}T}\left(\sum_{i=1}^{N}\left|\left|\widetilde{\Omega}_{v,i}\right|\right|\right)\left(\mathrm{E}\left|\left|\sum_{t_1=1}^{T}\eta^{c}_{it_1}\right|\right|^{2}\right)^{\frac{1}{2}}\left(\mathrm{E}\left|\left|\sum_{t_2=1}^{T}\widetilde{\xi}_{it_2}\right|\right|^{2}\right)^{\frac{1}{2}} \\
&\le \frac{C\sqrt{N}}{\sqrt{T^2h}}. \qquad \text{(E.77)}
\end{aligned}
$$

By Eqs. (E.73) to Eqs. (E.77), we have

$$\sup_{0\le r\le 1}||B_{NT,1,1}(r)||=O_P\left(\sqrt{\frac{N}{T^{1+\frac{\delta^*}{2+\delta^*/2}}}}\right). \tag{E.78}$$

We also have

$$\begin{aligned}B_{NT,1,2}(r)&=-\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T}\sum_{i=1}^{N}\sum_{t=1}^{T}\Omega_{v,i}\Delta_{Ti}\left(I_d+\widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it}\\&-\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T}\sum_{i=1}^{N}\sum_{t=1}^{T}\Omega_{v,i}\Delta_{Ti}\left(I_d+\widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{Q}_{Ti}\widetilde{\xi}_{it}\\&\equiv B_{NT,1,2,1}(r)+B_{NT,1,2,2}(r).\end{aligned} \tag{E.79}$$

In the proof of Proposition C.3(iv), we have already shown that

$$\sup_{0\le r\le 1}||B_{NT,1,2,1}(r)||=O_P\left(\sqrt{\frac{N}{T^2h^2}}\right), \tag{E.80}$$

$$\begin{aligned}B_{NT,1,2,2}(r)&=-\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\Delta_{Ti,1}\left(I_d+\widetilde{Q}_{Ti}\right)\eta_{it_1}\widetilde{Q}_{Ti}\xi_{it_2}\\&-\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\Delta_{Ti,1}\left(I_d+\widetilde{Q}_{Ti}\right)\eta_{it_1}\widetilde{Q}_{Ti}\xi_{it_2}^c\\&-\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\Delta_{Ti,1}\left(I_d+\widetilde{Q}_{Ti}\right)\eta_{it_1}^c\widetilde{Q}_{Ti}\widetilde{\xi}_{it_2}\\&-\frac{\sqrt{\gamma_{NT}}[rT]}{\sqrt{NT}T^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\Omega_{v,i}\Delta_{Ti,2}\left(I_d+\widetilde{Q}_{Ti}\right)\widetilde{\eta}_{it_1}\widetilde{Q}_{Ti}\widetilde{\xi}_{it_2}\\&\equiv B_{NT,1,2,2,1}(r)+\cdots+B_{NT,1,2,2,4}(r).\end{aligned} \tag{E.81}$$

Note that $\left|\left|I_d+\widetilde{Q}_{Ti}\right|\right|$ is bounded, and Lemma E.2, together with Cauchy Schwarz inequality, we have

$$\begin{aligned}&\mathrm{E}\left(\sup_{0\le r\le 1}||B_{NT,1,2,2,1}(r)||\right)\le\frac{C\sqrt{\gamma_{NT}}}{\sqrt{NT}T}\sum_{i=1}^{N}\left(\mathrm{E}\,||\Delta_{Ti,1}||^{4+\delta^*}\right)^{\frac{1}{4+\delta^*}}\left(\mathrm{E}\left|\left|\widetilde{Q}_{Ti}\right|\right|^{\frac{4+\delta^*}{1+\delta^*}}\right)^{\frac{1+\delta^*}{4+\delta^*}}\\&\times\left(\mathrm{E}\left|\left|\sum_{t_1=1}^{T}\eta_{it_1}\right|\right|^{4+\delta^*}\right)^{\frac{1}{4+\delta^*}}\left(\mathrm{E}\left|\left|\sum_{t_2=1}^{T}\xi_{it_2}\right|\right|^{4+\delta^*}\right)^{\frac{1}{4+\delta^*}}\le\frac{C\sqrt{N}}{T^{\frac{1}{2}+\frac{1+\delta^*}{4+\delta^*}}}.\end{aligned} \tag{E.82}$$

Using the facts that $\left|\left|I_d+\widetilde{Q}_{Ti}\right|\right|$ and $\left|\left|\widetilde{Q}_{Ti}\right|\right|$ are bounded, and by Lemma E.2, together with Cauchy Schwarz inequality, we have

$$\begin{aligned}&\mathrm{E}\left(\sup_{0\le r\le 1}||B_{NT,1,2,2,2}(r)||\right)\le\frac{C\sqrt{\gamma_{NT}}}{\sqrt{NT}T}\sum_{i=1}^{N}\left(\mathrm{E}\,||\Delta_{Ti,1}||^4\right)^{\frac{1}{4}}\left(\mathrm{E}\left|\left|\sum_{t_1=1}^{T}\eta_{it_1}\right|\right|^4\right)^{\frac{1}{4}}\left(\mathrm{E}\left|\left|\sum_{t_2=1}^{T}\xi_{it_2}^c\right|\right|^2\right)^{\frac{1}{2}}\\&\le\frac{C\sqrt{N}}{\sqrt{T^2h}}.\end{aligned} \tag{E.83}$$

Along with the results in Lemma E.2(vii), we have

$$\mathrm{E}\left(\sup_{0\le r\le 1}||B_{NT,1,2,2,3}(r)||\right) \le \frac{C\sqrt{\gamma_{NT}}}{\sqrt{NT}T}\sum_{i=1}^{N}\left(\mathrm{E}\,||\Delta_{Ti,1}||^4\right)^{\frac{1}{4}}\left(\mathrm{E}\left|\left|\sum_{t_1=1}^{T}\eta_{it_1}^{c}\right|\right|^4\right)^{\frac{1}{4}}\left(\mathrm{E}\left|\left|\sum_{t_2=1}^{T}\widetilde{\xi}_{it_2}\right|\right|^2\right)^{\frac{1}{2}}$$

$$\le \frac{C\sqrt{N}}{Th}. \tag{E.84}$$

Similarly

$$\mathrm{E}\left(\sup_{0\le r\le 1}||B_{NT,1,2,2,4}(r)||\right) \le \frac{C\sqrt{\gamma_{NT}}}{\sqrt{NT}T}\sum_{i=1}^{N}\left(\mathrm{E}\,||\Delta_{Ti,2}||^4\right)^{\frac{1}{4}}\left(\mathrm{E}\left|\left|\sum_{t_1=1}^{T}\widetilde{\eta}_{it_1}\right|\right|^4\right)^{\frac{1}{4}}\left(\mathrm{E}\left|\left|\sum_{t_2=1}^{T}\widetilde{\xi}_{it_2}\right|\right|^2\right)^{\frac{1}{2}}$$

$$\le \frac{C\sqrt{N}}{Th}. \tag{E.85}$$

By Eqs. (E.81) to Eqs. (E.85), we have shown

$$\sup_{0\le r\le 1}||B_{NT,1,2,2}(r)|| = O_P\left(\frac{\sqrt{N}}{Th}\right) \tag{E.86}$$

By Eqs. (E.80) and Eqs. (E.86), we then have

$$\sup_{0\le r\le 1}||B_{NT,1,2}(r)|| = O_P\left(\frac{\sqrt{N}}{Th}\right). \tag{E.87}$$

By going through the calculation for $B_{NT,1,1}(r)$ with taking $\widetilde{\Omega}_{v,i} = I_d$ and using $\Delta_{NT} = o_P\left(\frac{1}{\sqrt{\gamma_{NT}}}\right)$, it is straightforward to see

$$\sup_{0\le r\le 1}||B_{NT,1,3}(r)|| = O_P\left(\frac{\sqrt{N}}{Th}\right). \tag{E.88}$$

Then Eqs. (E.78), Eqs. (E.87) and Eqs. (E.88) together imply,

$$\sup_{0\le r\le 1}||B_{NT,1}(r)|| = O_P\left(\frac{\sqrt{N}}{Th}\right). \tag{E.89}$$

By the same calculation as for $B_{NT,1}(r)$, noting that $\Delta_{Ti} = \frac{1}{T}\sum_{t=1}^{T}\widetilde{\eta}_{it}$, it is easy to show that for any $1\le k_1 < k_2 \le T$,

$$\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}T}\mathrm{E}\left|\left|\sum_{i=1}^{N}\sum_{t_1=k_1+1}^{k_2}\sum_{t_2=1}^{T}\left(\widehat{\Omega}_{x,Ti}-\widehat{\Omega}_{x,NT}\right)\left(I_d+\widetilde{Q}_{Ti}\right)\widetilde{\eta}_{it_1}\left(I_d+\widetilde{Q}_{Ti}\right)\widetilde{\xi}_{it_2}\right|\right| \le \frac{C\sqrt{N(k_2-k_1)}}{\sqrt{T^3h^2}}.$$

Then applying the Lemma E.6 with $\gamma = 1$ and $g(i,j) = \frac{C\sqrt{N(j-i)}}{\sqrt{T^3h^2}}$, we have

$$\sup_{0\le r\le 1}||B_{NT,2}(r)|| = O_P\left(\frac{\sqrt{N\log(2T)}}{Th}\right). \tag{E.90}$$

Eqs. (E.90), along with Eqs. (E.89), suffices to prove Lemma E.5(i).

Proof of Lemma E.5(ii).

Recall that $\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT} = \widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti} - \Delta_{NT}$. Thus

$$\begin{aligned}
&\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \sum_{i=1}^{N}\sum_{t=1}^{[rT]} \left(\widehat{\Omega}_{x,Ti} - \widehat{\Omega}_{x,NT}\right)\left(I_d - \left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\right)\widetilde{\xi}_{it} - L_{NT}(r) \\
&= \left\{\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \sum_{i=1}^{N}\sum_{t=1}^{[rT]} \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti}\right)\left(I_d - \left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\right)\widetilde{\xi}_{it} - L_{NT}(r)\right\} \\
&\quad - \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}}\Delta_{NT} \sum_{i=1}^{N}\sum_{t=1}^{[rT]} \left(I_d - \left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\right)\widetilde{\xi}_{it} \\
&\equiv B_{NT,3}(r) + B_{NT,4}(r).
\end{aligned} \tag{E.91}$$

Note that

$$\begin{aligned}
B_{NT,3}(r) &= \left\{\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \sum_{i=1}^{N}\sum_{t=1}^{[rT]} \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti,1}\right)\xi_{it} - L_{NT}(r)\right\} \\
&\quad + \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \sum_{i=1}^{N}\sum_{t=1}^{[rT]} \Omega_{v,i}\Delta_{Ti,2}\xi_{it} + \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \sum_{i=1}^{N}\sum_{t=1}^{[rT]} \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti}\right)\xi_{it}^{c} \\
&\quad - \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \sum_{i=1}^{N}\sum_{t=1}^{[rT]} \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti}\right)\left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it} \\
&\equiv B_{NT,3,1}(r) + \cdots + B_{NT,3,4}(r).
\end{aligned} \tag{E.92}$$

By definition we have

$$B_{NT,3,1}(r) = 0. \tag{E.93}$$

By the same calculations as in the proofs of Propositions C.1(iii) and C.3(ii)(iii)(iv), for any $1 \leq k_1 < k_2 \leq T$, it is straightforward to have

$$\begin{aligned}
\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \mathrm{E}\left\|\sum_{i=1}^{N}\sum_{t=k_1+1}^{k_2} \Omega_{v,i}\Delta_{Ti,2}\xi_{it}\right\| &\leq \frac{C\sqrt{k_2 - k_1}}{Th} \\
\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \mathrm{E}\left\|\sum_{i=1}^{N}\sum_{t=k_1+1}^{k_2} \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti}\right)\xi_{it}^{c}\right\| &\leq \frac{C\sqrt{N\left(k_2 - k_1\right)}}{\sqrt{T^3h^2}} \\
\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}} \mathrm{E}\left\|\sum_{i=1}^{N}\sum_{t=k_1+1}^{k_2} \left(\widetilde{\Omega}_{v,i} + \Omega_{v,i}\Delta_{Ti}\right)\left(I_d + \widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it}\right\| &\leq \frac{C\sqrt{N\left(k_2 - k_1\right)}}{\sqrt{T^3h^2}}.
\end{aligned} \tag{E.94}$$

Then applying Lemma E.6 with $\gamma = 1$ and $g(i,j) = \frac{C\sqrt{j-i}}{Th}$, $\frac{C\sqrt{N(j-i)}}{\sqrt{T^3h^2}}$ and $\frac{C\sqrt{N(j-i)}}{\sqrt{T^3h^2}}$, respectively, we have

$$\mathrm{E}(\sup\|B_{NT,3,2}(r)\|) = O\left(\frac{\log(2T)}{\sqrt{Th^2}}\right), \tag{E.95}$$

$$\mathrm{E}(\sup\|B_{NT,3,3}(r)\|) = O\left(\frac{\log(2T)\sqrt{N}}{Th}\right), \tag{E.96}$$

$$\mathrm{E}(\sup\|B_{NT,3,4}(r)\|) = O\left(\frac{\log(2T)\sqrt{N}}{Th}\right). \tag{E.97}$$

Along with Eqs. (E.93), we have

$$\sup_{0\le r\le 1} ||B_{NT,3}(r)|| = o_P(1). \tag{E.98}$$

Meanwhile, we have

$$\begin{aligned}
B_{NT,4}(r) &= -\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}}\Delta_{NT}\sum_{i=1}^{N}\sum_{t=1}^{[rT]}\widetilde{\xi}_{it} - \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}}\Delta_{NT}\sum_{i=1}^{N}\sum_{t=1}^{[rT]}\xi_{it}^{c} \\
&+\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}}\Delta_{NT}\sum_{i=1}^{N}\sum_{t=1}^{[rT]}\left(I_d+\widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it} \\
&B_{NT,4,1}(r)+\cdots+B_{NT,4,3}(r).
\end{aligned} \tag{E.99}$$

By Assumption 2, and by Proposition C.1(ii)(iii) with $\widetilde{\Omega}_{v,i} = I_d$, we have

$$\begin{aligned}
\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}}\mathrm{E}\left\|\sum_{i=1}^{N}\sum_{t=k_1+1}^{k_2}\xi_{it}\right\| &\le \frac{\sqrt{\gamma_{NT}\,(k_2-k_1)}}{\sqrt{T}} \\
\frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}}\mathrm{E}\left\|\sum_{i=1}^{N}\sum_{t=k_1+1}^{k_2}\xi_{it}^{c}\right\| &\le \frac{CN^{1/4}\sqrt{\gamma_{NT}\,(k_2-k_1)}}{\sqrt{T^2h}} \\
\equiv \frac{\sqrt{\gamma_{NT}}}{\sqrt{NT}}\mathrm{E}\left\|\sum_{i=1}^{N}\sum_{t=k_1+1}^{k_2}\left(I_d+\widetilde{Q}_{Ti}\right)\Delta_{Ti}\widetilde{\xi}_{it}\right\| &\le \frac{C\sqrt{N\gamma_{NT}\,(k_2-k_1)}}{T^{1+\frac{\delta^*}{4+\delta^*}}}.
\end{aligned} \tag{E.100}$$

Then applying Lemma E.6 with $\gamma = 1$ and $g(i,j) = \frac{C\sqrt{\gamma_{NT}(j-i)}}{T}$, $\frac{CN^{1/4}\sqrt{\gamma_{NT}(j-i)}}{\sqrt{T^2h}}$ and $\frac{\sqrt{N\gamma_{NT}(j-i)}}{T^{1+\frac{\delta^*}{4+\delta^*}}}$ respectively, together with $\Delta_{NT} = O_P\left(\frac{1}{\sqrt{NT}}\right)$ in Lemma E.4(i), we have

$$\mathrm{E}(\sup ||B_{NT,4,2}(r)||) = O\left(\frac{\log(2T)}{\sqrt{N}}\right), \tag{E.101}$$

$$\mathrm{E}(\sup ||B_{NT,4,3}(r)||) = O\left(\frac{\log(2T)}{N^{1/4}\sqrt{Th}}\right), \tag{E.102}$$

$$\mathrm{E}(\sup ||B_{NT,4,4}(r)||) = O\left(\frac{\log(2T)}{T^{\frac{1}{2}+\frac{\delta^*}{4+\delta^*}}}\right), \tag{E.103}$$

which imply

$$\sup_{0\le r\le 1} ||B_{NT,4}(r)|| = o_P(1). \tag{E.104}$$

This completes the proof of Lemma E.5(ii).

Appendix F Additional results for the empirical application

In this section, we provide some additional tables and figures for the empirical example in Section 6. Figure 6 to Figure 9 give the estimates for the individual trending functions together with the 95% conference intervals calculated by the bootstrap standard deviations. These figures illustrate more explicitly the diversity of the trending functions among OECD countries. Figure 10 plots the mean group estimation of the average trending function across OECD countries. In comparison to the pool estimation in the right panel, the mean group estimation reveals that the averaging technology progress has experienced a slowdown after the Subprime crisis.

In Table 7, we provide the p-value (using bootstrap critical values) in the homogeneity test by including different combinations of explanatory. In particular, "GDP + DR + DH + GHE" means all four variables are included, and "GDP" means only GDP is used. Note that first of all, p-values are smaller when there are more covariates included (see the first 4 rows). A possible explanation would be an additional entry in β_i could introduce extra heterogeneity, which may increase the chance to reject the null. Secondly, in some combinations, the p-value are smaller in OECD level than those in groups (see the second to the forth rows). The means the parameter vectors could exhibit some group-specific features, thus it tends to reject the null across all groups, but accept the null within some groups. Thirdly, for the single variable cases, the small p-values with GDP and PH are consistent with the results in Figure 3. Finally, the large p-values with DR and GHE also agree the results in Figure 3, and small p-values are also found in certain group.

In addition, the estimations of β_i in each OECD countries with bootstrap standard error(in parentheses) are provided in Table 8. Similar to Table 4 and Table 5 in the paper, "*" here indicates a 5% level significance(double tail).

Table 7: Mean group estimators in each group

Covariates	OECD	G1	G2	G3	G4	G5	G6
GDP+DR+PH+GHE	0.01	0.57	0.07	0.00	0.00	0.99	0.03
GDP+DR+PH	0.05	0.48	0.01	0.08	0.06	0.76	0.08
GDP+DR	0.02	0.79	0.02	0.74	0.27	0.53	0.98
GDP	0.01	0.38	0.14	0.54	0.37	0.73	0.30
DR	0.74	0.88	0.38	0.03	0.15	0.65	0.96
PH	0.08	0.18	0.92	0.03	0.99	0.24	0.79
GHE	0.14	0.08	0.94	0.97	0.00	0.08	0.28

References

Baltagi, B. H. and Moscone, F. (2010). Health care expenditure and income in the oecd reconsidered: Evidence from panel data. *Economic Modelling*, 27(4):804–811.

Billingsley, P. (1968). *Convergence of probability Measures*. John Wiley, New York.

Boneva, L., Linton, O., and Vogt, M. (2015). A semiparametric model for heterogeneous panel data with fixed effects. *Journal of Econometrics*, 188(2):327–345.

Bonhomme, S. and Manresa, E. (2015). Grouped patterns of heterogeneity in panel data. *Econometrica*, 83(3):1147–1184.

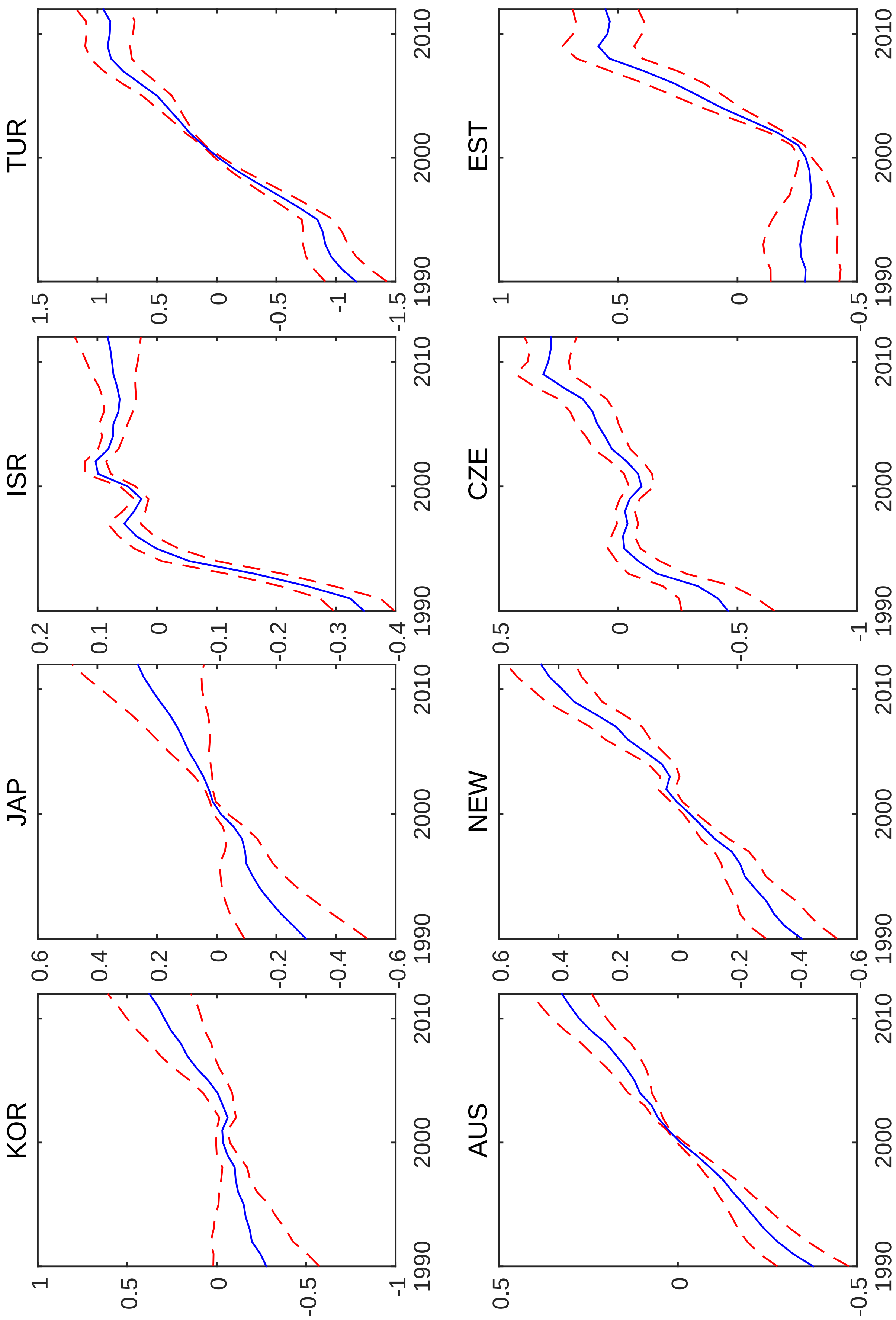

Figure 6: Estimation of individual trend in health expenditure

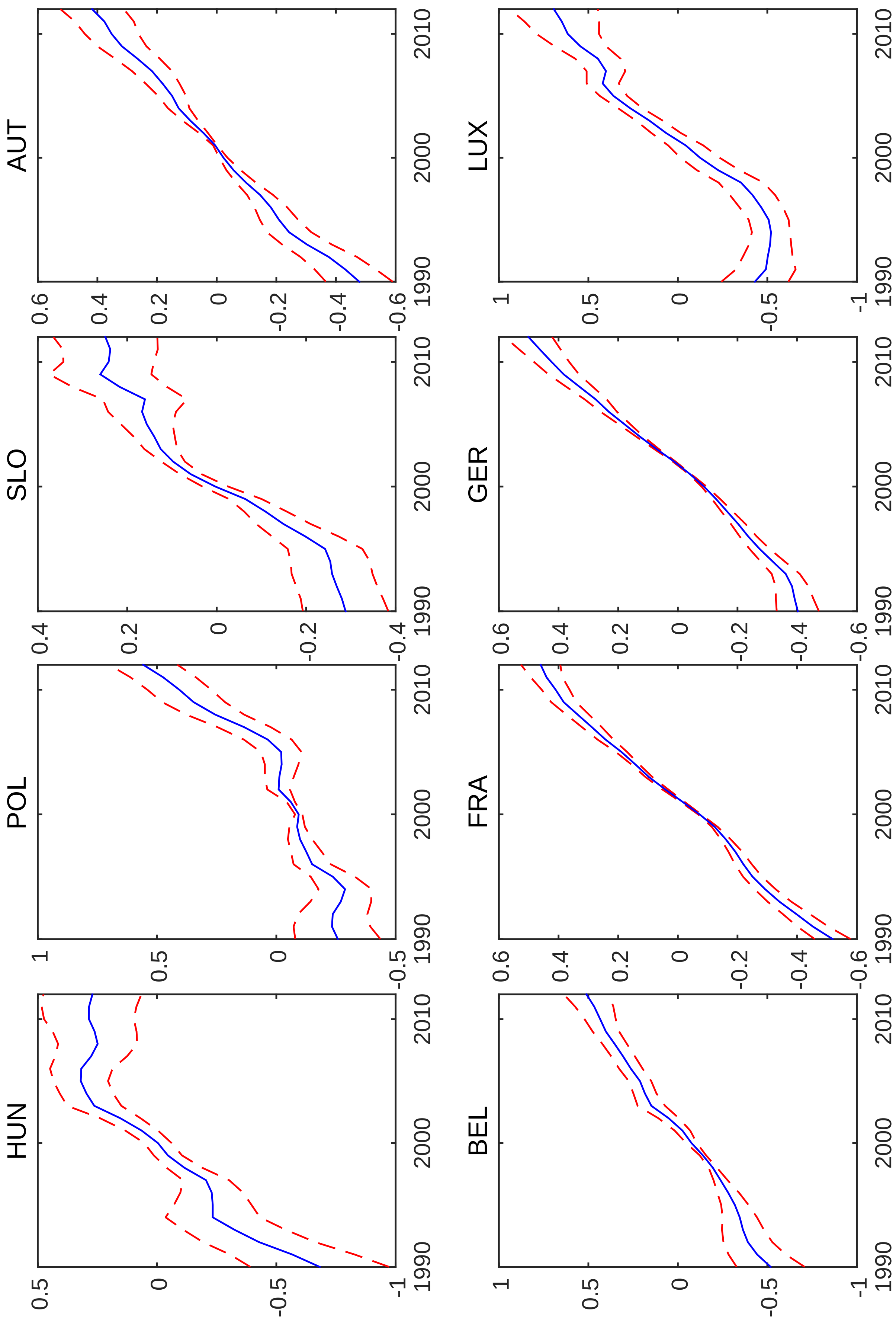

Figure 7: Estimation of individual trend in health expenditure

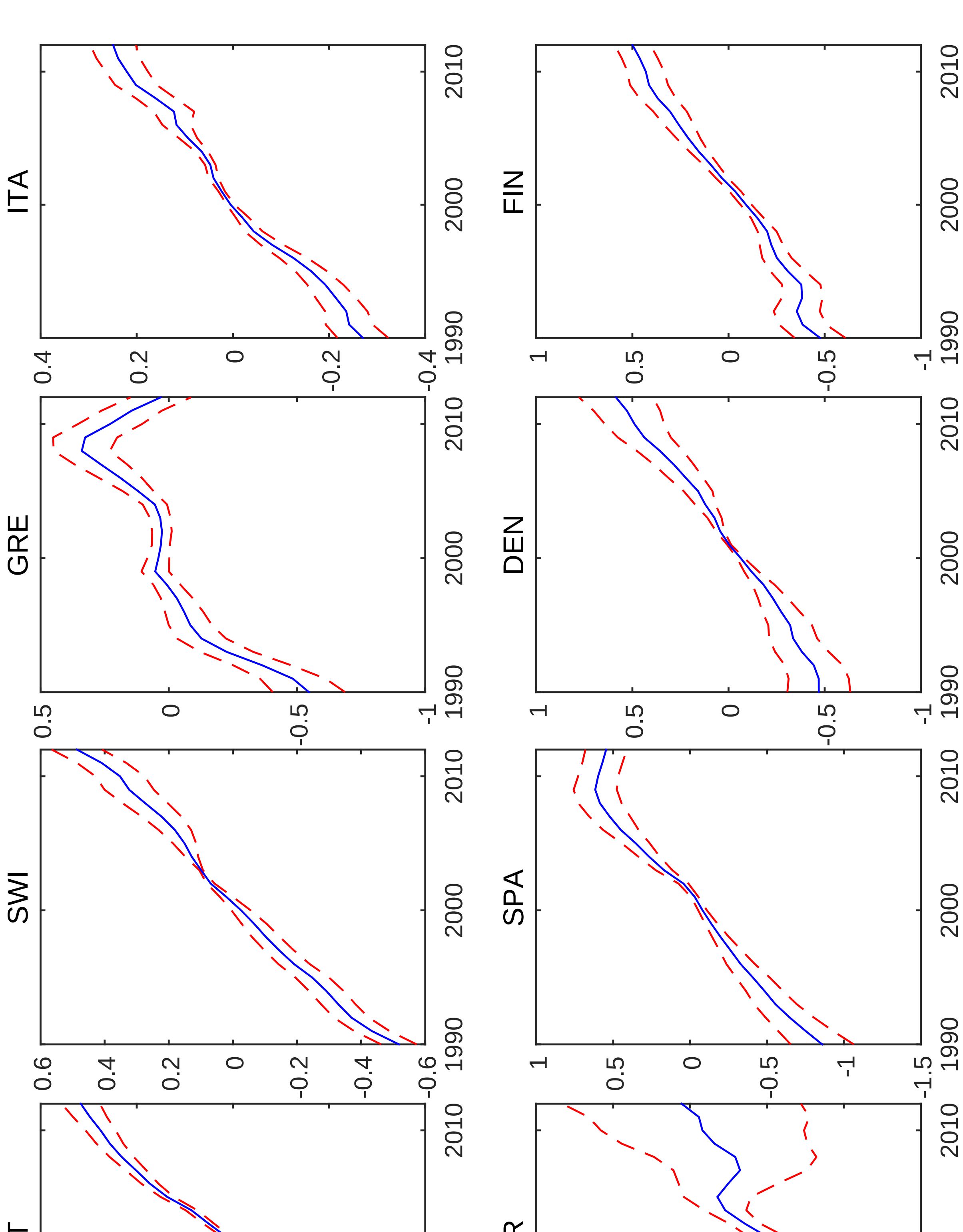

Figure 8: Estimation of individual trend in health expenditure

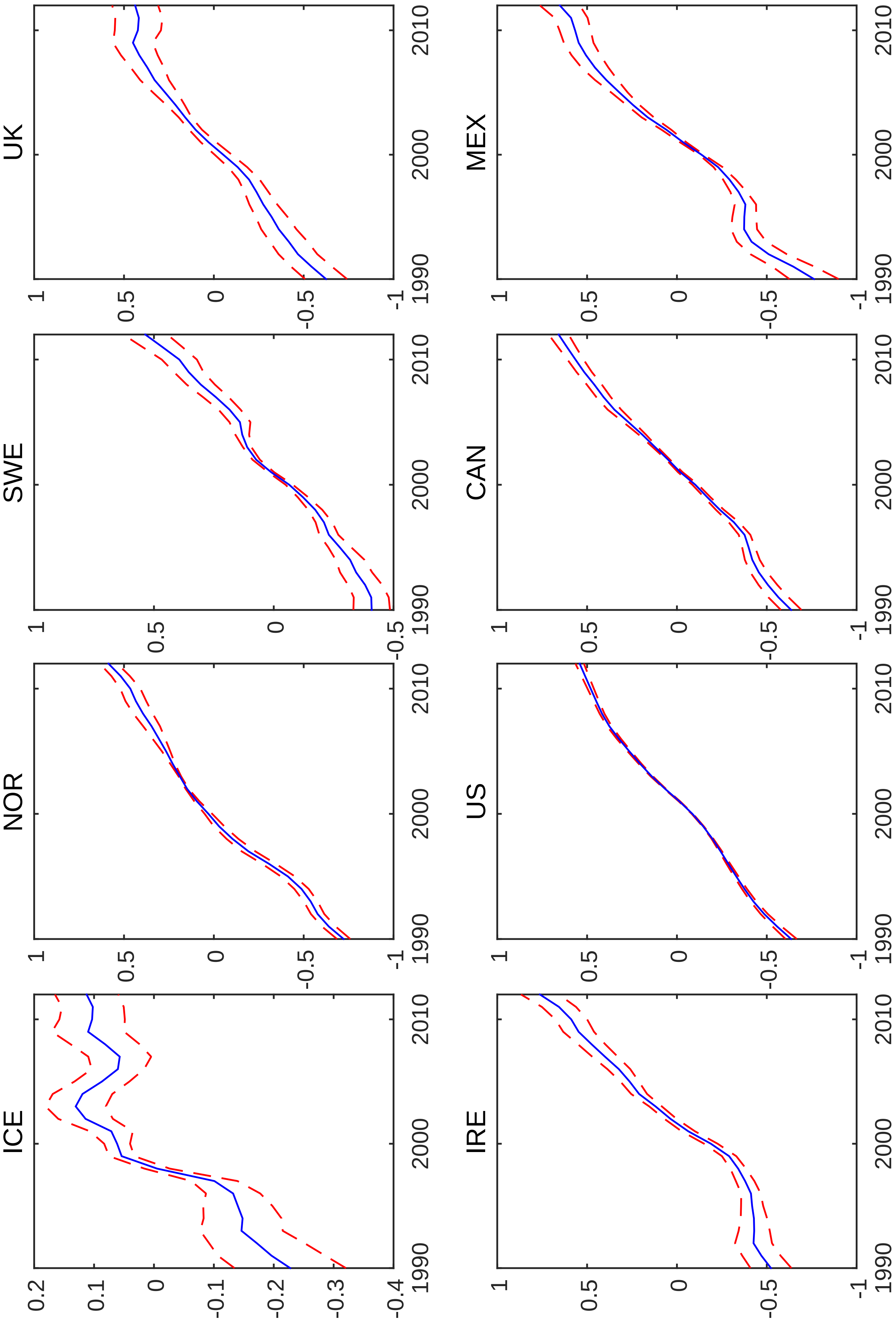

Figure 9: Estimation of individual trend in health expenditure

Table 8: $\widehat{\beta}_i$ for each OECD country

	KOR	JAP	ISR	TUR	AUS	NEW	CZE	EST	HUN	POL	SLO	AUT	BEL	FRA	GER	LUX
GDP	0.971 *	0.137 *	0.840 *	0.633 *	0.559 *	0.124	0.545 *	0.084	0.780 *	-0.161	0.514 *	0.245 *	0.203	0.149	0.091	0.526 *
	(0.178)	(0.062)	(0.148)	(0.103)	(0.148)	(0.221)	(0.205)	(0.163)	(0.284)	(0.266)	(0.192)	(0.100)	(0.167)	(0.104)	(0.097)	(0.160)
DR	0.063 *	-0.001	0.009	0.049 *	-0.034 *	-0.034	-0.027	0.035 *	0.057	-0.081 *	0.036 *	0.000	-0.004	-0.006	-0.009 *	0.053
	(0.019)	(0.004)	(0.006)	(0.022)	(0.009)	(0.019)	(0.015)	(0.015)	(0.041)	(0.022)	(0.014)	(0.007)	(0.025)	(0.012)	(0.003)	(0.040)
PHY	0.311	0.557	0.009	-0.030	-0.030 *	0.168 *	0.333 *	0.189 *	-0.001	-0.128	0.046	-0.053	-0.005	-0.026 *	0.039	-0.875 *
	(0.176)	(0.469)	(0.051)	(0.077)	(0.005)	(0.077)	(0.156)	(0.077)	(0.028)	(0.101)	(0.051)	(0.064)	(0.011)	(0.012)	(0.050)	(0.115)
GHE	0.010 *	0.002	-0.004	-0.011 *	0.000	-0.004	0.016	-0.004	0.008	-0.004 *	-0.008	0.010 *	0.010 *	0.026 *	0.020 *	0.021 *
	(0.003)	(0.001)	(0.003)	(0.001)	(0.002)	(0.005)	(0.011)	(0.004)	(0.009)	(0.002)	(0.007)	(0.003)	(0.002)	(0.002)	(0.002)	(0.007)
	NET	SWI	GRE	ITA	POR	SPA	DEN	FIN	ICE	NOR	SWE	UK	IRE	US	CAN	MEX
GDP	0.029	0.217 *	1.115 *	0.722 *	1.104 *	0.107	0.063	0.154	0.690 *	0.191 *	0.183	0.279	-0.451 *	-0.058	-0.326 *	0.100
	(0.093)	(0.105)	(0.235)	(0.118)	(0.243)	(0.205)	(0.176)	(0.111)	(0.171)	(0.050)	(0.124)	(0.233)	(0.155)	(0.072)	(0.122)	(0.158)
DR	-0.046 *	-0.031 *	0.072 *	-0.013	-0.056	-0.001	-0.054 *	-0.005	0.003	-0.032 *	-0.029 *	0.005	-0.060 *	0.000	0.002	0.005
	(0.011)	(0.013)	(0.016)	(0.009)	(0.041)	(0.019)	(0.018)	(0.009)	(0.015)	(0.014)	(0.005)	(0.008)	(0.012)	(0.004)	(0.011)	(0.004)
PHY	-0.032	-0.069	-0.172 *	-0.023 *	0.121	-0.110 *	0.166	-0.171	0.224	-0.125 *	-0.081	-0.018	-0.025	0.019	-0.097 *	-0.101
	(0.039)	(0.037)	(0.083)	(0.010)	(0.145)	(0.041)	(0.104)	(0.243)	(0.139)	(0.032)	(0.060)	(0.204)	(0.067)	(0.040)	(0.039)	(0.086)
GHE	0.000	0.001	-0.003	0.023 *	-0.005	0.000	0.031 *	0.009	0.013	0.006	0.000	0.008	-0.004	-0.002	0.006	0.018 *
	(0.001)	(0.001)	(0.003)	(0.002)	(0.005)	(0.005)	(0.009)	(0.008)	(0.010)	(0.005)	(0.004)	(0.004)	(0.006)	(0.002)	(0.007)	(0.004)

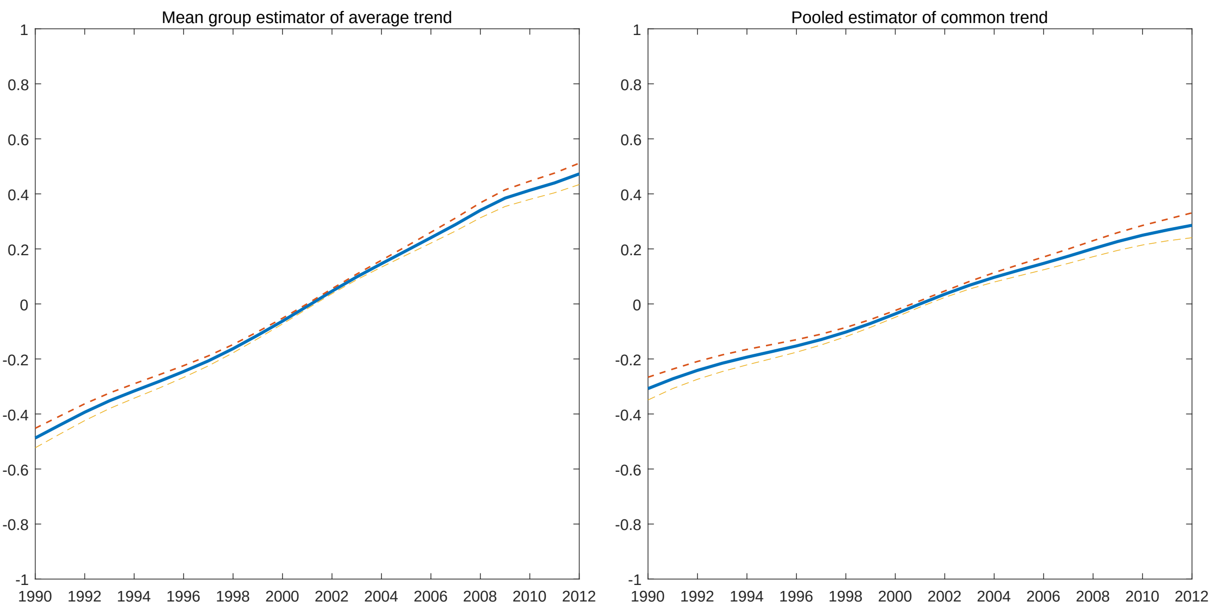

Figure 10: Mean group and pooled estimators of trends

Chandra, A. and Skinner, J. (2012). Technology growth and expenditure growth in health care. *Journal of Economic Literature*, 50(3):645–680.

Chen, J., Gao, J., and Li, D. (2012). Semiparametric trending panel data models with cross-sectional dependence. *Journal of Econometrics*, 171(1):71–85.

Chudik, A. and Pesaran, M. H. (2015). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics*, 188(2):393–420.

Davidson, J. (1994). *Stochastic Limit Theory: An Introduction for Econometricians: An Introduction for Econometricians*. Oxford University Press, UK.

Dong, C., Gao, J., and Peng, B. (2015a). Partially linear panel data models with cross-sectional dependence and nonstationarity. *Available at SSRN 2585334*.

Dong, C., Gao, J., and Peng, B. (2015b). Semiparametric single-index panel data models with cross-sectional dependence. *Journal of Econometrics*, 188(1):301–312.

Durlauf, S. N., Kourtellos, A., and Minkin, A. (2001). The local solow growth model. *European Economic Review*, 45(4):928–940.

Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer Series in Statistics. Springer, New York.

Gao, J. (2007). *Nonlinear Time Series: Semiparametric and Nonparametric Methods*. Chapman & Hall/CRC, London.

Heckman, J. J. (2001). Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of political Economy*, 109(4):673–748.

Hsiao, C. (2014). *Analysis of Panel Data*. Number 54. Cambridge university press, Cambridge.

Kiefer, N. M., Vogelsang, T. J., and Bunzel, H. (2000). Simple robust testing of regression hypotheses. *Econometrica*, 68(3):695–714.

Kim, M. S. and Sun, Y. (2013). Heteroskedasticity and spatiotemporal dependence robust inference for linear panel models with fixed effects. *Journal of Econometrics*, 177(1):85–108.

Lee, J. and Robinson, P. M. (2016). Series estimation under cross-sectional dependence. *Journal of Econometrics*, 190(1):1–17.

Lee, L.-F. and Yu, J. (2015). Estimation of fixed effects panel regression models with separable and nonseparable space–time filters. *Journal of Econometrics*, 184(1):174–192.

Li, Q. and Racine, J. S. (2007). *Nonparametric Econometrics: Theory and Practice*. Princeton University Press.

Lin, C.-C. and Ng, S. (2012). Estimation of panel data models with parameter heterogeneity when group membership is unknown. *Journal of Econometric Methods*, 1(1):42–55.

Magnus, J. R., Neudecker, H., et al. (1995). Matrix differential calculus with applications in statistics and econometrics.

Móricz, F., Serfling, R., and Stout, W. (1982). Moment and probability bounds with quasi-superadditive structure for the maximum partial sum. *The Annals of Probability*, pages 1032–1040.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4):967–1012.

Pesaran, M. H. and Yamagata, T. (2008). Testing slope homogeneity in large panels. *Journal of Econometrics*, 142(1):50–93.

Robinson, P. M. (2012). Nonparametric trending regression with cross-sectional dependence. *Journal of Econometrics*, 169(1):4–14.

Shao, Q.-M. and Yu, H. (1996). Weak convergence for weighted empirical processes of dependent sequences. *The Annals of Probability*, pages 2098–2127.

Su, L., Shi, Z., and Phillips, P. C. (2014). Identifying latent structures in panel data. *Available at SSRN 2448189*.

Su, L., Wang, X., and Jin, S. (2015). Sieve estimation of time-varying panel data models with latent structures.

Vogelsang, T. J. (2012). Heteroskedasticity, autocorrelation, and spatial correlation robust inference in linear panel models with fixed-effects. *Journal of Econometrics*, 166(2):303–319.

Vogt, M. and Linton, O. B. (2015). Classification of nonparametric regression functions in heterogeneous panels. *Available at SSRN 2567312*.

Zhang, Y., Su, L., and Phillips, P. C. (2012). Testing for common trends in semi-parametric panel data models with fixed effects. *The Econometrics Journal*, 15(1):56–100.