ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Partially Linear Panel Data Models with Cross-Sectional Dependence and Nonstationarity

Chaohua Dong, Jiti Gao and Bin Peng

March 2015

Working Paper 07/15

Partially Linear Panel Data Models with Cross–Sectional Dependence and Nonstationarity[1]

CHAOHUA DONG[⋆,∗], JITI GAO[⋆] AND BIN PENG[†]

[⋆]Monash University and [†]University of Technology Sydney
and [∗]Southwestern University of Finance and Economics

Abstract

In this paper, we consider a partially linear panel data model with cross–sectional dependence and non–stationarity. Meanwhile, we allow fixed effects to be correlated with the regressors to capture unobservable heterogeneity. Under a general spatial error dependence structure, we then establish some consistent closed–form estimates for both the unknown parameters and the unknown function for the case where N and T go jointly to infinity. Rates of convergence and asymptotic normality results are established for the proposed estimators. Both the finite–sample performance and the empirical applications show that the proposed estimation method works well when the cross–sectional dependence exists in the data set.

Keywords: Asymptotic theory; closed–form estimate; orthogonal series method; partially linear panel data model.

JEL classification: C13, C14, C23, C51

[1]The second author acknowledges comments and suggestions by Farshid Vahid. The authors of this paper also acknowledge the Australian Research Council Discovery Grants Program support under Grant numbers: DP130104229 and DP150101012.

1 Introduction

Nonlinear and nonstationary time series models have received considerable attention during the last thirty years. Nonlinearity and nonstationarity are dominant characteristics of many economic and financial data sets, for example, exchange rates and inflation rates. Many datasets, such as aggregate disposable income and consumption, are found to be integrated processes. With the development of asymptotic theory in recent years, researchers are able to construct econometric models using original data rather than a differenced version, while in the past one might need to use a differenced version to satisfy stationarity requirements. In a recent publication, Gao and Phillips (2013) consider a partially liner time series data model of the form:

$$
\begin{aligned}
Y_t &= AX_t + g(V_t) + e_t, \\
X_t &= H(V_t) + U_t, \quad t = 1, \dots, T,
\end{aligned}
$$

which extend existing partially linear models given in Härdle et al. (2000) and allow the integrated time series $V_t = V_{t-1} + \epsilon_t$ to be the driving force of the data set. Moreover, a semiparametric estimation method is provided in Gao and Phillips (2013) to recover the parameter A of interest and unknown function $g(\cdot)$ based on a kernel estimation technique. As a result, the relationship of some vital integrated economic and financial variables, like the impact of interest rates on private consumption, may be depicted directly in modelling. While the literature on nonstationary time series grows, very few nonlinear and nonstationary panel data models have been provided to accommodate nonstationarity.

Recent studies by Robinson (2012) and Chen et al. (2012*b*) involve the time trend to capture nonstationarity and extend the time series model in Gao and Phillips (2013) to the panel data setting:

$$
\begin{aligned}
y_{it} &= x_{it}'\beta_0 + g(t/T) + \omega_i + e_{it}, \\
x_{it} &= \phi(t/T) + \lambda_i + v_{it},
\end{aligned}
$$

for $i = 1, \dots, N$ and $t = 1, \dots, T$, where the relations $\sum_{i=1}^N \omega_i = 0$ and $\sum_{i=1}^N \lambda_i = 0$ are stipulated for the purpose of identification. Recently, Bai et al. (2009) and Kapetanios et al. (2011) extend the linear panel data models considered by Bai (2009) and Pesaran (2006) by allowing the factors (also often known as macro shocks in some basic economic concepts) to follow nonstationary time series processes. Meanwhile, Bai and Carrion-I-Silvestre (2009) study the problem of unit root testing in the presence of multiple structural changes and common dynamic factors, and Bai and Ng (2010) extend their earlier work in Bai and Ng (2004) to investigate the panel data unit root test with cross–sectional dependence.

Following the literature, it is necessary to establish some relevant asymptotic theory for panel data models when unit root processes are involved in the system. In this paper, one of our aims is to provide some new asymptotic theory for panel data models with the presence of integrated processes when N and T diverge jointly. These results can easily be employed to further studies on the panel data models. Due to the use of Hermite orthogonal functions, some results are also very useful to sieve–estimate–based studies. Moreover, taking into account the correlation among individuals has become an important topic when modelling panel data sets. One popular method is using a factor structure to mimic the strong correlation between individuals. Since Pesaran (2006) and Bai (2009), many extensions have been made. Another popular approach is measuring the correlation between individuals by geographical locations with a spatial error structure on the cross–section dimension. Many papers have adopted this approach, see, for example Pesaran and Tosetti (2011), Chen et al. (2012*b*) and Chen et al. (2012*a*). In this paper, the latter one is employed.

Based on the literature given above, we consider a partially linear panel data model with integrated time series. Specifically, the model is formulated as follows:

$$\begin{aligned} y_{it} &= x_{it}'\beta_0 + g(u_{it}) + \omega_i + e_{it}, \\ x_{it} &= \phi(u_{it}) + \lambda_i + v_{it}, \\ u_{it} &= u_{i,t-1} + \eta_{it}, \quad i = 1, \ldots, N \text{ and } t = 1, \ldots, T, \end{aligned} \tag{1.1}$$

where x_{it} and u_{it} are observable explanatory variables, u_{it} follows an integrated process on time dimension, $g(w)$ is an unknown function in $L^2(\mathbb{R})$, $\phi(w) = (\phi_1(w), \ldots, \phi_d(w))'$ is a vector of unknown integrable functions. Note that, under the current set–up, $\phi_j(w)$ for $j = 1, \ldots, d$ and $g(w)$ will not be constants and all the constant terms are absorbed in fixed effects ω_i and λ_i. Since we shall use the within transformation later on, all the fixed effects simply disappear from the system. Thereby, we do not require extra conditions on identifiability, which is similar to (3.2.5) on page 32 of Hsiao (2003). Accordingly, the fixed effects can capture unobservable heterogeneity and be correlated with the regressors. More detailed discussions and examples can be seen in Hsiao (2003). Note also that model (1.1) extends some time series models discussed in Härdle et al. (2000) to the panel data case.

One interesting finding is that for model (1.1), the within transformation does not affect the asymptotic theory to be established. This is different from those for panel data models with stationarity on the time dimension. A short explanation is that, for a stationary panel data set μ_{it}, $\frac{1}{T}\sum_{t=1}^T g(\mu_{it}) = E[g(\mu_{it})] + O_P\left(\frac{1}{\sqrt{T}}\right)$ under regular restrictions. However, for an integrated panel data regressor u_{it}, we have $\frac{1}{T}\sum_{t=1}^T g(u_{it}) = O_P\left(\frac{1}{\sqrt{T}}\right)$ due to the integrability of $g(\cdot)$. As a result, within transformation helps to remove the fixed effects without any cost. The detailed discussion will be seen in the rest of this paper.

Another crucial finding is that the joint divergence of $(N,T) \to (\infty,\infty)$ makes the asymptotic theory drastically different from that of the integrated time series case. As stated in Lemma B.5 below, when $(N,T) \to (\infty,\infty)$ jointly

$$L_{NT} - E[L_{NT}] \to_P 0, \quad \text{where } L_{NT} = \frac{1}{N\sqrt{T}} \sum_{i=1}^{N} \sum_{t=1}^{T} g(u_{it}). \tag{1.2}$$

However, if μ_t is a unit root process, we have $l_T = \frac{1}{\rho\sqrt{T}} \sum_{t=1}^{T} g(\mu_t) \to_D L_B(1,0) \int g(x)dx$ given some conditions on $g(x)$, where $\rho > 0$ is a constant, B stands for a standard Brownian motion generated by μ_t and $L_B(1,0)$ is the local process of B that measures the sojourning time of B at zero over the period $[0,1]$. To obtain the limit of L_{NT}, one naive thought would be that for each i,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} g(u_{it}) \to_D \rho \cdot L_{B_i}(1,0) \int g(x)dx, \tag{1.3}$$

as $T \to \infty$, where B_i is a standard Brownian motion generated by u_{it}, then by the law of large numbers, $L_{NT} \to_D \rho E[L_B(1,0)] \int g(x)dx$. Although $E[L_B(1,0)]$ does exist, this derivation contradicts the joint divergence of N and T, because (1.3) might not be true for $i = N$ when $(N,T) \to (\infty,\infty)$ jointly. On the other hand, the establishment of (1.2) does not need an expansion of probability space, while usually researchers have to do so in order to obtain a convergence in probability in the nonstationary context. See, for example, Park and Phillips (2001). This is extremely convenient for the establishment of our asymptotic theory.

In summary, we make the following contributions in this paper.

1. We extend the partially linear models given in Gao and Phillips (2013) and Chen et al. (2012*b*) and allow for the presence of nonstationarity processes on the time dimension.

2. The difference in asymptotic theory with the presence of nonstationarity for time series as $T \to \infty$ and for panel data as $(N,T) \to (\infty,\infty)$ jointly is phenomenal.

3. The sieve estimation method employed produces some simple closed–form estimators and the results in some new asymptotic properties for the estimators.

4. The results obtained under panel data setting are stronger than those achieved in the integrated time series setting due to the new limit of the type (1.2) that avoids the expansion of the original probability space in order to obtain a limit in probability.

The structure of this paper is as follows. Section 2 proposes the sieve–based estimation method and introduces the necessary assumptions for the establishment of an asymptotic theory in Section 3. Section 4 discusses some related extensions and limitations of our model. Section 5 evaluates the finite–sample performance by Monte Carlo simulation and a case study on Balassa–Samuelson model. Section 6 concludes. The proofs of the main results are given in Appendices A and B, while some proofs of the secondary results are provided in Appendix C of a supplementary document of this paper.

Throughout the paper, $1_d = (1, \cdots, 1)'$ is a $d \times 1$ vector; $M_P = I_n - P(P'P)^{-1}P'$ denotes the projection matrix generated by full column matrix $P_{n\times m}$; $\|\cdot\|$ denotes Euclidean norm; $\to_P$ and $\to_D$ stand for converging in probability and in distribution, respectively; $\lambda_{min}(A)$ and $\lambda_{max}(A)$ denote minimum and maximum eigenvalues of a $n \times n$ matrix A, respectively; $\lfloor a \rfloor \leq a$ means the largest integer part of a; $\int g(w)dw$ represents $\int_{-\infty}^{\infty} g(w)dw$ and similar notation applies to multiple integration.

2 Estimation method and assumptions

Let $\{H_i(w), i = 0, 1, 2, \ldots\}$ be the Hermite polynomial system orthogonal with respect to $\exp(-w^2)$, which is complete in the Hilbert space $L^2(\mathbb{R}, \exp(-w^2))$. The orthogonality of the system reads $\int H_i(w)H_j(w)\exp(-w^2)dw = \sqrt{\pi}2^i i!\delta_{ij}$, where δ_{ij} is the Kronecker delta. Correspondingly, the so–called Hermite functions are defined by $\mathscr{H}_i(w) = \frac{1}{\sqrt[4]{\pi}\sqrt{2^i i!}} H_i(w)\exp(-w^2/2)$ for $i \geq 0$, which is an orthonormal basis in the Hilbert space $L^2(\mathbb{R})$. Thus, the unknown function $g(w) \in L^2(\mathbb{R})$ can be expanded into the following orthogonal series:

$$g(w) = \sum_{j=0}^{\infty} c_j \mathscr{H}_j(w) = Z_k(w)'C + \gamma_k(w), \quad c_j = \int g(w)\mathscr{H}_j(w)dw, \tag{2.1}$$

where $Z_k(w) = (\mathscr{H}_0(w), \ldots, \mathscr{H}_{k-1}(w))'$, $C = (c_0, \ldots, c_{k-1})'$ and $\gamma_k(w) = \sum_{j=k}^{\infty} c_j \mathscr{H}_j(w)$.

Additionally, in order to remove fixed effects from the system, we take the within transformation and write the model as

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)'\beta_0 + (Z_k(u_{it}) - \bar{Z}_{k,i})'C + \gamma_k(u_{it}) - \bar{\gamma}_{k,i} + e_{it} - \bar{e}_i,$$

where $\bar{y}_i = \frac{1}{T}\sum_{t=1}^T y_{it}$, $\bar{x}_i = \frac{1}{T}\sum_{t=1}^T x_{it}$, $\bar{Z}_{k,i} = \frac{1}{T}\sum_{t=1}^T Z_k(u_{it})$, $\bar{\gamma}_{k,i} = \frac{1}{T}\sum_{t=1}^T \gamma_k(u_{it})$ and $\bar{e}_i = \frac{1}{T}\sum_{t=1}^T e_{it}$. For simplicity, let $\tilde{y}_{it} = y_{it} - \bar{y}_i$ and $\tilde{x}_{it}$, $\tilde{Z}_k(u_{it})$, $\tilde{\gamma}_k(u_{it})$ and $\tilde{e}_{it}$ be defined in the same fashion for $1 \leq i \leq N$ and $1 \leq t \leq T$. Then we rewrite (1.1) in matrix notation as

$$Y = X\beta_0 + \mathcal{Z}C + \gamma + \mathcal{E}, \tag{2.2}$$

where

$$\underset{NT\times 1}{Y} = (\tilde{y}_{11}, \ldots, \tilde{y}_{1T}, \ldots, \tilde{y}_{N1}, \ldots, \tilde{y}_{NT})',$$

$$\underset{NT\times d}{X} = (\tilde{x}_{11}, \ldots, \tilde{x}_{1T}, \ldots, \tilde{x}_{N1}, \ldots, \tilde{x}_{NT})',$$

$$\underset{NT\times k}{\mathcal{Z}} = (\tilde{Z}_k(u_{11}), \ldots, \tilde{Z}_k(u_{1T}), \ldots, \tilde{Z}_k(u_{N1}), \ldots, \tilde{Z}_k(u_{NT}))',$$

$$\underset{NT\times 1}{\gamma} = (\tilde{\gamma}_k(u_{11}), \ldots, \tilde{\gamma}_k(u_{1T}), \ldots, \tilde{\gamma}_k(u_{N1}), \ldots, \tilde{\gamma}_k(u_{NT}))',$$

$$\underset{NT\times 1}{\mathcal{E}} = (\tilde{e}_{11}, \ldots, \tilde{e}_{1T}, \ldots, \tilde{e}_{N1}, \ldots, \tilde{e}_{NT})'.$$

To simplify the proof and facilitate the discussion, we project out $\mathcal{Z}C$ and $X\beta_0$ respectively and focus on the next two equations in turn in the following sections:

$$M_{\mathcal{Z}}Y = M_{\mathcal{Z}}X\beta_0 + M_{\mathcal{Z}}\gamma + M_{\mathcal{Z}}\mathcal{E} \quad \text{and} \quad M_X Y = M_X\mathcal{Z}C + M_X\gamma + M_X\mathcal{E},$$

where $M_{\mathcal{Z}} = I_{NT} - \mathcal{Z}(\mathcal{Z}'\mathcal{Z})^{-1}\mathcal{Z}'$ and $M_X = I_{NT} - X(X'X)^{-1}X'$, giving the within OLS estimators of β_0 and C:

$$\hat{\beta} = (X'M_{\mathcal{Z}}X)^{-1}X'M_{\mathcal{Z}}Y \quad \text{and} \quad \hat{C} = (\mathcal{Z}'M_X\mathcal{Z})^{-1}\mathcal{Z}'M_XY. \tag{2.3}$$

The following assumptions are necessary for the theoretical development and their detailed discussion and some examples are provided in Appendix A.

Assumption 1

1. Let $\{\varepsilon_{ij}, i \in \mathbb{Z}^+, j \in \mathbb{Z}\}$ be a sequence of independent and identically distributed (i.i.d.) random variables across i and j. Moreover, $E[\varepsilon_{11}] = 0$, $E[\varepsilon_{11}^2] = 1$ and $E[|\varepsilon_{11}|^p] < \infty$ for some $p > 4$. In addition, ε_{11} has distribution absolutely continuous with respect to Lebesgue measure and characteristic function $c(r)$ satisfying $\int |rc(r)|dr < \infty$.

2. For $1 \leq i \leq N$ and $1 \leq t \leq T$, let $u_{it} = u_{i,t-1} + \eta_{it}$ with $u_{i0} = O_P(1)$, where η_{it} is a linear process of the form: $\eta_{it} = \sum_{j=0}^{\infty} \rho_j \varepsilon_{i,t-j}$, where $\{\rho_j\}$ is a scalar sequence, $\rho_0 = 1$, $\sum_{j=0}^{\infty} j|\rho_j| < \infty$ and $\rho := \sum_{j=0}^{\infty} \rho_j \neq 0$.

3. (a) Let $v_t = (v_{1t}, \ldots, v_{Nt})'$ be strictly stationary and α–mixing. Also, $E[v_{it}] = 0$ and $E[v_{it}v_{it}'] = \Sigma_v$ for all $1 \leq i \leq N$ and $1 \leq t \leq T$, where Σ_v is a positive definite matrix. Let $\alpha_{ij}(|t-s|)$ denote the α–mixing coefficient between v_{it} and v_{js}, such that for some $\delta > 0$, $\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\alpha_{ij}(|t-s|))^{\delta/(4+\delta)} = O(NT)$, and for the same δ, $E[\|v_{it}\|^{4+\delta}] < \infty$ uniformly in i and t.

   (b) Let $e_t = (e_{1t}, \ldots, e_{Nt})'$ be a martingale difference sequence. More precisely, with filtration $\mathcal{F}_{N,t} = \sigma(e_1, \ldots, e_t; v_1, \ldots, v_{t+1})$, suppose that $E[e_t|\mathcal{F}_{N,t-1}] = 0$ almost surely (a.s.) and $E[e_t e_t'|\mathcal{F}_{N,t-1}] = (\sigma_e(i,j))_{NN} =: \Sigma_e$ a.s., where Σ_e is a constant matrix independent of t, $\sum_{i=1}^N \sum_{j=1}^N |\sigma_e(i,j)| = O(N)$ and $\sigma_e(i,i) = \sigma_e^2$. Meanwhile, $\sup_{1\le i\le N, 1\le t\le T} E[e_{it}^4|\mathcal{F}_{N,t-1}] < \infty$. Let $\Sigma_{v,e} = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \sum_{j=1}^N E[v_{i1}v_{j1}']\sigma_e(i,j)$ and $\Sigma_{v,e}$ is positive definite.

(c) i. $\sum_{i=1}^N \sum_{j=1}^N \sum_{t_1=1}^T \sum_{t_2=1}^T \sum_{t_3=1}^T \sum_{t_4=1}^T E[v_{it_1} \otimes v_{it_2} \otimes v_{jt_3} \otimes v_{jt_4}] = O(NT^2)$.

ii. $\sum_{i=1}^N \sum_{j=1}^N \sum_{t_1=1}^T \sum_{t_2=1}^T \sum_{t_3=1}^T \sum_{t_4=1}^T E[v_{it_1}' e_{it_2} v_{jt_3} e_{jt_4}] = O(NT^2)$.

4. $\{\varepsilon_{ij}, i \in \mathbb{Z}^+, j \in \mathbb{Z}\}$ is independent of $\{(v_{i_1t_1}, e_{i_1t_1}), 1 \le i_1 \le N, 1 \le t_1 \le T\}$.

Assumption 2

1. There exists an integer $m > 1$, such that $x^{m-s}g^{(s)}(w) \in L^2(\mathbb{R})$ for $s = 0, 1, \ldots, m$. Moreover, for $j = 1, \ldots, d$, $\phi_j(w) \in L(\mathbb{R}) \cap L^2(\mathbb{R})$.

2. Let $k = \lfloor aT^\vartheta \rfloor$ with a constant $a > 0$ and $0 < \vartheta < \frac{1}{4}$. Also, $k/N \to 0$ as $(N,T) \to (\infty, \infty)$.

3 Asymptotic theory

We start from investigating $\hat{\beta}$. It follows from (2.3) that

$$\hat{\beta} - \beta_0 = (X'M_{\mathcal{Z}}X)^{-1} X'M_{\mathcal{Z}}\mathcal{E} + (X'M_{\mathcal{Z}}X)^{-1} X'M_{\mathcal{Z}}\gamma. \tag{3.1}$$

Observe that

$$\frac{1}{NT}X'M_{\mathcal{Z}}X = \frac{1}{NT}X'X - \frac{1}{N\sqrt{T}}X'\mathcal{Z}\left(\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z}\right)^{-1}\frac{1}{NT}\mathcal{Z}'X, \tag{3.2}$$

$$\frac{1}{NT}X'M_{\mathcal{Z}}\mathcal{E} = \frac{1}{NT}X'\mathcal{E} - \frac{1}{N\sqrt{T}}X'\mathcal{Z}\left(\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z}\right)^{-1}\frac{1}{NT}\mathcal{Z}'\mathcal{E}, \tag{3.3}$$

$$\frac{1}{NT}X'M_{\mathcal{Z}}\gamma = \frac{1}{NT}X'\gamma - \frac{1}{N\sqrt{T}}X'\mathcal{Z}\left(\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z}\right)^{-1}\frac{1}{NT}\mathcal{Z}'\gamma. \tag{3.4}$$

The consistency of $\hat{\beta} - \beta_0$ follows from Lemmas B.4–B.5 listed in Appendix B immediately and the normality can be achieved by further investigation on (3.2)–(3.4). We now state the first theorem of this paper.

Theorem 3.1. *Under Assumptions 1 and 2, as $(N,T) \to (\infty,\infty)$ jointly, $\hat{\beta}$ is consistent. If, in addition, $N/k^{m-1} \to 0$, then $\sqrt{NT}(\hat{\beta} - \beta_0) \to_D N(0, \Sigma_v^{-1}\Sigma_{v,e}\Sigma_v^{-1})$, where $\Sigma_{v,e}$ is defined in Assumption 1.3.b.*

Note that $\Sigma_{v,e}$ is the same as that in Theorem 1 of Chen et al. (2012*b*) and the discussion on the existence of $\Sigma_{v,e}$ can be found therein. Since our model is an extension of Gao and Phillips

(2013) to the panel data case, the rate of convergence given in Theorem 3.1 matches what Gao and Phillips (2013) obtain for the time series case. On the cross–sectional dimension, the optimal rate of convergence, $N^{-1/2}$, is also achieved. Thus, replacing the time trend in Chen et al. (2012*b*) with non–stationary time series processes do not affect the optimal rate of convergence of $\hat{\beta}$. Some other studies and discussions on panel data models including non–stationary time series (but not directly related to our model) can be seen in Bai et al. (2009) and Kapetanios et al. (2011). The condition N/k^{m-1} is similar to the one given in Theorem 2 of Newey (1997), Assumption 4.ii of Su and Jin (2012) and Assumption A5 of Chen et al. (2012*b*). The purpose of this restriction is to remove the truncation residual for us to establish the asymptotic normality. Since nonstationary times series regressors are introduced to our model, the proof of the asymptotic theory involves some new techniques, which are different from those used in the literature.

Before giving a consistent estimator for the asymptotic covariance matrix in Theorem 3.1, we show the consistency of $\hat{C}$ given in (2.3). Note that

$$\hat{C} - C = \left(\mathcal{Z}' M_X \mathcal{Z}\right)^{-1} \mathcal{Z}' M_X \gamma + \left(\mathcal{Z}' M_X \mathcal{Z}\right)^{-1} \mathcal{Z}' M_X \mathcal{E}. \tag{3.5}$$

In connection with Lemmas B.4–B.5 provided in the Appendix, we have the following lemma.

Lemma 3.1. *Under Assumptions 1 and 2, as $(N, T) \to (\infty, \infty)$ jointly*

$$\|\hat{C} - C\| = O_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T}}\right) + O_P\left(k^{-\frac{m-1}{2}}\right).$$

The proof is given in Appendix B. We now turn to consistent estimation on asymptotic covariance matrix in Theorem 3.1 in order to establish the confidence interval for $\hat{\beta}$. By (6) of Lemma B.4 below, $\hat{\Sigma}_v = \frac{1}{NT} X'X \to_P \Sigma_v$. Thus, we need only to focus on obtaining a consistent estimator for $\Sigma_{v,e}$. To do so, we have to impose some stronger assumptions, e.g. e_{it} is independent across i, which is in the same spirit as Corollary 3.1.ii and Theorem 3.3 of Gao and Phillips (2013) and will reduce $\Sigma_{v,e}$ to $\sigma_e^2 \Sigma_v^{-1}$. Define the estimator of σ_e^2 as

$$\hat{\sigma}_e^2 = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\tilde{Y}_{it} - \tilde{X}_{it}' \hat{\beta} - \tilde{Z}_k(u_{it})' \hat{C})^2. \tag{3.6}$$

Corollary 3.1. *Suppose that Assumptions 1 and 2 hold. (1) As $(N, T) \to (\infty, \infty)$ jointly, $\hat{\sigma}_e^2 \to_P \sigma_e^2$, where $\hat{\sigma}_e^2$ is denoted by (3.6). (2) Let e_{it} be independent across i. As $(N, T) \to (\infty, \infty)$ jointly, $\hat{\Sigma}_{v,e} \to_P \Sigma_{v,e}$, where $\hat{\Sigma}_{v,e} = \hat{\sigma}_e^2 \hat{\Sigma}_v^{-1}$ and $\hat{\Sigma}_v = \frac{1}{NT} X'X$.*

The proof of Corollary 3.1 is given in Appendix C of the supplementary document. Moreover, for $\forall w \in \mathbb{R}$, define the estimator of $g(w)$ as $\hat{g}(w) = Z_k(w)' \hat{C}$. After imposing some extra restrictions, the normality of $\hat{g}(w)$ can be achieved.

Theorem 3.2. *Under Assumptions 1 and 2,*

1. $\int(\hat{g}(w)-g(w))^2dw = O_P\left(\frac{k}{N\sqrt{T}}\right) + O_P\left(k^{-m+1}\right)$.

2. *Additionally, let*

   (1) $\sum_{i_1=1}^N\sum_{i_2=1}^N\sum_{i_3=1}^N\sum_{i_4=1}^N \left|E\left[e_{i_1t}e_{i_2t}e_{i_3t}e_{i_4t}|\mathcal{F}_{Nt-1}\right]\right| = O_P(N^2)$ *uniformly in* t, *and*

   (2) $k^2/N \to 0$ *and* $N^{1/2}T^{1/4}k^{-(m-1)/2} \to 0$.

   Then as $(N,T)\to(\infty,\infty)$ *jointly,* $\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}(\hat{g}(w)-g(w)) \to_D N(0,1)$, *where* $\sigma_k(w) = a_0^{-1}\sigma_e^2\|Z_k(w)\|^2$ *and* $a_0 = \sqrt{2/(\pi\rho^2)}(1+o(1))$ *with* $\rho = \sum_{j=0}^\infty \rho_j \neq 0$.

In the first result of Theorem 3.2, we establish a rate of convergence for the integrated mean squared error. For the second result of Theorem 3.2, two stronger restrictions are needed: Condition (1) is in the same spirit of (3.3) and (3.4) in Chen et al. (2012*a*), wherein all the relevant discussions and examples can be found; Condition (2) on the sharper bound for k is due to the development of (B.15) (see Appendix B for details). It is interesting to see that the cross–sectional dependence of the error terms does not play a role in the asymptotic variance (c.f. $\sigma_k(w) = a_0^{-1}\sigma_e^2\|Z_k(w)\|^2$). A short explanation is that in the derivation of the variance for the term on RHS of equation (B.15), $E[Z_k(w)'Z_k(u_{it})Z_k(u_{jt})'Z_k(w)]$ will attenuate at rate t^{-1} for $i\neq j$.

Moreover, notice that $\|Z_k(w)\|^2 = O(k)$ uniformly by Lemma B.1. Thus, the rate of convergence for the normality is essentially $\sqrt{k^{-1}N\sqrt{T}}$, which is equivalent to the rate obtained by using kernel estimation method $\sqrt{hN\sqrt{T}}$, where h is the bandwidth parameter. The condition $N^{1/2}T^{1/4}k^{-(m-1)/2}$ is in line with the same spirit of N/k^{m-1} provided in Theorem 3.1. The higher–order smoothness required here is due to the development of (B.15).

Notice based on the convergence that $\frac{1}{N\sqrt{T}}\sum_{i=1}^N\sum_{t=1}^T \mathscr{H}_0^2(x_{it}) \to_P a_0$ in Lemma B.5 and $\hat{\sigma}_e^2 \to_P \sigma_e^2$ in Corollary 3.1, $\hat{\sigma}_k(w)$, the estimator of $\sigma_k(w)$, is easily obtained and thus the hypothesis test on $\hat{g}(w)$ for $\forall w\in\mathbb{R}$ can be conducted from the second result of Theorem 3.2. In the next section, we provide some related discussion before presenting the finite sample studies using both simulated and real data examples.

4 Some extensions and discussions

In the above study, we have completely ruled out the cases where $\phi_j(w)$ for $j=1,\ldots,d$ and $g(w)$ are non–integrable. The study on (1.1) is fundamental and can provide many basic results for the cases where $g(w)$ includes a non–integrable term. For example, let $g(w) = w + g_1(w)$, where $g_1(w)$ is an integrable function on $\mathbb{R}$. In this case, the model becomes

$$\begin{aligned} y_{it} =& x_{it}'\beta_0 + u_{it} + g_1(u_{it}) + \omega_i + e_{it}, \\ x_{it} =& \phi(u_{it}) + \lambda_i + v_{it}. \end{aligned} \tag{4.1}$$

Then simple transformation shows that we can rewrite (4.1) as

$$\begin{aligned} y_{1,it} =& x_{it}'\beta_0 + g_1(u_{it}) + \omega_i + e_{it}, \\ x_{it} =& \phi(u_{it}) + \lambda_i + v_{it}, \end{aligned} \tag{4.2}$$

where $y_{1,it} = y_{it} - u_{it}$. Since both y_{it} and u_{it} are observable, $y_{1,it}$ can be treated as given. In this case, model (4.1) is reduced to (1.1).

We now turn to the structure of x_{it}. Consider a simple partially linear model of the form:

$$\begin{aligned} y_{it} =& x_{it}'\beta_0 + g(u_{it}) + \omega_i + e_{it}, \\ x_{it} =& u_{it} + \lambda_i + v_{it}. \end{aligned} \tag{4.3}$$

After taking first difference, it is easy to obtain that

$$\begin{aligned} \Delta y_{it} &= (\Delta x_{it})'\beta_0 + g(u_{it}) - g(u_{i,t-1}) + \Delta e_{it} \\ &= (\Delta x_{it})'\beta_0 + (Z_k(u_{it}) - Z_k(u_{i,t-1}))' C + \tilde{e}_{it} \end{aligned} \tag{4.4}$$

where $\tilde{e}_{it} = \gamma_k(u_{it}) - \gamma_k(u_{i,t-1}) + \Delta e_{it}$. Notice that (4.4) does not include fixed effects, so it is a simpler version of (1.1). In order to obtain consistent estimators for β_0 and C, we can carry out the similar procedure as the previous sections without using within transformation.

There are also some limitations in this study. Assumption 1.3.b has excluded the case where $E[e_{it}v_{it}] \neq 0$. For example, we cannot allow the error term to have a form like $e_{it} = \psi(v_{it}) + \epsilon_{it}$. This is in the same spirit as Assumptions 2 and 4 of Pesaran (2006), Assumption D of Bai (2009) and Assumption A.4 of Chen et al. (2012*b*). To introduce some endogeneity between e_{it} and v_{it}, new techniques similar to those developed by Dong and Gao (2014) may be needed. When $E[e_{it}v_{it}] = 0$, we can allow $e_{it} = \psi(v_{it}) \cdot \epsilon_{it}$, where ϵ_{it} is independent of v_{it}. In this sense, v_{it} can partially be the driving force of e_{it} by having an impact on its variance. A detailed example is given in the Monte Carlo study below.

5 Numerical Study

This section provides the results of a simple Monte Carlo study and an empirical case study by looking into Balassa–Samuelson model. In the simulation study, the biases and root of mean squared errors (RMSEs) are reported. As we can see, biases are quite small and RMSEs decrease as both N and T increase. The empirical case study suggests that model (1.1) outperforms the traditional panel data model used for investigating Balassa–Samuelson model.

5.1 Monte Carlo simulation

In Monte Carlo study, the data generating process (DGP) is as follows.

$$\begin{aligned} y_{it} &= x_{it}'\beta_0 + (1+u_{it}^2)\exp(-u_{it}^2) + \omega_i + e_{it}, \\ x_{it} &= \big((1+u_{it}+u_{it}^2)\exp(-u_{it}^2)\big)\otimes 1_d + \lambda_i + v_{it}, \\ e_{it} &= \gamma_i f_t(1+f_{t-1}) + v_{it}'\beta_0\varepsilon_{it} \end{aligned}$$

where $\beta_0 = (1,2)'$ and $d = 2$. In this DGP, the error term e_{it} depends on the information from the past, f_{t-1}, and the information from the current time period, v_{it}.

For each i, $u_{i1} \sim$ i.i.d. $N(0,1)$ and $u_{it} = u_{i,t-1} +$ i.i.d. $N(0,1)$ for $t = 2,\dots,T$. For the factor loadings, $(\gamma_1,\dots,\gamma_N)' \sim N(0,\Sigma_\gamma)$, where the $(i,j)^{th}$ element of Σ_γ is $0.5^{|i-j|}$. For the factors, $f_t \sim$ i.i.d. $N(0,1)$ for each t. The error terms $\varepsilon_{it} \sim$ i.i.d. $N(0,1)$. For each t,

$$(v_{1t},\dots,v_{Nt})' = 0.5(v_{1,t-1},\dots,v_{N,t-1})' + N(0,\Sigma_v),$$

where the $(i,j)^{th}$ element of Σ_v is $0.3^{|i-j|}$. For the fixed effects, $\omega_i \sim N((1+u_{i1}+u_{i1}^2),1)$ and $\lambda_i \sim N(1_d, I_d)$, so w_i is certainly correlated with the regressor x_{it}.

Based on the above, the cross–sectional dependence comes into the system through both the error terms e_{it} and v_{it}. e_{it} certainly satisfies the martingale condition and slightly violates the requirements of Assumption 1.3.b on covariances, but it does not affect the accuracy of the estimators as shown later. In order to make sure that the Assumption 2.2 is satisfied, the truncation parameter is chosen as $k = \lfloor 3.3\cdot T^{1/7}\rfloor$. For each replication, we record the bias and squared error as: $bias = \hat\beta_j - \beta_{j0}$ and $se = (\hat\beta_j - \beta_{0j})^2$ for $j = 1,2$, where $\hat\beta_j$ denotes the estimate of β_{0j} and β_{0j} is the j^{th} element of β_0. After 1000 replications, we report the mean of these biases and the root of the mean of these squared errors, which are labeled as Bias and RMSE in Table 1. It is evident that in Table 1 the biases decrease to zero very quick, and the RMSEs decrease as both N and T increase. Though both N and T start from 10, the sample size given by the product of NT is sufficient to obtain accurate estimation for the parameters.

	$T \setminus N$	$\hat{\beta}_1$ 10	40	80	$\hat{\beta}_2$ 10	40	80
Bias	10	0.005	0.003	0.001	0.017	0.002	0.004
	40	-0.005	0.004	0.000	-0.010	0.002	-0.002
	80	-0.003	0.001	-0.001	0.004	0.000	0.000
RMSE	10	0.334	0.162	0.115	0.404	0.194	0.141
	40	0.150	0.077	0.056	0.199	0.099	0.069
	80	0.107	0.052	0.037	0.136	0.069	0.049

Table 1: Bias and RMSE

5.2 Empirical study

The Balassa–Samuelson model implies that countries with a relatively low ratio of tradables to nontradables productivity will have a depreciated real exchange rate, which can be evaluated by calculating the gap between a purchasing power parity (PPP)–based U.S. dollar exchange rate and the nominal U.S. dollar exchange rate. The PPP–based exchange rate measures how many goods the domestic currency buys within the country relative to the U.S. as numéraire country, while the nominal U.S. dollar exchange rate measures how many U.S. dollars the domestic currency buys in the foreign exchange market. Specifically, we consider equation (1) of de Boeck and Slok (2006), i.e. (5.1) provided below. A very detailed description can be found therein.

$$\ln\left(\frac{ppp_{it}}{ne_{it}}\right) = \beta \cdot \ln pgp_{it} + \gamma_i + \varepsilon_{it}, \tag{5.1}$$

where ppp_{it}=PPP–based U.S. dollar exchange rate at (i, t), ne_{it}=nominal U.S. dollar exchange rate at (i, t), pgp_{it}=PPP GDP per capita at (i, t).

However, running OLS regression on the above linear model by using the data set provided below gives a R^2 smaller than 15%. A modified form of the linear model (5.1) is given by

$$\ln ppp_{it} = \beta \cdot \ln pgp_{it} + \alpha \cdot \ln ne_{it} + \gamma_i + \varepsilon_{it}. \tag{5.2}$$

This section proposes a partially linear model of the form:

$$\ln ppp_{it} = \beta \cdot \ln pgp_{it} + g(ne_{it}) + \gamma_i + \varepsilon_{it}, \tag{5.3}$$

where $g(\cdot)$ is an unknown function. We therefore compare models (5.1)–(5.3), referred to as LM1, LM2 and PM, respectively, for brevity.

For this study, the yearly data is collected from Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 7.1, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, July 2012. We choose the time period 1950–2010 and focus on OECD countries only. Since not all OECD countries have the data recorded for the whole period, we simply remove those countries that have the missing data to ensure a balanced panel data set. Notice that most of the countries left have reasonable exchange rates during this period, which vary between 0 to 5. However, some countries' exchange rates have dramatic changes and have a clear signal on structural break, for example the exchange rate of Iceland is always less than 1 before 1975 and increases dramatically to 122 after that. Thus, we also remove some countries, whose exchange rates act as outliers to rest of the countries. It then leaves us with 17 countries, which are Australia, Austria, Belgium, Canada, Finland, France, Ireland, Israel, Italy, Luxembourg, Mexico, Netherlands, New Zealand, Portugal, Spain, Switzerland, Turkey and United Kingdom.

Before any further investigation, we examine if the nominal U.S. dollar exchange rates of all these countries have unit roots. To do so, we carry on the Augmented Dickey–Fuller (ADF) test for the time series $\{ne_{i1}, \cdots, ne_{iT}\}$ for $i = 1, \ldots, 17$ and report the p–values of the ADF tests below. According to the report, except for Switzerland all other 16 countries have unit root in nominal U.S. dollar exchange rates based on 5% significant level, so we remove Switzerland from the data set.

	Australia	Austria	Belgium	Canada	Finland	France
p–value	0.61	0.29	0.29	0.53	0.70	0.60
	Ireland	Israel	Italy	Luxembourg	Netherlands	NZ
p–value	0.68	0.97	0.77	0.29	0.11	0.68
	Portugal	Spain	Switzerland	Turkey	UK	
p–value	0.85	0.78	0.02	0.99	0.83	

Table 2: P-value of Augmented Dickey-Fuller Test

For the data set used in this study, ppp_{it} mainly varies between 2.5 and 5.5; pgp_{it} roughly varies between 5 and 11; except Israel, ne_{it} fluctuates between 0 and 2. Due to the limitation of space, we use Canada as an example to illustrate how these three variables change across time.

To get In–MSE, all the data collected above ($i = 1, \ldots, 16$ and $t = 1, \ldots, 61$) are used to run the regression in order to get $\hat{\beta}_{\text{In}}$ and $\hat{C}_{\text{In}}$. Then In–MSE is given by

$$\text{In–MSE} = \frac{1}{16 \times 61} \sum_{i=1}^{16} \sum_{t=1}^{61} (\tilde{Y}_{it} - \tilde{X}_{it}' \hat{\beta}_{\text{In}} - \tilde{Z}_k(u_{it})' \hat{C}_{\text{In}})^2,$$

where $\tilde{Y}_{it} = Y_{it} - \frac{1}{61} \sum_{t=1}^{61} Y_{it}$; $\tilde{X}_{it}$ and $\tilde{Z}_k(u_{it})$ are defined in the same fashion.

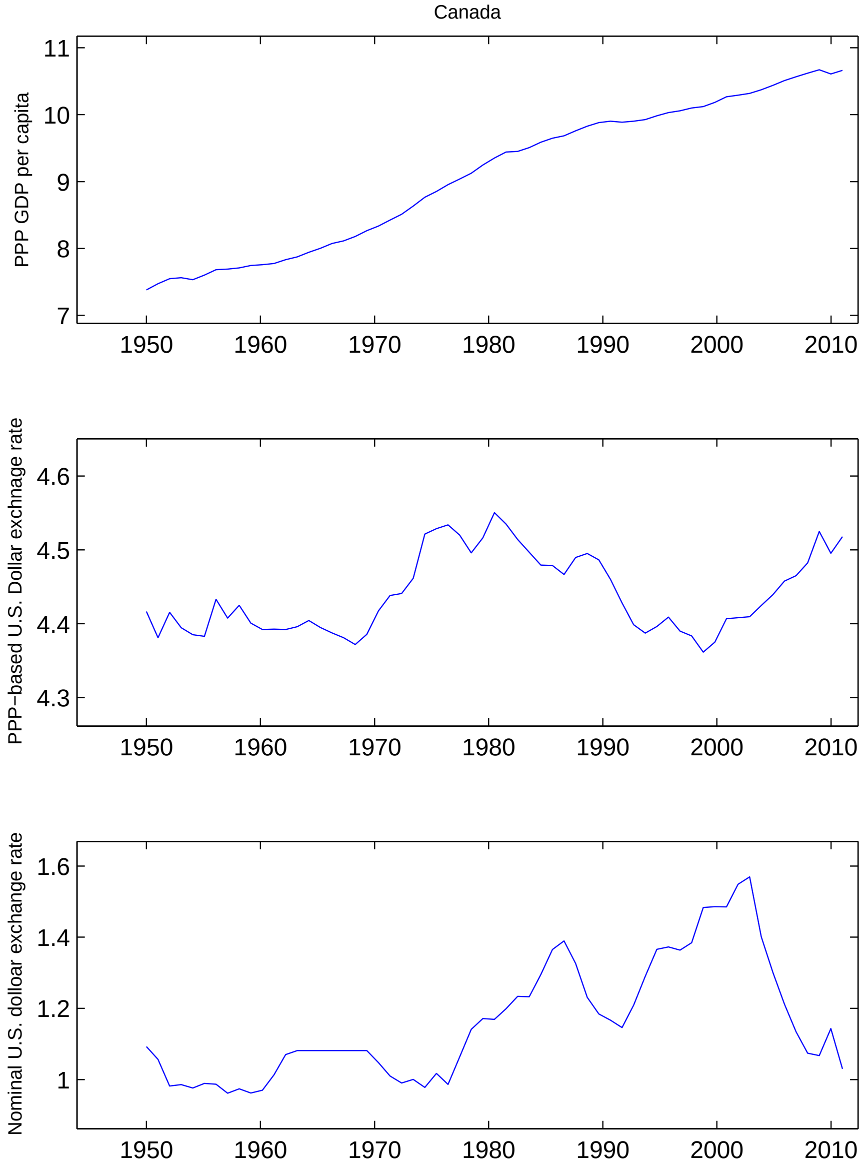

Figure 1: Canada

To get Out–MSE, only part of data collected above ($i = 1, \ldots, 16$, $t = 1, \ldots, \tilde{T}$ and $\tilde{T} = 56, \ldots, 60$) are used to estimate $\hat{\beta}_{\text{Out},\tilde{T}}$ and $\hat{C}_{\text{Out},\tilde{T}}$ in order to forecast $\tilde{Y}_{i,\tilde{T}+1}$. Then Out–MSE is obtained as

$$\text{Out–MSE} = \frac{1}{16 \times (61-56)} \sum_{i=1}^{16} \sum_{\tilde{T}=56}^{60} (\tilde{Y}_{i,\tilde{T}+1} - \tilde{X}'_{i,\tilde{T}+1}\hat{\beta}_{\text{Out},\tilde{T}} - \tilde{Z}_k(u_{i,\tilde{T}+1})'\hat{C}_{\text{Out},\tilde{T}})^2,$$

where $\tilde{Y}_{i,\tilde{T}+1} = Y_{i,\tilde{T}+1} - \frac{1}{\tilde{T}+1}\sum_{t=1}^{\tilde{T}+1} Y_{it}$; $\tilde{X}_{it}$ and $\tilde{Z}_k(u_{it})$ are defined in the same fashion.

Even though we treat $g(w)$ as an unknown function and have less information for (5.3),

Table 4 shows that the estimates from partially linear model by taking into account the non–stationarity of nominal U.S. dollar exchange rates outperform the estimates from linear models.

PM	LM1	LM2	
$\hat{\beta}$ = 0.075	$\hat{\beta}$ = -0.484	$\hat{\beta}$ = 0.106	$\hat{\alpha}$ = -0.011
(0.004)	(0.044)	(0.004)	(0.002)

Table 3: Estimation results for models (5.1)–(5.3)

	PM ($k = 6$)	LM1	LM2
In–MSE	0.01180	2.49943	0.01449
Out–MSE	0.01273	5.50020	0.01421

Table 4: In–MSE and Out–MSE

For the partially linear panel data model (5.3), our comparisons based on the in sample mean squared errors (In–MSE) and rolling out sample mean squared errors (Out–MSE) suggest using $k = 6$ as the truncation parameter. As a comparison, the estimates of within OLS estimates for (5.1) and (5.2) are also reported. We now use the partially linear model as an example to demonstrate how to calculate these two values.

The coefficients of the basis functions are (7.78, -19.58, 26.81, -22.84, 11.85, -3.01), which imply that the estimated unknown function is

$$\hat{g}(w) = 7.78\mathscr{H}_0(w) - 19.58\mathscr{H}_1(w) + 26.81\mathscr{H}_2(w) - 22.84\mathscr{H}_3(w) + 11.85\mathscr{H}_4(w) - 3.01\mathscr{H}_5(w).$$

Moreover, we plot $\hat{g}(w)$ and its confidence interval in Figure 2. Since most of ne_{it}'s are between 0 and 2, we only report $\hat{g}(w)$ on the interval $[0, 2]$. The dash–dot line in the mid represents the estimated unknown function, $\hat{g}(w)$, and the two dashed lines represent the 95% confidence interval curves.

Due to the limit of space, we report the estimated PPP–based U.S. dollar exchange rate for Belgium only in Figure 3. The dash–dot line includes the observable values and the solid line includes the estimated values. Figure 3 indicates that including more relevant explanatory variables may be necessary for improving the performance of Balassa–Samuelson model. However, this is beyond the scope of this paper. We will leave this for future research.

6 Conclusions

In this paper, we have established the estimate for a group of partially linear panel data models with non–stationarity and cross–sectional dependence. Spatial error structure analysis

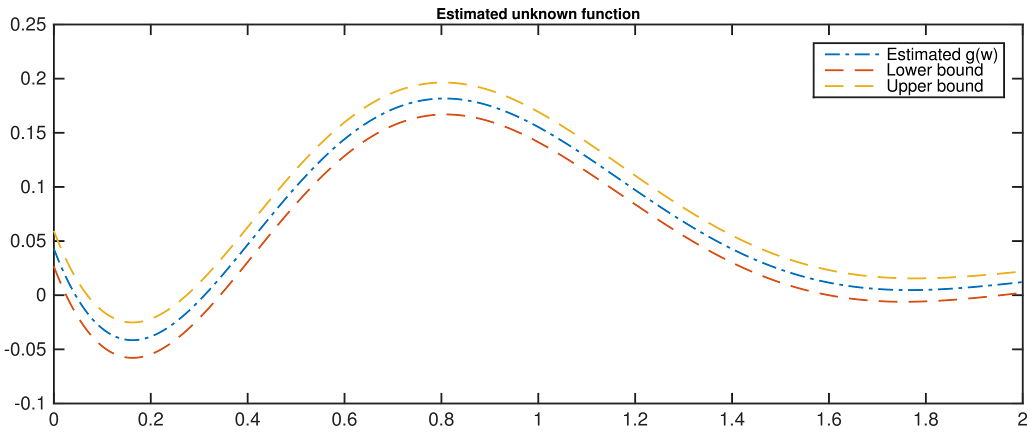

Figure 2: Estimated unknown function

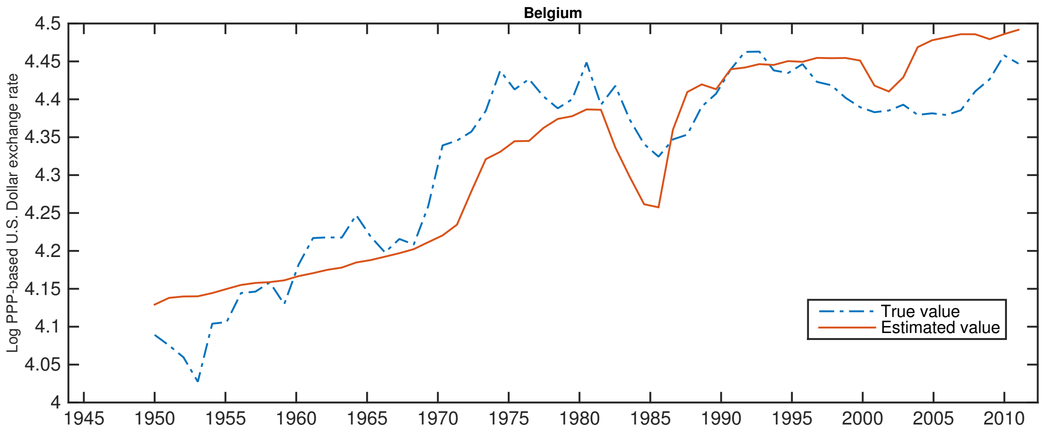

Figure 3: Estimated PPP-based U.S. dollar exchange rate (Belgium)

technique has been used to measure the correlation among individuals. An asymptotic theory has been established for the estimators. Particularly, we do not impose any assumptions on the fixed effects, so they can be used to deal with the models with unobservable heterogeneity. More importantly, new findings include the significant difference in asymptotical theory for times series and panel data sets. The finite sample properties are demonstrated through Monte Carlo experiment and a real data example of Balassa–Samuelson Model. The possible extensions and limitations of our model have been discussed in details and they will guide our future research projects.

Appendix A: Discussion of assumptions

Assumption 1: Assumptions 1.1 and 1.2 are standard in the time series literature and imply that the non–stationary time series processes $\{u_{i1}, \ldots, u_{iT}\}$'s are i.i.d. across i (c.f. Assumption 1.i of Phillips and Moon (1999), where the coefficients of ε_{ij} are treated as i.i.d. random variables.). Notice that the coefficients ρ_j in Assumption 1.2 can also be written in a matrix form if u_{it} is chosen as a vector.

Assumption 1.3.a is in the same spirit as Assumption C of Bai (2009) and Assumption A2 and A4 of Chen et al. (2012*b*). On the cross–section dimension, it is also similar to the set–up on spatial error structure in Pesaran and Tosetti (2011). Therefore, it certainly allows the cross–sectional dependence of the error terms to come in the model. On the time dimension, it entails that only the stationary case is considered. Specifically, the mixing coefficient $\alpha_{ij}(|t-s|)$ is used to measure the relationship between individuals at different time periods, i.e. the relationship between v_{it} and v_{js}. Two examples are given below to demonstrate this assumption is reasonable:

- It can easily be seen that Assumption 1.3.a holds if v_{it} is i.i.d. over i and t.

- We now use a factor model structure to show that Assumption 1.3.a is verifiable. Suppose that $v_{it} = \gamma_i f_t + \varepsilon_{it}$, where all variables are scalars and ε_{it} is i.i.d. sequence over i and t with mean zero. Simple algebra shows that the coefficient $\alpha_{ij}(|t-s|)$ reduces to $\alpha_{ij} \cdot b(|t-s|)$, in which $\alpha_{ij} = E[\gamma_i \gamma_j]$ and $b(|t-s|)$ is the α–mixing coefficient of the factor time series $\{f_1, \ldots, f_T\}$. If f_t is a strictly stationary α–mixing process and γ_i is a functional coefficient which converges to 0 at a certain rate as i increases, Assumption 1.3.a can easily be verified. More details and useful empirical examples can be found under Assumption A2 in Chen et al. (2012*b*).

Assumption 1.3.b is similar to Assumption 1.3.a, but focuses on the cross–sectional dimension of the error term e_{it}. It is the same as Assumption A.4 of Chen et al. (2012*b*) and allows for weak endogeneity between regressors and error terms through v_{it} and e_{it}.

The Assumption 1.3.c is a simpler version of (A.18) in Chen et al. (2012*a*). For the first equation of Assumption 1.3.c, a very detailed proof and relevant discussion can be found on page 17 of Chen et al. (2012*a*). The second equation is in line with the spirit of Assumption 3.2.ii of Gao and Phillips (2013) and can be easily verified if e_{it} is independent of v_{js}.

Within transformation allows us to relax the identification restrictions (1.3) of Chen et al. (2012*b*) and (1.2) of Chen et al. (2013), i.e. $\sum_{i=1}^N \omega_i = 0$ and $\sum_{i=1}^N \lambda_i = 0$. Notice that we do not impose any conditions on ω_i and λ_i, so they can be correlated with any variables arbitrarily.

Notice that the results of this paper are not achieved in the richer probability space (c.f. Park and Phillips (2000) and Park and Phillips (2001) for the discussion on the richer probability space) due to that we avoid using the local time process in the development of Lemma B.5. In this sense, the results under the panel data setting are stronger than those achieved in the time series setting.

Assumption 2: Assumption 2.1 (c.f. Dong et al. (2014)) ensures that the approximation of the unknown functions $g(w)$ by an orthogonal expansion can have a fast rate. Assumption 2.2 puts restrictions on the truncated parameter k, N and T, so that they go to infinity at appropriate rates. The requirement of $k = \lfloor aT^{\vartheta} \rfloor$ for $0 < \vartheta < \frac{1}{4}$ is consistent with the set–up for time series data (c.f. Dong et al. (2014)) and similar to Assumption A3 of Chen et al. (2012*a*). The requirement of $k/N \to 0$ is consistent with the case that sieve estimation is used in panel data literature (c.f. Su and Jin (2012)). These two restrictions further imply that $T^{\vartheta}/N \to 0$ for the ϑ given above. The similar conditions and more discussions under panel data settings can be seen in Su and Jin (2012), Chen et al. (2012*b*) and Chen et al. (2012*a*).

Appendix B: Proof of the main results

We first give some necessary lemmas for the proofs of the main results before the proofs of the lemmas are given in Appendix C of the supplementary document.

Lemma B.1. *Suppose that $g(w)$ is differentiable on $\mathbb{R}$ and $x^{m-j}g^{(j)}(w) \in L^2(\mathbb{R})$ for $j = 0, 1, \ldots, m$ and $m \geq 1$. For the expansion (2.1), the following results hold:*
(1) $\int w^2 \mathscr{H}_n^2(w)dw = n + 1/2$; (2) $\max_w |\gamma_k(w)| = O(1)k^{-(m-1)/2-1/12}$; (3) $\int \gamma_k^2(w)dw = O(1)k^{-m}$; (4) $\int \|Z_k(w)\|dw = O(1)k^{11/12}$; (5) $\int \|Z_k(w)\|^2 dw = k$; (6) $\|Z_k(w)\|^2 = O(1)k$ uniformly on $\mathbb{R}$; (7) $\int |\gamma_k(w)|dw = O(1)k^{-m/2+11/12}$; (8) $\int |\mathscr{H}_n(w)|dw = O(1)n^{5/12}$; (9) $\int |x|^2\|Z_k(x)\|^2 dx = O(1)k^2$.

The proof of Lemma B.1 is exactly the same as that in Lemma A.1 of Dong et al. (2014).

Lemma B.2. *Under Assumptions 1 and 2, as $(N, T) \to (\infty, \infty)$ jointly,*
(1) $\left\|\frac{1}{N\sqrt{T}}\sum_{i=1}^N\sum_{t=1}^T Z_k(u_{it})\gamma_k(u_{it})\right\| = O_P(k^{-(m-1)/2})$; (2) $\left\|\frac{1}{N\sqrt{T}}\sum_{i=1}^N\sum_{t=1}^T \phi(u_{it})Z_k(u_{it})'\right\| = O_P(1)$; (3) $\left\|\frac{1}{N\sqrt{T}}\sum_{i=1}^N\sum_{t=1}^T v_{it}Z_k(u_{it})'\right\| = O_P\left(\sqrt{\frac{k}{N}}\right)$; (4) $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T v_{it}v_{it}' = \Sigma_v + O_P\left(\frac{1}{\sqrt{NT}}\right)$; (5) $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T v_{it}\phi(u_{it})' = O_P\left(\frac{1}{\sqrt{NT}}\right)$; (6) $\frac{1}{N\sqrt{T}}\sum_{i=1}^N\sum_{t=1}^T Z_k(u_{it})e_{it} = O_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T}}\right)$; (7) $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \phi(u_{it})e_{it} = O_P\left(\frac{1}{\sqrt{N}\sqrt[4]{T^3}}\right)$; (8) $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \phi(u_{it})\gamma_k(u_{it}) = O_P\left(\frac{1}{\sqrt{k^mT}}\right)$; (9) $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T v_{it}\gamma_k(u_{it}) = O_P\left(\frac{k^{-m/2+5/12}}{\sqrt{NT}}\right)$.

The proof of Lemma B.2 is provided later in Appendix C.

Lemma B.3. *For two non–singular symmetric matrices A, B with same dimensions $k \times k$, where k tends to ∞. Suppose that their minimum eigenvalues satisfy that $\lambda_{min}(A) > 0$ and $\lambda_{min}(B) > 0$ uniformly in k. Then $\left\|A^{-1} - B^{-1}\right\|^2 \leq \lambda_{min}^{-2}(A) \cdot \lambda_{min}^{-2}(B)\|A - B\|^2$.*

The proof of Lemma B.3 is provided later in Appendix C.

Lemma B.4. *Let Assumptions 1 and 2 hold. As $(N, T) \to (\infty, \infty)$ jointly, (1) $\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}\mathcal{E}\right\| = O_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T}}\right)$; (2) $\left\|\frac{1}{N\sqrt{T}}X'\mathcal{Z}\right\| = O_P(1)$; (3) $\frac{1}{NT}X'\mathcal{E} = O_P\left(\frac{1}{\sqrt{NT}}\right)$; (4) $\frac{1}{NT}X'\gamma = O_P\left(\frac{1}{\sqrt{k^mT}}\right)$; (5) $\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'\gamma\right\| = O_P(k^{-(m-1)/2})$; (6) $\frac{1}{NT}X'X \to_P \Sigma_v$.*

The proof of Lemma B.4 is provided later in Appendix C.

Lemma B.5. *Suppose that Assumptions 1.1, 1.2 and 2.2 hold. As (N, T) go to (∞, ∞) jointly,*

(1) $\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z} - a_0I_k\right\| \to_P 0$, where $a_0 = \sqrt{\frac{2}{\pi|\rho|^2}}(1 + o(1))$.

(2) Suppose further that $\frac{k^2}{N} \to 0$. Then $\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z} - a_0I_k\right\| = o_P(k^{-1/2})$.

In the above lemma, the first result is of general interest and can be used in sieve estimation for panel data models where nonstationary time series are involved, while the second one establishes the convergence rate with a harsher requirement on k and N, which will be used in the proof of Theorem 3.2. The proof of Lemma B.5 is provided later in Appendix C.

We are now ready to provide the proofs for the mains results of this paper.

Proof of Theorem 3.1: By Lemma B.5, we have uniformly in k

$$\begin{aligned}\lambda_{min}\left(\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z}\right) &= \min_{\|\mu\|=1}\left\{\mu' a_0 I_k \mu + \mu'\left(\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z} - a_0 I_k\right)\mu\right\} \\ &\geq a_0 - \left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z} - a_0 I_k\right\| \geq \frac{1}{2}a_0(1+o_P(1)). \end{aligned} \tag{B.1}$$

Therefore, by Lemma B.3

$$\left\|\left(\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z}\right)^{-1} - \frac{1}{a_0}I_k\right\| \leq \frac{2(1+o_P(1))}{a_0^2}\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z} - a_0 I_k\right\| = o_P(1). \tag{B.2}$$

For consistency, we consider (3.2)–(3.4) respectively below. Start from (3.2).

$$\begin{aligned}\frac{1}{NT}X'M_{\mathcal{Z}}X &= \frac{1}{NT}X'X - \frac{1}{N\sqrt{T}}X'\mathcal{Z}a_0^{-1}I_k\frac{1}{NT}\mathcal{Z}'X \\ &\quad + \frac{1}{N\sqrt{T}}X'\mathcal{Z}\left[a_0^{-1}I_k - \left(\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z}\right)^{-1}\right]\frac{1}{NT}\mathcal{Z}'X. \end{aligned} \tag{B.3}$$

Notice that

$$\begin{aligned}&\left\|\frac{1}{N\sqrt{T}}X'\mathcal{Z}\left[a_0^{-1}I_k - \left(\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z}\right)^{-1}\right]\frac{1}{NT}\mathcal{Z}'X\right\| \\ &\leq \frac{1}{\sqrt{T}}\left\|\frac{1}{N\sqrt{T}}X'\mathcal{Z}\right\|^2\left\|a_0^{-1}I_k - \left(\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z}\right)^{-1}\right\| = o_P\left(\frac{1}{\sqrt{T}}\right),\end{aligned}$$

where the last line follows from (B.2) and (2) of Lemma B.4. Similarly, by (2) of Lemma B.4,

$$\left\|\frac{1}{N\sqrt{T}}X'\mathcal{Z}a_0^{-1}I_k\frac{1}{NT}\mathcal{Z}'X\right\| \leq O(1)\left\|\frac{1}{N\sqrt{T}}X'\mathcal{Z}\right\|\left\|\frac{1}{NT}\mathcal{Z}'X\right\| = O_P\left(\frac{1}{\sqrt{T}}\right).$$

In connection with (6) of Lemma B.4, we can further write

$$\frac{1}{NT}X'M_{\mathcal{Z}}X = \frac{1}{NT}X'X + O_P\left(\frac{1}{\sqrt{T}}\right) \to_P \Sigma_v. \tag{B.4}$$

For (3.3), write

$$\begin{aligned}\frac{1}{NT}X'M_{\mathcal{Z}}\mathcal{E} &= \frac{1}{NT}X'\mathcal{E} - \frac{1}{N\sqrt{T}}X'\mathcal{Z}a_0^{-1}I_k\frac{1}{NT}\mathcal{Z}'\mathcal{E} \\ &\quad + \frac{1}{N\sqrt{T}}X'\mathcal{Z}\left[a_0^{-1}I_k - \left(\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z}\right)^{-1}\right]\frac{1}{NT}\mathcal{Z}'\mathcal{E}. \end{aligned} \tag{B.5}$$

Notice that

$$\left\|\frac{1}{N\sqrt{T}}X'\mathcal{Z}\left[a_0^{-1}I_k - \left(\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z}\right)^{-1}\right]\frac{1}{NT}\mathcal{Z}'\mathcal{E}\right\|$$

$$\leq \left\| \frac{1}{N\sqrt{T}} X'\mathcal{Z} \right\| \left\| a_0^{-1} I_k - \left(\frac{1}{N\sqrt{T}} \mathcal{Z}'\mathcal{Z} \right)^{-1} \right\| \left\| \frac{1}{NT} \mathcal{Z}'\mathcal{E} \right\| = o_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T^3}} \right),$$

where the last line follows from (B.2) and (1)–(2) of Lemma B.4. Similarly, by (1)–(2) of Lemma B.4,

$$\left\| \frac{1}{N\sqrt{T}} X'\mathcal{Z} a_0^{-1} I_k \frac{1}{NT} \mathcal{Z}'\mathcal{E} \right\| \leq O(1) \left\| \frac{1}{N\sqrt{T}} X'\mathcal{Z} \right\| \left\| \frac{1}{NT} \mathcal{Z}'\mathcal{E} \right\| = O_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T^3}} \right).$$

Then we can further write (B.5) as

$$\frac{1}{NT} X' M_{\mathcal{Z}} \mathcal{E} = \frac{1}{NT} X'\mathcal{E} + O_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T^3}} \right) = O_P\left(\frac{1}{\sqrt{NT}} \right), \tag{B.6}$$

where the second equality follows from (3) of Lemma B.4 and Assumption 2.2.

For (3.4), write

$$\begin{aligned} \frac{1}{NT} X' M_{\mathcal{Z}} \gamma = & \frac{1}{NT} X'\gamma - \frac{1}{N\sqrt{T}} X'\mathcal{Z} a_0^{-1} I_k \frac{1}{NT} \mathcal{Z}'\gamma \\ & + \frac{1}{N\sqrt{T}} X'\mathcal{Z} \left[a_0^{-1} I_k - \left(\frac{1}{N\sqrt{T}} \mathcal{Z}'\mathcal{Z} \right)^{-1} \right] \frac{1}{NT} \mathcal{Z}'\gamma. \end{aligned} \tag{B.7}$$

Notice that

$$\begin{aligned} & \left\| \frac{1}{N\sqrt{T}} X'\mathcal{Z} \left[a_0^{-1} I_k - \left(\frac{1}{N\sqrt{T}} \mathcal{Z}'\mathcal{Z} \right)^{-1} \right] \frac{1}{NT} \mathcal{Z}'\gamma \right\| \\ \leq & \left\| \frac{1}{N\sqrt{T}} X'\mathcal{Z} \right\| \left\| a_0^{-1} I_k - \left(\frac{1}{N\sqrt{T}} \mathcal{Z}'\mathcal{Z} \right)^{-1} \right\| \left\| \frac{1}{NT} \mathcal{Z}'\gamma \right\| = o_P\left(\frac{1}{\sqrt{k^{m-1}T}} \right), \end{aligned}$$

where the last line follows from (B.2), and (2) and (5) of Lemma B.4. Similarly, by (2) and (5) of Lemma B.4,

$$\left\| \frac{1}{N\sqrt{T}} X'\mathcal{Z} a_0^{-1} I_k \frac{1}{NT} \mathcal{Z}'\gamma \right\| \leq O(1) \left\| \frac{1}{N\sqrt{T}} X'\mathcal{Z} \right\| \left\| \frac{1}{NT} \mathcal{Z}'\gamma \right\| = O_P\left(\frac{1}{\sqrt{k^{m-1}T}} \right).$$

Then we can further write (B.7) as

$$\frac{1}{NT} X' M_{\mathcal{Z}} \gamma = \frac{1}{NT} X'\gamma + O_P\left(\frac{1}{\sqrt{k^{m-1}T}} \right) = O_P\left(\frac{1}{\sqrt{k^{m-1}T}} \right), \tag{B.8}$$

where the second equality follows from (4) of Lemma B.4.

The consistency follows from (B.4), (B.6) and (B.8) immediately.

Below, we focus on the normality.

$$\sqrt{NT}(\hat{\beta} - \beta_0) = \left(\frac{1}{NT} X' M_{\mathcal{Z}} X \right)^{-1} \frac{1}{\sqrt{NT}} X' M_{\mathcal{Z}} (\gamma + \mathcal{E}) \tag{B.9}$$

By (B.4) and (B.8), it is straightforward to obtain that

$$\left(\frac{1}{NT}X'M_{\mathcal{Z}}X\right)^{-1}\frac{1}{\sqrt{NT}}X'M_{\mathcal{Z}}\gamma = O_P\left(N^{\frac{1}{2}}k^{-\frac{m-1}{2}}\right) = o_P(1),$$

where the second equality follows from the assumption $N/k^{m-1} \to 0$.

Therefore, we need only to consider the next term

$$\sqrt{NT}(\hat{\beta}-\beta_0) = \left(\frac{1}{NT}X'M_{\mathcal{Z}}X\right)^{-1}\frac{1}{\sqrt{NT}}X'M_{\mathcal{Z}}\mathcal{E} + o_P(1).$$

By (B.4), $\frac{1}{NT}X'M_{\mathcal{Z}}X \to_P \Sigma_v$. We then focus on $\frac{1}{\sqrt{NT}}X'M_{\mathcal{Z}}\mathcal{E}$ below. Further expand (B.6)

$$\begin{aligned}\frac{1}{\sqrt{NT}}X'M_{\mathcal{Z}}\mathcal{E} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{x}_{it}\tilde{e}_{it} + O_P\left(\frac{\sqrt{k}}{\sqrt[4]{T}}\right)\\ &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}(\phi(u_{it})+v_{it})e_{it} - \frac{\sqrt{NT}}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\phi(u_{it})+v_{it})e_{is} + O_P\left(\frac{\sqrt{k}}{\sqrt[4]{T}}\right).\end{aligned}$$

In the proof for (3) of Lemma B.4, we have shown that $\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\phi(u_{it})e_{is} = o_P\left(\frac{1}{\sqrt{NT}}\right)$ and $\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}v_{it}e_{is} = o_P\left(\frac{1}{\sqrt{NT}}\right)$. Thus, it is straightforward to obtain that

$$\frac{\sqrt{NT}}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\phi(u_{it})+v_{it})e_{is} = o_P(1).$$

Hence, we can further write

$$\frac{1}{\sqrt{NT}}X'M_{\mathcal{Z}}\mathcal{E} = \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}(\phi(u_{it})+v_{it})e_{it} + o_P(1).$$

By (7) of Lemma B.2, it is easy to know that $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\phi(u_{it})e_{it} = o_P(1)$. Therefore, we can write $\frac{1}{\sqrt{NT}}X'M_{\mathcal{Z}}\mathcal{E}$ as

$$\frac{1}{\sqrt{NT}}X'M_{\mathcal{Z}}\mathcal{E} = \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}v_{it}e_{it} + o_P(1).$$

Chen et al. (2012*b*) have shown that $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}v_{it}e_{it} \to_D N(0,\Sigma_{v,e})$ in their formula (A.44). In connection with $\frac{1}{NT}X'M_{\mathcal{Z}}X \to_P \Sigma_v$, the normality follows. ■

Proof of Lemma 3.1: Note that

$$\hat{C} - C = \left(\mathcal{Z}'M_X\mathcal{Z}\right)^{-1}\mathcal{Z}'M_X\gamma + \left(\mathcal{Z}'M_X\mathcal{Z}\right)^{-1}\mathcal{Z}'M_X\mathcal{E},$$

and we normalize each term as

$$\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\mathcal{Z} = \frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z} - \frac{1}{N\sqrt{T}}\mathcal{Z}'X\left(\frac{1}{NT}X'X\right)^{-1}\frac{1}{NT}X'\mathcal{Z} \tag{B.10}$$

$$\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\gamma = \frac{1}{N\sqrt{T}}\mathcal{Z}'\gamma - \frac{1}{N\sqrt{T}}\mathcal{Z}'X\left(\frac{1}{NT}X'X\right)^{-1}\frac{1}{NT}X'\gamma \tag{B.11}$$

$$\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\mathcal{E} = \frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{E} - \frac{1}{N\sqrt{T}}\mathcal{Z}'X\left(\frac{1}{NT}X'X\right)^{-1}\frac{1}{NT}X'\mathcal{E}. \tag{B.12}$$

We now consider (B.10)–(B.12) respectively. Firstly, notice that

$$\begin{aligned}&\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\mathcal{Z} - a_0I_k\right\|\\ \le&\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z} - a_0I_k\right\| + \frac{1}{\sqrt{T}}\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'X\right\|^2\left\|\left(\frac{1}{NT}X'X\right)^{-1}\right\| = o_P(1),\end{aligned}$$

where the last line follows from Lemma B.5 and (2) and (6) of Lemma B.4 in this paper. Consequently, we obtain that

$$\begin{aligned}\lambda_{min}\left(\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\mathcal{Z}\right) &= \min_{\|\mu\|=1}\left\{\mu'a_0I_k\mu + \mu'\left(\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\mathcal{Z} - a_0I_k\right)\mu\right\}\\ &\ge a_0 - \left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\mathcal{Z} - a_0I_k\right\| \ge \frac{1}{2}a_0 + o_P(1).\end{aligned}$$

For (B.11),

$$\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\gamma\right\| \le \left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'\gamma\right\| + \left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'X\right\|\left\|\left(\frac{1}{NT}X'X\right)^{-1}\right\|\left\|\frac{1}{NT}X'\gamma\right\| = O_P\left(k^{-(m-1)/2}\right),$$

where the equality follows from (2), (4), (5) and (6) of Lemma B.4. According to the above, it is easy to obtain that

$$\left\|\left(\mathcal{Z}'M_X\mathcal{Z}\right)^{-1}\mathcal{Z}'M_X\gamma\right\|^2 \le \lambda_{min}^{-2}\left(\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\mathcal{Z}\right)\cdot\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\gamma\right\|^2 = O_P(k^{-m+1}). \tag{B.13}$$

For (B.12),

$$\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\mathcal{E}\right\| \le \left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{E}\right\| + \left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'X\right\|\left\|\left(\frac{1}{NT}X'X\right)^{-1}\right\|\left\|\frac{1}{NT}X'\mathcal{E}\right\| = O_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T}}\right),$$

where the equality follows from (1), (2), (3) and (6) of Lemma B.4 in this paper. Similar to (B.13), it is straightforward to obtain that

$$\left\|\left(\mathcal{Z}'M_X\mathcal{Z}\right)^{-1}\mathcal{Z}'M_X\mathcal{E}\right\|^2 \le \lambda_{min}^{-2}\left(\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\mathcal{Z}\right)\cdot\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\mathcal{E}\right\|^2 = O_P\left(\frac{k}{N\sqrt{T}}\right). \tag{B.14}$$

Therefore, the result follows from (B.13) and (B.14) immediately. ■

Proof of Theorem 3.2: 1) It follows from the orthogonality of the Hermite sequence that

$$\begin{aligned}\int (\hat{g}(w)-g(w))^2\,dw =&(\hat{C}-C)'\int Z_k(w)Z_k(w)'dw(\hat{C}-C)+\int \gamma_k^2(w)dw\\ =&\|\hat{C}-C\|^2+\int \gamma_k^2(w)dw = O_P\left(\frac{k}{N\sqrt{T}}\right)+O_P\left(k^{-m+1}\right),\end{aligned}$$

where Lemmas 3.1 and B.1 are used.

2) We now focus on the normality. We can write

$$\begin{aligned}&\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}(\hat{g}(w)-g(w))\\ =&\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}Z_k(w)'\left(\hat{C}-C\right)-\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}Z_k(w)'\gamma_k(w)\\ =&\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}Z_k(w)'(\mathcal{Z}'M_X\mathcal{Z})^{-1}\mathcal{Z}'M_X\mathcal{E}\\ &+\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}Z_k(w)'(\mathcal{Z}'M_X\mathcal{Z})^{-1}\mathcal{Z}'M_X\gamma-\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}Z_k(w)'\gamma_k(w)\\ =&\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}Z_k(w)'(\mathcal{Z}'M_X\mathcal{Z})^{-1}\mathcal{Z}'M_X\mathcal{E}+O_P(N^{\frac{1}{2}}T^{\frac{1}{4}}k^{-\frac{m-1}{2}})+O_P(N^{\frac{1}{2}}T^{\frac{1}{4}}k^{-\frac{m}{2}+\frac{5}{12}})\\ =&\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}Z_k(w)'\left(\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\mathcal{Z}\right)^{-1}\frac{1}{N\sqrt{T}}\left(\mathcal{Z}'\mathcal{E}-\mathcal{Z}'X(X'X)^{-1}X'\mathcal{E}\right)+o_P(1)\\ =&\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}Z_k(w)\left(\left(\frac{1}{N\sqrt{T}}\mathcal{Z}'M_X\mathcal{Z}\right)^{-1}-a_0^{-1}I_k\right)\cdot\frac{1}{N\sqrt{T}}\left(\mathcal{Z}'\mathcal{E}-\mathcal{Z}'X(X'X)^{-1}X'\mathcal{E}\right)\\ &+\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}Z_k(w)'a_0^{-1}I_k\frac{1}{N\sqrt{T}}\left(\mathcal{Z}'\mathcal{E}-\mathcal{Z}'X(X'X)^{-1}X'\mathcal{E}\right)+o_P(1)\\ =&\frac{1}{\sqrt{N\sigma_k(w)a_0^2}\sqrt[4]{T}}Z_k(w)'\mathcal{Z}'\mathcal{E}+o_P(1)\\ =&\frac{1}{\sqrt{N\sigma_k(w)a_0^2}\sqrt[4]{T}}\sum_{i=1}^N\sum_{t=1}^T Z_k(w)'Z_k(u_{it})e_{it}+o_P(1),\end{aligned}\tag{B.15}$$

where the third equality follows from $Z_k(w)=O(\sqrt{k})$, (2) of Lemma B.1 and (B.13); the fourth equality follows from the assumption in the body of this theorem; the sixth equality follows from (2) of Lemme B.5, (2), (3) and (6) of Lemma B.4 of this paper and Lemma B.3; the last equality follows from the proof for (1) of Lemma B.4.

For notation simplicity, denote $V_{Nk}(t;w)=\frac{1}{\sqrt{N\|Z_k(w)\|^2}}\sum_{i=1}^N Z_k(w)'Z_k(u_{it})e_{it}$ and $\tilde{\sigma}=\sqrt{a_0\sigma_e^2}$. We further write

$$\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}(\hat{g}(w)-g(w))=\frac{1}{\tilde{\sigma}}\sum_{t=1}^T\frac{1}{\sqrt[4]{T}}V_{Nk}(t;w)+o_P(1).\tag{B.16}$$

Notice that $V_{Nk}(t;w)$ is a martingale difference array by Assumption 1. We then use the central limit theorem for martingale difference array to show the normality. See Lemma B.1 of Chen et al. (2012*b*) and Corollary 3.1 of Hall and Heyde (1980, p. 58) for reference. Firstly, we verify the conditional

Lindeberg condition, i.e. as $(N,T) \to (\infty,\infty)$, for $\forall \epsilon > 0$,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} E\left[V_{Nk}^2(t;w) I\left(|V_{Nk}(t;w)| \geq \epsilon \sqrt[4]{T}\right) |\mathcal{F}_{Nt-1}\right] = o_P(1). \tag{B.17}$$

To this end, write

$$\begin{aligned}
&\frac{1}{\sqrt{T}} \sum_{t=1}^{T} E\left[V_{Nk}^2(t;w) I\left(|V_{Nk}(t;w)| \geq \epsilon \sqrt[4]{T}\right) |\mathcal{F}_{Nt-1}\right] \leq \frac{1}{\epsilon^2 T} \sum_{t=1}^{T} E\left[V_{Nk}^4(t;w)|\mathcal{F}_{Nt-1}\right] \\
&\leq \frac{1}{\epsilon^2 T} \sum_{t=1}^{T} \frac{1}{N^2 \|Z_k(w)\|^4} E[|Z_k(w)' Z_k(u_{1t})|^4] \sum_{i_1=1}^{N} \sum_{i_2=1}^{N} \sum_{i_3=1}^{N} \sum_{i_4=1}^{N} |E\left[e_{i_1 t} e_{i_2 t} e_{i_3 t} e_{i_4 t} |\mathcal{F}_{Nt-1}\right]| \\
&\leq \frac{1}{\epsilon^2 T} \sum_{t=1}^{T} E[\|Z_k(u_{1t})\|^4] \cdot \frac{1}{N^2} \sum_{i_1=1}^{N} \sum_{i_2=1}^{N} \sum_{i_3=1}^{N} \sum_{i_4=1}^{N} |E\left[e_{i_1 t} e_{i_2 t} e_{i_3 t} e_{i_4 t} |\mathcal{F}_{Nt-1}\right]| \\
&\leq O_P(1) \frac{1}{\epsilon^2 T} \sum_{t=1}^{T} \frac{1}{d_t} \int \|Z_k(x)\|^4 dx = O_P(1) \frac{k^2}{\epsilon^2 \sqrt{T}} = o_P(1),
\end{aligned} \tag{B.18}$$

due to the independence of u_{it} and u_{jt} for $i \neq j$, where the first inequality follows from Hölder inequality; the second inequality follows from Markovs inequality; the last line follows from the assumption in the body of this theorem, and $\|Z_k(\cdot)\|^2 = O(k)$, $\int \|Z_k(x)\|^2 dx = k$ as well as the density $f_t(x)$ of $d_t^{-1} u_{1t}$ being bounded uniformly (note that $d_t = |\rho|\sqrt{t}(1 + o(1))$ and see the proof of Lemma B.5 in the supplement of this paper for more details).

Next, we verify the convergence of the conditional variance of $V_{Nk}(t;w)$. Again, by the independence of u_{it} and u_{jt} for $i \neq j$,

$$\begin{aligned}
\sum_{t=1}^{T} \frac{E[V_{Nk}^2(t;w)|\mathcal{F}_{Nt-1}]}{\sqrt{T}} &= \frac{1}{\|Z_k(w)\|^2} \frac{1}{N\sqrt{T}} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t=1}^{T} Z_k(w)' E\left[Z_k(u_{it}) Z_k(u_{jt})'\right] Z_k(w) \sigma_e(i,j) \\
&= \frac{1}{\|Z_k(w)\|^2} \frac{1}{N\sqrt{T}} \sum_{i=1}^{N} \sum_{t=1}^{T} Z_k(w)' E\left[Z_k(u_{it}) Z_k(u_{it})'\right] Z_k(w) \sigma_e^2 \\
&\quad + \frac{1}{\|Z_k(w)\|^2} \frac{1}{N\sqrt{T}} \sum_{i \neq j} \sum_{t=1}^{T} Z_k(w)' E\left[Z_k(u_{it})\right] E\left[Z_k(u_{jt})'\right] Z_k(w) \sigma_e(i,j) \\
&\equiv A_{NT1} + A_{NT2}.
\end{aligned}$$

By (1) of Lemma B.5, we have $A_{NT1} \to_P a_0 \sigma_e^2$, and we may show that $A_{NT2} = o_P(1)$. In fact,

$$\begin{aligned}
|A_{NT2}| &\leq \frac{1}{N\sqrt{T}} \sum_{i \neq j} \sum_{t=1}^{T} E\left[\|Z_k(u_{it})\|\right] E\left[\|Z_k(u_{jt})\|\right] \cdot |\sigma_e(i,j)| \\
&= \frac{1}{N} \sum_{i \neq j} |\sigma_e(i,j)| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\int \|Z_k(d_t x)\| f_t(x) dx \right)^2 \\
&\leq O(1) \frac{1}{N} \sum_{i \neq j} |\sigma_e(i,j)| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{1}{d_t^2} \left(\int \|Z_k(x)\| dx \right)^2
\end{aligned}$$

$$= O(1)\frac{k^{11/6}}{\sqrt{T}}\sum_{t=1}^{T}\frac{1}{t} \leq O(1)\frac{k^{11/6}\ln(T)}{\sqrt{T}} = o(1),$$

where the first inequality follows from that submultiplicativity of Euclidean norm; the second inequality follows from the uniformly boundedness of $f_t(x)$; the last line follows from (4) of Lemma B.1 and Assumption 1.3.b.

Therefore, in connection with (B.16), $\sqrt{N\sigma_k^{-1}(w)}\sqrt[4]{T}(\hat{g}(w) - g(w)) \to_D N(0,1)$. ■

Appendix C below is a supplementary document for the proofs of the lemmas and corollary.

Appendix C: Proofs of Lemmas and Corollary

We start from the proof of Lemma B.5, which provides some fundamental results and notations used throughout this document.

Proof of Lemma B.5

1) It suffices to show that as $(N, T) \to (\infty, \infty)$ jointly,

$$\left\| \frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z} - \frac{1}{N\sqrt{T}}E[\mathcal{Z}'\mathcal{Z}] \right\| \to_P 0 \quad \text{and} \quad \frac{1}{N\sqrt{T}}E[\mathcal{Z}'\mathcal{Z}] = a_0 I_k.$$

Notice that

$$\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z} = \frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=1}^{T} Z_k(u_{it})Z_k(u_{it})' - \frac{\sqrt{T}}{N}\sum_{i=1}^{N}\bar{Z}_{k,i}\bar{Z}_{k,i}'$$

$$= \frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=1}^{T} Z_k(u_{it})Z_k(u_{it})' - \frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} Z_k(u_{it})Z_k(u_{is})' \equiv A_{NT} - B_{NT}. \quad \text{(C.1)}$$

Stage One. Calculate the expectation. Note that $\{u_{it}\}$ is i.i.d sequence across i. Therefore, the distribution of u_{it} does not depend on i. Let $d_t = (E[u_{it}^2])^{1/2} = |\rho|\sqrt{t}(1 + o(1))$, where $\rho \neq 0$ is given in Assumption 1. Hence, $d_t^{-1}u_{it}$ has a density $f_t(x)$, which is uniformly bounded over x and large t. Meanwhile, as $t \to \infty$, $\max_x |f_t(x) - \varphi(x)| \leq C d_t^{-1}$ for some $C > 0$, where $\varphi(x)$ is the density of a standard normal variable (see Dong and Gao (2014) for more details on the properties of $f_t(x)$). Let $\nu = \nu(T)$ be a function of T such that $\nu \to \infty$ and $k\nu/\sqrt{T} \to 0$ as $T \to \infty$.

$$\begin{aligned}
E[A_{NT}] =& \frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=1}^{T} E[Z_k(u_{it})Z_k(u_{it})'] \\
=& \frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=1}^{\nu} E[Z_k(u_{it})Z_k(u_{it})'] + \frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=\nu+1}^{T} E[Z_k(u_{it})Z_k(u_{it})'] \\
=& \frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=1}^{\nu} E[Z_k(u_{it})Z_k(u_{it})'] + \frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=\nu+1}^{T} d_t^{-1}\int Z_k(x)Z_k(x)' f_t(d_t^{-1}x)dx \\
=& \frac{1}{\sqrt{T}}\sum_{t=1}^{\nu} E[Z_k(u_{1t})Z_k(u_{1t})'] + \frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T} d_t^{-1}\int Z_k(x)Z_k(x)' f_t(d_t^{-1}x)dx = A_{NT,1} + A_{NT,2}.
\end{aligned}$$

By the construction, it is easy to obtain that for $A_{NT,1}$

$$\frac{1}{\sqrt{T}}\left\|\sum_{t=1}^{\nu}E[Z_k(u_{1t})Z_k(u_{1t})']\right\|\leq\frac{1}{\sqrt{T}}\sum_{t=1}^{\nu}E[\|Z_k(u_{1t})\|^2]=O(1)\frac{\nu k}{\sqrt{T}}\to 0,$$

where the equality follows from (6) of Lemma B.1. We then consider $A_{NT,2}$

$$\begin{aligned}
A_{NT,2}=&\frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T}d_t^{-1}\int Z_k(x)Z_k(x)'f_t(d_t^{-1}x)dx\\
=&\frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T}d_t^{-1}\int Z_k(x)Z_k(x)'\left(f_t(d_t^{-1}x)-\varphi(d_t^{-1}x)\right)dx+\frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T}d_t^{-1}\int Z_k(x)Z_k(x)'\varphi(d_t^{-1}x)dx\\
=&o(1)+\varphi(0)\frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T}d_t^{-1}\int Z_k(x)Z_k(x)'dx+\frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T}d_t^{-1}\int Z_k(x)Z_k(x)'\left(\varphi(d_t^{-1}x)-\varphi(0)\right)dx\\
=&o(1)+2\varphi(0)/|\rho|(1+o(1))\cdot I_k+\frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T}d_t^{-1}\int_{|x|<\varepsilon d_t}Z_k(x)Z_k(x)'\left(\varphi(d_t^{-1}x)-\varphi(0)\right)dx\\
&+\frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T}d_t^{-1}\int_{|x|\geq\varepsilon d_t}Z_k(x)Z_k(x)'\left(\varphi(d_t^{-1}x)-\varphi(0)\right)dx
\end{aligned}$$

where $\varepsilon>0$ can be as small as we wish; and the second equality follows from

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\left\|\sum_{t=\nu+1}^{T}d_t^{-1}\int Z_k(x)Z_k(x)'\left(f_t(d_t^{-1}x)-\varphi(d_t^{-1}x)\right)dx\right\|\\
&\leq O(1)\frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T}d_t^{-2}\int\|Z_k(x)\|^2dx=O(1)\frac{k\ln T}{\sqrt{T}}=o(1).
\end{aligned}$$

Notice also that

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\left\|\sum_{t=\nu+1}^{T}d_t^{-1}\int_{|x|<\varepsilon d_t}Z_k(x)Z_k(x)'\left(\varphi(d_t^{-1}x)-\varphi(0)\right)dx\right\|\\
\leq&\frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T}d_t^{-1}\int_{|x|<\varepsilon d_t}\|Z_k(x)Z_k(x)'\|\cdot|\varphi(d_t^{-1}x)-\varphi(0)|dx\\
\leq&O(1)\frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T}d_t^{-2}\int_{|x|<\varepsilon d_t}\|Z_k(x)Z_k(x)'\|\cdot|x|dx\\
\leq&O(1)\frac{\ln(T)}{\sqrt{T}}\left(\int|x|^2\|Z_k(x)\|^2dx\int\|Z_k(x)\|^2dx\right)^{1/2}\\
=&O(1)\frac{\ln(T)}{\sqrt{T}}\left(k^2\cdot k\right)^{1/2}=O(1)\frac{k^{3/2}\ln(T)}{\sqrt{T}},
\end{aligned}\tag{C.2}$$

where the last line follows from (5) and (9) of Lemma B.1. Moreover,

$$\frac{1}{\sqrt{T}}\left\|\sum_{t=\nu+1}^{T}d_t^{-1}\int_{|x|\geq\varepsilon d_t}Z_k(x)Z_k(x)'\left(\varphi(d_t^{-1}x)-\varphi(0)\right)dx\right\|$$

$$
\begin{aligned}
&\leq O(1)\frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T} d_t^{-1}\int_{|x|\geq\varepsilon d_t}\left\|Z_k(x)Z_k(x)'\right\|dx\\
&\leq O(1)\frac{1}{\sqrt{T}}\sum_{t=\nu+1}^{T}\varepsilon^{-1}d_t^{-2}\int_{|x|\geq\varepsilon d_t}\|Z_k(x)Z_k(x)'\|\cdot|x|dx\\
&\leq O(1)\varepsilon^{-1}\frac{\ln(T)}{\sqrt{T}}\left(\int|x|^2\|Z_k(x)\|^2dx\int\|Z_k(x)\|^2dx\right)^{1/2}=O(1)\frac{k^{3/2}\ln(T)}{\sqrt{T}}. \qquad \text{(C.3)}
\end{aligned}
$$

In view of Assumption 2, (C.2) and (C.3), we obtain that $E[A_{NT}]=2\varphi(0)/|\rho|\cdot I_k(1+o(1))$.

Next, we will show that $E[B_{NT}]=o(1)$. For $t>s$ and $t-s$ is large, note that, without loss of generality letting $u_{i0}=0$ a.s.

$$
\begin{aligned}
u_{it}=&\sum_{\ell=1}^{t}\eta_{i\ell}=\sum_{\ell=1}^{t}\sum_{j=-\infty}^{\ell}\rho_{t-j}\epsilon_{ij}=\sum_{j=-\infty}^{t}b_{t,j}\epsilon_{ij}\\
=&\sum_{j=s+1}^{t}b_{t,j}\epsilon_{ij}+\sum_{j=-\infty}^{s}b_{t,j}\epsilon_{ij}:=u_{i,ts}+u_{i,ts}^*,
\end{aligned}
$$

where $b_{t,j}=\sum_{\ell=\max(1,j)}^{t}\rho_{\ell-j}$.

Similar to the proof of Lemma A.4 of Dong et al. (2014), $\frac{1}{d_{ts}}u_{i,ts}$ has uniformly bounded densities $f_{ts}(w)$ over all t and s, where $d_{ts}=O(1)\sqrt{t-s}$. Without loss of generality, in what follows we abuse the density by neglecting the argument on $\nu=\nu(T)$ as we did before. Let $\mathcal{R}_{is}=\sigma(\ldots,\varepsilon_{i,s-1},\varepsilon_{is})$ be the sigma field generated by $\varepsilon_{ij},j\leq s$. Then,

$$
\begin{aligned}
E[B_{NT}]=&\frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left[Z_k(u_{it})Z_k(u_{is})'\right]\\
=&\frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}E\left[Z_k(u_{it})Z_k(u_{it})'\right]+\frac{2}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E\left[E[Z_k(u_{it})Z_k(u_{is})'|\mathcal{R}_{is}]\right]\\
=&\frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\int Z_k(d_tw)Z_k(d_tw)'f_t(w)dw\\
&+\frac{2}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E\int Z_k(d_{ts}w_1+u_{i,ts}^*)Z_k(u_{is})'f_{ts}(w_1)dw_1\\
=&\frac{1}{T^{3/2}}\sum_{t=1}^{T}\frac{1}{d_t}\int Z_k(w)Z_k(w)'f_t(w/d_t)dw\\
&+\frac{2}{T^{3/2}}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}E\int Z_k(w_1)Z_k(u_{is})'f_{ts}\left(\frac{w_1-u_{i,ts}^*}{d_{ts}}\right)dw_1.
\end{aligned}
$$

The first term is confined by

$$
\frac{1}{T^{3/2}}\sum_{t=1}^{T}\frac{1}{d_t}\int\left\|Z_k(w)Z_k(w)'\right\|f_t(w/d_t)dw\leq O(1)\frac{1}{T^{3/2}}\sum_{t=1}^{T}\frac{1}{d_t}\int\left\|Z_k(w)\right\|^2dw=O(1)\frac{k}{T}
$$

while the second term is bounded by

$$
\begin{aligned}
&\frac{2}{T^{3/2}}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}E\int \|Z_k(w_1)Z_k(u_{is})'\|f_{ts}\left(\frac{w_1-u_{i,ts}^*}{d_{ts}}\right)dw_1\\
\le\, & O(1)\frac{1}{T^{3/2}}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}\int \|Z_k(w_1)\|\,dw_1 E\|Z_k(u_{is})\|\\
\le\, & O(1)\frac{1}{T^{3/2}}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}\frac{1}{d_s}\left(\int \|Z_k(w_1)\|\,dw_1\right)^2=O(1)\frac{k^{11/6}}{T^{1/2}}=o(1),
\end{aligned}
$$

where the last equality follows from (4) of Lemma B.1. The calculation yields $\frac{1}{N\sqrt{T}}E[\mathcal{Z}'\mathcal{Z}]=a_0I_k(1+o(1))$.

Stage Two. We shall show that as $(N,T)\to(\infty,\infty)$ jointly, $E\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z}-\frac{1}{N\sqrt{T}}E[\mathcal{Z}'\mathcal{Z}]\right\|^2\to 0$. To do so, $N\to\infty$ and u_{it} being independent with respect to (w.r.t.) i are important. By (C.1) again,

$$
\begin{aligned}
&E\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z}-\frac{1}{N\sqrt{T}}E[\mathcal{Z}'\mathcal{Z}]\right\|^2 \qquad \text{(C.4)}\\
\le&\frac{2}{N^2T}E\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}\{Z_k(u_{it})Z_k(u_{it})'-E[Z_k(u_{it})Z_k(u_{it})']\}\right\|^2\\
&+\frac{2}{N^2T^3}E\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\{Z_k(u_{it})Z_k(u_{is})'-E[Z_k(u_{it})Z_k(u_{is})']\}\right\|^2\equiv \bar{A}_{NT}+\bar{B}_{NT}.
\end{aligned}
$$

We now consider $\bar{A}_{NT}$ and $\bar{B}_{NT}$ respectively.

$$
\begin{aligned}
\bar{A}_{NT}&=\frac{2}{N^2T}E\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}\{Z_k(u_{it})Z_k(u_{it})'-E[Z_k(u_{it})Z_k(u_{it})']\}\right\|^2\\
&=\frac{2}{N^2T}\sum_{i=1}^{N}E\left\|\sum_{t=1}^{T}\{Z_k(u_{it})Z_k(u_{it})'-E[Z_k(u_{it})Z_k(u_{it})']\}\right\|^2\\
&\le\frac{2}{N^2T}\sum_{i=1}^{N}E\left\|\sum_{t=1}^{T}Z_k(u_{it})Z_k(u_{it})'\right\|^2\\
&=\frac{2}{N^2T}\sum_{n=0}^{k-1}\sum_{m=0}^{k-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left[\mathscr{H}_n(u_{it})\mathscr{H}_m(u_{it})\mathscr{H}_n(u_{is})\mathscr{H}_m(u_{is})\right]\\
&=\frac{2}{N^2T}\sum_{n=0}^{k-1}\sum_{m=0}^{k-1}\sum_{i=1}^{N}\sum_{t=1}^{T}E\left[\mathscr{H}_n^2(u_{it})\mathscr{H}_m^2(u_{it})\right]\\
&\quad+\frac{4}{N^2T}\sum_{n=0}^{k-1}\sum_{m=0}^{k-1}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E\left[\mathscr{H}_n(u_{it})\mathscr{H}_m(u_{it})\mathscr{H}_n(u_{is})\mathscr{H}_m(u_{is})\right]\\
&\equiv\bar{A}_{NT,1}+\bar{A}_{NT,2}.
\end{aligned}
$$

For $\bar{A}_{NT,1}$, write

$$
\bar{A}_{NT,1}=\frac{2}{NT}\sum_{n=0}^{k-1}\sum_{m=0}^{k-1}\sum_{t=1}^{T}\int \mathscr{H}_n^2(d_tw)\mathscr{H}_m^2(d_tw)f_t(w)dw
$$

$$\leq O(1)\frac{1}{NT}\sum_{n=0}^{k-1}\sum_{m=0}^{k-1}\sum_{t=1}^{T}\frac{1}{d_t}\int \mathscr{H}_n^2(w)f_t(w/d_t)dw$$

$$\leq O(1)\frac{1}{NT}\sum_{n=0}^{k-1}\sum_{m=0}^{k-1}\sum_{t=1}^{T}\frac{1}{d_t}\int \mathscr{H}_n^2(w)dw = O(1)\frac{k^2}{N\sqrt{T}},$$

where the first inequality follows from $\mathscr{H}_n(w)$ being bounded uniformly.

For $\bar{A}_{NT,2}$, write

$$\bar{A}_{NT,2} = \frac{4}{N^2T}\sum_{n=0}^{k-1}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E\left[\mathscr{H}_n^2(u_{it})\mathscr{H}_n^2(u_{is})\right]$$

$$+\frac{8}{N^2T}\sum_{n=1}^{k-1}\sum_{m=0}^{n-1}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E\left[\mathscr{H}_n(u_{it})\mathscr{H}_m(u_{it})\mathscr{H}_n(u_{is})\mathscr{H}_m(u_{is})\right] \equiv \bar{A}_{NT,21}+\bar{A}_{NT,22}.$$

For $\bar{A}_{NT,21}$, using conditional argument again we have

$$\bar{A}_{NT,21} \leq O(1)\frac{1}{NT}\sum_{n=0}^{k-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}\frac{1}{d_s}\iint \mathscr{H}_n^2(w_1)\mathscr{H}_n^2(w_2)dw_1dw_2 = O(1)\frac{k}{N}.$$

For $\bar{A}_{NT,22}$, we use the decomposition $u_{it} = u_{i,ts} + u_{i,ts}^*$ again. Note that for $1 \leq i \leq N$ and $s < t$, $u_{i,ts}$ includes all the information between time periods $s+1$ and t and $u_{i,ts}^*$ includes all the information up to time period s. As Dong and Gao (2014) show, $\frac{1}{d_{ts}}u_{i,ts}$ has a density $f_{ts}(w)$, which is uniformly bounded on $\mathbb{R}$ and satisfies uniform Lipschitz condition on $\mathbb{R}$, i.e. $\sup_w |f_{ts}(w+v) - f_{ts}(w)| \leq C|v|$ for some absolutely constant C. Then we can write

$$\bar{A}_{NT,22} = O(1)\frac{1}{NT}\sum_{n=1}^{k-1}\sum_{m=0}^{n-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E\left[\mathscr{H}_n(u_{it})\mathscr{H}_m(u_{it})\mathscr{H}_n(u_{is})\mathscr{H}_m(u_{is})\right]$$

$$= O(1)\frac{1}{NT}\sum_{n=1}^{k-1}\sum_{m=0}^{n-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E\left[E\left[\mathscr{H}_n(u_{i,ts}+u_{i,ts}^*)\mathscr{H}_m(u_{i,ts}+u_{i,ts}^*)|\mathcal{R}_{is}\right]\mathscr{H}_n(u_{is})\mathscr{H}_m(u_{is})\right]$$

$$= O(1)\frac{1}{NT}\sum_{n=1}^{k-1}\sum_{m=0}^{n-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E\left[\int \mathscr{H}_n(d_{ts}w+u_{is}^*)\mathscr{H}_m(d_{ts}w+u_{is}^*)f_{ts}(w)dw\cdot\mathscr{H}_n(u_{is})\mathscr{H}_m(u_{is})\right]$$

$$= O(1)\frac{1}{NT}\sum_{n=1}^{k-1}\sum_{m=0}^{n-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}E\left[\int \mathscr{H}_n(w)\mathscr{H}_m(w)f_{ts}\left(\frac{w-u_{is}^*}{d_{ts}}\right)dw\cdot\mathscr{H}_n(u_{is})\mathscr{H}_m(u_{is})\right]$$

$$= O(1)\frac{1}{NT}\sum_{n=1}^{k-1}\sum_{m=0}^{n-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}$$

$$\cdot E\left[\int \mathscr{H}_n(w)\mathscr{H}_m(w)\left[f_{ts}\left(\frac{w-u_{is}^*}{d_{ts}}\right)-f_{ts}\left(\frac{-u_{is}^*}{d_{ts}}\right)\right]dw\cdot\mathscr{H}_n(u_{is})\mathscr{H}_m(u_{is})\right],$$

where the last line follows from the truth that $\int \mathscr{H}_n(w)\mathscr{H}_m(w)dw = 0$ for $m \neq n$. By the uniform Lipschitz condition of f_{ts}, we then obtain that

$$|\bar{A}_{NT,22}| \leq O(1)\frac{1}{NT}\sum_{n=1}^{k-1}\sum_{m=0}^{n-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}^2}\int |w\mathscr{H}_n(w)\mathscr{H}_m(w)|dw\cdot E\left[|\mathscr{H}_n(u_{is})\mathscr{H}_m(u_{is})|\right]$$

$$
\begin{aligned}
&= O(1)\frac{1}{NT}\sum_{n=1}^{k-1}\sum_{m=0}^{n-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}^2}\int |w\mathscr{H}_n(w)\mathscr{H}_m(w)|dw\cdot\int |\mathscr{H}_n(d_s w)\mathscr{H}_m(d_s w)|f_s(w)dw\\
&\le O(1)\frac{1}{NT}\sum_{n=1}^{k-1}\sum_{m=0}^{n-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}^2}\frac{1}{d_s}\int |w\mathscr{H}_n(w)\mathscr{H}_m(w)|dw\cdot\int |\mathscr{H}_n(w)\mathscr{H}_m(w)|dw\\
&\le O(1)\frac{1}{NT}\sum_{n=1}^{k-1}\sum_{m=0}^{n-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}^2}\frac{1}{d_s}\left\{\int \mathscr{H}_n^2(w)dw\int w^2\mathscr{H}_m^2(w)dw\right\}^{1/2}\\
&\quad\cdot\left\{\int \mathscr{H}_n^2(w)dw\int \mathscr{H}_m^2(w)dw\right\}^{1/2}\\
&\le O(1)\frac{1}{NT}\sum_{n=1}^{k-1}\sum_{m=0}^{n-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}^2}\frac{1}{d_s}\sqrt{m} = O\left(\frac{k^{5/2}\ln T}{N\sqrt{T}}\right) = o(1).
\end{aligned}
$$

By the calculation of $\bar{A}_{NT,1}$ and $\bar{A}_{NT,2}$, we have shown that $\bar{A}_{NT} = o(1)$.

For $\bar{B}_{NT}$, by the independence across i of $\{u_{i1},\ldots,u_{iT}\}$, write

$$
\begin{aligned}
\bar{B}_{NT} &= \frac{2}{N^2T^3}\sum_{i=1}^{N}E\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\{Z_k(u_{it})Z_k(u_{is})' - E[Z_k(u_{it})Z_k(u_{is})']\}\right\|^2\\
&\le \frac{O(1)}{N^2T^3}\sum_{i=1}^{N}E\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}Z_k(u_{it})Z_k(u_{is})'\right\|^2 \le \frac{O(1)}{N^2T^3}\sum_{i=1}^{N}E\left[\sum_{t=1}^{T}\sum_{s=1}^{T}\|Z_k(u_{it})\|\|Z_k(u_{is})\|\right]^2\\
&= O(1)\frac{1}{N^2T^3}\sum_{i=1}^{N}E\left[\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T}\|Z_k(u_{it_1})\|\|Z_k(u_{it_2})\|\|Z_k(u_{it_3})\|\|Z_k(u_{it_4})\|\right]\\
&= O(1)\frac{1}{N^2T^3}\sum_{i=1}^{N}E\left[\sum_{\text{all } t_1,t_2,t_3,t_4 \text{different}}\|Z_k(u_{it_1})\|\|Z_k(u_{it_2})\|\|Z_k(u_{it_3})\|\|Z_k(u_{it_4})\|\right]\\
&\quad +O(1)\frac{1}{N^2T^3}\sum_{i=1}^{N}E\left[\sum_{\text{two of } t_1,t_2,t_3,t_4 \text{same}}\|Z_k(u_{it_1})\|\|Z_k(u_{it_2})\|\|Z_k(u_{it_3})\|\|Z_k(u_{it_4})\|\right]\\
&\quad +O(1)\frac{1}{N^2T^3}\sum_{i=1}^{N}E\left[\sum_{\text{three of } t_1,t_2,t_3,t_4 \text{same}}\|Z_k(u_{it_1})\|\|Z_k(u_{it_2})\|\|Z_k(u_{it_3})\|\|Z_k(u_{it_4})\|\right]\\
&\quad +O(1)\frac{1}{N^2T^3}\sum_{i=1}^{N}E\left[\sum_{t=1}^{T}\|Z_k(u_{it})\|^4\right]\\
&\equiv \bar{B}_{NT,1}+\bar{B}_{NT,2}+\bar{B}_{NT,3}+\bar{B}_{NT,4}.
\end{aligned}
$$

For $\bar{B}_{NT,1}$, without losing generality, assume that $t_1 > t_2 > t_3 > t_4$. Then, by the conditional argument,

$$
\begin{aligned}
\bar{B}_{NT,1} &= \frac{O(1)}{N^2T^3}\sum_{i=1}^{N}\sum_{t_1=4}^{T}\sum_{t_2=3}^{t_1-1}\sum_{t_3=2}^{t_2-1}\sum_{t_4=1}^{t_3-1}E[\|Z_k(u_{it_1})\|\|Z_k(u_{it_2})\|\|Z_k(u_{it_3})\|\|Z_k(u_{it_4})\|]\\
&\le \frac{O(1)}{N^2T^3}\sum_{i=1}^{N}\sum_{t_1=4}^{T}\sum_{t_2=3}^{t_1-1}\sum_{t_3=2}^{t_2-1}\sum_{t_4=1}^{t_3-1}\frac{1}{d_{t_1t_2}}\frac{1}{d_{t_2t_3}}\frac{1}{d_{t_3t_4}}\frac{1}{d_{t_4}}\\
&\quad\cdot\iiiint \|Z_k(w_1)\|\|Z_k(w_2)\|\|Z_k(w_3)\|\|Z_k(w_4)\|dw_1dw_2dw_3dw_4
\end{aligned}
$$

$$= \frac{O(1)}{NT}\left(\int \|Z_k(w)\|dw\right)^4 = O\left(\frac{k^{11/3}}{NT}\right) = o(1),$$

where the last line follows from (4) of Lemma B.1 and Assumption 2.2.

For $\bar{B}_{NT,2}$, without losing generality, assume that $t_1 = t_2 > t_3 > t_4$. Then write

$$\begin{aligned}
\bar{B}_{NT,2} &= \frac{O(1)}{N^2T^3}\sum_{i=1}^{N}\sum_{t_1=3}^{T}\sum_{t_3=2}^{t_1-1}\sum_{t_4=1}^{t_3-1} E[\|Z_k(u_{it_1})\|^2\|Z_k(u_{it_3})\|\|Z_k(u_{it_4})\|] \\
&\leq \frac{O(1)}{N^2T^3}\sum_{i=1}^{N}\sum_{t_1=3}^{T}\sum_{t_3=2}^{t_1-1}\sum_{t_4=1}^{t_3-1} \frac{1}{d_{t_1t_3}}\frac{1}{d_{t_3d_4}}\frac{1}{d_{t_4}} \\
&\cdot \iiint \|Z_k(w_1)\|^2\|Z_k(w_2)\|\|Z_k(w_3)\|dw_1dw_2dw_3 \\
&\leq \frac{O(1)}{NT^{3/2}}\int \|Z_k(w)\|^2dw\left(\int \|Z_k(w)\|dw\right)^2 = O\left(\frac{k^{17/6}}{NT^{3/2}}\right) = o(1),
\end{aligned}$$

where the last line follows from (4)–(5) of Lemma B.1 and Assumption 2.2.

For $\bar{B}_{NT,3}$, without losing generality, assume that $t_1 = t_2 = t_3 > t_4$. Then write

$$\begin{aligned}
\bar{B}_{NT,3} &= \frac{O(1)}{N^2T^3}\sum_{i=1}^{N}\sum_{t_1=2}^{T}\sum_{t_4=1}^{t_1-1} E[\|Z_k(u_{it_1})\|^3\|Z_k(u_{it_4})\|] \\
&\leq \frac{O(1)}{N^2T^3}\sum_{i=1}^{N}\sum_{t_1=2}^{T}\sum_{t_4=1}^{t_1-1} \frac{1}{d_{t_1d_4}}\frac{1}{d_{t_4}}\iint \|Z_k(w_1)\|^3\|Z_k(w_2)\|dw_1dw_2 \\
&\leq O\left(\frac{k}{NT^2}\right)\left(\int \|Z_k(w)\|dw\right)^2 = O\left(\frac{k^{17/6}}{NT^2}\right) = o(1),
\end{aligned}$$

where the last line follows from (4) and (6) of Lemma B.1 and Assumption 2.2.

For $\bar{B}_{NT,4}$, write

$$\bar{B}_{NT,4} = \frac{1}{N^2T^3}\sum_{i=1}^{N}\sum_{t=1}^{T} E\|Z_k(w)\|^4 = \frac{1}{N^2T^3}\sum_{i=1}^{N}\sum_{t=1}^{T}\int \|Z_k(d_tw)\|^4 f_t(w)dw$$

$$\leq O\left(\frac{1}{N^2T^3}\right)\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{1}{d_t}\int \|Z_k(w)\|^4dw \leq O\left(\frac{k}{NT^{5/2}}\right)\int \|Z_k(w)\|^2dw = O\left(\frac{k^2}{NT^{5/2}}\right) = o(1).$$

Combining $\bar{B}_{NT,1}$, $\bar{B}_{NT,2}$, $\bar{B}_{NT,3}$ and $\bar{B}_{NT,4}$ together, we know that $\bar{B}_{NT} = o(1)$.

Therefore, we have shown that $\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{Z} - \sqrt{\frac{2}{\pi|\rho|^2}}(1+o(1))I_k\right\| = o_P(1)$. We now complete the proof for the first result of this lemma.

2) Noticing that $k^2/N \to 0$ and that (C.4) in **Stage Two**, particularly $\bar{A}_{NT} = k/N$ and $\bar{B}_{NT} = k^{11/3}/(NT)$, the second result of this lemma follows immediately. ■

Lemma B.2. *Under Assumptions 1 and 2, as $(N,T) \to (\infty,\infty)$ jointly,*

1. $\left\|\frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=1}^{T} Z_k(u_{it})\gamma_k(u_{it})\right\| = O_P(k^{-(m-1)/2})$;

2. $\left\|\frac{1}{N\sqrt{T}}\sum_{i=1}^N\sum_{t=1}^T \phi(u_{it})Z_k(u_{it})'\right\| = O_P(1)$;

3. $\left\|\frac{1}{N\sqrt{T}}\sum_{i=1}^N\sum_{t=1}^T v_{it}Z_k(u_{it})'\right\| = O_P\left(\sqrt{\frac{k}{N}}\right)$;

4. $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T v_{it}v_{it}' = \Sigma_v + O_P\left(\frac{1}{\sqrt{NT}}\right)$;

5. $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T v_{it}\phi(u_{it})' = O_P\left(\frac{1}{\sqrt{NT}}\right)$;

6. $\frac{1}{N\sqrt{T}}\sum_{i=1}^N\sum_{t=1}^T Z_k(u_{it})e_{it} = O_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T}}\right)$;

7. $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \phi(u_{it})e_{it} = O_P\left(\frac{1}{\sqrt{N}\sqrt[4]{T^3}}\right)$;

8. $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \phi(u_{it})\gamma_k(u_{it}) = O_P\left(\frac{1}{\sqrt{k^mT}}\right)$;

9. $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T v_{it}\gamma_k(u_{it}) = O_P\left(\frac{k^{-m/2+5/12}}{\sqrt{NT}}\right)$.

Proof of Lemma B.2:

1) Write

$$\begin{aligned}
&E\left\|\frac{1}{N\sqrt{T}}\sum_{i=1}^N\sum_{t=1}^T Z_k(u_{it})\gamma_k(u_{it})\right\|^2 \\
&= \frac{1}{N^2T}E\left[\sum_{i=1}^N\left(\sum_{t=1}^T\|Z_k(u_{it})\|^2\gamma_k^2(u_{it}) + 2\sum_{t=2}^T\sum_{s=1}^{t-1}Z_k(u_{it})'Z_k(u_{is})\gamma_k(u_{it})\gamma_k(u_{is})\right)\right] \\
&\quad + \frac{2}{N^2T}E\left[\sum_{i=2}^N\sum_{j=1}^{i-1}\sum_{t=1}^T Z_k(u_{it})'Z_k(u_{jt})\gamma_k(u_{it})\gamma_k(u_{jt})\right] \\
&\quad + \frac{4}{N^2T}E\left[\sum_{i=2}^N\sum_{j=1}^{i-1}\sum_{t=2}^T\sum_{s=1}^{t-1} Z_k(u_{it})'Z_k(u_{js})\gamma_k(u_{it})\gamma_k(u_{js})\right] \equiv A_1 + 2A_2 + 4A_3.
\end{aligned}$$

Notice that

$$\begin{aligned}
A_1 = &\frac{1}{N^2T}\sum_{i=1}^N\sum_{t=1}^T E\left[\|Z_k(u_{it})\|^2\gamma_k^2(u_{it})\right] \\
&+\frac{2}{N^2T}\sum_{i=1}^N\sum_{t=2}^T\sum_{s=1}^{t-1} E\left[Z_k(u_{it})'Z_k(u_{is})\gamma_k(u_{it})\gamma_k(u_{is})\right]. \qquad \text{(C.5)}
\end{aligned}$$

The first term on RHS of (C.5) can be written as

$$\begin{aligned}
&\frac{1}{N^2T}\sum_{i=1}^N\sum_{t=1}^T E[\|Z_k(u_{it})\|^2\gamma_k^2(u_{it})] \le O\left(\frac{k^{-m+5/6}}{N^2T}\right)\sum_{i=1}^N\sum_{t=1}^T\int\|Z_k(d_tw)\|^2f_t(w)dw \\
&= O\left(\frac{k^{-m+5/6}}{N^2T}\right)\sum_{i=1}^N\sum_{t=1}^T\int\frac{1}{d_t}\|Z_k(w)\|^2f_t(w/d_t)dw \\
&\le O\left(\frac{k^{-m+5/6}}{N^2T}\right)\sum_{i=1}^N\sum_{t=1}^T\frac{1}{d_t}\int\|Z_k(w)\|^2dw \le O\left(\frac{k^{-m+11/6}}{N\sqrt{T}}\right),
\end{aligned}$$

where the first inequality follows from (2) of Lemma B.1 and the second inequality follows from $f_t(w)$ being bounded uniformly.

For the second term on RHS of (C.5),

$$\left|\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E[Z_k(u_{it})'Z_k(u_{is})\gamma_k(u_{it})\gamma_k(u_{is})]\right|$$
$$\leq \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E[\|Z_k(u_{it})\|\|Z_k(u_{is})\||\gamma_k(u_{it})||\gamma_k(u_{is})|]$$
$$\leq O\left(\frac{1}{N^2T}\right)\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\iint\frac{1}{d_{ts}}\frac{1}{d_s}\|Z_k(w_1)\|\|Z_k(w_2)\||\gamma_k(w_1)||\gamma_k(w_2)|dw_1dw_2$$
$$\leq O\left(\frac{1}{NT}\right)\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}\frac{1}{d_s}\iint\|Z_k(w_1)\|\|Z_k(w_2)\||\gamma_k(w_1)||\gamma_k(w_2)|dw_1dw_2$$
$$\leq O\left(\frac{1}{N}\right)\left(\int\|Z_k(w)\||\gamma_k(w)|dw\right)^2\leq O\left(\frac{1}{N}\right)\int\|Z_k(w)\|^2dw\int|\gamma_k(w)|^2dw=O\left(\frac{k^{-m+1}}{N}\right).$$

Therefore, $A_1=O\left(\frac{k^{-m+1}}{N}\right)$.

For A_2, by virtue of $Z_k(w)=(\mathscr{H}_0(w),\ldots,\mathscr{H}_{k-1}(w))'$

$$|A_2|=\left|\frac{1}{N^2T}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t=1}^{T}\sum_{n=0}^{k-1}E\left[\mathscr{H}_n(u_{it})\gamma_k(u_{it})\right]\cdot E\left[\mathscr{H}_n(u_{jt})\gamma_k(u_{jt})\right]\right|$$
$$=\left|\frac{1}{N^2T}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t=1}^{T}\sum_{n=0}^{k-1}\int\mathscr{H}_n(d_tw)\gamma_k(d_tw)f_t(w)dw\int\mathscr{H}_n(d_tw)\gamma_k(d_tw)f_t(w)dw\right|$$
$$\leq O\left(\frac{1}{N^2T}\right)\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t=1}^{T}\frac{1}{d_t^2}\sum_{n=0}^{k-1}\left(\int|\mathscr{H}_n(w)\gamma_k(w)|dw\right)^2$$
$$\leq O\left(\frac{\ln T}{T}\right)\int\|Z_k(w)\|^2dw\int|\gamma_k(w)|^2dw=o\left(k^{-m+1}\right).$$

Similar to A_2, for A_3 we write

$$|A_3|=\left|\frac{1}{N^2T}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\sum_{n=0}^{k-1}E\left[\mathscr{H}_n(u_{it})\gamma_k(u_{it})\right]\cdot E\left[\mathscr{H}_n(u_{js})\gamma_k(u_{js})\right]\right|$$
$$\leq O\left(\frac{1}{T}\right)\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_t}\frac{1}{d_s}\sum_{n=0}^{k-1}\left(\int|\mathscr{H}_n(w)\gamma_k(w)|dw\right)^2$$
$$\leq O(1)\int\|Z_k(w)\|^2dw\int|\gamma_k(w)|^2dw=O(k^{-m+1}).$$

Thus, the result follows. ■

2) Write

$$E\left\|\frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=1}^{T}\phi(u_{it})Z_k(u_{it})'\right\|^2$$
$$=\frac{1}{N^2T}\sum_{m=1}^{d}\sum_{n=0}^{k-1}\sum_{i=1}^{N}\sum_{t=1}^{T}E[\phi_m^2(u_{it})\mathscr{H}_n^2(u_{it})]$$

$$
+\frac{2}{N^2T}\sum_{m=1}^{d}\sum_{n=0}^{k-1}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E[\phi_m(u_{it})\mathscr{H}_n(u_{it})\phi_m(u_{is})\mathscr{H}_n(u_{is})]
$$

$$
+\frac{2}{N^2T}\sum_{m=1}^{d}\sum_{n=0}^{k-1}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t=1}^{T}\sum_{s=1}^{T}E[\phi_m(u_{it})\mathscr{H}_n(u_{it})]E[\phi_m(u_{js})\mathscr{H}_n(u_{js})]
$$

$$
\leq O\left(\frac{1}{N^2T}\right)\sum_{m=1}^{d}\sum_{n=0}^{k-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{1}{d_t}\int\mathscr{H}_n^2(w)dw
$$

$$
+O\left(\frac{2}{N^2T}\right)\sum_{m=1}^{d}\sum_{n=0}^{k-1}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}\frac{1}{d_s}\iint\phi_m(w_1)\mathscr{H}_n(w_1)\phi_m(w_2)\mathscr{H}_n(w_2)dw_1dw_2
$$

$$
+O\left(\frac{2}{N^2T}\right)\sum_{m=1}^{d}\sum_{n=0}^{k-1}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{d_t}\frac{1}{d_s}\left(\int\phi_m(w)\mathscr{H}_n(w)dw\right)^2
$$

$$
\leq O\left(\frac{k}{N\sqrt{T}}\right)+\frac{2}{N^2}\sum_{m=1}^{d}\sum_{n=0}^{k-1}\sum_{i=1}^{N}\left(\int\phi_m(w)\mathscr{H}_n(w)dw\right)^2
$$

$$
+\frac{2}{N^2}\sum_{m=1}^{d}\sum_{n=0}^{k-1}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\left(\int\phi_m(w)\mathscr{H}_n(w)dw\right)^2
$$

$$
=o(1)+\frac{2}{N^2}\sum_{m=1}^{d}\sum_{n=0}^{k-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\int\phi_m(w)\mathscr{H}_n(w)dw\right)^2
$$

$$
\leq o(1)+\frac{2}{N^2}\sum_{m=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\int\phi_m^2(w)dw=O(1),
$$

where the first equality is due to Assumption 1.4; the first inequality follows from that $f_t(w)$ being bounded uniformly and $\phi_m(w)$ being bounded uniformly on $\mathbb{R}$ for $m=1,\ldots,d$; the last inequality follows from that $\phi_m(w)\in L^2(\mathbb{R})$ (such that $\phi_m(w)=\sum_{n=0}^{\infty}c_{m,n}\mathscr{H}_n(w)$ for $m=1,\ldots,d$, $c_{m,n}=\int\phi_m(w)\mathscr{H}_n(w)dw$ for $n=0,\ldots,\infty$ and $\sum_{n=0}^{\infty}c_{m,n}^2=\int\phi_m^2(w)dw$).

The proof is then complete. ■

3) Let v_{it,n_1} denote the n_1^{th} element of v_{it}. Write

$$
E\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}v_{it}Z_k(u_{it})'\right\|^2=\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}E\left[\sum_{i=1}^{N}\sum_{t=1}^{T}v_{it,n_1}\mathscr{H}_{n_2}(u_{it})\right]^2
$$

$$
=\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left[v_{it,n_1}v_{js,n_1}\right]E\left[\mathscr{H}_{n_2}(u_{it})\mathscr{H}_{n_2}(u_{js})\right]
$$

$$
\leq O(k)\sum_{n_1=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}c_\delta(\alpha_{ij}(|t-s|))^{\delta/(4+\delta)}\left(E|v_{it,n_1}|^{2+\delta/2}\right)^{2/(4+\delta)}\left(E|v_{js,n_1}|^{2+\delta/2}\right)^{2/(4+\delta)}
$$

$$
\leq O(k)\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\alpha_{ij}(|t-s|))^{\delta/(4+\delta)}=O(kNT),
$$

where $c_\delta=2^{(4+2\delta)/(4+\delta)}\cdot(4+\delta)/\delta$; the second equality follows from Assumption 1.4; the first inequality follows from Davydov inequality (c.f. pages 19–20 in Bosq (1996) and supplementary of Su and Jin (2012)) and the truth that $\mathscr{H}_n(w)$ is bounded uniformly (c.f. Nevai (1986)); the second inequality follows from Assumption 1.3.a. Thus, the result follows. ■

4) Let Σ_{v,n_1n_2} denote the $(n_1,n_2)^{th}$ element of Σ_v. Write

$$E\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}(v_{it}v_{it}'-\Sigma_v)\right\|^2=\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}E\left[\sum_{i=1}^{N}\sum_{t=1}^{T}(v_{it,n_1}v_{it,n_2}-\Sigma_{v,n_1n_2})\right]^2$$

$$\leq\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}c_\delta(\alpha_{ij}(|t-s|))^{\delta/(4+\delta)}$$

$$\cdot\left(E|v_{it,n_1}v_{it,n_2}|^{2+\delta/2}\right)^{2/(4+\delta)}\left(E|v_{js,n_1}v_{js,n_2}|^{2+\delta/2}\right)^{2/(4+\delta)}$$

$$\leq\frac{c_\delta}{2}\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\alpha_{ij}(|t-s|))^{\delta/(4+\delta)}\cdot\left(E|v_{it,n_1}v_{it,n_2}|^{2+\delta/2}\right)^{4/(4+\delta)}$$

$$+\frac{c_\delta}{2}\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\alpha_{ij}(|t-s|))^{\delta/(4+\delta)}\cdot\left(E|v_{js,n_1}v_{js,n_2}|^{2+\delta/2}\right)^{4/(4+\delta)}$$

$$\leq O(1)\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\alpha_{ij}(|t-s|))^{\delta/(4+\delta)}\cdot\left(E|v_{it,n_1}|^{4+\delta}\cdot E|v_{it,n_2}|^{4+\delta}\right)^{2/(4+\delta)}$$

$$\leq O(1)\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\alpha_{ij}(|t-s|))^{\delta/(4+\delta)}=O(NT),$$

where $c_\delta=2^{(4+2\delta)/(4+\delta)}\cdot(4+\delta)/\delta$; the first inequality follows from Davydov inequality; the third inequality follows from Cauchy–Schwarz inequality; the last line follows from Assumption 1.3.a. Therefore, the result follows. ■

5) Write

$$E\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}v_{it}\phi(u_{it})'\right\|^2=\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}E\left[\sum_{i=1}^{N}\sum_{t=1}^{T}v_{it,n_1}\phi_{n_2}(u_{it})\right]^2$$

$$=\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left[v_{it,n_1}v_{js,n_1}\right]E\left[\phi_{n_2}(u_{it})\phi_{n_2}(u_{js})\right]$$

$$\leq O(1)\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\alpha_{ij}(|t-s|))^{\delta/(4+\delta)}\left(E|v_{it,n}|^{2+\delta/2}\right)^{2/(4+\delta)}\left(E|v_{js,n}|^{2+\delta/2}\right)^{2/(4+\delta)}$$

$$\leq O(1)\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\alpha_{ij}(|t-s|))^{\delta/(4+\delta)}=O(NT),$$

where the second equality follows from Assumption 1.4; the first inequality follows from Davydov inequality and the uniform boundedness of $\phi_n(w)$ on $\mathbb{R}$ for $n=1,\ldots,d$. Therefore, the result follows immediately. ■

6) By Assumptions 1.1 and 1.4,

$$E\left\|\frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=1}^{T}Z_k(u_{it})e_{it}\right\|^2=\frac{1}{N^2T}E\left[\left(\sum_{i=1}^{N}\sum_{t=1}^{T}Z_k(u_{it})e_{it}\right)'\left(\sum_{i=1}^{N}\sum_{t=1}^{T}Z_k(u_{it})e_{it}\right)\right]$$

$$= \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{t=1}^{T} E\left[\|Z_k(u_{it})\|^2\right] E\left[e_{it}^2\right] + \frac{2}{N^2T}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t=1}^{T} E\left[Z_k(u_{it})'\right] E\left[Z_k(u_{jt})\right] E[e_{it}e_{jt}]$$

$$\equiv B_1 + 2B_2.$$

For B_1, write

$$\begin{aligned} B_1 &= \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{t=1}^{T}\int \|Z_k(d_t w)\|^2 f_t(w)dw \cdot \sigma_e^2 \\ &\leq O\left(\frac{1}{NT}\right)\sum_{t=1}^{T}\frac{1}{d_t}\int \|Z_k(w)\|^2 dw = O\left(\frac{k}{N\sqrt{T}}\right), \end{aligned}$$

where the second line follow from that $f_t(w)$ being bounded uniformly.

For B_2,

$$\begin{aligned} |B_2| &= \left|\frac{1}{N^2T}\sum_{n=0}^{k-1}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t=1}^{T}\int \mathscr{H}_n(d_t w)f_t(w)dw \int \mathscr{H}_n(d_t w)f_t(w)dw \cdot \sigma_e(i,j)\right| \\ &\leq \frac{1}{N^2T}\sum_{n=0}^{k-1}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t=1}^{T}\int \mathscr{H}_n^2(d_t w)dw \int f_t^2(w)dw\, |\sigma_e(i,j)| \\ &\leq O\left(\frac{1}{N^2T}\right)\sum_{n=0}^{k-1}\sum_{t=1}^{T}\frac{1}{d_t}\int \mathscr{H}_n^2(w)dw \sum_{i=2}^{N}\sum_{j=1}^{i-1}|\sigma_e(i,j)| \leq O\left(\frac{k}{N\sqrt{T}}\right), \end{aligned}$$

where the first inequality follows from Cauchy–Schwarz inequality; the second inequality follows from $f_t(w)$ being bounded uniformly; the third inequality follows from Assumption 1.3.b. In connection with $B_1 = O\left(\frac{k}{N\sqrt{T}}\right)$, we obtain that $\frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=1}^{T} Z_k(u_{it})e_{it} = O_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T}}\right)$. ■

7) Write

$$\begin{aligned} E\left\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\phi(u_{it})e_{it}\right\|^2 &= \frac{1}{N^2T^2}\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T} E\left[\phi_n(u_{it})\phi_n(u_{jt})\right] E\left[e_{it}e_{jt}\right] \\ = \frac{\sigma_e^2}{N^2T^2}&\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t=1}^{T}\int \phi_n^2(d_t w)f_t(w)dw \\ &+ \frac{2}{N^2T^2}\sum_{n=1}^{d}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t=1}^{T}\int \phi_n(d_t w)f_t(w)dw \int \phi_n(d_t w)f_t(w)dw \cdot \sigma_e(i,j) \\ \leq O\left(\frac{1}{NT^2}\right)&\sum_{n=1}^{d}\sum_{t=1}^{T}\frac{1}{d_t}\int \phi_n^2(w)dw \\ &+ O\left(\frac{2}{N^2T^2}\right)\sum_{n=1}^{d}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t=1}^{T}\int \phi_n^2(d_t w)dw \int f_t^2(w)dw|\sigma_e(i,j)| \\ \leq O\left(\frac{1}{NT^{3/2}}\right) &+ O\left(\frac{1}{NT^{3/2}}\right) = O\left(\frac{1}{NT^{3/2}}\right), \end{aligned}$$

where the first equality follows from Assumption 1.4; the first inequality follows from $f_t(w)$ being bounded uniformly and $\phi_n(w) \in L^2(\mathbb{R})$; the second inequality follows from $\phi_n(w)$ being integrable and Assumption 1.3. Therefore, the result follows. ■

8) Write

$$E\left\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\phi(u_{it})\gamma_k(u_{it})\right\|^2 = \frac{1}{N^2T^2}\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left[\phi_n(u_{it})\gamma_k(u_{it})\phi_n(u_{js})\gamma_k(u_{js})\right]$$

$$= \frac{1}{N^2T^2}\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left[\phi_n(u_{it})\gamma_k(u_{it})\phi_n(u_{is})\gamma_k(u_{is})\right]$$

$$+\frac{2}{N^2T^2}\sum_{n=1}^{d}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left[\phi_n(u_{it})\gamma_k(u_{it})\right]E\left[\phi_n(u_{js})\gamma_k(u_{js})\right]$$

$$= \frac{1}{N^2T^2}\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t=1}^{T}E\left[\phi_n^2(u_{it})\gamma_k^2(u_{it})\right] + \frac{2}{N^2T^2}\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E\left[\phi_n(u_{it})\gamma_k(u_{it})\phi_n(u_{is})\gamma_k(u_{is})\right]$$

$$+\frac{2}{N^2T}\sum_{n=1}^{d}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}E\left[\phi_n(u_{it})\gamma_k(u_{it})\right]\frac{1}{\sqrt{T}}\sum_{s=1}^{T}E\left[\phi_n(u_{js})\gamma_k(u_{js})\right]$$

$$\equiv C_1 + 2C_2 + 2C_3.$$

By (2) of Lemma B.1,

$$C_1 = O\left(\frac{k^{-m+5/6}}{N^2T^2}\right)\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t=1}^{T}\int \phi_n^2(d_tw)f_t(w)dw$$

$$\leq O\left(\frac{k^{-m+5/6}}{N^2T^2}\right)\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{1}{d_t}\int \phi_n^2(w)dw \leq O\left(\frac{k^{-m+5/6}}{NT^{3/2}}\right).$$

For C_2, write

$$|C_2| = \left|\frac{1}{N^2T^2}\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E\left[\phi_n(u_{it})\gamma_k(u_{it})\phi_n(u_{is})\gamma_k(u_{is})\right]\right|$$

$$\leq \frac{1}{N^2T^2}\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\iint |\phi_n(d_{ts}w_1 + d_sw_2)\gamma_k(d_{ts}w_1 + d_sw_2)\phi_n(d_sw_2)\gamma_k(d_sw_2)|$$

$$\cdot f_{ts}(w_1)f_s(w_2)dw_1dw_2$$

$$\leq O\left(\frac{1}{N^2T^2}\right)\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}\frac{1}{d_s}\iint|\phi_n(w_1)\gamma_k(w_1)\phi_n(w_2)\gamma_k(w_2)|dw_1dw_2$$

$$\leq O\left(\frac{1}{NT}\right)\sum_{n=1}^{d}\left(\int|\phi_n(w)\gamma_k(w)|dw\right)^2 \leq O\left(\frac{1}{NT}\right)\sum_{n=1}^{d}\int\phi_n^2(w)dw\int\gamma_k^2(w)dw = O\left(\frac{k^{-m}}{NT}\right).$$

For C_3, write

$$\left|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}E\left[\phi_n(u_{it})\gamma_k(u_{it})\right]\right| \leq \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\int|\phi_n(d_tw)\gamma_k(d_tw)|f_t(w)dw$$

$$\leq O(1)\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{1}{d_t}\int|\phi_n(w)\gamma_k(w)|dw \leq O(1)\left\{\int\phi_n^2(w)dw\int\gamma_k^2(w)dw\right\}^{1/2} = O(k^{-m/2}).$$

Thus, $|C_3| \leq \frac{1}{N^2T}\sum_{n=1}^{d}\sum_{i=2}^{N}\sum_{j=1}^{i-1}o(k^{-m}) = O\left(\frac{1}{k^mT}\right)$. In connection with the analysis for C_1 and C_2, $\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\phi(u_{it})\gamma_k(u_{it}) = O_P\left(\frac{1}{\sqrt{k^mT}}\right)$. Then the proof is complete. ■

9) Write

$$E\left\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}v_{it}\gamma_k(u_{it})\right\|^2=\frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left[v_{it}'v_{js}\gamma_k(u_{it})\gamma_k(u_{js})\right]$$

$$=\frac{1}{N^2T^2}\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left[v_{it,n}v_{js,n}\right]E\left[\gamma_k(u_{it})\gamma_k(u_{js})\right]$$

$$\leq O\left(\frac{k^{-m+5/6}}{N^2T^2}\right)\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\alpha_{ij}(|t-s|))^{\delta/(4+\delta)}\left(E\left[|v_{it,n}|^{2+\delta/2}\right]\cdot E\left[|v_{js,n}|^{2+\delta/2}\right]\right)^{2/(4+\delta)}$$

$$\leq O\left(\frac{k^{-m+5/6}}{N^2T^2}\right)\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\alpha_{ij}(|t-s|))^{\delta/(4+\delta)}=O\left(\frac{k^{-m+5/6}}{NT}\right),$$

where the second equality follows from Assumption 1.4; the first inequality follows from (2) of Lemma B.1 and Davydov inequality; the last line follows from Assumption 1.3.a. Then the result follows. ■

Lemma B.3. *For two non–singular symmetric matrices A, B with same dimensions $k\times k$, where k tends to ∞. Suppose that their minimum eigenvalues satisfy that $\lambda_{min}(A)>0$ and $\lambda_{min}(B)>0$ uniformly in k. Then $\left\|A^{-1}-B^{-1}\right\|^2\leq\lambda_{min}^{-2}(A)\cdot\lambda_{min}^{-2}(B)\left\|A-B\right\|^2$.*

Proof of Lemma B.3:

For two non-singular symmetric matrices A and B with same dimensions, we observe that

$$\begin{aligned}\left\|A^{-1}-B^{-1}\right\|^2&=\left\|B^{-1}(B-A)A^{-1}\right\|^2=\left\|vec\left(B^{-1}(B-A)A^{-1}\right)\right\|^2\\&=\left\|\left(A^{-1}\otimes B^{-1}\right)vec(B-A)\right\|^2\leq\lambda_{min}^{-2}(A\otimes B)\left\|vec(B-A)\right\|^2\\&=\lambda_{min}^{-2}(A)\cdot\lambda_{min}^{-2}(B)\left\|A-B\right\|^2.\end{aligned}$$

The above calculation is straightforward and all the necessary theorems can be found in Magnus and Neudecker (2007). ■

Lemma B.4. *Let Assumptions 1 and 2 hold. As $(N,T)\to(\infty,\infty)$ jointly, (1) $\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}\mathcal{E}\right\|=O_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T}}\right)$; (2) $\left\|\frac{1}{N\sqrt{T}}X'\mathcal{Z}\right\|=O_P(1)$; (3) $\frac{1}{NT}X'\mathcal{E}=O_P\left(\frac{1}{\sqrt{NT}}\right)$; (4) $\frac{1}{NT}X'\gamma=O_P\left(\frac{1}{\sqrt{k^mT}}\right)$; (5) $\left\|\frac{1}{N\sqrt{T}}\mathcal{Z}'\gamma\right\|=O_P(k^{-(m-1)/2})$; (6) $\frac{1}{NT}X'X\to_P\Sigma_v$.*

Proof of Lemma B.4:

1)

$$\frac{1}{N\sqrt{T}}\mathcal{Z}'\mathcal{E}=\frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=1}^{T}Z_k(u_{it})e_{it}-\frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}Z_k(u_{it})e_{is}\equiv A_1-A_2$$

We have shown that $A_1=O_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T}}\right)$ in (6) of Lemma B.2. We then just focus on A_2. By Assumptions 1.1, 1.3.b and 1.4, write

$$
\begin{aligned}
E\left\|A_2\right\|^2 =& \frac{1}{N^2T^3}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}E\left[Z_k(u_{it_1})'Z_k(u_{it_2})\right]\sum_{s=1}^{T}E[e_{is}^2] \\
&+\frac{2}{N^2T^3}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}E\left[Z_k(u_{it_1})'\right]E\left[Z_k(u_{jt_2})\right]\sum_{s=1}^{T}E[e_{is}e_{js}] \\
=& \frac{\sigma_e^2}{N^2T^2}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}E\left[Z_k(u_{it_1})'Z_k(u_{it_2})\right] \\
&+\frac{2}{N^2T^2}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}E\left[Z_k(u_{it_1})'\right]E\left[Z_k(u_{jt_2})\right]\sigma_e(i,j) \equiv A_{21}+2A_{22}.
\end{aligned}
$$

For A_{21}, write

$$
|A_{21}| \le O\left(\frac{1}{N^2T^2}\right)\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}E\left[\|Z_k(u_{it_1})\|\|Z_k(u_{it_2})\|\right] \le O\left(\frac{Tk^{\frac{11}{6}}}{NT^2}\right) = o\left(\frac{k}{N\sqrt{T}}\right),
$$

where the second inequality has been provided in the proof of Lemma B.5 in this paper and the last equality follows from Assumption 2.2.

For A_{22}, write

$$
\begin{aligned}
|A_{22}| \le& \frac{1}{N^2T^2}\sum_{n=0}^{k-1}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\left|\int \mathscr{H}_n(d_{t_1}w)f_{t_1}(w)dw\right|\left|\int \mathscr{H}_n(d_{t_2}w)f_{t_2}(w)dw\right||\sigma_e(i,j)| \\
\le& O\left(\frac{1}{N^2T^2}\right)\sum_{n=0}^{k-1}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\int\frac{1}{d_{t_1}}\left|\mathscr{H}_n(w)\right|dw\int\frac{1}{d_{t_2}}\left|\mathscr{H}_n(w)\right|dw\sum_{i=2}^{N}\sum_{j=1}^{i-1}|\sigma_e(i,j)| \\
\le& O\left(\frac{\sum_{n=0}^{k-1}n^{\frac{5}{6}}}{NT}\right) \le O\left(\frac{k^2}{NT}\right) = o\left(\frac{k}{N\sqrt{T}}\right),
\end{aligned}
$$

where the last line follows from (8) of Lemma B.1 and Assumption 2.2.

By the analysis for A_{21} and A_{22}, we obtain that $A_2 = o_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T}}\right)$. In connection with that $A_1 = O_P\left(\frac{\sqrt{k}}{\sqrt{N}\sqrt[4]{T}}\right)$, the result follows. ■

2)

$$
\begin{aligned}
\frac{1}{N\sqrt{T}}X'\mathcal{Z} =& \frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=1}^{T}(\phi(u_{it})+v_{it})Z_k(u_{it})' - \frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\phi(u_{it})+v_{it})Z_k(u_{is})' \\
\equiv& B_1 - B_2
\end{aligned}
$$

$\|B_1\| = O_P(1)$ follows from (2) and (3) of Lemma B.2 of this paper immediately. Then we just need to focus on B_2 below.

$$
B_2 = \frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\phi(u_{it})Z_k(u_{is})' + \frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}v_{it}Z_k(u_{is})' \equiv B_{21}+B_{22}
$$

For B_{21}, write

$$
E\|B_{21}\|^2 = \frac{1}{N^2T^3}\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T}E[\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_2})\phi_{n_1}(u_{jt_3})\mathscr{H}_{n_2}(u_{jt_4})]
$$

$$
\begin{aligned}
&= \frac{1}{N^2T^3} \sum_{n_1=1}^{d} \sum_{n_2=0}^{k-1} \sum_{i=1}^{N} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \sum_{t_3=1}^{T} \sum_{t_4=1}^{T} E[\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\mathscr{H}_{n_2}(u_{it_4})] \\
&\quad + \frac{2}{N^2T^3} \sum_{n_1=1}^{d} \sum_{n_2=0}^{k-1} \sum_{i=2}^{N} \sum_{j=1}^{i-1} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \sum_{t_3=1}^{T} \sum_{t_4=1}^{T} E[\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_2})]E[\phi_{n_1}(u_{jt_3})\mathscr{H}_{n_2}(u_{jt_4})] \\
&\equiv B_{211} + 2B_{212}.
\end{aligned}
$$

For B_{211}, we write

$$
\begin{aligned}
B_{211} &= \frac{1}{N^2T^3} \sum_{n_1=1}^{d} \sum_{n_2=0}^{k-1} \sum_{i=1}^{N} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \sum_{t_3=1}^{T} \sum_{t_4=1}^{T} E[\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\mathscr{H}_{n_2}(u_{it_4})] \\
&= \frac{1}{N^2T^3} \sum_{n_1=1}^{d} \sum_{n_2=0}^{k-1} \sum_{i=1}^{N} \sum_{\text{all } t_1,t_2,t_3,t_4 \text{ are different}} E[\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\mathscr{H}_{n_2}(u_{it_4})] \\
&\quad + \frac{1}{N^2T^3} \sum_{n_1=1}^{d} \sum_{n_2=0}^{k-1} \sum_{i=1}^{N} \sum_{\text{only two of } t_1,t_2,t_3,t_4 \text{ are same}} E[\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\mathscr{H}_{n_2}(u_{it_4})] \\
&\quad + \frac{1}{N^2T^3} \sum_{n_1=1}^{d} \sum_{n_2=0}^{k-1} \sum_{i=1}^{N} \sum_{\text{only three of } t_1,t_2,t_3,t_4 \text{ are same}} E[\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\mathscr{H}_{n_2}(u_{it_4})] \\
&\quad + \frac{1}{N^2T^3} \sum_{n_1=1}^{d} \sum_{n_2=0}^{k-1} \sum_{i=1}^{N} \sum_{\text{four of } t_1,t_2,t_3,t_4 \text{ are same}} E[\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\mathscr{H}_{n_2}(u_{it_4})] \\
&\equiv B_{2111} + B_{2112} + B_{2113} + B_{2114}.
\end{aligned}
$$

For B_{2111}, without losing generality, assume that $t_1 > t_2 > t_3 > t_4$. For other cases, for example $t_2 > t_3 > t_1 > t_4$, the analysis will be same and the order will remain same. Then

$$
\begin{aligned}
&\frac{1}{N^2T^3} \sum_{n_1=1}^{d} \sum_{n_2=0}^{k-1} \sum_{i=1}^{N} \sum_{t_1=4}^{T} \sum_{t_2=3}^{t_1-1} \sum_{t_3=2}^{t_2-1} \sum_{t_4=1}^{t_3-1} |E[\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\mathscr{H}_{n_2}(u_{it_4})]| \\
&\leq O\left(\frac{1}{N^2T^3}\right) \sum_{n_1=1}^{d} \sum_{n_2=0}^{k-1} \sum_{i=1}^{N} \sum_{t_1=4}^{T} \sum_{t_2=3}^{t_1-1} \sum_{t_3=2}^{t_2-1} \sum_{t_4=1}^{t_3-1} \frac{1}{d_{t_1t_2}} \frac{1}{d_{t_2t_3}} \frac{1}{d_{t_3t_4}} \frac{1}{d_{t_4}} \\
&\cdot \iiiint |\phi_{n_1}(w_1)\mathscr{H}_{n_2}(w_2)\phi_{n_1}(w_3)\mathscr{H}_{n_2}(w_4)| dw_1 dw_2 dw_3 dw_4 \\
&\leq O\left(\frac{1}{N^2T^3}\right) \sum_{n_1=1}^{d} \sum_{n_2=0}^{k-1} \sum_{i=1}^{N} \sum_{t_1=4}^{T} \sum_{t_2=3}^{t_1-1} \sum_{t_3=2}^{t_2-1} \sum_{t_4=1}^{t_3-1} \frac{1}{d_{t_1t_2}} \frac{1}{d_{t_2t_3}} \frac{1}{d_{t_3t_4}} \frac{1}{d_{t_4}} \\
&\cdot \left(\int |\phi_{n_1}(w)| dw\right)^2 \left(\int |\mathscr{H}_{n_2}(w)| dw\right)^2 \\
&\leq O\left(\frac{1}{NT}\right) \sum_{n_2=0}^{k-1} n_2^{5/6} \leq O\left(\frac{k^2}{NT}\right) = o(1),
\end{aligned}
$$

where the last line follows from (8) of Lemma B.1, $\phi_n(w)$ being integrable function on $\mathbb{R}$ for $n = 1, \ldots, d$ and Assumption 2.2.

For B_{2112}, without losing generality, assume that $t_1 = t_2 > t_3 > t_4$. For other cases, for example $t_1 = t_3 > t_2 > t_4$, the analysis will be even simpler and the order will remain same. Then write

$$
\begin{aligned}
&\frac{1}{N^2T^3}\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{t_1=3}^{T}\sum_{t_3=2}^{t_1-1}\sum_{t_4=1}^{t_3-1}|E[\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_1})\phi_{n_1}(u_{it_3})\mathscr{H}_{n_2}(u_{it_4})]|\\
&\leq O\left(\frac{1}{N^2T^3}\right)\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{t_1=3}^{T}\sum_{t_3=2}^{t_1-1}\sum_{t_4=1}^{t_3-1}\frac{1}{d_{t_1t_3}}\frac{1}{d_{t_3t_4}}\frac{1}{d_{t_4}}\\
&\cdot\iiint|\phi_{n_1}(w_1)\mathscr{H}_{n_2}(w_1)\phi_{n_1}(w_2)\mathscr{H}_{n_2}(w_3)|dw_1dw_2dw_3\\
&\leq O\left(\frac{1}{NT^{3/2}}\right)\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\left\{\int\phi_{n_1}^2(w)dw\int\mathscr{H}_{n_2}^2(w)dw\right\}^{1/2}\int|\phi_{n_1}(w)|dw\int|\mathscr{H}_{n_2}(w)|dw\\
&\leq O\left(\frac{1}{NT^{3/2}}\right)\sum_{n_2=0}^{k-1}n_2^{5/12}\leq O\left(\frac{k^2}{NT^{3/2}}\right)=o(1),
\end{aligned}
$$

where the last line follows from (8) of Lemma B.1, $\phi_n(w)$ being integrable function on $\mathbb{R}$ for $n=1,\dots,d$ and Assumption 2.2.

For B_{2113}, without losing generality, assume that $t_1=t_2=t_3>t_4$. For other cases, for example $t_1=t_3=t_4>t_2$, the analysis will be same and the order will remain same. Then write

$$
\begin{aligned}
&\frac{1}{N^2T^3}\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{t_1=2}^{T}\sum_{t_4=1}^{t_1-1}|E[\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_1})\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_4})]|\\
&\leq O\left(\frac{1}{N^2T^3}\right)\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{t_1=2}^{T}\sum_{t_4=1}^{t_1-1}\frac{1}{d_{t_1t_4}}\frac{1}{d_{t_4}}\\
&\cdot\iint|\phi_{n_1}(w_1)\mathscr{H}_{n_2}(w_1)\phi_{n_1}(w_1)\mathscr{H}_{n_2}(w_2)|dw_1dw_2\\
&\leq O\left(\frac{1}{NT^2}\right)\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\int|\phi_{n_1}(w)|dw\int|\mathscr{H}_{n_2}(w)|dw=o(1),
\end{aligned}
$$

where the last line follows from $\mathscr{H}_j(w)$ and $\phi_j(w)$ being bounded uniformly, (8) of Lemma B.1 and Assumption 2.2.

For B_{2114}, write

$$
\begin{aligned}
B_{2114}&=\frac{1}{N^2T^3}\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{t=1}^{T}E[\phi_{n_1}^2(u_{it})\mathscr{H}_{n_2}^2(u_{it})]\\
&=\frac{1}{N^2T^3}\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\int\phi_{n_1}^2(d_tw)\mathscr{H}_{n_2}^2(d_tw)f_t(w)dw\\
&\leq O\left(\frac{1}{N^2T^3}\right)\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{1}{d_t}\int\phi_{n_1}^2(w)dw\leq O\left(\frac{k}{NT^{5/2}}\right)=o(1).
\end{aligned}
$$

Combining B_{2111}, B_{2112}, B_{2113} and B_{2114} together, we obtain that $B_{211}=o(1)$.

For B_{212}, write

$$
\begin{aligned}
&\left|E\left[\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\phi_{n_1}(u_{it_1})\mathscr{H}_{n_2}(u_{it_2})\right]\right|\\
&\leq\sum_{t=1}^{T}E[|\phi_{n_1}(u_{it})\mathscr{H}_{n_2}(u_{it})|]+\sum_{t=2}^{T}\sum_{s=1}^{t-1}E[|\phi_{n_1}(u_{it})\mathscr{H}_{n_2}(u_{is})|]+\sum_{t=2}^{T}\sum_{s=1}^{t-1}E[|\mathscr{H}_{n_2}(u_{it})\phi_{n_1}(u_{is})|]
\end{aligned}
$$

$$\le O(1)\sum_{t=1}^{T}\frac{1}{d_t}\int|\mathscr{H}_{n_2}(w)|dw+O(1)\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}\frac{1}{d_s}\int|\mathscr{H}_{n_2}(w)|dw$$
$$\le O(\sqrt{T}n_2^{5/12})+O(Tn_2^{5/12})=O(Tn_2^{5/12}),$$

where the last line follows from (8) of Lemma B.1. Therefore,

$$|B_{212}|\le O\left(\frac{1}{N^2T^3}\right)\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=2}^{N}\sum_{j=1}^{i-1}T^2n_2^{5/6}\le O\left(\frac{k^2}{T}\right)=o(1).$$

Since $|B_{211}|=o(1)$ and $|B_{212}|=o(1)$, then $B_{21}=o_P(1)$.

Below, we focus on B_{22}.

$$E\|B_{22}\|^2=E\left\|\frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}v_{it}Z_k(u_{is})'\right\|^2$$
$$=\frac{1}{N^2T^3}\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T}E[v_{it_1,n_1}\mathscr{H}_{n_2}(u_{it_2})v_{jt_3,n_1}\mathscr{H}_{n_2}(u_{jt_4})]$$
$$=\frac{1}{N^2T^3}\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T}E[v_{it_1,n_1}v_{it_3,n_1}]E[\mathscr{H}_{n_2}(u_{it_2})\mathscr{H}_{n_2}(u_{it_4})]$$
$$+\frac{2}{N^2T^3}\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T}E[v_{it_1,n_1}v_{jt_3,n_1}]E[\mathscr{H}_{n_2}(u_{it_2})]E[\mathscr{H}_{n_2}(u_{jt_4})]$$
$$\equiv B_{221}+2B_{222}.$$

For B_{221}, write

$$|B_{221}|\le\frac{1}{N^2T^3}\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T}|E[v_{it_1,n_1}v_{it_3,n_1}]E[\mathscr{H}_{n_2}(u_{it_2})\mathscr{H}_{n_2}(u_{it_4})]|$$
$$\le O\left(\frac{1}{N^2T^3}\right)\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_3=1}^{T}|E[v_{it_1,n_1}v_{it_3,n_1}]|\sum_{t_2=1}^{T}\frac{1}{d_{t_2}}\int\mathscr{H}_{n_2}^2(w)dw$$
$$+O\left(\frac{1}{N^2T^3}\right)\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_3=1}^{T}|E[v_{it_1,n_1}v_{it_3,n_1}]|\sum_{t_2=2}^{T}\sum_{t_4=1}^{t_2-1}\frac{1}{d_{t_2t_4}}\frac{1}{d_{t_4}}\left(\int|\mathscr{H}_{n_2}(w)|dw\right)^2$$
$$\le O\left(\frac{k}{N^2T^3}\right)\sum_{i=1}^{N}T^{3/2}+O\left(\frac{1}{N^2T^3}\right)\sum_{n_2=0}^{k-1}\sum_{i=1}^{N}T^2n_2^{5/6}\le O\left(\frac{k}{NT^{3/2}}\right)+O\left(\frac{k^2}{NT}\right)=o(1),$$

where the last line follows from Davydov inequality, Assumption 1.3 and (8) of Lemma B.1.

For B_{222}, write

$$|B_{222}|\le\frac{1}{N^2T^3}\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t_1=1}^{T}\sum_{t_3=1}^{T}|E[v_{it_1,n_1}v_{jt_3,n_1}]|$$
$$\cdot\sum_{t_2=1}^{T}\int|\mathscr{H}_{n_2}(d_{t_2}w)|f_{t_2}(w)dw\sum_{t_4=1}^{T}\int|\mathscr{H}_{n_2}(d_{t_4}w)|f_{t_4}(w)dw$$
$$\le O\left(\frac{1}{N^2T^3}\right)\sum_{n_1=1}^{d}\sum_{n_2=0}^{k-1}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t_1=1}^{T}\sum_{t_3=1}^{T}c_\delta(\alpha_{ij}(|t_1-t_3|))^{\delta/(4+\delta)}\cdot\left(E[|v_{it_1,n_1}|^{2+\delta/2}]\right)^{2/(4+\delta)}$$

$$\cdot \left(E[|v_{jt_3,n_1}|^{2+\delta/2}]\right)^{2/(4+\delta)} \sum_{t_2=1}^{T} \frac{1}{d_{t_2}} \int |\mathscr{H}_{n_2}(w)|dw \sum_{t_4=1}^{T} \frac{1}{d_{t_4}} \int |\mathscr{H}_{n_2}(w)|dw$$

$$\leq O\left(\frac{1}{N^2T^2}\right) \sum_{n_2=0}^{k-1} n_2^{5/6} \sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t_1=1}^{T}\sum_{t_3=1}^{T} (\alpha_{ij}(|t_1-t_3|))^{\delta/(4+\delta)} \leq O\left(\frac{k^2}{NT}\right) = o(1),$$

where the second inequality follows form Davydov inequality and $c_\delta = 2^{(4+2\delta)/(4+\delta)} \cdot (4+\delta)/\delta$; the third inequality follows from Assumption 1.3 and (8) of Lemma B.1.

Since $|B_{221}| = o(1)$ and $|B_{222}| = o(1)$, we know that $B_{22} = o_P(1)$.

Based on the analysis for B_1, B_{21} and B_{22}, the result follows. ■

3)

$$\frac{1}{NT}X'\mathcal{E} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\phi(u_{it})+v_{it})e_{it} - \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\phi(u_{it})+v_{it})e_{is} \equiv C_1 - C_2$$

Expand C_1 as

$$C_1 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\phi(u_{it})e_{it} + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} v_{it}e_{it} \equiv C_{11} + C_{12}.$$

We have shown that $C_{11} = O_P\left(\frac{1}{\sqrt{N}\sqrt[4]{T^3}}\right)$ in (7) of Lemma B.2. Moreover, by Assumption 1.3.b

$$E\|C_{12}\|^2 = \frac{1}{N^2T^2}\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T} E\left[v_{it,n}v_{jt,n}E[e_{it}e_{jt}|\mathcal{F}_{Nt-1}]\right] = O\left(\frac{1}{NT}\right).$$

Thus, $C_{12} = O_P\left(\frac{1}{\sqrt{NT}}\right)$. In connection with $C_{11} = O_P\left(\frac{1}{\sqrt{N}\sqrt[4]{T^3}}\right)$, we obtain $C_1 = O_P\left(\frac{1}{\sqrt{NT}}\right)$. Then we focus on C_2 below and write

$$C_2 = \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\phi(u_{it})e_{is} + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} v_{it}e_{is} = C_{21} + C_{22}.$$

For C_{21},

$$\begin{aligned}
E\left\|C_{21}\right\|^2 &= \frac{1}{N^2T^4}\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t=1}^{T} E[\phi_n^2(u_{it})]\sum_{s=1}^{T}E[e_{is}^2] \\
&\quad + \frac{2}{N^2T^4}\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t_1=2}^{T}\sum_{t_2=1}^{t_1-1} E[\phi_n(u_{it_1})\phi_n(u_{it_2})]\sum_{s=1}^{T}E[e_{is}^2] \\
&\quad + \frac{2}{N^2T^4}\sum_{n=1}^{d}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T} E[\phi_n(u_{it_1})\phi_n(u_{jt_2})]\sum_{s=1}^{T}E[e_{is}^2] \\
&\equiv C_{211} + 2C_{212} + 2C_{213}.
\end{aligned}$$

For C_{211},

$$C_{211} = O\left(\frac{1}{N^2T^3}\right)\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t=1}^{T}\int \phi_n^2(d_tw)f_t(w)dw \leq O\left(\frac{1}{NT^3}\right)\sum_{n=1}^{d}\sum_{t=1}^{T}\frac{1}{d_t}\int \phi_n^2(w)dw$$

$$= O\left(\frac{1}{NT^{5/2}}\right) = o\left(\frac{1}{NT}\right).$$

Similarly, for C_{212}

$$\begin{aligned}
|C_{212}| &\leq O\left(\frac{1}{N^2T^3}\right)\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t_1=2}^{T}\sum_{t_2=1}^{t_1-1} E[|\phi_n(u_{it_1})\phi_n(u_{it_2})|] \\
&\leq O\left(\frac{1}{N^2T^3}\right)\sum_{n=1}^{d}\sum_{i=1}^{N}\sum_{t_1=2}^{T}\sum_{t_2=1}^{t_1-1} \frac{1}{d_{t_1t_2}}\frac{1}{d_{t_2}}\int |\phi_n(w)|dw \int |\phi_n(w)|dw \\
&\leq O\left(\frac{1}{NT^2}\right) = o\left(\frac{1}{NT}\right),
\end{aligned}$$

where the last inequality follows from $f_{t_1t_2}(w)$ and $f_{t_2}(w)$ being bounded uniformly and $\phi_n(w)$ being integrable for $n = 1, \ldots, d$.

For C_{213}, note that

$$\begin{aligned}
|C_{213}| &\leq \frac{1}{N^2T^3}\sum_{n=1}^{d}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{i=2}^{N}\sum_{j=1}^{i-1} |E[\phi_n(u_{it_1})]E[\phi_n(u_{jt_2})]| \cdot |\sigma_e(i,j)| \\
&\leq \frac{1}{N^2T^3}\sum_{n=1}^{d}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{i=2}^{N}\sum_{j=1}^{i-1} \frac{1}{d_{t_1}}\int |\phi_n(w)|dw \cdot \frac{1}{d_{t_2}}\int |\phi_n(w)|dw \cdot |\sigma_e(i,j)| \\
&\leq O\left(\frac{1}{N^2T^3}\right)\sum_{n=1}^{d}\sum_{t_1=1}^{T}\frac{1}{d_{t_1}}\sum_{t_2=1}^{T}\frac{1}{d_{t_2}}\sum_{i=2}^{N}\sum_{j=1}^{i-1}|\sigma_e(i,j)| = O\left(\frac{1}{NT^2}\right) = o\left(\frac{1}{NT}\right),
\end{aligned}$$

where the last line follows form f_{t_1} and f_{t_2} being bounded uniformly, $\phi_n(w)$ being integrable for $n = 1, \ldots, d$ and Assumption 1.3.b.

Since $|C_{211}| = o\left(\frac{1}{NT}\right)$, $|C_{212}| = o\left(\frac{1}{NT}\right)$ and $|C_{213}| = o\left(\frac{1}{NT}\right)$, we obtain that $C_{21} = o_P\left(\frac{1}{\sqrt{NT}}\right)$.

By Assumption 1.3.c, it is straightforward to obtain that

$$E\|C_{22}\|^2 = \frac{1}{N^2T^4}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T} E[v_{it_1}'e_{it_2}v_{jt_3}e_{jt_4}] = O\left(\frac{1}{NT^2}\right).$$

Thus, $C_{22} = o_P\left(\frac{1}{\sqrt{NT}}\right)$. Since $C_{21} = o_P\left(\frac{1}{\sqrt{NT}}\right)$ and $C_{22} = o_P\left(\frac{1}{\sqrt{NT}}\right)$, then we have $C_2 = o_P\left(\frac{1}{\sqrt{NT}}\right)$. In connection with that $C_1 = O_P\left(\frac{1}{\sqrt{NT}}\right)$, the result follows. ■

4)

$$\frac{1}{NT}X'\gamma = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\phi(u_{it}) + v_{it})\gamma_k(u_{it}) - \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\phi(u_{it}) + v_{it})\gamma_k(u_{is}) \equiv D_1 - D_2$$

By (8) and (9) of Lemma B.2, $D_1 = O_P\left(\frac{1}{\sqrt{k^mT}}\right)$ follows immediately. D_2 can be expanded as

$$D_2 = \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\phi(u_{it})\gamma_k(u_{is}) + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}v_{it}\gamma_k(u_{is}) \equiv D_{21} + D_{22}.$$

For D_{21},

$$\|D_{21}\| \leq \frac{1}{NT^2}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T \|\phi(u_{it})\gamma_k(u_{is})\| \leq O(1)\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T |\gamma_k(u_{it})|,$$

where the second inequality follows from $\phi_n(w)$ being bounded uniformly for $n = 1, \ldots, d$. For the summation on RHS above,

$$\begin{aligned}
&E\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T |\gamma_k(u_{it})|\right|^2 = \frac{1}{N^2T^2}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T\sum_{s=1}^T E[|\gamma_k(u_{it})||\gamma_k(u_{js})|] \\
&= \frac{1}{N^2T^2}\sum_{i=1}^N\sum_{t=1}^T E[\gamma_k^2(u_{it})] + \frac{2}{N^2T^2}\sum_{i=1}^N\sum_{t=2}^T\sum_{s=1}^{t-1} E[|\gamma_k(u_{it})||\gamma_k(u_{is})|] \\
&+\frac{2}{N^2T^2}\sum_{i=2}^N\sum_{j=1}^{i-1}\sum_{t=1}^T\sum_{s=1}^T E[|\gamma_k(u_{it})|]E[|\gamma_k(u_{js})|] \\
&\leq O\left(\frac{1}{N^2T^2}\right)\sum_{i=1}^N\sum_{t=1}^T \frac{1}{d_t}\int \gamma_k^2(w)dw \\
&+O\left(\frac{1}{N^2T^2}\right)\sum_{i=1}^N\sum_{t=2}^T\sum_{s=1}^{t-1}\frac{1}{d_{ts}}\frac{1}{d_s}\int |\gamma_k(w)|^2 dw\left\{\int f_{ts}^2(w)dw\right\}^{1/2}\left\{\int f_s^2(w)dw\right\}^{1/2} \\
&+O\left(\frac{1}{N^2T^2}\right)\sum_{i=2}^N\sum_{j=1}^{i-1}\sum_{t=1}^T\sum_{s=1}^T\frac{1}{d_t}\frac{1}{d_s}\int |\gamma_k(w)|^2 dw\left\{\int f_t^2(w)dw\right\}^{1/2}\left\{\int f_s^2(w)dw\right\}^{1/2} \\
&\leq O\left(\frac{1}{NT^{3/2}}\right) + O\left(\frac{k^{-m}}{NT}\right) + O\left(\frac{k^{-m}}{T}\right), \qquad \text{(C.6)}
\end{aligned}$$

where the first inequality follows from $f_t(w)$ being bounded uniformly and Cauchy–Schwarz inequality; the second inequality follows from (3) of Lemma B.1 and $f_t(w)$ and $f_{ts}(w)$ being bounded uniformly. Then we have shown that $D_{21} = O_P\left(\frac{1}{\sqrt{k^m T}}\right)$.

We now focus on D_{22}.

$$\begin{aligned}
E\|D_{22}\|^2 &= \frac{1}{N^2T^4}\sum_{n=1}^d\sum_{i=1}^N\sum_{j=1}^N\sum_{t_1=1}^T\sum_{t_2=1}^T\sum_{t_3=1}^T\sum_{t_4=1}^T E[v_{it_1,n}v_{jt_3,n}]E[\gamma_k(u_{it_2})\gamma_k(u_{jt_4})] \\
&\leq O(1)\frac{1}{N^2T^2}\sum_{n=1}^d\sum_{i=1}^N\sum_{j=1}^N\sum_{t_2=1}^T\sum_{t_4=1}^T |E[\gamma_k(u_{it_2})\gamma_k(u_{jt_4})]| \leq O\left(\frac{k^{-m}}{T}\right),
\end{aligned}$$

where the first equality follows from Assumption 1.4; the first inequality follows from Assumption 1.3.a; the second inequality follows from (C.6). Then $D_{22} = O_P\left(\frac{1}{\sqrt{k^m T}}\right)$.

Based on the above, the result follows. ■

5)

$$\frac{1}{N\sqrt{T}}\mathcal{Z}'\gamma = \frac{1}{N\sqrt{T}}\sum_{i=1}^N\sum_{t=1}^T Z_k(u_{it})\gamma_k(u_{it}) - \frac{\sqrt{T}}{N}\sum_{i=1}^N \bar{Z}_{k,i}\bar{\gamma}_{k,i} \equiv E_1 - E_2$$

In (1) of Lemma B.2, we have shown that $\|E_1\| = O_P(k^{-(m-1)/2})$. Then we just need to focus on E_2 below and write

$$\|E_2\| \le \frac{1}{NT^{3/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \|Z_k(u_{it})\gamma_k(u_{is})\| \le O(k^{1/2}) \frac{1}{N\sqrt{T}} \sum_{i=1}^{N} \sum_{t=1}^{T} |\gamma_k(u_{it})|,$$

where the second inequality follows from (6) of Lemma B.1. In (C.6) of this lemma, we have shown that $\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}|\gamma_k(u_{it})| = O_P\left(\frac{k^{-m/2}}{\sqrt{T}}\right)$, so we easily obtain $\frac{1}{N\sqrt{T}}\sum_{i=1}^{N}\sum_{t=1}^{T}|\gamma_k(u_{it})| = O_P\left(k^{-m/2}\right)$. Based on the above, it further implies that $\|E_2\| = O_P(k^{-(m-1)/2})$. Then the result follows. ■

6) For the first result, write

$$\begin{aligned}
\frac{1}{NT}X'X &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\phi(u_{it}) + v_{it})(\phi(u_{it}) + v_{it})' \\
&\quad - \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}(\phi(u_{it}) + v_{it})(\phi(u_{is}) + v_{is})' \equiv F_1 - F_2.
\end{aligned}$$

By going a procedure similar to (C.6), it is easy to show that

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\phi(u_{it})\phi(u_{it})' \to_P 0.$$

In connection with (4) and (5) of Lemma B.2, we obtain that $F_1 \to_P \Sigma_v$ immediately.

We just need to focus on F_2 below.

$$\begin{aligned}
F_2 &= \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\phi(u_{it})\phi(u_{is})' + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\phi(u_{it})v_{is}' \\
&\quad + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}v_{it}\phi(u_{is})' + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}v_{it}v_{is}' \\
&\equiv F_{21} + F_{22} + F_{23} + F_{24}.
\end{aligned}$$

Notice that $F_{24} = o_P(1)$ follows from Assumption 1.3.c straightaway. We then focus on F_{21} below and write

$$\begin{aligned}
E\|F_{21}\|^2 &= \frac{1}{N^2T^4}\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T} E[\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_2})\phi_{n_1}(u_{jt_3})\phi_{n_2}(u_{jt_4})] \\
&= \frac{1}{N^2T^4}\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T} E[\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\phi_{n_2}(u_{it_4})] \\
&\quad + \frac{2}{N^2T^4}\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T} E[\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_2})\phi_{n_1}(u_{jt_3})\phi_{n_2}(u_{jt_4})] \\
&= F_{211} + 2F_{212}.
\end{aligned}$$

For F_{211}, write

$$F_{211} = \frac{1}{N^2T^4}\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}\sum_{i=1}^{N}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\sum_{t_3=1}^{T}\sum_{t_4=1}^{T} E[\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\phi_{n_2}(u_{it_4})]$$

$$
\begin{aligned}
&= \frac{1}{N^2T^4} \sum_{n_1=1}^{d} \sum_{n_2=1}^{d} \sum_{i=1}^{N} \sum_{\text{all } t_1,t_2,t_3,t_4 \text{ are different}} E[\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\phi_{n_2}(u_{it_4})] \\
&\quad + \frac{1}{N^2T^4} \sum_{n_1=1}^{d} \sum_{n_2=1}^{d} \sum_{i=1}^{N} \sum_{\text{only two of } t_1,t_2,t_3,t_4 \text{ are same}} E[\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\phi_{n_2}(u_{it_4})] \\
&\quad + \frac{1}{N^2T^4} \sum_{n_1=1}^{d} \sum_{n_2=1}^{d} \sum_{i=1}^{N} \sum_{\text{only three of } t_1,t_2,t_3,t_4 \text{ are same}} E[\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\phi_{n_2}(u_{it_4})] \\
&\quad + \frac{1}{N^2T^4} \sum_{n_1=1}^{d} \sum_{n_2=1}^{d} \sum_{i=1}^{N} \sum_{\text{four of } t_1,t_2,t_3,t_4 \text{ are same}} E[\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\phi_{n_2}(u_{it_4})] \\
&\equiv F_{2111} + F_{2112} + F_{2113} + F_{2114}.
\end{aligned}
$$

For F_{2111}, without losing generality, assume that $t_1 > t_2 > t_3 > t_4$. For other cases, for example $t_2 > t_3 > t_1 > t_4$, the analysis will be same and the order will remain same. Then

$$
\begin{aligned}
&\frac{1}{N^2T^4} \sum_{n_1=1}^{d} \sum_{n_2=1}^{d} \sum_{i=1}^{N} \sum_{t_1=4}^{T} \sum_{t_2=3}^{t_1-1} \sum_{t_3=2}^{t_2-1} \sum_{t_4=1}^{t_3-1} |E[\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_2})\phi_{n_1}(u_{it_3})\phi_{n_2}(u_{it_4})]| \\
&\leq O\left(\frac{1}{N^2T^4}\right) \sum_{n_1=1}^{d} \sum_{n_2=1}^{d} \sum_{i=1}^{N} \sum_{t_1=4}^{T} \sum_{t_2=3}^{t_1-1} \sum_{t_3=2}^{t_2-1} \sum_{t_4=1}^{t_3-1} \frac{1}{d_{t_1t_2}} \frac{1}{d_{t_2t_3}} \frac{1}{d_{t_3t_4}} \frac{1}{d_{t_4}} \\
&\quad \cdot \iiiint |\phi_{n_1}(w_1)\phi_{n_2}(w_2)\phi_{n_1}(w_3)\phi_{n_2}(w_4)| dw_1 dw_2 dw_3 dw_4 \\
&\leq O\left(\frac{1}{NT^2}\right) \sum_{n_1=1}^{d} \sum_{n_2=1}^{d} \left(\int |\phi_{n_1}(w)| dw\right)^2 \left(\int |\phi_{n_2}(w)| dw\right)^2 \leq O\left(\frac{1}{NT^2}\right) = o(1),
\end{aligned}
$$

where the last inequality follows from $\phi_n(w)$ being integrable for $n = 1, \ldots, d$.

For F_{2112}, without losing generality, assume that $t_1 = t_2 > t_3 > t_4$. For other cases, for example $t_1 = t_3 > t_2 > t_4$, the analysis will be even simpler and the order will remain same. Then write

$$
\begin{aligned}
&\frac{1}{N^2T^4} \sum_{n_1=1}^{d} \sum_{n_2=1}^{d} \sum_{i=1}^{N} \sum_{t_1=3}^{T} \sum_{t_3=2}^{t_1-1} \sum_{t_4=1}^{t_3-1} |E[\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_1})\phi_{n_1}(u_{it_3})\phi_{n_2}(u_{it_4})]| \\
&\leq O\left(\frac{1}{N^2T^4}\right) \sum_{n_1=1}^{d} \sum_{n_2=1}^{d} \sum_{i=1}^{N} \sum_{t_1=3}^{T} \sum_{t_3=2}^{t_1-1} \sum_{t_4=1}^{t_3-1} \frac{1}{d_{t_1t_3}} \frac{1}{d_{t_3t_4}} \frac{1}{d_{t_4}} \\
&\quad \cdot \iiint |\phi_{n_1}(w_1)\phi_{n_2}(w_1)\phi_{n_1}(w_2)\phi_{n_2}(w_3)| dw_1 dw_2 dw_3 \\
&\leq O\left(\frac{1}{NT^{5/2}}\right) \sum_{n_1=1}^{d} \sum_{n_2=1}^{d} \left\{\int \phi_{n_1}^2(w) dw \int \phi_{n_2}^2(w) dw\right\}^{1/2} \int |\phi_{n_1}(w)| dw \int |\phi_{n_2}(w)| dw \\
&\leq O\left(\frac{1}{NT^{5/2}}\right),
\end{aligned}
$$

where the last line follows from $\phi_n(w)$ being integrable and $\phi_n(w) \in L^2(\mathbb{R})$ for $n = 1, \ldots, d$.

For F_{2113}, without losing generality, assume that $t_1 = t_2 = t_3 > t_4$. For other cases, for example $t_1 = t_3 = t_4 > t_2$, the analysis will be same and the order will remain same. Then write

$$
\frac{1}{N^2T^4} \sum_{n_1=1}^{d} \sum_{n_2=1}^{d} \sum_{i=1}^{N} \sum_{t_1=2}^{T} \sum_{t_4=1}^{t_1-1} |E[\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_1})\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_4})]|
$$

$$\leq O\left(\frac{1}{N^2T^4}\right)\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}\sum_{i=1}^{N}\sum_{t_1=2}^{T}\sum_{t_4=1}^{t_1-1}\frac{1}{d_{t_1t_4}}\frac{1}{d_{t_4}}$$
$$\cdot\iint|\phi_{n_1}(w_1)\phi_{n_2}(w_1)\phi_{n_1}(w_1)\phi_{n_2}(w_2)|dw_1dw_2$$
$$\leq O\left(\frac{1}{NT^3}\right)\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}\int\phi_{n_1}^2(w)dw\int|\phi_{n_2}(w)|dw=O\left(\frac{1}{NT^3}\right),$$

where the last line follows from $\phi_n(w)$ being integrable and $\phi_n(w)\in L^2(\mathbb{R})$ for $n=1,\ldots,d$.

For F_{2114}, write

$$F_{2114}=\frac{1}{N^2T^4}\sum_{n_1=1}^{d}\sum_{n_2=1}^{d}\sum_{i=1}^{N}\sum_{t=1}^{T}\int\phi_{n_1}^2(d_tw)\phi_{n_2}^2(d_tw)f_t(w)dw$$
$$\leq O\left(\frac{1}{N^2T^4}\right)\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{1}{d_t}\leq O\left(\frac{1}{NT^{7/2}}\right),$$

where the first inequality follows from $\phi_n(w)$ being bounded uniformly and $\phi_n(w)\in L^2(\mathbb{R})$ for $n=1,\ldots,d$.

Combining F_{2111}, F_{2112}, F_{2113} and F_{2114} together, we obtain that $F_{211}=o(1)$.

We now turn to F_{212} and write

$$\left|E\left[\sum_{t_1=1}^{T}\sum_{t_2=1}^{T}\phi_{n_1}(u_{it_1})\phi_{n_2}(u_{it_2})\right]\right|$$
$$\leq\sum_{t=1}^{T}E[|\phi_{n_1}(u_{it})\phi_{n_2}(u_{it})|]+\sum_{t=2}^{T}\sum_{s=1}^{t-1}E[|\phi_{n_1}(u_{it})\phi_{n_2}(u_{is})|]+\sum_{t=2}^{T}\sum_{s=1}^{t-1}E[|\phi_{n_2}(u_{it})\phi_{n_1}(u_{is})|]$$
$$\leq O(1)\sum_{t=1}^{T}\frac{1}{d_t}\int|\phi_{n_1}(w)|dw+O(1)\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}\frac{1}{d_s}\int|\phi_{n_1}(w)|dw$$
$$\leq O(1)\sqrt{T}+O(1)T=O(T),$$

where the last inequality follows from $\phi_n(w)$ being integrable for $n=1,\ldots,d$. Therefore,

$$|F_{212}|\leq O\left(\frac{1}{N^2T^4}\right)\sum_{n=1}^{d}\sum_{i=2}^{N}\sum_{j=1}^{i-1}T^2=o(1).$$

Since $F_{211}=o(1)$ and $F_{212}=o(1)$, we have shown that $\|F_{21}\|=o_P(1)$. Similarly, we can show that $\|F_{22}\|=o_P(1)$ and $\|F_{23}\|=o_P(1)$. Therefore, the result follows. ■

Proof of Corollary 3.1:

We need only to verify the first result of this corollary. The second result then follows immediately.

1) By (6) of Lemma B.4, $\hat{\Sigma}_v=\frac{1}{NT}X'X\to_P\Sigma_v$. Thus, we just need to focus on $\hat{\sigma}_e^2$, where

$$\hat{\sigma}_e^2=\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\tilde{X}_{it}'(\beta_0-\hat{\beta})+\tilde{Z}_k(u_{it})'(C-\hat{C})+\tilde{\gamma}_k(u_{it})+\tilde{e}_{it})^2.\tag{C.7}$$

Now denote that $A_1 = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T(\tilde{X}_{it}'(\beta_0-\hat{\beta}))^2$, $A_2 = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T(\tilde{Z}_k(u_{it})'(C-\hat{C}))^2$, $A_3 = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\tilde{\gamma}_k^2(u_{it})$ and $A_4 = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\tilde{e}_{it}^2$.

For A_1, write

$$|A_1| \le \left\|\beta_0-\hat{\beta}\right\|^2 \cdot \left\|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\tilde{X}_{it}\tilde{X}_{it}'\right\| = o_P(1),$$

where the last equality follows from Theorem 3.1 and (6) of Lemma B.4.

For A_2, write

$$|A_2| \le \left\|C-\hat{C}\right\|^2 \cdot \frac{1}{\sqrt{T}}\left\|\frac{1}{N\sqrt{T}}\sum_{i=1}^N\sum_{t=1}^T\tilde{Z}_k(u_{it})\tilde{Z}_k(u_{it})'\right\| = o_P(1),$$

where the last equality follows from Lemma 3.1, Lemma B.5 and Assumption 2.2.

For A_3, by (2) of Lemma B.1, $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\tilde{\gamma}_k^2(u_{it}) = O(k^{-m+5/6}) = o(1)$.

For A_4, write

$$A_4 = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T e_{it}^2 - \frac{1}{NT^2}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T e_{it}e_{is} \equiv A_{41} - A_{42}.$$

For A_{41}, Assumption 1.3.(b),

$$E\left[A_{41}^2-\sigma_e^2\right]^2 = \frac{1}{N^2T^2}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T\sum_{s=1}^T E[(e_{it}^2-\sigma_e^2)(e_{js}^2-\sigma_e^2)] = o\left(1\right).$$

For A_{42}

$$E\left[A_{42}^2\right] = \frac{1}{N^2T^4}\sum_{i=1}^N\sum_{j=1}^N\sum_{t_1=1}^T\sum_{t_2=1}^T\sum_{t_3=1}^T\sum_{t_4=1}^T E[e_{it_1}e_{it_2}e_{jt_3}e_{jt_4}] = o\left(1\right),$$

where the RHS follows from e_t being martingale difference sequence (c.f. Assumption 1.3.b).

Therefore, we have shown that $A_1 \to_P 0$, $A_2 \to_P 0$, $A_3 \to_P 0$ and $A_4 \to_P \sigma_e^2$. Based on the above, all the interaction terms generated by $\tilde{X}_{it}'(\beta_0-\hat{\beta})$, $\tilde{Z}_k(u_{it})'(C-\hat{C})$ and $\tilde{\gamma}_k(u_{it})$ from the expansion of (C.7) can be shown converging to 0 in probability easily. For example,

$$\left|\frac{2}{NT}\sum_{i=1}^N\sum_{t=1}^T\tilde{X}_{it}'(\beta_0-\hat{\beta})\tilde{Z}_k(u_{it})'(C-\hat{C})\right| \le \frac{2}{NT}\sum_{i=1}^N\sum_{t=1}^T\left|\tilde{X}_{it}'(\beta_0-\hat{\beta})\tilde{Z}_k(u_{it})'(C-\hat{C})\right|$$
$$\le \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\left|\tilde{X}_{it}'(\beta_0-\hat{\beta})\right|^2 + \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\left|\tilde{Z}_k(u_{it})'(C-\hat{C})\right|^2 = A_1 + A_2 = o_P(1).$$

We now focus on the interaction terms generated by $\tilde{e}_{it}$.

Firstly,

$$\left|\frac{2}{NT}\sum_{i=1}^N\sum_{t=1}^T\tilde{X}_{it}'(\beta_0-\hat{\beta})\tilde{e}_{it}\right| \le \left\|\beta_0-\hat{\beta}\right\| \cdot \left\|\frac{2}{NT}\sum_{i=1}^N\sum_{t=1}^T\tilde{X}_{it}\tilde{e}_{it}\right\| = o_P(1),$$

where the last equality follows from Theorem 3.1 and (3) of Lemma B.4.

Secondly,

$$\left|\frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{Z}_k(u_{it})'(C-\hat{C})\tilde{e}_{it}\right| \leq \left\|C-\hat{C}\right\| \cdot \left\|\frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{Z}_k(u_{it})\tilde{e}_{it}\right\| = o_P(1),$$

where the last equality follows from Lemma 3.1 and (1) of Lemma B.4.

Thirdly, by similar approach to (9) of Lemma B.2, $\left|\frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{\gamma}_k(u_{it})\tilde{e}_{it}\right| = o_P(1)$.

Therefore, based on the above, the result follows. ■

References

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

Bai, J. and Carrion-I-Silvestre, J. L. (2009), 'Structural changes, common stochastic trends, and unit roots in panel data', *The Review of Economic Studies* pp. 471–501.

Bai, J., Kao, C. and Ng, S. (2009), 'Panel cointegration with global stochastic trends', *Journal of Econometrics* pp. 82–99.

Bai, J. and Ng, S. (2004), 'A panic attack on unit root and cointegration', *Econometrica* pp. 1127–1177.

Bai, J. and Ng, S. (2010), 'Panel unit root tests with cross–sectional dependence: A further investigation', *Econometric Theory* pp. 1088–1114.

Bosq, D. (1996), *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction*, Springer, New York.

Chen, J., Gao, J. and Li, D. (2012*a*), 'A new diagnostic test for cross-section uncorrelatedness in nonparametric panel data models', *Econometric Theory* pp. 1–20.

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

Chen, J., Gao, J. and Li, D. (2013), 'Estimation in partially linear single-index panel data models with fixed effects', *Journal of Business & Economic Statistics* pp. 315–330.

de Boeck, M. and Slok, T. (2006), 'Interpreting real exchange rate movements in transition countries', *Journal of International Economics* pp. 368–383.

Dong, C. and Gao, J. (2014), 'Specification testing in structural nonparametric cointegration', *https://ideas.repec.org/p/msh/ebswps/2014-2.html* .

Dong, C., Gao, J. and Tjøstheim, D. (2014), 'Estimation for single-index and partially linear single-index nonstationary time series models', *https://ideas.repec.org/p/msh/ebswps/2014-7.html* .

Gao, J. and Phillips, P. (2013), 'Semiparametric estimation in triangular system equations with nonstationarity', *Journal of Econometrics* pp. 59–79.

Hall, P. and Heyde, C. (1980), *Martingale Limit Theory and Its Applications*, Academic Press, New York.

Härdle, W., Liang, H. and Gao, J. (2000), *Partially Linear Models*, Springer, Heidelberg.

Hsiao, C. (2003), *Analysis of Panel Data*, Cambridge University Press.

Kapetanios, G., Pesaran, M. H. and Yamagata, T. (2011), 'Panels with non-stationary multifactor error structures', *Journal of Econometrics* pp. 326–348.

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

Nevai, P. (1986), 'Géza fredu, orthogonal polynomials and christoffel functions. A case study', *Journal of Approximation Theory* pp. 3–167.

Newey, W. K. (1997), 'Covergece rates and asymptotic normality for series estimators', *Journal of Econometrics* pp. 147–168.

Park, J. and Phillips, P. (2000), 'Nonstationary binary choice', *Econometrica* pp. 1249–1280.

Park, J. and Phillips, P. (2001), 'Nonlinear regressions with integrated time series', *Econometrica* pp. 117–161.

Pesaran, M. H. (2006), 'Estimation and inference in large heterogeneous panels with a multifactor error structure', *Econometrica* pp. 967–1012.

Pesaran, M. H. and Tosetti, E. (2011), 'Large panels with common factors and spatial correlation', *Journal of Econometrics* pp. 182–202.

Phillips, P. C. and Moon, H. R. (1999), 'Linear regression limit theory for nonstationary panel data', *Econometrica* pp. 1057–1111.

Robinson, P. (2012), 'Nonparametric trending regression with cross-sectional dependence', *Journal of Econometrics* pp. 4–14.

Su, L. and Jin, S. (2012), 'Sieve estimation of panel data models with cross section dependence', *Journal of Econometrics* pp. 34–47.