Journal of Economic Theory and Econometrics, Vol. 27, No. 4, Dec. 2016, 25–45

Common Correlated Effects Estimation of Unbalanced Panel Data Models with Cross-Sectional Dependence*

Qiankun Zhou† Yonghui Zhang‡

Abstract We consider the estimation and inference of unbalanced panel data models with cross-sectional dependence with a large number of individual units in a relatively short time period. By following the common correlated effects (CCE) approach of Pesaran (2006), we propose a CCE estimator for unbalanced panels (CCE-UB). The asymptotics of the CCE-UB estimator is developed in the paper. Small scale Monte Carlo simulation is conducted to examine the finite sample properties of the proposed estimator.

Keywords Common correlated effects, Cross-sectional dependence, Multifactor error structure, Unbalanced panel data model

JEL Classification C01, C12, C13, C33

*We are most grateful to the co-editor Heejoon Han and two anonymous referees for their helpful suggestions and constructive comments on the previous version of this paper.

†Department of Economics, State University of New York at Binghamton, Binghamton, NY, 13902, USA. Email: qzhou@binghamton.edu.

‡Address correspondence. School of Economics, Renmin University of China, 59 Zhongguancun Street Road, Haidian District, Beijing, China, 100872; Phone: +86 8250 0732; Email: yonghui.zhang@hotmail.com. Zhang gratefully acknowledges the financial support from the National Natural Science Foundation of China (Project No.71401166).

Received May 12, 2016, Revised September 20, 2016, Accepted December 21, 2016

1. INTRODUCTION

As discussed by Hsiao (2014), panel data provides several benefits for econometric estimation such as increasing degrees of freedom, alleviating the problem of data multicollinearity, and eliminating or reducing the estimation bias for a more robust inference, etc. Over the past few decades, econometric analysis of panel data models has grown into a major subfield of econometrics and gained increasing attentions both empirically and theoretically.

One major issue that arises in almost every research of panel data models with potential implications on parameter estimation and inference is the possible interdependence among different individual units. How to characterize or capture cross-sectional dependence (CSD for short) has attracted considerable interests among researchers over the years, see Sarafidis and Wansbeek (2012) for an overview and reference therein. A prominent approach of dealing with CSD is the factor structure approach,[1] which assumes the error term contains a finite number of unobserved factors that affect each individual with individual-specific factor loadings (e.g., Bai (2009), Bai and Li (2012, 2014) and Pesaran (2006)). For this approach, three main methods, namely, the principal component (PC) method by Bai (2009), the maximum likelihood estimation (MLE) of Bai and Li (2012, 2014) and the common correlated effects (CCE) approach of Pesaran (2006), have been developed in large panels where both cross-sectional and time series dimensions tend to infinity. The first two methods require to estimate the common factors and factor loadings, while the CCE approach focus on the estimation of slope parameters by using the cross-sectional averages of observables to approximate the unknown common factors. Furthermore, since researchers especially microeconometricians generally deal with panels involving a large number of individual units (N) in a relative short time period (T),[2] it is more challenging to deal with the presence of CSD in panel data models with a short/fixed time period. For some recent works on balanced panel with CSD , see Ahn, Lee, and Schmidt (2013), Juodis and Sarafidis (2014) and Hayakawa (2012) for GMM approach, and Bai (2013) and Hayakawa (2014) for maximum likelihood method.

Another important issue for panel data models is to take into account of unbalancedness of the data structure. Unbalanced panels can arise for various

[1]There is also arising literature on spatial approach to deal with CSD, which is developed primarily for cross-sectional data using a concept of a distance metric. For more details on spatial econometrics, see Lee and Yu (2010) for overview and the reference therein.

[2]For instance, the data come from surveys where a large group of people or households has been followed over a few years, e.g., the NLS and PSID dataset.

reasons. For example, a variable is unobserved in certain time periods due to some pre-specified rules, or individuals initially participating in the panel may not be willing or able to participate it anymore. Regarding this, it could be of crucial importance to take the feature of unbalancedness into account in the estimation of panel models. For more works on unbalanced panels, see Baltagi and Chang (1994), Wansbeek and Kapteyn (1989), and Baltagi and Song (2006). Note that almost all the previous works on unbalanced panels in the literature focus on models without CSD.

To our best knowledge, the only exception is Bai, Liao and Yang (2015) who consider an unbalanced panel data model with interactive fixed effects when both N and T are large. They propose an LS-EM-PCA algorithm to estimate the parameters, which combines the EM algorithm with the least squares (LS) method and principal component analysis (PCA). The LS-EM-PCA algorithm consists of two loops. The inner loop carries out the EM, while the outer loop estimates the slope parameters. Such iterative algorithm may be time-consuming and instable due to the possible existence of local optimizers. In addition, they only use simulation studies to show that the EM-type estimators are consistent and converge rapidly when both individual and time dimensions are large, say $N,T \geq 50$. No asymptotic analysis is provided due to the technically difficulty in proving consistency and further deriving the inferential theory of the proposed estimators.

In this paper, we consider the estimation and inference of unbalanced panel data models with CSD when N is large and T is small. To our best knowledge, it is the first paper to study the CCE estimator for unbalanced panel data. Also our paper contributes the literature on panel data model with cross-sectional dependence when T is small. To be specific, we modify the CCE estimate of Pesaran (2006) to our unbalanced panel and propose two methods to take cross-sectional averages for unbalanced data. We focus more on taking the cross-sectional averages of all available observations for each time period due to the efficiency consideration. Following Pesaran (2006), we derive the CCE estimate for unbalanced panel data models (CCE-UB for short). We show that the CCE-UB is consistent and asymptotically normally distributed under some regular conditions when N is large and T is small. A set of Monte Carlo simulations is conducted to investigate the finite sample performance of the proposed estimator in this paper. From the simulation results, we can observe that the CCE-UB is indeed consistent and leads to valid statistical inference for unbalanced panels with CSD. In all, we can conclude that the CCE-UB is suitable for estimation and inference for unbalanced panel with CSD when N is large and T is small.

The rest of the paper is organized as follows. In Section 2, we present the models with CSD and list the main assumptions. In Section 3, we introduce our CCE-UB estimator for the unbalanced panel and derive its asymptotics. In Section 4, we present Monte Carlo evidence for the finite sample performance of the proposed estimator. Conclusion is made at Section 5.

NOTATION. Throughout the paper, let C signifies a generic constant whose exact value may vary from case to case. "IID" refers to "independently and identically distributed". $\|A\| = [\mathtt{tr}(A'A)]^{1/2}$ denotes the Frobenius norm of matrix A. The operators $\xrightarrow{p}$ and $\xrightarrow{d}$ denote the convergences in probability and distribution, respectively.

2. MODEL AND ASSUMPTIONS

To begin with, let's assume the unbalanced panel data models are given by

$$\begin{aligned} y_{it} &= \alpha_{yi} + \mathbf{x}_{it}'\beta + e_{it}, & (1)\\ e_{it} &= \lambda_i'\mathbf{f}_t + v_{it}, & (2)\\ i &= 1,\ldots,N, t = t_i \in \mathscr{T}_i \equiv \{t_i(1),\ldots,t_i(T_i)\}, & (3) \end{aligned}$$

where $\mathscr{T}_i$ is the set included time indices of the observed observations for the ith individual, α_{yi} represents individual-specific effect, $\mathbf{x}_{it}$ is a $k \times 1$ vector of observables which is strict exogenous, λ_i and $\mathbf{f}_t$ are $r \times 1$ unobservables, and v_{it} is the idiosyncratic error. Usually, $\mathbf{f}_t$ refers to unobservable factors and λ_i refers to factor loadings, and r is unknown to researchers.

Throughout this paper, we assume the panels are unbalanced due to *randomly missing* observations. For each individual i, there are T_i observations available at times $(t_i(1),\ldots,t_i(T_i))$, and T_i can be different across i. Let $T = \max_{i=1,2,\ldots,N}\{T_i\}$. For each t, let $N_t = \sum_{i=1}^N 1(t \in \mathscr{T}_i)$ denote the number of observations observed at time period t. In this paper, we are interested in the estimation and inference of β when N is large while T is fixed.

In addition, we also assume that $\mathbf{x}_{it}$ is correlated with $\mathbf{f}_t$ as

$$\mathbf{x}_{it} = \alpha_{xi} + \Gamma_i\mathbf{f}_t + \varepsilon_{it}, \tag{4}$$

where $\alpha_{x,i}$ is a $k \times 1$ vector of individual-specific effects, Γ_i is a $k \times r$ matrix of factor loadings, and ε_{it} is a $k \times 1$ vector of idiosyncratic errors. The above setup is similar to that of Bai and Li (2014) and Pesaran (2006).

For models (1)-(4), we make the following assumptions for asymptotic analysis in the next section.

Assumption 1. $\mathbf{v}_i = (v_{i1},...,v_{iT})' \sim$IID$(\mathbf{0},\Omega_v)$ across i, and $\max_t Ev_{it}^4 < C < \infty$.

Assumption 2. $\varepsilon_i = (\varepsilon_{i1},...,\varepsilon_{iT})'$ are IID across i and ε_{it} has finite fourth moment for each t.

Assumption 3. The individual-specific errors $\mathbf{v}_i$ and ε_j are distributed independently for all i and j.

Assumption 4. The individual-specific effects α_{yi} and α_{xi} are IID across i and independent of v_{jt}, ε_{jt} and $\mathbf{f}_t$ for all j and t.

Assumption 5. (λ_i, Γ_i) are IID across i with finite 4th moment, and independent of v_{jt}, ε_{jt} and $\mathbf{f}_t$ for all j and t.

Assumption 6. The factors $\mathbf{f}_t$ have finite fourth moment and are independent of v_{is} and ε_{is} for all i,s and t.

Assumption 7. $\tilde{N} \equiv \min_{t=1,...,T}\{N_t\} \to \infty$ and $N/\tilde{N}^2 \to 0$, as $N \to \infty$.

Remark 1. *Assumptions 1-2 impose IID structure across i while allow for general nonstationarity along time. Assumptions 3-6 impose some dependent structure on the data generating process, which ensure that the regressors $\mathbf{x}_{it}$ are strictly exogenous and allows for the correlation between unobserved factors and $\mathbf{x}_{it}$. It is possible to relax Assumptions 1-2 to allow for weak CSD in $\{v_{it}\}$ or $\{\varepsilon_{it}\}$ as Bai (2009) with more complicated arguments. The existence of 4th moment is usually imposed to apply the LLN and CLT for independent but not identically distributed (inid) sequence. Assumption 7 requires that the minimum number of observed observations along t, $\tilde{N}$ should tend to infinity at a rate not slower than $N^{1/2}$, which is used in the establishing the limiting distribution of CCE-UB estimator. The consistency of CCE-UB estimator only requires $\tilde{N} \to \infty$.*

3. CCE ESTIMATOR FOR UNBALANCE PANEL WITH CSD

In this paper, we take the CCE approach of Pesaran (2006), which uses observed variables as the proxies for unknown factors, and is less affected by the problem of missing data once N is large. Another reason for choosing CCE is the fact, which is pointed out in a recent study by Westerlund and Urbain (2015), that the CCE estimators of slope coefficients generally perform the best in the case of homogeneous slopes and known number of unobserved common factors although the PC estimates of factors are more efficient than the cross-sectional averages. Also, we need to mention that the LS-EM-PCA algorithm in Bai, Liao and Yang (2015) does not work due to the fixed T in our setup.

The basic idea of CCE approach in Pesaran (2006) is to approximate the unobservable $\mathbf{f}_t$ by the linear combination of cross-sectional averages, $\bar{y}_i$ and

$\bar{\mathbf{x}}_i$. In our setup of unbalanced panel, we notice that models (1) and (4) can be rewritten in a compact form as

$$\begin{pmatrix} y_{it} \\ \mathbf{x}_{it} \end{pmatrix} = \begin{pmatrix} \alpha_{yi} + \beta' \alpha_{xi} \\ \alpha_{x,i} \end{pmatrix} + \begin{pmatrix} \beta' \Gamma_i + \lambda_i' \\ \Gamma_i \end{pmatrix} \mathbf{f}_t + \begin{pmatrix} v_{it} + \beta' \varepsilon_{it} \\ \varepsilon_{it} \end{pmatrix}, \quad (5)$$

where $t = t_i(s)$, $s = 1, \ldots, T_i$ and $i = 1, \ldots, N$. By letting

$$\underset{(k+1)\times 1}{\mathbf{z}_{it}} = \begin{pmatrix} y_{it} \\ \mathbf{x}_{it} \end{pmatrix}, \quad \underset{(k+1)\times 1}{\mu_i} = \begin{pmatrix} \alpha_{y,i} + \beta' \alpha_{xi} \\ \alpha_{xi} \end{pmatrix},$$

$$\underset{(k+1)\times r}{\mathbf{C}_i} = \begin{pmatrix} \beta' \Gamma_i + \lambda_i' \\ \Gamma_i' \end{pmatrix}, \quad \underset{(k+1)\times 1}{\mathbf{u}_{it}} = \begin{pmatrix} v_{it} + \beta' \varepsilon_{it} \\ \varepsilon_{it} \end{pmatrix},$$

(5) can be rewritten as

$$\mathbf{z}_{it} = \mu_i + \mathbf{C}_i \mathbf{f}_t + \mathbf{u}_{it}. \quad (6)$$

For unbalanced panel, there are two typical ways to take cross-sectional average. The first one is to use the same set of individual units for different time periods, and the second one is based on the actual number of observed individual units at each time period. For the first approach, let $\mathscr{S}_t = \{i : 1 \leq i \leq N, t \in \mathscr{T}_i\}$ denote the indicator set of observed individuals at the tth period. We take the cross-sectional average over the same set $\mathscr{S} = \cap_{t=1}^{T} \mathscr{S}_t$ for each t. That is

$$\bar{\mathbf{z}}_{St} = \bar{\mu}_S + \bar{\mathbf{C}}_S \mathbf{f}_t + \bar{\mathbf{u}}_{St}, \quad (7)$$

where $\bar{\mathbf{z}}_{St} = n^{-1} \sum_{i=1}^{N} \mathbf{z}_{it} 1_i, \bar{\mu}_S = n^{-1} \sum_{i=1}^{N} \mu_i 1_i, \bar{\mathbf{C}}_S = n^{-1} \sum_{i=1}^{N} \mathbf{C}_i 1_i,$ and $\bar{\mathbf{u}}_{St} = n^{-1} \sum_{i=1}^{N} \mathbf{u}_{it} 1_i$ with $n = |\mathscr{S}|$ being the cardinality of set $\mathscr{S}$ and $1_i \equiv 1(i \in \mathscr{S})$. In this case, the coefficient matrix for $\mathbf{f}_t$ in (7) is $\bar{\mathbf{C}}_S$, which is time-invariant in contrast with that in (8). This feature facilitates the asymptotic study for the estimates since it induces the same structure as Pesaran (2006). However, there are some obvious efficiency loss in the approximation of factors because many observations are not used in the cross-sectional averaging. In addition, to make CCE work, the number of individuals in the common set $\mathscr{S}$ should go to infinity at a rate faster than $N^{1/2}$. The requirement can be too restrictive in empirical applications and often breaks down. Therefore we prefer to the second approach. For each t, taking the cross-sectional average over all observed individuals leads to,

$$\bar{\mathbf{z}}_t = \bar{\mu}_t + \bar{\mathbf{C}}_t \mathbf{f}_t + \bar{\mathbf{u}}_t, \quad (8)$$

where $\bar{\mathbf{z}}_t = N_t^{-1}\sum_{i=1}^{N}\mathbf{z}_{it}1_{it}$, $\bar{\mu}_t = N_t^{-1}\sum_{i=1}^{N}\mu_i 1_{it}$, $\bar{\mathbf{C}}_t = N_t^{-1}\sum_{i=1}^{N}\mathbf{C}_i 1_{it}$, $\bar{\mathbf{u}}_t = N_t^{-1}\sum_{i=1}^{N}\mathbf{u}_{it}1_{it}$ with $1_{it} \equiv 1\left(t \in \mathscr{T}_i\right)$. However, the coefficient matrix in (8) for $\mathbf{f}_t$ is $\bar{\mathbf{C}}_t$, which is not time-invariant and different from Pesaran (2006). If we construct the approximator of $\mathbf{f}_t$ based on (8) directly, we cannot obtain time-invariant coefficients for $\bar{\mathbf{z}}_t$ in (10). Fortunately, by Assumptions 6 and 7 and Chebyshev's inequality, we can show that as $N_t \to \infty$, $\bar{\mathbf{C}}_t = \mathbf{C} + O_P(N_t^{-1/2})$ and $\bar{\mu}_t = \mu + O_P(N_t^{-1/2})$, where $\mathbf{C} = E\left(\mathbf{C}_i\right)$ and $\mu = E\left(\mu_i\right)$. Then we have

$$\bar{\mathbf{z}}_t = \mu + \mathbf{C}\mathbf{f}_t + \bar{\mathbf{u}}_t^* \tag{9}$$

where $\bar{\mathbf{u}}_t^* = \bar{\mathbf{u}}_t + \left(\bar{\mu}_t - \mu\right) + \left(\bar{\mathbf{C}}_t - \mathbf{C}\right)\mathbf{f}_t$ includes additional two terms: $\bar{\mu}_t - \mu$ and $\left(\bar{\mathbf{C}}_t - \mathbf{C}\right)\mathbf{f}_t$. Now we have

$$\mathbf{f}_t = \left(\mathbf{C}'\mathbf{C}\right)^{-1}\mathbf{C}'\left(\bar{\mathbf{z}}_t - \mu\right) + O_P(N_t^{-1/2}),$$

by the fact $\texttt{rank}\left(\mathbf{C}\right) = r \le k+1$ and $\bar{\mathbf{u}}_t^* = O_P(N_t^{-1/2})$ as $N_t \to \infty$. Then, for any t, $\mathbf{f}_t$ can be approximated by the linear combination of μ and $\bar{\mathbf{z}}_t$ by using Frish-Waugh Theorem (Sarafidis and Wansbeek, 2012).

As a result, replacing $\mathbf{f}_t$ by $\left(\mathbf{C}'\mathbf{C}\right)^{-1}\mathbf{C}'\left(\bar{\mathbf{z}}_t - \mu\right)$ leads to the augmented regression model for (1) as follows

$$y_{it} = \alpha_{yi}^* + \mathbf{x}_{it}'\beta + \mathbf{b}_i'\bar{\mathbf{z}}_t + \varepsilon_{it}, \tag{10}$$

where $\alpha_{yi}^* = \alpha_{yi} + \mathbf{a}_i'\mu$ with $\mathbf{a}_i$ and $\mathbf{b}_i$ are nuisance parameters. Rewrite model (10) in vector form we have

$$\underset{T_i\times 1}{\mathbf{y}_{iT_i}} = \alpha_{yi}^*\underset{T_i\times 1}{\mathbf{1}_{T_i}} + \underset{T_i\times k}{\mathbf{X}_{iT_i}}\underset{k\times 1}{\beta} + \underset{T_i\times(k+1)}{\bar{\mathbf{Z}}_{T_i}}\underset{(k+1)\times 1}{\mathbf{b}_i} + \underset{T_i\times 1}{\varepsilon_{iT_i}}, \tag{11}$$

where $\mathbf{y}_{iT_i} = \left(y_{it_i(1)}, \ldots, y_{it_i(T_i)}\right)'$, $\mathbf{X}_{i,T_i} = \left(\mathbf{x}_{it_i(1)}, \ldots, \mathbf{x}_{it_i(T_i)}\right)'$, $\bar{\mathbf{Z}}_{T_i} = \left(\bar{\mathbf{z}}_1, \ldots, \bar{\mathbf{z}}_{T_i}\right)'$, $\varepsilon_{iT_i} = \left(\varepsilon_{it_i(1)}, \ldots, \varepsilon_{it_i(T_i)}\right)'$, and $\mathbf{1}_{T_i}$ is a $T_i \times 1$ vector of ones.

The CCE-UB estimator of β based on model (11) is given by

$$\hat{\beta}_{CCE}^{UB} = \left(\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{M}_{H_i}\mathbf{X}_{iT_i}\right)^{-1}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{M}_{H_i}\mathbf{y}_{iT_i}, \tag{12}$$

where the superscript "UB" stands for unbalanced panel and $\mathbf{M}_{H_i} = \mathbf{I}_{T_i} - \bar{\mathbf{H}}_{T_i}(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1}\bar{\mathbf{H}}_{T_i}'$ with $\bar{\mathbf{H}}_i = \left(\mathbf{1}_{T_i}, \bar{\mathbf{Z}}_{T_i}\right)$.

Let $\mathbf{F}_{T_i} = (\mathbf{f}_{t_i(1)}, \ldots, \mathbf{f}_{t_i(T_i)})'$, $\mathbf{G}_{T_i} = (\mathbf{1}_{T_i}, \mathbf{F}_{T_i})$, and $\mathbf{M}_{G_i} = \mathbf{I}_{T_i} - \mathbf{G}_{T_i}(\mathbf{G}_{T_i}'\mathbf{G}_{T_i})^{-1}\mathbf{G}_{T_i}'$. Define

$$\mathbf{D}_N\left(\mathbf{G}_{T_i}\right) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{M}_{G_i}\mathbf{X}_{iT_i} \text{ and } \mathbf{D} = \operatorname*{plim}_{N\to\infty}\mathbf{D}_N\left(\mathbf{G}_{T_i}\right).$$

Similarly, we write $\mathbf{D}_N\left(\bar{\mathbf{H}}_{T_i}\right)=\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{M}_{H_i}\mathbf{X}_{iT_i}$.

We establish the asymptotic properties for $\hat{\beta}_{CCE}^{UB}$ in the following theorem.

Theorem 1. *Suppose Assumptions 1-6 and rank*$(\mathbf{C})=r\leq k+1$ *hold. If* $\tilde{N}\to\infty$ *as* $N\to\infty$,

$$\hat{\beta}_{CCE}^{UB}\xrightarrow{p}\beta.$$

Further, if Assumption 7 also holds, then

$$\sqrt{N}\left(\hat{\beta}_{CCE}^{UB}-\beta\right)\xrightarrow{d}N\left(\mathbf{0},\mathbf{V}\right)$$

where $\mathbf{V}=\mathbf{D}^{-1}\Sigma\mathbf{D}^{-1}$ *and* $\Sigma=\operatorname{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{M}_{G_i}\Omega_v\mathbf{M}_{G_i}\mathbf{X}_{iT_i}$.

Theorem 1 establishes the consistency of CCE-UB estimate and derives its limiting distribution. The proof is tedious and is therefore relegated to Appendix. Note that the condition on the number $\tilde{N}$ to ensure consistency is much weaker than that used in deriving the asymptotic distribution.

To carry out the statistical inference, we have to construct the estimator for the asymptotic variance matrix $\mathbf{V}$ of $\hat{\beta}_{CCE}^{UB}$. A consistent variance estimator can be reached by

$$\hat{\mathbf{V}}=\mathbf{D}_N^{-1}\left(\bar{\mathbf{H}}_{T_i}\right)\hat{\Omega}\mathbf{D}_N^{-1}\left(\bar{\mathbf{H}}_{T_i}\right)\tag{13}$$

where $\mathbf{D}$ and Σ in $\mathbf{V}$ are replaced by $\mathbf{D}_N\left(\bar{\mathbf{H}}_{T_i}\right)$ and

$$\hat{\Sigma}=\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{M}_{H_i}\hat{\mathbf{e}}_{iT_i}\hat{\mathbf{e}}'_{iT_i}\mathbf{M}_{H_i}\mathbf{X}_{iT_i}\tag{14}$$

respectively, where $\hat{\mathbf{e}}_{iT_i}=\mathbf{M}_{H_i}\mathbf{y}_{iT_i}-\mathbf{M}_{H_i}\mathbf{X}_{iT_i}\hat{\beta}_{CCE}^{UB}$. When T is not small, a consistent variance estimator can be constructed by following the argument of White (2001) and Pesaran (2006).

Remark 2. *In the case when* $T_i=T_j$ *for* $i\neq j$, *i.e., the balanced panel case, Sarafidis and Wansbeek (2012, P497) argue that when* T *is fixed, the limiting distribution of pooled CCE (12) for heterogeneous panel is nonstandard. As we demonstrated, for homogenous panel, even* T *is fixed, the pooled CCE-UB (12) still converges to a normal distribution asymptotically which is free of nuisance parameters. As shown in the simulation below, the pooled CCE (12) is consistent, and the* t*-test based on our asymptotic distribution has appropriate size and power performance in finite sample.*

4. SIMULATIONS

In this section, we investigate the finite sample performance of the CCE-UB estimator for unbalanced panels discussed in the previous section. The data generating process (DGP) is given by

$$y_{it} = 1 + \alpha_{yi} + x_{1,it}\beta_1 + x_{2,it}\beta_2 + \lambda_{1i}f_{1t} + \lambda_{2i}f_{2t} + u_{it}, \tag{15}$$

and $\mathbf{x}_{it}$*'s* are generated according to

$$x_{k,it} = 1 + c_{k1}\lambda_{1i} + c_{k2}\lambda_{2i} + \pi_{i,k1}f_{1t} + \pi_{i,k2}f_{2t} + \eta_{k,it},\ k = 1,2;$$

where

$$\eta_{k,it} = \rho_{ki}\eta_{k,it-1} + v_{k,it},$$

with ρ_{ki} are IID drawn from $U[0.1, 0.9]$ for $k = 1,2$ and $i = 1,\ldots,N$. Let $\alpha_{yi} \sim$IID$N(0,1)$, $u_{it} \sim$IID$N\left(0, \sigma^2_{u,i}\right)$, $v_{k,it} \sim$IID $N\left(0, \sigma^2_{vk,i}\right)$ for $k = 1,2$, and $\sigma^2_{u,i}$, $\sigma^2_{v1,i}$, $\sigma^2_{v2,i}$ are independently drawn from $0.5\left(1 + 0.5\chi^2(2)\right)$. For the factors, let $f_{jt} \sim$IID$N(0,1)$, $j = 1,2$. For the factor loadings, λ_{ri} are IID drawn from $N(1,1)$ and $\pi_{i,kr}$ are IID drawn from $U[0,2]$ for $i = 1,2,\ldots,N$, $k = 1,2$, and $r = 1,2$. We set $c_{11} = 0.5$, $c_{12} = 2$, $c_{21} = 2$, and $c_{22} = 0.5$.

The true value of β_1 and β_2 are set at $\beta_1 = 1$ and $\beta_2 = 2$. We let $N \in \{50, 100, 200, 400\}$ and T_i are the integers uniformly drawn from $[5, 20]$ in each replication. We consider two patterns of unbalanced data:

1. **UB1**: Consecutive observations with common initial observed period ($t_i(1) = 1$ for all i). The observed time periods are $1, 2, ..., T_i$ for each i.

2. **UB2**: Nonconsecutive observations with different initial observed periods. For each i, we randomly draw T_i time periods from $\{1, 2, ..., T_{\max}\}$.

The number of replications is 1000. We report the bias, absolute bias (Abias) and RMSE. The simulation results are given by in Table 1. From Table 1, we can observe that the CCE-UB estimator is consistent for large N and small T. The bias and RMSE of the CCE-UB estimators decrease with the increase of cross-sectional dimension N for both unbalanced patterns, which suggests that our CCE-UB estimator has good finite sample performance in estimating the unknown slope coefficients.

To examine the statistical inference performance of the CCE-UB estimators, for UB1, we draw the rejection frequencies plots for tests $\mathbb{H}_{01} : \beta_1 = 1$ vs $\mathbb{H}_{11}$:

Table 1: Estimation results for DGP (15)

		UB1		UB2	
N		β_1	β_2	β_1	β_2
50	Bias	0.0194	0.0196	0.0093	0.0167
	Abias	0.1088	0.1122	0.0996	0.1022
	RMSE	0.1411	0.1441	0.1273	0.1309
100	Bias	0.0088	0.0109	0.0093	0.0099
	Abias	0.0780	0.0853	0.0754	0.0723
	RMSE	0.1008	0.1100	0.0951	0.0935
200	Bias	0.0041	0.0026	0.0024	0.0029
	Abias	0.0554	0.0587	0.0538	0.0515
	RMSE	0.0742	0.0757	0.0685	0.0663
400	Bias	0.0027	-0.0006	0.0008	0.0029
	Abias	0.0411	0.0407	0.0384	0.0392
	RMSE	0.0526	0.0524	0.0486	0.0500

Note: Abias refers to absolute bias and RMSE is the root MSE.

$\beta_1 \neq 1$ and $\mathbb{H}_{02} : \beta_2 = 2$ vs $\mathbb{H}_{12} : \beta_2 \neq 2$ when $N = 100$ in Figures 1 and 2, separately, where we use the variance estimator proposed in (13). We can see from the figures that both tests have correct size under the null hypotheses, and their powers increase fast as the parameters are getting away from the values under $\mathbb{H}_{01}$ and $\mathbb{H}_{02}$, respectively.

In all, we can conclude that the CCE-UB estimator for unbalanced panel data models with CSD has desirable finite sample properties and is suitable for statistical inference purposes.

5. CONCLUSION

In this paper, we consider the estimation and inference of unbalanced panel data models with cross-sectional dependence when cross-sectional dimension is large and the time series dimension is small or fixed. We adapt the CCE approach of Pesaran (2006) and propose the CCE-UB estimator for unbalanced panels.

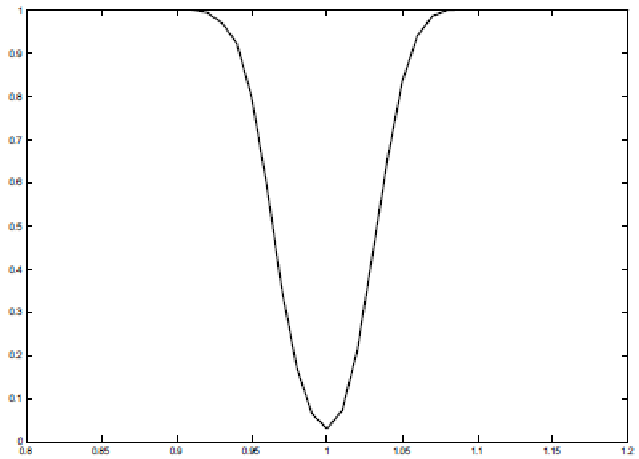

Figure 1: The plots of rejection frequencies of test $\mathbb{H}_{01}$: $\beta_1 = 1$

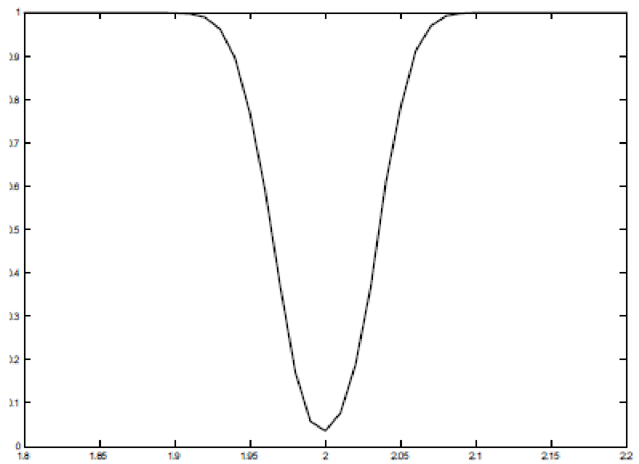

Figure 2: The plots of rejection frequencies of test $\mathbb{H}_{02}$: $\beta_2 = 2$

The asymptotics of the CCE-UB is developed in the paper and it is shown to be consistent and asymptotically normally distributed. Finite sample properties of the CCE-UB is investigated by simulation and it is shown the CCE-UB has desirable finite sample performance.

Appendix

A. PROOF OF THEOREM 1

Before analyze the asymptotic properties of $\hat{\beta}^{UB}_{CCE}$, we derive some equations for the unobservable factors. Let $\mathbf{A} = (\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'$. Recall that

$$\bar{\mathbf{z}}_t = \mu + \mathbf{C}\mathbf{f}_t + \bar{\mathbf{u}}_t^* \text{ and } \mathbf{f}_t = \mathbf{A}\bar{\mathbf{z}}_t - \mathbf{A}\bar{\mu} - \mathbf{A}\bar{\mathbf{u}}_t^*,$$

where $\bar{\mathbf{u}}_t^* = \bar{\mathbf{u}}_t + (\bar{\mu}_t - \mu) + \left(\bar{\mathbf{C}}_t - \mathbf{C}\right)\mathbf{f}_t$. Then

$$\mathbf{F}_{T_i} = (\mathbf{f}_{t_i(1)}, \ldots, \mathbf{f}_{t_i(T_i)})' = \bar{\mathbf{Z}}_{T_i}\mathbf{A}' - \left(\mathbf{1}_{T_i} \otimes \mu'\right)\mathbf{A}' + \mathbf{U}^*_{T_i}\mathbf{A}' \tag{16}$$

where $\bar{\mathbf{Z}}_{T_i} \equiv \left(\bar{\mathbf{z}}_{t_i(1)}, ..., \bar{\mathbf{z}}_{t_i(T_i)}\right)'$, $\mathbf{U}^*_{T_i} = (\bar{\mathbf{u}}^*_{t_i(1)}, ..., \bar{\mathbf{u}}^*_{t_i(T_i)})'$ and $\mathbf{1}_{T_i}$ is a $T_i \times 1$ vector of ones. Let $\bar{\mu}_{iT_i} = (\bar{\mu}_{t_i(1)}, ..., \bar{\mu}_{t_i(T_i)})'$ and $\bar{\mathbf{C}}_{iT_i} = \left(\bar{\mathbf{C}}_{t_i(1)}, ..., \bar{\mathbf{C}}_{t_i(T_i)}\right)'$. Then rewrite the models in (4) and (8) as follows

$$\underset{T_i\times k}{\mathbf{X}_{iT_i}} = \underset{T_i\times(r+1)}{\mathbf{G}_{T_i}} \underset{(r+1)\times k}{\Pi_i} + \underset{T_i\times k}{\varepsilon_{iT_i}}, \tag{17}$$

$$\underset{T_i\times(k+2)}{\bar{\mathbf{H}}_{T_i}} = \underset{T_i\times(r+1)}{\mathbf{G}_{T_i}} \underset{(r+1)\times(k+2)}{\mathbf{P}} + \underset{T_i\times(k+2)}{\mathbf{U}^\dagger_{T_i}} \tag{18}$$

where $\Pi_i = (\alpha_{xi}, \Gamma_i)'$, $\varepsilon_{iT_i} = \left(\varepsilon_{it_i(1)}, \ldots, \varepsilon_{it_i(T_i)}\right)'$, $\mathbf{P} = \begin{pmatrix} 1 & \mu' \\ \mathbf{0} & \mathbf{C}' \end{pmatrix}$, and $\mathbf{U}^\dagger_{T_i} = (\mathbf{0}, \mathbf{U}^*_{T_i})$ with $\mathbf{U}^*_{T_i} = \bar{\mathbf{U}}_{T_i} + \mathbf{U}^\mu_{T_i} + \mathbf{U}^c_{T_i}$, $\bar{\mathbf{U}}_{T_i} = \left(\bar{\mathbf{u}}_{t_i(1)}, \ldots, \bar{\mathbf{u}}_{t_i(T_i)}\right)'$, $\mathbf{U}^\mu_{T_i} = \bar{\mu}_{T_i} - (\mathbf{1}_{T_i} \otimes \mu')$, and $\mathbf{U}^C_{T_i} = [\mathbf{C}_{T_i} - (\mathbf{1}_{T_i} \otimes \mathbf{C}')]\mathbf{F}_{T_i}$.

We summarize some preliminary results in the following lemma.

Lemma 2. *Under Assumptions 1-7, when T is fixed and $N \to \infty$, we have uniformly in i,*

(i) $\left\|\bar{\mathbf{H}}_{T_i} - \mathbf{G}_{T_i}\mathbf{P}\right\| = O_p(\tilde{N}^{-1/2})$;

(ii) $\left\|(\bar{\mathbf{H}}'_{T_i}\bar{\mathbf{H}}_{T_i})^{-1} - (\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}\right\| = O_p(\tilde{N}^{-1/2})$;

(iii) $\|\mathbf{M}_{H_i} - \mathbf{M}_{G_i}\| = O_p(\tilde{N}^{-1/2})$;

Proof. Let $\bar{\mathbf{H}}_T = \left(\mathbf{1}_T, \bar{\mathbf{Z}}_T\right)$, where $\mathbf{1}_T$ is $T \times 1$ vector of ones and $\bar{\mathbf{Z}}_T = (\bar{\mathbf{z}}_1, ..., \bar{\mathbf{z}}_T)'$. Define $\mathbf{G}_T = (\mathbf{1}_T, \mathbf{F})$, $\mathbf{U}^\dagger_T = (\mathbf{0}, \mathbf{U}^*_T)$ and $\mathbf{U}^*_T = \bar{\mathbf{U}}_T + \mathbf{U}^\mu_T + \mathbf{U}^c_T$, where $\bar{\mathbf{U}}_T = (\bar{\mathbf{u}}_1, \ldots, \bar{\mathbf{u}}_T)'$ and $\mathbf{U}^\mu_T$ and $\mathbf{U}^c_T$ are defined similarly. Noting that

$$E\left\|\bar{\mathbf{U}}_T\right\|^2 = \sum_{t=1}^T E\left\|\bar{\mathbf{u}}_t\right\|^2 = \sum_{t=1}^T \frac{1}{N_t^2}\sum_{i=1}^N E(\|\mathbf{u}_{it}\|^2 1_{it}) = \sum_{t=1}^T O\left(\frac{1}{N_t}\right) = O\left(\frac{1}{\tilde{N}}\right),$$

we have $\left\|\bar{\mathbf{U}}_T\right\| = O_p(\tilde{N}^{-1/2})$ by the Markov inequality. Similarly, we can show that $\left\|\mathbf{U}_T^{\mu}\right\| = O_p(\tilde{N}^{-1/2})$ and $\|\mathbf{U}_T^{c}\| = O_p(\tilde{N}^{-1/2})$. It follows that $\left\|\bar{\mathbf{H}}_T - \mathbf{G}_T\mathbf{P}\right\| = \left\|\mathbf{U}_T^{\dagger}\right\| \leq \left\|\bar{\mathbf{U}}_T\right\| + \left\|\mathbf{U}_T^{\mu}\right\| + \|\mathbf{U}_T^{c}\| = O_p(\tilde{N}^{-1/2})$. Noting that $\bar{\mathbf{H}}_{T_i} - \mathbf{G}_{T_i}\mathbf{P}$ are sub-vector of $\bar{\mathbf{H}}_T - \mathbf{G}_T\mathbf{P}$, so we have $\left\|\bar{\mathbf{H}}_{T_i} - \mathbf{G}_{T_i}\mathbf{P}\right\| = O_p(\tilde{N}^{-1/2})$ uniformly in i.

(ii) Using $A^{-1} - B^{-1} = A^{-1}(B-A)B^{-1}$ and $a'a - b'b = a'(a-b) + (a-b)'b$, we have

$$\begin{aligned}
& \left\|(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1} - (\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\right\| \\
= & \left\|(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1}(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i} - \mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\right\| \\
\leq & \left\|(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1}\right\| \left\|(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i} - \mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})\right\| \left\|(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\right\| \\
\leq & \left\|(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1}\right\| \left(\left\|\bar{\mathbf{H}}_{T_i}'\right\| + \|\mathbf{G}_{T_i}\mathbf{P}\|\right) \left\|\bar{\mathbf{H}}_{T_i} - \mathbf{G}_{T_i}\mathbf{P}\right\| \left\|(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\right\| \\
= & O_p(\tilde{N}^{-1/2})
\end{aligned}$$

by (i) and the fact that $\left\|\bar{\mathbf{H}}_{T_i}'\right\| = O_p(1)$, $\|\mathbf{G}_{T_i}\mathbf{P}\| = O_P(1)$, $\left\|(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1}\right\| = O_P(1)$ and $\left\|(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\right\| = O_p(1)$.

(iii) By the definition of $\mathbf{M}_{H_i}$ and $\mathbf{M}_{G_i}$, we have

$$\begin{aligned}
\|\mathbf{M}_{H_i} - \mathbf{M}_{G_i}\| = & \left\|\bar{\mathbf{H}}_{T_i}(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1}\bar{\mathbf{H}}_{T_i}' - \mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{P}'\mathbf{G}_{T_i}'\right\| \\
\leq & \left\|\bar{\mathbf{H}}_{T_i}(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1}(\bar{\mathbf{H}}_{T_i}' - \mathbf{P}'\mathbf{G}_{T_i}')\right\| \\
& + \left\|(\bar{\mathbf{H}}_{T_i} - \mathbf{G}_{T_i}\mathbf{P})(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}]\mathbf{P}'\mathbf{G}_{T_i}'\right\| \\
& + \left\|\bar{\mathbf{H}}_{T_i}[(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1} - (\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}]\mathbf{P}'\mathbf{G}_{T_i}'\right\| \\
\leq & \left(\left\|\bar{\mathbf{H}}_{T_i}\right\| + \left\|\mathbf{P}'\mathbf{G}_{T_i}'\right\|\right) \left\|(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1}\right\| \left\|\bar{\mathbf{H}}_{T_i}' - \mathbf{P}'\mathbf{G}_{T_i}'\right\| \\
& + \left\|\bar{\mathbf{H}}_{T_i}\right\| \left\|\mathbf{P}'\mathbf{G}_{T_i}'\right\| \left\|(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1} - (\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\right\| \\
= & O_p(\tilde{N}^{-1/2}),
\end{aligned}$$

by (i) and (ii). □

Lemma 3. *Under Assumption,*

(i) $N^{-1}\sum_{i=1}^N \mathbf{X}_{iT_i}'\mathbf{M}_{H_i}\mathbf{X}_{iT_i} = N^{-1}\sum_{i=1}^N \mathbf{X}_{iT_i}'\mathbf{M}_{G_i}\mathbf{X}_{iT_i} + O_p(\tilde{N}^{-1/2})$

(ii) $N^{-1}\sum_{i=1}^N \mathbf{X}_{iT_i}'\mathbf{M}_{H_i}\mathbf{F}_{T_i}\lambda_i = O_p[(N\tilde{N})^{-1/2}] + O_p(\tilde{N}^{-1})$;

(iii) $N^{-1}\sum_{i=1}^N \mathbf{X}_{iT_i}'\mathbf{M}_{H_i}\mathbf{v}_{iT_i} = N^{-1}\sum_{i=1}^N \mathbf{X}_{iT_i}'\mathbf{M}_{G_i}\mathbf{v}_{iT_i} + O_p(\tilde{N}^{-1})$

Proof. (i) We have $\left\|\frac{1}{N}\sum_{i=1}^N \mathbf{X}_{iT_i}'(\mathbf{M}_{H_i} - \mathbf{M}_{G_i})\mathbf{X}_{iT_i}\right\| \leq \frac{1}{N}\sum_{i=1}^N \left\|\mathbf{X}_{iT_i}'(\mathbf{M}_{H_i} - \mathbf{M}_{G_i})\mathbf{X}_{iT_i}\right\|$ $\leq \frac{1}{N}\sum_{i=1}^N \|\mathbf{X}_{iT_i}\|^2 \|\mathbf{M}_{H_i} - \mathbf{M}_{G_i}\| \leq \frac{1}{N}\sum_{i=1}^N \|\mathbf{X}_{iT_i}\|^2 \max_i \|\mathbf{M}_{H_i} - \mathbf{M}_{G_i}\| = O_p(\tilde{N}^{-1/2})$ by Lemma 2 (iii).

(ii) By (16), (17), $\mathbf{M}_{H_i}\bar{\mathbf{Z}}_{T_i}=\mathbf{0}$ and $\mathbf{M}_{H_i}(\mathbf{1}_{T_i}\otimes\mu')=0$, we have

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{M}_{H_i}\mathbf{F}_{T_i}\lambda_i &= \frac{1}{N}\sum_{i=1}^{N}\varepsilon'_{iT_i}\mathbf{M}_{H_i}\mathbf{U}^*_{T_i}\mathbf{A}'\lambda_i+\frac{1}{N}\sum_{i=1}^{N}\Pi'_i\mathbf{G}'_{T_i}\mathbf{M}_{H_i}\mathbf{U}^*_{T_i}\mathbf{A}'\lambda_i\\ &\equiv J_{N1}+J_{N2}, \text{ say.}\end{aligned}$$

For J_{N1}, we have

$$\begin{aligned}J_{N1} &= \frac{1}{N}\sum_{i=1}^{N}\varepsilon'_{iT_i}\mathbf{U}^\dagger_{T_i}\mathbf{A}'\lambda_i-\frac{1}{N}\sum_{i=1}^{N}\varepsilon'_{iT_i}\bar{\mathbf{H}}_{T_i}(\bar{\mathbf{H}}'_{T_i}\bar{\mathbf{H}}_{T_i})^{-1}\bar{\mathbf{H}}'_{T_i}\mathbf{U}^\dagger_{T_i}\mathbf{A}'\lambda_i\\ &= J_{N11}-J_{N12}, \text{ say.}\end{aligned}$$

Noting that $\mathbf{U}^\dagger_{T_i}=(\mathbf{0},\bar{\mathbf{U}}_{T_i}+\mathbf{U}^\mu_{T_i}+\mathbf{U}^c_{T_i})$, we have

$$J_{N11}=\frac{1}{N}\sum_{i=1}^{N}\varepsilon'_{iT_i}(\mathbf{0},\bar{\mathbf{U}}_{T_i}+\mathbf{U}^\mu_{T_i}+\mathbf{U}^c_{T_i})\mathbf{A}'\lambda_i=(\mathbf{0},J_{N11a}+J_{N11b}+J_{N11c}), \text{ say,}$$

where $J_{N11a}=\frac{1}{N}\sum_{i=1}^{N}\varepsilon'_{iT_i}\bar{\mathbf{U}}_{T_i}\mathbf{A}'\lambda_i$, $J_{N11b}=\frac{1}{N}\sum_{i=1}^{N}\varepsilon'_{iT_i}\mathbf{U}^\mu_{T_i}\mathbf{A}'\lambda_i$ and $J_{N11c}=\frac{1}{N}\sum_{i=1}^{N}\varepsilon'_{iT_i}\mathbf{U}^c_{T_i}\mathbf{A}'\lambda_i$. For the last two terms, noting that $\mathbf{U}^\mu_{T_i}\mathbf{A}'\lambda_i$ and $\mathbf{U}^c_{T_i}\mathbf{A}'\lambda_i$ are independent with ε_{iT_i}, by the Chebyshev's inequality and the fact that $E\left\|\mathbf{U}^\mu_{T_i}\right\|^2=O(\tilde{N}^{-1})$ and $E\left\|\mathbf{U}^c_{T_i}\right\|^2=O(\tilde{N}^{-1})$, we readily show that $J_{N11b}=O_p((N\tilde{N})^{-1/2})$. Noting that $\bar{\mathbf{U}}_{T_i}=(\bar{\mathbf{V}}_{T_i}+\bar{\varepsilon}_{iT_i}\beta,\bar{\varepsilon}_{iT_i})$ where $\bar{\mathbf{V}}_{T_i}=\left(\bar{\mathbf{v}}_{t_i(1)},\ldots,\bar{\mathbf{v}}_{t_i(T_i)}\right)'$ and $\bar{\varepsilon}_{iT_i}=\left(\bar{\varepsilon}_{t_i(1)},\ldots,\bar{\varepsilon}_{t_i(T_i)}\right)'$, we have

$$\begin{aligned}J_{N11a} &= \frac{1}{N}\sum_{i=1}^{N}(\varepsilon'_{iT_i}\bar{\mathbf{V}}_{T_i}+\varepsilon'_{iT_i}\bar{\varepsilon}_{iT_i}\beta,\varepsilon'_{iT_i}\bar{\varepsilon}_{iT_i})\mathbf{A}'\lambda_i\\ &= \frac{1}{N}\sum_{i=1}^{N}(\varepsilon'_{iT_i}\bar{\mathbf{V}}_{T_i}+\varepsilon'_{iT_i}\bar{\varepsilon}_{iT_i}\beta,\varepsilon'_{iT_i}\bar{\varepsilon}_{iT_i})\mathbf{A}'\lambda_i\\ &= \frac{1}{N}\sum_{i=1}^{N}\varepsilon'_{iT_i}\bar{\mathbf{V}}_{T_i}(\beta'\Gamma+\lambda')\lambda_i+\frac{1}{N}\sum_{i=1}^{N}\varepsilon'_{iT_i}\bar{\varepsilon}_{iT_i}[\beta(\beta'\Gamma+\lambda')+\Gamma']\lambda_i\\ &= J_{N11a}(1)+J_{N11a}(2), \text{ say,}\end{aligned}$$

where we use the definition of $\mathbf{A}$ in the third equation. Noting that ε_{iT_i} are independent of $\bar{\mathbf{V}}_{T_i}$ and λ_i and $E\left\|\bar{\mathbf{V}}_{T_i}\right\|^2=O(\tilde{N}^{-1})$, we can show that $J_{N11a}(1)=O_p((N\tilde{N})^{-1/2})$ by the Chebyshev's inequality again. For $J_{N11a}(2)$, let $\mathbf{L}=[\beta(\beta'\Gamma+$

$\lambda')+\Gamma']$, we have

$$\begin{aligned}
J_{N11a}(2) &= \frac{1}{N}\sum\nolimits_{i=1}^{N}\varepsilon'_{iT_i}\bar{\varepsilon}_{iT_i}\mathbf{L}\lambda_i = \frac{1}{N}\sum\nolimits_{i=1}^{N}\sum\nolimits_{l=1}^{T_i}\varepsilon_{it_i(l)}\bar{\varepsilon}_{it_i(l)}\mathbf{L}\lambda_i \\
&= \frac{1}{NN_{t_i(l)}}\sum\nolimits_{i=1}^{N}\sum\nolimits_{j=1}^{N}\sum\nolimits_{l=1}^{T_i}\frac{1}{N_{t_j(l)}}\varepsilon_{it_i(l)}\varepsilon_{jt_j(l)}1_{jt_j(l)}\mathbf{L}\lambda_i \\
&= \frac{1}{N}\sum\nolimits_{i=1}^{N}\sum\nolimits_{l=1}^{T_i}\frac{1}{N_{t_i(l)}}\varepsilon^2_{it_i(l)}1_{it_i(l)}\mathbf{L}\lambda_i \\
&\quad +\frac{1}{N}\sum\nolimits_{i=1}^{N}\sum\nolimits_{j=1,\neq i}^{N}\sum\nolimits_{l=1}^{T_i}\frac{1}{N_{t_j(l)}}\varepsilon_{it_i(l)}\varepsilon_{jt_j(l)}1_{jt_j(l)}\mathbf{L}\lambda_i \\
&= O_p(\tilde{N}^{-1})+O_p[(N\tilde{N})^{-1/2}] = O_p(\tilde{N}^{-1})
\end{aligned}$$

where the last equation comes from the Chebyshev's inequality. For J_{N12}, we have

$$\begin{aligned}
J_{N12} &= \frac{1}{N}\sum\nolimits_{i=1}^{N}\varepsilon'_{iT_i}\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{U}^{\dagger}_{T_i}\mathbf{A}'\lambda_i \\
&\quad +\frac{1}{N}\sum\nolimits_{i=1}^{N}\varepsilon'_{iT_i}\left(\mathbf{M}_{H_i}-\mathbf{M}_{G_i}\right)\mathbf{U}^{\dagger}_{T_i}\mathbf{A}'\lambda_i \\
&= J_{N12a}+J_{N12b}, \text{ say.}
\end{aligned}$$

We can follow the determination of the probability order of J_{N11} to show that $J_{N12a}=O_p(\tilde{N}^{-1})$, and use Lemma 2 to show that
$\|J_{N12b}\| \le \frac{1}{N}\sum_{i=1}^{N}\|\varepsilon_{iT_i}\|\left(\|\mathbf{M}_{H_i}-\mathbf{M}_{G_i}\|\right)\left\|\mathbf{U}^{\dagger}_{T_i}\right\|\|\mathbf{A}\|\,\|\lambda_i\| = O_p(\tilde{N}^{-1})$.

Now we turn to the term J_{N2}. Noting that $\mathbf{G}_{T_i}\Pi_i = \mathbf{1}_{T_i}\otimes\alpha'_{xi}+\mathbf{F}_{T_i}\Gamma_i$, we have

$$\begin{aligned}
J_{N2} &= \frac{1}{N}\sum\nolimits_{i=1}^{N}\left(\alpha_{xi}\otimes\iota'_{T_i}\right)\mathbf{M}_{H_i}\mathbf{U}^{\dagger}_{T_i}\mathbf{A}'\lambda_i+\frac{1}{N}\sum\nolimits_{i=1}^{N}\Gamma'_i\mathbf{F}'_{T_i}\mathbf{M}_{H_i}\mathbf{U}^{\dagger}_{T_i}\mathbf{A}'\lambda_i \\
&= \frac{1}{N}\sum\nolimits_{i=1}^{N}\Gamma_i\mathbf{A}\mathbf{U}^{\dagger\prime}_{T_i}\mathbf{M}_{H_i}\mathbf{U}^{\dagger}_{T_i}\mathbf{A}'\lambda_i
\end{aligned}$$

where we use $\left(\alpha_{xi}\otimes\mathbf{1}'_{T_i}\right)\mathbf{M}_{H_i}=0$ and $\mathbf{M}_{H_i}\mathbf{F}_{T_i}=\mathbf{M}_{H_i}\mathbf{U}^{\dagger}_{T_i}\mathbf{A}$. Then we have

$$\|J_{N2}\| \le \frac{1}{N}\sum\nolimits_{i=1}^{N}\|\mathbf{A}\|^2\|\Gamma_i\|\left\|\mathbf{U}^{\dagger}_{T_i}\right\|\|\mathbf{M}_{H_i}\|\left\|\mathbf{U}^{\dagger}_{T_i}\right\|\|\lambda_i\| = O_p(\tilde{N}^{-1})$$

because of $\left\|\mathbf{U}^{\dagger}_{T_i}\right\| = O_p(\tilde{N}^{-1/2})$ and $\|\mathbf{M}_{H_i}\| \le \sqrt{\mathrm{tr}\left(\mathbf{M}_{H_i}\mathbf{M}_{H_i}\right)} \le \sqrt{T_i} \le T^{1/2} < \infty$.

(iii) Write

$$
\begin{aligned}
& \frac{1}{N}\sum\nolimits_{i=1}^{N}\mathbf{X}'_{iT_i}\left(\mathbf{M}_{H_i}-\mathbf{M}_{G_i}\right)\mathbf{v}_{iT_i} \\
= & \frac{1}{N}\sum\nolimits_{i=1}^{N}\mathbf{X}'_{iT_i}\left[\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{P}'\mathbf{G}'_{T_i}-\bar{\mathbf{H}}_{T_i}(\bar{\mathbf{H}}'_{T_i}\bar{\mathbf{H}}_{T_i})^{-1}\bar{\mathbf{H}}'_{T_i}\right]\mathbf{v}_{iT_i} \\
= & \frac{-1}{N}\sum\nolimits_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}(\bar{\mathbf{H}}'_{T_i}-\mathbf{P}'\mathbf{G}'_{T_i})\mathbf{v}_{iT_i} \\
& +\frac{1}{N}\sum\nolimits_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{G}_{T_i}\mathbf{P}[(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}-(\bar{\mathbf{H}}'_{T_i}\bar{\mathbf{H}}_{T_i})^{-1}]\bar{\mathbf{H}}'_{T_i}\mathbf{v}_{iT_i} \\
& \frac{-1}{N}\sum\nolimits_{i=1}^{N}\mathbf{X}'_{iT_i}\left(\bar{\mathbf{H}}_{T_i}-\mathbf{G}_{T_i}\mathbf{P}\right)(\bar{\mathbf{H}}'_{T_i}\bar{\mathbf{H}}_{T_i})^{-1}\bar{\mathbf{H}}'_{T_i}\mathbf{v}_{iT_i} \\
= & \Delta_{N1}+\Delta_{N2}+\Delta_{N3}, \text{ say.}
\end{aligned}
$$

For Δ_{N1}, we have

$$
\begin{aligned}
\Delta_{N1} &= \frac{1}{N}\sum\nolimits_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{U}^{\dagger}_{T_i}\mathbf{v}_{iT_i} \\
&= \frac{1}{N}\sum\nolimits_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}(\mathbf{0}'\mathbf{v}_{iT_i},\bar{\mathbf{U}}'_{T_i}\mathbf{v}_{iT_i}+\mathbf{U}^{\mu\prime}_{T_i}\mathbf{v}_{iT_i}+\mathbf{U}^{c\prime}_{T_i}\mathbf{v}_{iT_i})' \\
&= \left(\mathbf{0},\Delta_{N1a}+\Delta_{N1b}+\Delta_{N1c}\right), \text{ say,}
\end{aligned}
$$

where $\Delta_{N1a}=\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}\bar{\mathbf{U}}'_{T_i}\mathbf{v}_{iT_i}$, $\Delta_{N1b}=\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}$ $\times\mathbf{U}^{\mu\prime}_{T_i}\mathbf{v}_{iT_i}$, and $\Delta_{N1c}=\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{U}^{c\prime}_{T_i}\mathbf{v}_{iT_i}$. It is straightforward to show that $\Delta_{N1b}=O_p((N\tilde{N})^{-1/2})$ and $\Delta_{N1c}=O_p((N\tilde{N})^{-1/2})$ by the Chebyshev's inequality and the fact that $E\left\|\mathbf{U}^{c}_{T_i}\right\|^2=O(\tilde{N}^{-1})$ and $E\left\|\mathbf{U}^{\mu}_{T_i}\right\|^2=O((\tilde{N})^{-1})$. For the first term Δ_{N1a}, we can further decompose it as follows:

$$
\begin{aligned}
\Delta_{N1a} &= \frac{1}{N}\sum\nolimits_{i=1}^{N}\Pi'_i\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}\bar{\mathbf{U}}'_{T_i}\mathbf{v}_{iT_i} \\
&\quad +\frac{1}{N}\sum\nolimits_{i=1}^{N}\varepsilon'_{iT_i}\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}\bar{\mathbf{U}}'_{T_i}\mathbf{v}_{iT_i}
\end{aligned}
$$

We can show the both terms in Δ_{N1a} are $O_p((N\tilde{N})^{-1/2})$ as the proof of $J_{N11a}=O_p((N\tilde{N})^{-1/2})$.

For Δ_{N2}, we have

$$
\begin{aligned}
\Delta_{N2} &= \frac{1}{N}\sum\nolimits_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{G}_{T_i}\mathbf{P}[(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}-(\bar{\mathbf{H}}'_{T_i}\bar{\mathbf{H}}_{T_i})^{-1}]\mathbf{G}_{T_i}\mathbf{P}\mathbf{v}_{iT_i} \\
&\quad +\frac{1}{N}\sum\nolimits_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{G}_{T_i}\mathbf{P}[(\mathbf{P}'\mathbf{G}'_{T_i}\mathbf{G}_{T_i}\mathbf{P})^{-1}-(\bar{\mathbf{H}}'_{T_i}\bar{\mathbf{H}}_{T_i})^{-1}]\left(\bar{\mathbf{H}}_{T_i}-\mathbf{G}_{T_i}\mathbf{P}\right)\mathbf{v}_{iT_i} \\
&= \Delta_{N2a}+\Delta_{N2b}, \text{ say.}
\end{aligned}
$$

We can see that $\|\Delta_{N2b}\| \leq \frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{X}_{iT_i}'\right\|\|\mathbf{G}_{T_i}\mathbf{P}\|\|\mathbf{v}_{iT_i}\|\left\|(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}-(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1}\right\|$ $\left\|\left(\bar{\mathbf{H}}_{T_i}-\mathbf{G}_{T_i}\mathbf{P}\right)\right\| = O_p(\tilde{N}^{-1})$ by Lemma 2. We rewrite Δ_{N2a} as follows

$$
\begin{aligned}
\Delta_{N2a} &= \frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}[\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i}-\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P}](\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1}\mathbf{G}_{T_i}\mathbf{P}\mathbf{v}_{iT_i} \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}[\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i}-\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P}](\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{G}_{T_i}\mathbf{P}\mathbf{v}_{iT_i} \\
&\quad +\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}[\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i}-\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P}] \\
&\quad \times\left[(\bar{\mathbf{H}}_{T_i}'\bar{\mathbf{H}}_{T_i})^{-1}-(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\right]\mathbf{G}_{T_i}\mathbf{P}\mathbf{v}_{iT_i} \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{U}_{T_i}^{\dagger}(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{G}_{T_i}\mathbf{P}\mathbf{v}_{iT_i} \\
&\quad +\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{U}_{T_i}^{\dagger\prime}\mathbf{U}_{T_i}^{\dagger}(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{G}_{T_i}\mathbf{P}\mathbf{v}_{iT_i} \\
&\quad +\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{U}_{T_i}^{\dagger\prime}\mathbf{G}_{T_i}\mathbf{P}(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{G}_{T_i}\mathbf{P}\mathbf{v}_{iT_i}+O_p(\tilde{N}^{-1}) \\
&= O_p(\tilde{N}^{-1})
\end{aligned}
$$

where we show that the first and third term are $O_p(\tilde{N}^{-1})$ by following the proof of $\Delta_{N1} = O_p(\tilde{N}^{-1})$, and bound the second term by $O_p(\tilde{N}^{-1})$ using the fact that $\left\|\mathbf{U}_{T_i}^{\dagger}\right\|^2 = O_p(\tilde{N}^{-1})$.

Lastly, by Lemma 2, we have $\Delta_{N3} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{U}_{T_i}^{\dagger}(\mathbf{P}'\mathbf{G}_{T_i}'\mathbf{G}_{T_i}\mathbf{P})^{-1}\mathbf{G}_{T_i}\mathbf{P}\mathbf{v}_{iT_i}$ $+O_p(\tilde{N}^{-1})$. It is easy to prove that the first term is $O_p(\tilde{N}^{-1})$ using the similar arguments in the proof of Δ_{N1}. It follows that $\Delta_{N3} = O_p(\tilde{N}^{-1})$. □

Proof of Theorem 1. (i) Note that model (1) can be rewritten in vector form as

$$
\mathbf{y}_{iT_i} = \alpha_{yi}\mathbf{1}_{T_i}+\mathbf{X}_{iT_i}\beta+\mathbf{F}_{T_i}\lambda_i+\mathbf{v}_{iT_i}, \tag{19}
$$

then the CCE-UB estimator (12) can be written as

$$
\begin{aligned}
\hat{\beta}_{CCE}^{UB} &= \left(\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{M}_{H_i}\mathbf{X}_{iT_i}\right)^{-1}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{M}_{H_i}\left(\alpha_{yi}\mathbf{1}_{T_i}+\mathbf{X}_i\beta+\mathbf{F}_{T_i}\lambda_i+\mathbf{v}_{iT_i}\right) \\
&= \beta+\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{M}_{H_i}\mathbf{X}_{iT_i}\right)^{-1}\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{M}_{H_i}\left(\mathbf{F}_{T_i}\lambda_i+\mathbf{v}_{iT_i}\right) \\
&= \beta+\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{M}_{G_i}\mathbf{X}_{iT_i}\right)^{-1}\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{iT_i}'\mathbf{M}_{G_i}\mathbf{v}_{iT_i}+O_p(\tilde{N}^{-1}) \qquad (20)
\end{aligned}
$$

where we use the fact that $\mathbf{M}_{H_i}\mathbf{1}_{T_i} = 0$ at the second equation and Lemma (3) in the last equation. Conditional on $\mathscr{F} \equiv \sigma\{\mathbf{f}_1,...,\mathbf{f}_T\}$,[3] the sigma field generated

[3] Alternatively, we can treat F as fixed.

by $\mathbf{f}_1,...,\mathbf{f}_T$, we can show that $N^{-1}\sum_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{M}_{G_i}\mathbf{X}_{iT_i}\xrightarrow{p}\mathbf{D}$ and $N^{-1}\sum_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{M}_{G_i}\mathbf{v}_{iT_i}\xrightarrow{p}$ 0 by the Law of Large Number (LLN) for inid sequence under Assumptions 1-6. Then $\hat{\beta}_{CCE}^{UB}\xrightarrow{p}\beta$ as $N\to\infty$ by Continuous Mapping Theorem.

(ii) Under Assumptions 1-7, we can show that $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{M}_{G_i}\mathbf{v}_{iT_i}\xrightarrow{d}N(\mathbf{0},\Sigma)$ by central limiting theorem (CLT) for inid observations conditional on $\mathscr{F}$. Noting that $N^{-1}\sum_{i=1}^{N}\mathbf{X}'_{iT_i}\mathbf{M}_{G_i}\mathbf{X}_{iT_i}\xrightarrow{p}\mathbf{D}$, we have

$$\sqrt{N}(\hat{\beta}_{CCE}^{UB}-\beta)\xrightarrow{d}N(\mathbf{0},\mathbf{D}^{-1}\mathbf{\Sigma}\mathbf{D}^{-1}),$$

by the Slutsky lemma. □

References

Ahn, S. C., Y. H. Lee, and P. Schmidt (2001). Panel Data Models with Multiple Time-varying Individual Effects, *Journal of Econometrics* 174, 1-14.

Bai, J. (2009). Panel Data Models with Interactive Effects, *Econometrica* 77, 1229-1279.

Bai, J. (2013). Likelihood Approach to Dynamic Panel Models with Interactive Effects, *Working Paper*. Columbia University.

Bai, J. and K. Li (2012). Statistical Analysis of Factor Models of High Dimension, *Annals of Statistics* 40, 436-465.

Bai, J. and K. Li (2014). Theory and Methods of Panel Data Models with Interactive Effects, *Annals of Statistics* 42, 142-170.

Bai, J., Y. Liao, and J. Yang (2015). Unbalanced Panel Data Models with Interactive Effects, In Baltagi, B. H. (Ed.) *The Oxford Handbook of Panel Data*, Oxford University Press, Chapter 5, 149-170.

Baltagi, B. H. and Y. J. Chang (1994). Incomplete Panels: A Comparative Study of Alternative Estimators for the Unbalanced One-way Error Component Regression Model, *Journal of Econometrics* 62, 67-89.

Baltagi, B. H. and S. H. Song, (2006), Unbalanced Panel Data: A Survey, *Statistical Papers* 47, 493-523.

Hayakawa, K. (2012). GMM Estimation of Short Dynamic Panel Data Model with Interactive Fixed Effects, *Journal of the Japan Statistical Society* 42, 109-123.

Hayakawa, K., H. M. Pesaran, and L. V. Smith (2014). Transformed Maximum Likelihood Estimation of Short Dynamic Panel Data Models with Interactive Effectects, *Working Paper,* USC.

Hsiao, C. (2014). *Analysis of Panel Data,* 3nd edition, Cambridge University Press.

Juodis, A. and V. Sarafidis, V. (2014). Fixed *T* Dynamic Panel Data Estimators with Multi-Factor Errors, *MPRA Paper* 57659, University Library of Munich, Germany.

Lee, L. F. and J. Yu (2010). Estimation of Spatial Panels, in *Foundations and Trends in Econometrics,* Vol. 4, 1-164.

Pesaran, M. H. (2006). Estimation and Inference in Large Heterogenous Panels with Multifactor Error Structure, *Econometrica* 74, 967-1012.

Sarafidis, V. and T. Wansbeek (2012). Cross-Sectional Dependence in Panel Data Analysis, *Econometric Reviews* 31, 483-531.

Wansbeek, T. and A. Kapteyn (1989). Estimation of the Error-Components Model with Incomplete Panels, *Journal of Econometrics* 41, 341-361.

Westerlund, J.. and J. Urbain (2015). Cross-Sectional Averages versus Principal Components, *Journal of Econometrics* 185, 372-377.

White, H. (2001). *Asymptotic Theory for Econometricians, Revised Edtion*, Academic Press.