Avoiding the Rank Condition in the CCE Estimator by the Orthogonal Projection

Wei Wang* Cheng Fang† Yanyang Wei‡

January 12, 2024

Abstract

The common correlated effects (CCE) is one of the most popular estimating approaches for the panel data model with the interactive fixed effects. However, one of the restrictions of CCE is that the rank condition affects the asymptotic properties of the pooled CCE estimator. In this paper, we proposed the two-steps estimation, combining the orthogonal projection and CCE, to avoiding the rank condition. The asymptotic properties of the proposed estimators are proved and show in the experiments.

Keywords: Panel Data, Interactive Fixed Effects, Common Correlated Effects, Orthogonal Projection.

JEL Classification: C23, C33.

1 Introduction

Panel data models with the interactive fixed effects have become popular in the econometric model (Sarafidis and Wansbeek, 2012; Hsiao, 2018) and economic applications, such as the international trade (Serlenga and Shin, 2007), finance (Gagliardini et al., 2020), environments (Khan et al., 2020). To estimate the slopes, Pesaran (2006) proposed the CCE estimators. This estimation is easy to implement, which has been extended into the endogenous model (Harding and Lamarche, 2011), dynamic model (Chudik and Pesaran, 2015) and others. The CCE used the cross-sectional average of observed data to estimating the unobserved common factors, leading to the discusses

*School of Statistics and Mathematics, Shandong University of Finance and Economics, No.7366 East Erhuan Road, 250014, Jinan, Shandong, China. Email Address: wangwei_0115@outlook.com.

†Corresponding Author. School of Data Sciences, Zhejiang University of Finance & Economics, Hangzhou 310018, PR China. E-mail address: fangcheng12@zufe.edu.cn.

‡The National Financial Regulatory Administration, China. E-mail address: 4912618@qq.com.

of the rank condition in the pooled CCE estimator (Karabiyik et al., 2017; Karabiyik et al., 2019; Juodis et al., 2021). Karabiyik et al. (2017) show that, while the number of factors is strictly less than the number of observed data, the pooled CCE estimator is biased, also show in Westerlund and Urbain (2015).

In this paper, we aim to apply the orthogonal projection of the CCE to avoid the rank condition, inspired by the transformed estimation of Hsiao et al. (2022). Specifically, we find the null space of the observable data and then use the pooled regression to estimate the interested parameters. Under the CCE, the two-steps estimation by the orthogonal projection do not need iteration, as the transformation estimation of Hsiao et al. (2022). The rest of this paper is organized as follows. Section 2 introduce the model and the proposed estimation are presented. Section 3 conducts the Monte Carlo experiments and Section 4 conclude this paper.

2 Model

For each individual $i \in \{1, \cdots, N\}$, $Y_i = (y_{i1}, \cdots, y_{iT})'$ denotes individual i's dependent variable with the time span $t = \{1, \cdots, T\}$, and $X_i = (x_{i1}, \cdots, x_{iT})'$ is the individual i's covariates with $x_{it} = (x_{it,1}, \cdots, x_{it,k})'$. The panel data model with the interactive fixed effect or factor structure is

$$Y_i = X_i \beta + F\lambda_i + \varepsilon_i, \tag{1}$$

where β is $k \times 1$ vector of slopes, $F = (f_1, \cdots, f_T)$ is an $T \times m$ dimensional unobservable common factors with factor loadings λ_i, and $\varepsilon_i = (\varepsilon_{i1}, \cdots, \varepsilon_{iT})'$ is the individual i's idiosyncratic error.

In the setup of Pesaran (2006)'s CCE, the regressors are also driven by the common factors,

$$X_i = F\Lambda_i + V_i, \tag{2}$$

where Λ_i is $m \times k$ dimensional factor loadings of common factors F, $V_i = (v_{it}, \cdots, v_{iT})'$ is the idiosyncratic error of the equation (2).

Let $Z_i = (Y_i, X_i)$ denotes all the observable variables, and combining the equations (1) and (2) gives

$$Z_i = FC_i + U_i, \tag{3}$$

where $C_i = (\Lambda_i'\beta + \lambda_i, \Lambda_i')$ and $U_i = (V_i\beta + \varepsilon_i, V_i)$. Equation (3) is always called the common correlated effects, showing that the observable data are both correlated by the common factor F. The following Assumption are general in the CCE's setup (Juodis et al., 2021).

Assumption 1 *(i) The disturbances $e_{it} = (\varepsilon_{it}, v'_{it})'$ is a stationary process that is independent across i with $\Gamma_{e,i}(h) = E(e_{it}e'_{i,t-h})$ absolutely summable, $E(e_{it}) = 0_{(k+1)\times 1}$, $E||e_{it}|| < \infty$. (ii) For each series i, v_{it} is independent of $\varphi_{t'}$ for all t and t'; (iii) v_{it} are linear stationary processes with zero mean and absolute summable autocovariances, $v_{it} = \sum_{l=0}^{\infty} \Xi_{il} v_{i,t-l}$, where $(\zeta_{it}, v'_{it}\ )'$ are $(p+1)\times 1$ vectors of IID random variables with variance-covariance matrix I_{p+1} and has a finite fourth-order moments, and $Var(v_{it}) = \sum_{l=0}^{\infty} \Xi_{il}\Xi'_{il} = \Sigma_{v,i}$, and $0 < \|\Sigma_{v,i}\| < \infty$.*

Assumption 2 *(i) Factor loadings C_i is independent across i, with $E||C_i||^2 < \infty$, and $E(C_i) = C$; (ii) C_i and u_j are independent for all i and j; (iii) $E||\lambda_i||^4 < \infty$, and $plim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \lambda_i \lambda'_i = \Sigma_\lambda$, with positive definite matrix Σ_λ.*

Assumption 3 *(i) f_t is a stationary process with $E(f_t) = 0_{m\times 1}$,$\Gamma_f(h) = \lim_{T\to\infty} \frac{1}{T} f_t f'_{t-h}$, and $\Sigma_f = \Gamma_f(0)$ is a positive definite matrix, and $\Gamma_f(h)$ absolutely summable; (ii) e_{it} and f_s are mutually independent for all i, t and s.*

According to equation (3), the common factors are estimated by the cross-sectional average of observable variables, $\hat{F} = \bar{Z} = \frac{1}{N}\sum_{i=1}^{N} Z_i = FC + \bar{U}$, as $N \to \infty$. Thus, the pooled CCE estimator (CCEP) of Pesaran (2006) is defined as,

$$\widehat{\beta}_{CCEP} = \left(\sum_{i=1}^{N} X'_i M_{\bar{Z}} X_i \right)^{-1} \left(\sum_{i=1}^{N} X'_i M_{\bar{Z}} Y_i \right)^{-1}. \tag{4}$$

However, the rank condition affects the asymptotic properties of $\widehat{\beta}_{CCEP}$, as show in Karabiyik et al. (2017), Juodis et al. (2021) and others. They distinguished different cases, including $m < k+1$, $m = k+1$, and $m > k+1$, and show that the CCEP is consistent but not asymptotic normal for some case under mild conditions. It's natural to ask how to avoid the CCE's rank condition under the common correlated equation (3). In this paper, we aim to provide an alternative method, based on the common correlated effects of equation (3), to avoid the regular rank condition and then obtain the consistent estimation.

2.1 Two-step Estimation

According to equation (3), $\bar{Z} = FC + \bar{U}$. Inspired by Hsiao et al. (2022), we proposed the following alternative method to handle the common correlated effects $\bar{Z} = FC + \bar{U}$.

Step 1: we find the null space of $\bar{Z}$, that is $\{\underset{T\times 1}{w} \in \mathbb{R}^T : w'\bar{Z} = 0_{1\times(k+1)}\}$, by minimizing $w'(\bar{Z}\bar{Z}'/T)w$,under the constraint that $w'w = 1$. According to the first order conditions, $\hat{w}$ is the eigenvalue of the matrix $\bar{Z}\bar{Z}'/T$, corresponding the zero eigenvalues. Since $\bar{U} \xrightarrow{p} 0_{T\times(k+1)}$ under Assumption 1, as $N \to \infty$, thus, $\hat{w}'\bar{Z} = 0_{1\times(k+1)}$ implies that $\hat{w}'FC \xrightarrow{p} 0_{1\times(k+1)}$, as $N \to \infty$.

Remark: In the step 1, we aim to find the null space of $\bar{Z}$ based on the common correlated effects of equation (3), and the null space of $\bar{Z}$ asymptotically equals the null space of the common factors F. In this step, we do not directly estimate the common factors by the observable $\bar{Z}$, and then estimate the null space of F by $M_{\bar{Z}}$, as the traditional CCE approach of equation (4). Thus, it do not need the rank condition in the estimation procedures. In addition, the function of $\hat{w}$ is similar to $M_{\bar{Z}}$, eliminating the interactive fixed effects.

Step 2: Pre-multipling the models (1) by $\hat{w}'$ gives

$$\hat{w}'Y_i \quad = \quad \hat{w}'X_i\beta + \hat{w}'F\lambda_i + \hat{w}'\varepsilon_{it}. \tag{5}$$

Last, the transformed pooled CCE estimator $\widehat{\beta}_{TP}$ is defined as

$$\widehat{\beta}_{TP} = \left(\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'X_i\right)^{-1} \left(\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'Y_i\right). \tag{6}$$

Similar to the transformed estimation in Hsiao et al. (2022)that the rank of $\bar{Z}\bar{Z}'/T$ equals $k+1$. , there exists multiple estimated $\hat{w}$ that line in the null space of $\bar{Z}$. For $j \in \{1, \cdots, J\}$,the j-th estimator is

$$\widehat{\beta}_{TP,j} = \left(\sum_{i=1}^{N} X_i'\hat{w}_j\hat{w}_j'X_i\right)^{-1} \left(\sum_{i=1}^{N} X_i'\hat{w}_j\hat{w}_j'Y_i\right), \tag{7}$$

and the average transformed pooled CCE estimator is more efficient than $\widehat{\beta}_{TP}$, and is defined as

$$\widehat{\beta}_{ATP} = \frac{1}{J}\sum_{j=1}^{J} \widehat{\beta}_{TP,j}.$$

Let S_1 denotes vector space spanned by $\frac{1}{NT}\sum_{i=1}^{N} X_i' X_i$, S_2 denote the vector space spanned by F, S_3 denote the vector space orthogonal to the space spanned by F in $\mathbb{R}^T$. To show the asymptotic properties of the estimators, we give another Assumptions,

Assumption 4 *For any nonzero vector* b, *such that* $b'b = O(1)$, $b'\left(\frac{1}{NT}\sum_{i=1}^{N} X_i' X_i\right) b \xrightarrow{p} \Phi_X < \infty$, *as* $(N,T) \to \infty$.

Assumption 5 $S_1 \cap S_3 = \emptyset$.

Assumption 4 is the identification condition and Assumption 5 make sure w is an nonzero vectors. According to equation (7) and Assumptions 1-5, we obain

$$\sqrt{NT}(\widehat{\beta}_{TP} - \beta) = (\frac{1}{NT}\sum_{i=1}^{N} X_i' \hat{w}\hat{w}' X_i)^{-1}(\frac{1}{\sqrt{NT}}\sum_{i=1}^{N} X_i' \hat{w}\hat{w}' \varepsilon_i) + o_p(1),$$

Thus, we obtain the following Theorem 1.

Theorem 1 *Under Assumptions 1-5, then as* $(N,T) \to \infty$, *(i)*

$$\sqrt{NT}(\widehat{\beta}_{TP} - \beta) \xrightarrow{d} \mathcal{N}(0, \Omega_{TP}).$$

where $\Omega_{TP} = \Phi_X^{-1}\Psi\Phi_X^{-1}$, *with* $\Phi_X = \lim_{(N,T)\to\infty} \frac{1}{NT}\sum_{i=1}^{N} X_i' w w' X_i$, *and*

$$\Psi = \lim_{(N,T)\to\infty} \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N} X_i' w w' E(\varepsilon_i \varepsilon_j') w w' X_j';$$

(ii)

$$\sqrt{NT}(\widehat{\beta}_{ATP} - \beta) \xrightarrow{d} \mathcal{N}(0, \Omega_{ATP}).$$

where $\Omega_{ATP} = \frac{1}{J}\sum_{j=1}^{J} \Omega_{TP,j}$ *with* $\Omega_{TP,j}$ *denotes the asymptotic covariance matrix of* $\widehat{\beta}_{TP,j}$.

In the transformation estimation of Hsiao et al. (2020), the relative growth rate of the samples is restrictive such that $\frac{N}{T} \to a < \infty$ and we do not need it in the CCE's setup. Last, we also allow for the case of T is fixed and $N \to \infty$.

For the case of random individual slope $\beta_i = \beta + \eta_i$, under the following Assumption 6,

Assumption 6 *the random elements $\eta_i \sim iid(0, \Sigma_\eta)$ with $||\Sigma_\eta|| < \infty$, and η_i is independent of ε_i and v_i.*

Plugging $\beta_i = \beta + \eta_i$ into the asymptotic expression of $\sqrt{NT}(\widehat{\beta}_{TP} - \beta)$, under Assumption 1-6, we directly obtain,

$$\sqrt{NT}\left(\widehat{\beta}_{TP} - \beta\right) = \left(\frac{1}{NT}\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'X_i\right)^{-1}\left(\frac{1}{\sqrt{NT}}\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'X_i\eta_i\right) + \left(\frac{1}{NT}\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'X_i\right)^{-1}\left(\frac{1}{\sqrt{NT}}\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'\varepsilon_i\right) + o_p(1).$$

Thus, we obtain the following Proposition 1:

Proposition 1 *Under Assumptions 1-6, then, as $(N,T) \to \infty$, (i)*

$$\sqrt{NT}\left(\widehat{\beta}_{TP} - \beta\right) \xrightarrow{d} N(0, \Omega_{TP}^{+}),$$

where $\Omega_{TP}^{+} = \Omega_{TP} + \Phi_X^{-1}\Psi^{+}\Phi_X^{-1}$, with $\Psi^{+} = \lim_{(N,T)\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N} X_i'\hat{w}\hat{w}'X_iE(\eta_i\eta_j')X_jww'X_j'$;
(ii)

$$\sqrt{NT}\left(\widehat{\beta}_{ATP} - \beta\right) \xrightarrow{d} \mathcal{N}(0, \Omega_{ATP}^{+}),$$

where $\Omega_{ATP}^{+} = \frac{1}{J}\sum_{j=1}^{J}\Omega_{TP}^{+}$ with Ω_{TP}^{+} denotes the asymptotic covariance matrix of $\widehat{\beta}_{TP,j}$.

3 Experiments

In this paper, we consider four data generating processes (DGP), as following.

DGP 1(Rank condition: $k + 1 = m$): the observed dependent variable y_{it}, and one-dimensional regressor are generated by

$$y_{it} = x_{it}\beta + \lambda_i'f_t + \varepsilon_{it},$$
$$x_{it} = \Lambda_i'f_t + v_{it}.$$

where the true slopes $\beta = 1$, the factor loadings $\lambda_i = (\lambda_{i1}, \lambda_{i2})'$ are set as $\lambda_{i1} \sim iid\chi^2(1)$ and $\lambda_{i2} \overset{iid}{\sim} N(0.2, 0.2)$. $\Lambda_i = (\Lambda_{i1}, \Lambda_{i2})$ with $\Lambda_{i1}, \Lambda_{i2} \sim iidN(0.5, 1)$. For $t = \{-49, ...0, ..., T\}$, let $v_{ft} \sim iid\chi^2(3) - 3$ with $\rho_f = 0.5$, $f_{i,-50} = 0$. the common factors $f_t = (f_{t1}, f_{t2})'$ follows AR(1) process,

$$f_t = \rho_f f_{t-1} + v_{ft}.$$

For the errors, ε_{it} follows stationary AR(1) process with heteroskedasticity across each i. Let $\rho_{i\varepsilon} \sim iidU[0.5, 0.9], \sigma_i \sim iidU[0.8, 1.8], \zeta_{it} \sim iid\chi^2(3) - 3$ for both i and $t = \{-49, ...0, ..., T\}$,

$$\varepsilon_{it} = \rho_{i\varepsilon}\varepsilon_{i,t-1} + \sigma_i(1 - \rho_{i\varepsilon}^2)^{0.5}\zeta_{it},$$

where $\varepsilon_{i,-50} = 0$. $v_{it} \sim iidN(0, 1)$.

DGP 2 (Rank condition: $k + 1 < m$): We add one another common factor $f_{t3} \sim iidN(0, 1)$ and factor loadings $\lambda_{i3} \sim iidN(0, 1)$ and $\Lambda_{i3} \sim iidN(0.5, 1)$ into the case of DGP 1.

DGP 3 (Rank deficiency): this DGP is similar to that of DGP 1, exception that $\Lambda_{i1}, \Lambda_{i2} \sim iidN(0, 1)$.

DGP 4 (Heterogeneous slope): For the DGP 1 withe heterogeneous slope, the model becomes

$$y_{it} = x_{it}\beta_i + \lambda_i' f_t + \varepsilon_{it},$$

where $\beta_i = 1 + \eta_i$, with $\eta_i \sim iidN(0, 1)$.

We replicate each experiments 1000 replications and report the root mean squared errors ($RMSE \times 100$) and mean bias ($Bias \times 100$) of $\hat{\beta}$, for the pooled CCE estimator $\hat{\beta}_{CCEP}$ and our proposed estimators $\hat{\beta}_{TP}$ and $\hat{\beta}_{ATP}$. The samples vary along the individual dimension $N = \{20, 50, 100, 200\}$ and time dimension $T = \{20, 50, 100, 200\}$.

3.1 Results

Table 1 shows that the $RMSE$ and $Bias$ of all the three estimators decrease as the samples become larger when the rank condition is satisfied. Overall, the $RMSE$ of $\hat{\beta}_{ATP}$ is smaller than that of $\hat{\beta}_{CCEP}$,and $\hat{\beta}_{TP}$ is less efficient. For example, when $N = 200$ and $T = 200$, the $RMSE$ of $\hat{\beta}_{ATP}$ equals 0.025 and that of $\hat{\beta}_{CCEP}$ equals 0.031. We also find that the $RMSE$ of $\hat{\beta}_{ATP}$ decreases faster than that of $\hat{\beta}_{CCEP}$ as T increase. From the viewpoint of bias, $\hat{\beta}_{CCEP}$ has smaller bias than $\hat{\beta}_{TP}$ and Table 2 considers the case that rank condition is not satisfied such that $k + 1 < m$, and the $RMSE$ of $\hat{\beta}_{CCEP}$ increase to 0.064 when $N = 200$ and $T = 200$, Table 3 considers the case of rank deficiency with mean zero of Λ_i. We could also obtain the similar results as Table 1 and Table 2 and the $RMSE$ of $\hat{\beta}_{CCEP}$ and $\hat{\beta}_{ATP}$ increase a little. For the case of random individual coefficient, Table 4 show that the $RMSE$ of all the three estimators increase a bitter than homogeneous case and the $RMSE$ of $\hat{\beta}_{ATP}$ is smaller that of $\hat{\beta}_{CCEP}$.

Last, we also consider the Example 3.1 of Juodis et al. (2021) that if λ_i is correlated with Λ_i, the simulated distribution of $\sqrt{N}(\hat{\beta}_{CCEP} - \beta)$ deviated from the normal distribution, show in Figure 1 of Juodis et al. (2021). We also plot the simulated distribution of $\sqrt{N}(\hat{\beta}_{CCEP} - \beta)$ in Figure 1 and show that it does not deviate from the normal distribution.

Table 1: The $RMSE$ and $Bias$ of β in DGP 1

		$RMSE \times 100$				$Bias \times 100$			
Methods	T\N	20	50	100	200	20	50	100	200
$\hat{\beta}_{CCEP}$	20	17.67	10.82	7.12	4.70	-1.25	-0.36	0.23	-0.13
	50	16.94	9.57	6.04	3.59	1.13	-0.33	-0.43	0.01
	100	15.27	8.97	5.61	3.18	0.01	-0.07	-0.41	0.02
	200	15.22	9.01	5.07	3.05	0.03	-0.29	-0.02	0.30
$\hat{\beta}_{TP}$	20	47.14	30.70	23.14	16.72	-0.30	-0.65	0.06	0.001
	50	50.06	31.91	24.52	16.83	1.05	-0.31	-0.20	0.34
	100	50.57	33.11	23.31	17.55	-0.34	-0.47	-1.36	0.36
	200	52.59	32.35	23.54	17.88	0.22	0.59	0.85	0.50
$\hat{\beta}_{ATP}$	20	17.60	11.64	9.15	7.41	5.08	5.54	6.14	5.66
	50	14.44	8.18	5.47	3.90	2.98	1.75	1.73	2.10
	100	12.17	7.09	4.48	2.87	1.09	0.99	0.72	1.06
	200	11.59	6.77	3.92	2.54	0.51	0.19	0.54	0.77

Table 2: The $RMSE$ and $Bias$ of β in DGP 2

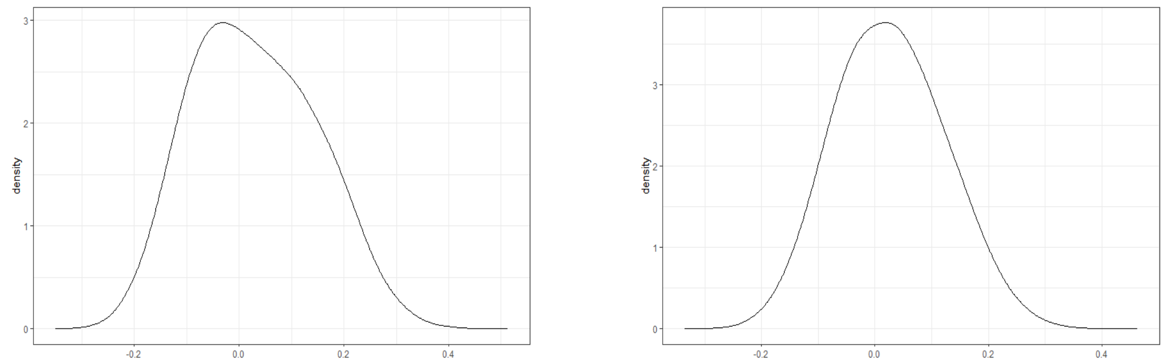
		$RMSE \times 100$				$Bias \times 100$			
Methods	T\N	20	50	100	200	20	50	100	200
$\hat{\beta}_{CCEP}$	20	19.32	12.90	8.98	6.36	0.73	-0.20	-0.20	-0.18
	50	19.09	12.24	9.17	6.44	0.22	0.78	0.22	-0.14
	100	19.37	12.47	8.74	6.17	0.07	0.45	-0.08	-0.01
	200	18.58	12.18	8.61	6.37	-0.42	-0.33	-0.09	0.18
$\hat{\beta}_{TP}$	20	37.36	24.47	16.98	12.71	2.55	-1.16	-0.55	0.57
	50	37.86	23.65	16.82	12.77	0.49	1.16	0.43	-0.53
	100	38.50	24.55	16.99	12.56	-0.65	0.15	0.85	-0.35
	200	37.10	23.82	17.53	12.62	-2.01	-0.73	0.14	0.50
$\hat{\beta}_{ATP}$	20	18.63	12.49	9.62	8.03	6.63	5.52	5.65	5.84
	50	15.67	10.02	7.49	5.45	2.18	2.93	2.13	2.04
	100	15.30	9.67	6.77	4.82	1.11	1.46	0.97	1.08
	200	14.39	9.17	6.46	4.83	0.12	0.29	0.43	0.64

Table 3: The $RMSE$ and $Bias$ of β in DGP 3

		$RMSE \times 100$				$Bias \times 100$			
Methods	T\N	20	50	100	200	20	50	100	200
$\hat{\beta}_{CCEP}$	20	17.33	11.17	7.91	5.45	-0.54	-0.04	0.11	0.11
	50	16.87	9.80	6.98	5.07	0.10	0.04	-0.26	-0.28
	100	15.67	9.13	6.85	5.22	0.09	-0.35	0.19	-0.05
	200	15.38	9.28	7.05	4.69	-0.52	-0.02	0.26	-0.31
$\hat{\beta}_{TP}$	20	47.28	27.75	21.68	15.14	-0.81	-0.28	0.29	-0.73
	50	51.07	29.25	21.44	15.43	-0.64	0.89	-0.02	-0.45
	100	49.36	29.38	22.12	16.15	-3.30	-0.37	-0.22	0.29
	200	49.63	31.97	21.59	16.11	-1.50	0.90	0.41	0.43
$\hat{\beta}_{ATP}$	20	17.37	11.98	9.42	7.78	5.65	5.97	6.07	5.89
	50	13.90	8.41	5.94	4.48	2.04	2.07	1.86	1.97
	100	12.36	6.96	5.34	4.06	1.24	0.85	1.11	0.98
	200	11.67	6.90	5.20	3.51	0.15	0.46	0.73	0.27

Table 4: The $RMSE$ and $Bias$ of β in DGP 4

		$RMSE \times 100$				$Bias \times 100$			
Methods	T\N	20	50	100	200	20	50	100	200
$\hat{\beta}_{CCEP}$	20	33.71	21.04	14.40	9.57	-1.07	-0.16	0.01	0.23
	50	34.35	20.74	13.97	8.66	0.83	0.01	0.52	-0.14
	100	31.84	19.51	13.47	8.52	-2.69	-0.27	0.26	0.01
	200	33.23	19.29	12.94	8.62	0.85	-0.58	0.21	-0.18
$\hat{\beta}_{TP}$	20	61.50	38.31	29.27	20.37	0.83	-1.41	1.01	-0.63
	50	60.97	39.38	29.54	21.26	-0.81	-2.04	-1.05	0.12
	100	60.74	39.47	29.33	20.84	-0.69	-1.34	1.94	-0.06
	200	61.48	40.75	28.14	20.64	0.57	-0.22	-0.19	-0.65
$\hat{\beta}_{ATP}$	20	32.86	21.03	15.41	11.37	5.05	5.77	6.10	6.03
	50	30.61	19.06	13.15	8.41	2.81	2.04	2.48	1.95
	100	27.81	17.37	12.43	8.11	-1.61	0.82	1.26	0.97
	200	28.80	16.95	11.53	8.08	1.44	0.13	0.69	0.32



(a) $\sqrt{N}(\hat{\beta}_{CCEP} - \beta)$ (b) $\sqrt{N}(\hat{\beta}_{ATP} - \beta)$

Figure 1: Simulated finite sample distribution

4 Conclusion

In this paper, we proposed the transformed pooled CCE estimator of the panel data model with the common factors, which avoiding the rank condition in the traditional CCE estimators. The proposed method is also easy to implement and computationally efficient. Asymptotic properties of the proposed estimators are developed. Numerical results show the well performance of the estimators in finite samples.

Acknowledgements

Wei Wang was supported by the Shandong Provincial Natural Science Foundation, China (ZR2022QA109, ZR2020MG035), the National Social Science Foundation of China (21BTJ046) and the Major Project of Humanities and Social Sciences Research of Shandong University (22RWZD16) . Cheng Fang was supported by the Statistical Science Research Project of Zhejiang (23TJQN12).

References

[1] Chudik, A., & Pesaran, M. H. (2015). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics*, 188(2), 393-420.

[2] Gagliardini, P., Ossola, E., & Scaillet, O. (2020). Estimation of large dimensional conditional factor models in finance. *Handbook of Econometrics*, 7, 219-282.

[3] Harding, M., & Lamarche, C. (2011). Least squares estimation of a panel data model with multi-factor error structure and endogenous covariates. *Economics Letters*, 111(3), 197-199.

[4] Hsiao, C. (2018). Panel models with interactive effects. *Journal of Econometrics*, 206(2), 645-673.

[5] Hsiao, C., Shi, Z., & Zhou, Q. (2022). Transformed Estimation for Panel Interactive Effects Models. *Journal of Business & Economic Statistics*, 40(4), 1831-1848.

[6] Juodis, A., Karabiyik, H., & Westerlund, J. (2021). On the robustness of the pooled CCE estimator. *Journal of Econometrics*, 220(2), 325-348.

[7] Karabiyik, H., Reese, S., & Westerlund, J. (2017). On the role of the rank condition in CCE estimation of factor-augmented panel regressions. *Journal of Econometrics*, 197(1), 60-64.

[8] Karabiyik, H., Urbain, J. P., & Westerlund, J. (2019). CCE estimation of factor-augmented regression models with more factors than observable. *Journal of Applied Econometrics*, 34(2), 268-284.

[9] Khan, S. A. R., Yu, Z., Belhadi, A., & Mardani, A. (2020). Investigating the effects of renewable energy on international trade and environmental quality. *Journal of Environmental management*, 272, 111089.

[10] Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multi-factor error structure. *Econometrica*, 74(4), 967-1012.

[11] Sarafidis, V., & Wansbeek, T. (2012). Cross-sectional dependence in panel data analysis. *Econometric Reviews*, 31(5), 483-531.

[12] Serlenga, L., & Shin, Y. (2007). Gravity models of intra-EU trade: application of the CCEP-HT estimation in heterogeneous panels with unobserved common time-specific factors. *Journal of Applied Econometrics*, 22(2), 361-381.

[13] Westerlund, J., & Urbain, J. P. (2015). Cross-sectional averages versus principal components. *Journal of Econometrics*, 185(2), 372-377.

Appendix

A The Proofs

Proof of Theorem 1: (i) For the step 1 of the two-steps estimation:

Similar to equation (23) and (24) of Hsiao et al. (2022), the first order condition of minimizing $w'(\bar{Z}\bar{Z}'/T)w$,under the constraint that $w'w = 1$, is that

$$(\frac{1}{T}\bar{Z}\bar{Z}')w = \delta^* w$$
$$w'(\frac{1}{T}\bar{Z}\bar{Z}')w = \delta^*$$

where δ^* is the Lagrange multiplier. Thus, $\hat{w}$ is the eigenvalue of the matrix $\frac{1}{T}\bar{Z}\bar{Z}'$, corresponding the zero eigenvalues, such that $\hat{w}'(\frac{1}{T}\bar{Z}\bar{Z}')\hat{w} = 0$ and then $\hat{w}'\bar{Z} = 0_{1\times(k+1)}$.

Step 2. According to Assumption 1 that $\bar{U} \xrightarrow{p} 0_{T\times(k+1)}$ as $N \to \infty$,and $||C|| < \infty$, $w'\bar{Z} = 0_{1\times(k+1)}$ implies that $w'F \xrightarrow{p} 0_{1\times r}$, as $N \to \infty$. Thus, we obtain

$$\sqrt{NT}(\widehat{\beta}_{TP} - \beta) = \left(\frac{1}{NT}\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'X_i\right)^{-1}\left(\frac{1}{\sqrt{NT}}\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'\varepsilon_i\right) + o_p(1).$$

Under the Assumption 1-5 and the same arguments as Hsiao et al. (2022), we obtain the following Lemmas 1 and 2:

Lemma 1 *Let $\xi = \hat{w} - w$, and as $T \to \infty$, $\xi'\xi \leq O_p\left(\frac{1}{NT^{1-\alpha}}\right)$, where $0 \leq \alpha < 1$.*

Lemma 2 *(i) $\frac{1}{N}\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'X_i \xrightarrow{p} \Phi_X$, as $N \to \infty$, with Φ_X is positive-definite matrix; (ii) $\frac{1}{\sqrt{N}}\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'\varepsilon_i \xrightarrow{d} \mathcal{N}(0, \Psi)$,with $\Psi = \lim_{(N,T)\to\infty} \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N} X_i'ww'E(\varepsilon_i\varepsilon_j')ww'X_i$.*

Thus, according to Lemma 1 and 2:

$$\frac{1}{NT}\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'X_i \xrightarrow{p} \frac{1}{\sqrt{NT}}\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'X_i$$

and

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N} X_i'\hat{w}\hat{w}'\varepsilon_i \xrightarrow{p} \frac{1}{\sqrt{NT}}\sum_{i=1}^{N} X_i'ww'\varepsilon_i$$

and $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N} X_i'ww'\varepsilon_i \xrightarrow{d} \mathcal{N}(0, \Psi)$. Thus, we conclude it.

(ii) The proof of the asymptotic normality is merely same as the proof of Proposition 6 in Hsiao et al. (2022) and then is omitted.