Dipartimento di Economia e Finanza

Southern Europe Research in Economic Studies

Testing for Correlated Factor Loadings in Cross Sectionally Dependent Panels

George Kapetanios

Laura Serlenga

Yongcheol Shin

SERIES Working Papers n. 02/2019

Testing for Correlated Factor Loadings in Cross Sectionally Dependent Panels*

George Kapetanios
King's College London

Laura Serlenga
University of Bari

Yongcheol Shin
University of York

June 2019

Abstract

A large strand of the literature on panel data models has focused on explicitly modelling the cross-section dependence between panel units. Factor augmented approaches have been proposed to deal with this issue. Under a mild restriction on the correlation of the factor loadings, we show that factor augmented panel data models can be encompassed by a standard two-way fixed effect model. This highlights the importance of verifying whether the factor loadings are correlated, which, we argue, is an important hypothesis to be tested, in practice. As a main contribution, we propose a Hausman-type test that determines the presence of correlated factor loadings in panels with interactive effects. Furthermore, we develop two nonparametric variance estimators that are robust to the presence of heteroscedasticity, autocorrelation as well as slope heterogeneity. Via Monte Carlo simulations, we demonstrate desirable size and power performance of the proposed test, even in small samples. Finally, we provide extensive empirical evidence in favour of uncorrelated factor loadings in panels with interactive effects.

JEL Classification: C13, C33.

Key Words: Panel Data Models, Cross-sectional Error Dependence, Unobserved Heterogeneous Factors, Factor Correlated Loadings.

*We are mostly grateful to Jia Chen, In Choi, Young Hoon Lee, Rui Lin, Ron Smith, Joakim Westerlund, Takashi Yamagata, Chaowen Zheng, and seminar participants at King's College London and University of York for their helpful comments. The usual disclaimer applies.

1 Introduction

Panel data models have been increasingly popular in applied economics and finance, due to their ability to model various sources of heterogeneity. A standard practice is to impose strong restrictions on error cross-section dependence (CSD). This takes the form of independence across individual units under the fixed effects model whilst a common time effect severely restricts the nature of CSD under the random effects specification.

The pervasive evidence detecting the presence of strong CSD in panels over the last decade (e.g., Pesaran (2015)), has prompted a large number of studies to develop proper econometric methodologies for modelling CSD, mainly through the structure of interactive effects. This introduces heterogeneous unobserved factors into the error components, allowing for a richer cross-sectional covariance structure.

Currently, there are two leading approaches that have received considerable attention in the literature, see Chudik and Pesaran (2015) for a survey. The first, based on principal component (PC) estimation, estimates the factors jointly with the main slope parameters. This approach has been exhaustively analysed by Bai (2009), and extended by, e.g., Charbonneau (2017), Fernandez-Val and Weidner (2016), and Moon and Weidner (2015). The second approach, advanced by Pesaran (2006), treats factors as nuisance terms, and removes their effects through proxying them by the cross-section averages of regressors as well as the dependent variable. This is referred to as the common correlated effects (CCE) estimator. A growing number of extensions have also been developed by, e.g., Chudik and Pesaran (2015), Kapetanios, Pesaran, and Yamagata (2011), and Westerlund and Urbain (2015). The finite sample performance of the two approaches has been intensively investigated. The earlier studies by Kapetanios and Pesaran (2005) and Chudik, Pesaran, and Tosetti (2011), provide Monte Carlo evidence in favour of the CCE estimator, which is partly due to uncertainty associated with estimating the true number of unobserved factors, see also Moon and Weidner (2015). Westerlund and Urbain (2015) provide an insightful summary arguing that the PC estimator performs better if the coefficients on regressors are zero while the CCE estimator tends to be superior otherwise, albeit in a restricted context.

The conventional wisdom has so far been that the standard two-way fixed effects (FE) estimator would be inappropriate and inconsistent in the presence of interactive effects, due to ignoring the potential endogeneity arising from the correlation between regressors and factors and/or factor loadings (e.g. Bai (2009)). In this paper, we first highlight a simple fact that has been almost neglected in this literature. We note that the FE estimator is not always inconsistent even in the presence of unobserved factor structure. In the case where the factor loadings are uncorrelated,[1] the FE estimator is shown to be consistent, albeit inefficient.[2] Furthermore, we provide

[1] Notice that this is the maintained assumption in the CCE literature, see Pesaran (2006).

[2] We find that only Coakely, Fuertes, and Smith (2006) and Sarafidis and Wansbeek (2012) briefly mention this fact. Recently, Westerlund (2018) argues that the FE estimator can be consistent in the presence of interactive

two nonparameteric variance estimators that are robust to the presence of heteroscedasticity, autocorrelation and slope heterogeneity.

Via Monte Carlo studies, we find that the FE and CCE estimators display a similar and satisfactory performance when factor loadings satisfy an uncorrelatedness condition whereas the performance of PC is more or less satisfactory except in small samples. As expected, the performance of both CCE and FE estimators worsens significantly under correlated factor loadings, which is in line with Westerlund and Urbain (2013). By contrast, the performance of the PC estimator is not unduly affected by correlated factor loadings.

A number of specification tests have been proposed to testing the validity of the cross section dependence or the presence of interactive effects, e.g. Pesaran (2015), Sarafidis, Yamagata, and Robertson (2009), Bai (2009) and Castagnetti, Rossi, and Trapani (2015). Notice, however, that the rejection of the null hypothesis by these tests does not determine whether the FE estimator is consistent or not under the alternative model with unobserved factors. For instance, Sarafidis, Yamagata, and Robertson (2009) maintain an assumption that factor loadings are uncorrelated under the alternative whilst the Hausman test proposed by Bai (2009) would have no power when factor loadings are uncorrelated, under the alternative hypothesis. This suggests that the presence of correlated loadings emerges as an influential but under-appreciated feature of the panel data model with interactive effects, see also Westerlund and Urbain (2013).

In retrospect, it is rather surprising to find that the literature has been silent on investigating the important issue of testing the validity of uncorrelated factor loadings. In order to fill this gap, as the main contribution of this paper, we proceed to develop a Hausman-type test that determines the validity of correlated factor loadings in cross-sectionally correlated panels. Both the FE and PC estimators are consistent under the null hypothesis of uncorrelated factor loadings whilst only the latter is consistent under the alternative hypothesis of correlated factor loadings. Further, the PC estimator is more efficient even under the null. Based on this observation, we develop two nonparametric variance estimators for the difference between the FE and PC estimators, that are shown to be robust to the presence of heteroscedasticity, autocorrelation and slope heterogeneity. We then show that the proposed test statistic follows the χ^2 distribution asymptotically. Monte Carlo simulation results confirm that the size and the power of the test is quite satisfactory even in small samples. Given that the (estimated) number of factors can make a considerable difference in the performance of the PC estimator, we also propose a pretest estimator which selects either the FE estimator if the null hypothesis of uncorrelated factor loadings is not rejected, or the PC estimator if the null is rejected. We find that the pretest estimator performs well, irrespective of whether factor loadings are correlated or not.

Finally, and crucially, we provide extensive empirical evidence, suggesting the lack of factor

effects, because both FE and CEE estimators belong to a class of estimators that satisfy a zero sum restriction. However, he still maintains the crucial assumption that factor loadings are uncorrelated.

loadings correlation in a number of datasets considered. We emphasize that the FE estimator is simpler and does not involve any complex issues related to selecting the correct number of unobserved factors, which has been shown to significantly affect the performance of PC estimators. This suggests that the standard FE estimator can still be of considerable applicability in a wide variety of cross-sectionally correlated panel datasets.

The paper proceeds as follows. Section 2 describes the model setup and derives the simple fact that the FE estimator is still consistent in the presence of interactive effects under the assumptions maintained in the CCE literature. Section 3 develops the Hausman-type test for the validity of uncorrelated factor loadings, which is the crucial condition, for the FE estimator to be consistent. Section 4 employs a range of Monte Carlo simulations to investigate the finite sample properties of the alternative estimators and the proposed test. Section 5 presents empirical evidence documenting that the null of uncorrelated factor loadings is not rejected for many dataset we investigate. Section 6 offers some concluding remarks, while mathematical proofs and data descriptions are collected in the Appendices.

2 The Model and the Simple Fact

Consider the following panel data model with interactive effects:

$$y_{it} = \alpha_i + \boldsymbol{\beta}_i' \boldsymbol{x}_{it} + \boldsymbol{\gamma}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{1}$$

$$\boldsymbol{x}_{it} = \boldsymbol{b}_i + \boldsymbol{\Gamma}_i' \boldsymbol{f}_t + \boldsymbol{v}_{it} \tag{2}$$

where y_{it} is the dependent variable of the $i-$th cross-sectional unit in period t, $\boldsymbol{x}_{it}$ is the $k \times 1$ vector of covariates with $\boldsymbol{\beta}_i$ the $k \times 1$ vector of parameters. α_i and $\boldsymbol{b}_i$ are unobserved individual effects, and ε_{it} and $\boldsymbol{v}_{it}$ are idiosyncratic errors. $\boldsymbol{f}_t$ is an $r \times 1$ vector of unobserved common factors while $\boldsymbol{\gamma}_i$ and $\boldsymbol{\Gamma}_i$ are random heterogenous loadings.

Following Pesaran (2006) and Karabiyik, Reese, and Westerlund (2017), we make the following assumptions:

Assumption A. (i) ε_{it} is independently distributed across i with $E(\varepsilon_{it}) = 0$, $E\left(\varepsilon_{it}^2\right) = \sigma_{\varepsilon_i}^2$ and $E\left(\varepsilon_{it}^{8+\delta}\right) < \infty$ for some $\delta > 0$. Each ε_{it} follows a linear process with absolutely summable autocovariances.

(ii) $\boldsymbol{u}_{it}$ is independently distributed across i with $E(\boldsymbol{v}_{it}) = \boldsymbol{0}$, $E(\boldsymbol{v}_{it}\boldsymbol{v}_{it}') = \boldsymbol{\Sigma}_{vi}$ and $E\left(\|\boldsymbol{v}_{it}\|^{8+\delta}\right) < \infty$ for some $\delta > 0$, where $\boldsymbol{\Sigma}_{vi}$ is a $k \times k$ positive definite matrix and $\|\boldsymbol{A}\| = \sqrt{tr(\boldsymbol{A}'\boldsymbol{A})}$ is the Frobenius norm. Further, $\boldsymbol{\Sigma}_v = \lim_{N\to\infty} N^{-1} \sum_{i=1}^N \boldsymbol{\Sigma}_{vi}$ is a positive definite matrix. Each $\boldsymbol{v}_{it}$ follows a vector linear process with absolutely summable autocovariance matrix norms.

(iii) $\boldsymbol{f}$ is covariance stationary such that $E\left(\|\boldsymbol{f}_t\|^4\right) < \infty$ and $E\left(\boldsymbol{f}_t\boldsymbol{f}_t'\right) = \boldsymbol{\Sigma}_f$ where $\boldsymbol{\Sigma}_f$ is an $r \times r$ positive definite matrix.

(iv) ε_{it}, $\boldsymbol{v}_{js}$ and f_ℓ are mutually independent for all i, j, t, s and ℓ.

(v) $\boldsymbol{\gamma}_i$ and $\boldsymbol{\Gamma}_i$ are *iid* across i and mutually uncorrelated, with finite means, $\bar{\boldsymbol{\gamma}}$ and $\bar{\boldsymbol{\Gamma}}$ and finite variances, $\boldsymbol{\Sigma}_\gamma$ and $\boldsymbol{\Sigma}_\Gamma$, respectively. They are independent of ε_{jt}, $\boldsymbol{v}_{jt}$ and $\boldsymbol{f}_t$ for all i and j.

(vi) The $k \times 1$ vector of heterogeneous parameters, $\boldsymbol{\beta}_i$ are generated as $\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\eta}_i$. $\boldsymbol{\eta}_i$ is independent across i, $E(\boldsymbol{\eta}_i) = 0$, $E(\boldsymbol{\eta}_i\boldsymbol{\eta}_i') = \Omega_{\eta\eta,i}$ which is a positive definite matrix uniformly for every i, $E\|\boldsymbol{\eta}_i\|^4 \leq \Delta < \infty$ and $\|\boldsymbol{\beta}\| < \infty$, and $\boldsymbol{\eta}_i$ is group-wise independent of ε_{it}, $\boldsymbol{v}_{it}$, γ_i and $\boldsymbol{\Gamma}_i$.

Assumption A is standard in the literature. For simplicity we assume that both ε_{it} and $\boldsymbol{v}_{it}$ are *iid*. But, we allow them to be serially correlated and conditionally heteroscedastic as well as weakly cross-sectionally correlated as in Assumption C in Bai (2009) and Assumptions B1 and B2 in Hayakawa, Nagata, and Yamagata (2018), hereafter, HNY. We then develop the nonparametric variance estimators, which are shown to be robust to heteroscedasticity, serial correlation and slope heterogeneity.

Combining (1) and (2), we have the system representation:

$$\boldsymbol{z}_{it} = \boldsymbol{\mu}_i + \boldsymbol{\Phi}_i \boldsymbol{f}_t + \boldsymbol{e}_{it} \tag{3}$$

where

$$\boldsymbol{z}_{it} = \begin{pmatrix} y_{it} \\ \boldsymbol{x}_{it} \end{pmatrix}, \boldsymbol{\mu}_i = \begin{pmatrix} \alpha_i + \boldsymbol{\beta}_i'\boldsymbol{b}_i \\ \boldsymbol{b}_i \end{pmatrix}, \boldsymbol{\Phi}_i = \begin{pmatrix} \boldsymbol{\Gamma}_i' + \boldsymbol{\beta}_i'\boldsymbol{\Gamma}_i' \\ \boldsymbol{\Gamma}_i' \end{pmatrix}, \boldsymbol{e}_{it} = \begin{pmatrix} \varepsilon_{it} + \boldsymbol{\beta}_i'\boldsymbol{v}_{it} \\ \boldsymbol{v}_{it} \end{pmatrix} \tag{4}$$

where the covariance matrix of $\boldsymbol{e}_{it}$ is given by

$$\boldsymbol{\Sigma}_{ei} = \begin{bmatrix} \sigma_{\varepsilon_i}^2 + \boldsymbol{\beta}_i'\boldsymbol{\Sigma}_{vi}\boldsymbol{\beta}_i' & \boldsymbol{\beta}_i'\boldsymbol{\Sigma}_{vi} \\ \boldsymbol{\Sigma}_{vi}\boldsymbol{\beta}_i' & \boldsymbol{\Sigma}_{vi} \end{bmatrix}$$

For a consistent estimation of the parameters in (1), we need to first account for the unobserved factors, and then estimate $\boldsymbol{\beta}$ by applying panel estimators to (1) with defactored variables. On the basis of this idea, two popular approaches have been proposed. The first approach, advanced by Pesaran (2006) and referred to as the common correlated effects (CCE) estimator, proxies factors by the cross-section averages of the dependent variable and regressors. The second is the principal component (PC) approach, that estimates the factors jointly with the parameters.

The validity of both approaches depends crucially upon whether the appropriate rank condition holds. Westerlund and Urbain (2015), Remark 4 on p.374, argue that the issue of correctly selecting the number of factors, r in the PC estimation, is essentially the same as the problem of satisfying the rank condition, $Rank\left(\frac{1}{N}\sum_{i=1}^N \boldsymbol{\Phi}_i\right) = r \leq k+1$ in CCE estimation.[3] Further, it is shown that both estimators involve bias terms, which do not disappear unless $N/T \to 0$. Monte Carlo simulations suggest that the performance of the PC estimator is sensitive to the value of $\boldsymbol{\beta}$. For $\boldsymbol{\beta} = 0$, the PC estimator outperforms CCE, while for $\boldsymbol{\beta} \neq 0$, the CCE estimator tends to outperform.

[3]Karabiyik, Reese, and Westerlund (2017) also investigate the importance of the rank condition in deriving the asymptotic distribution of the CCE estimator.

Surprisingly, however, we find that the performance of the two-way Fixed Effect (FE) estimator has not been explicitly investigated in the presence of unobserved multifactor structures, except for the studies by Coakely, Fuertes, and Smith (2006) and by Sarafidis and Wansbeek (2012). Such an omission simply reflects the conventional view that the FE estimator would be inappropriate or inconsistent in the presence of interactive effects, due to ignoring endogeneity stemming from the correlation between regressors and factors and/or factor loadings.

We aim to challenge this maintained view. To this end we assume $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all $i = 1, ..., N$ without loss of generality and represent (1) as the two-way error component model:

$$y_{it} = \alpha_i + \theta_t + \boldsymbol{x}_{it}'\boldsymbol{\beta} + u_{it} \tag{5}$$

where $\theta_t = \boldsymbol{f}_t'\boldsymbol{\gamma}$ and $u_{it} = \varepsilon_{it} + \boldsymbol{f}_t'(\boldsymbol{\gamma}_i - \boldsymbol{\gamma})$ with $\boldsymbol{\gamma} = E(\boldsymbol{\gamma}_i)$. Under Assumptions A (especially, A(v)), it is easily seen by the independence of $\boldsymbol{\gamma}_i - \boldsymbol{\gamma}$ from all other random quantities in the model that

$$E(u_{it}\theta_t) = E\left(\varepsilon_{it} + \boldsymbol{f}_t'(\boldsymbol{\gamma}_i - \boldsymbol{\gamma})\right)\boldsymbol{f}_t'\boldsymbol{\gamma} = E\left(\varepsilon_{it}\boldsymbol{f}_t'\boldsymbol{\gamma}\right) + E\left(\boldsymbol{f}_t'(\boldsymbol{\gamma}_i - \boldsymbol{\gamma})\boldsymbol{f}_t'\boldsymbol{\gamma}\right) = 0, \tag{6}$$

$$E(u_{it}\boldsymbol{x}_{it}) = E\left(\varepsilon_{it} + \boldsymbol{f}_t'(\boldsymbol{\gamma}_i - \boldsymbol{\gamma})\right)\left(\boldsymbol{b}_i + \boldsymbol{\Gamma}_i'\boldsymbol{f}_t + \boldsymbol{v}_{it}\right) = 0. \tag{7}$$

In this situation the two-way FE estimation can still produce an unbiased estimator of $\boldsymbol{\beta}$. Applying the two-way within transformation to y_{it} and $\boldsymbol{x}_{it}$ in (1), we obtain the transformed model as

$$\dot{y}_{it} = \dot{\boldsymbol{x}}_{it}'\boldsymbol{\beta} + \dot{u}_{it}, \tag{8}$$

where

$$\dot{y}_{it} = y_{it} - y_{i.} - y_{.t} + y_{..},\ \dot{\boldsymbol{x}}_{it} = \boldsymbol{x}_{it} - \boldsymbol{x}_{i.} - \boldsymbol{x}_{.t} + \boldsymbol{x}_{..}, \dot{u}_{it} = u_{it} - u_{i.} - u_{.t} + u_{..},$$

and

$$z_{i.} = T^{-1}\sum_{t=1}^{T} z_{it},\ z_{.t} = N^{-1}\sum_{i=1}^{N} z_{it},\ z_{..} = (NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} z_{it} \text{ for } z = y, \boldsymbol{x}, u$$

Under (6) and (7), it follows that $E(\dot{u}_{it}\dot{\boldsymbol{x}}_{it}) = 0$ since $E(u_{i.}\dot{\boldsymbol{x}}_{it}) = E(u_{.t}\dot{\boldsymbol{x}}_{it}) = E(u_{..}\dot{\boldsymbol{x}}_{it}) = 0$. Therefore, when factor loadings, $\boldsymbol{\gamma}_i$ and $\boldsymbol{\Gamma}_i$, are uncorrelated under Assumption A(v), we can apply FE estimation to obtain a consistent (albeit inefficient) estimator of $\boldsymbol{\beta}$ in (8). Conversely, if $\boldsymbol{\gamma}_i$ and $\boldsymbol{\Gamma}_i$ are correlated, it is clear that $E(u_{it}\boldsymbol{x}_{it}) \neq 0$ so that the FE estimator is inconsistent. Notice that the consistency of the FE estimator requires only $\boldsymbol{\gamma}_i$ to be uncorrelated with $\boldsymbol{\Gamma}_i$ (and $\varepsilon_{it}, \boldsymbol{u}_{it}, \boldsymbol{f}_t$), but this is a maintained assumption in the CCE literature, see Westerlund and Urbain (2015).[4]

[4] Pesaran (2006) implicitly assumes that the factor loadings are uncorrelated. Bai (2009) discusses this implication in detail, and shows via simulations that the CCE estimator is biased when $\boldsymbol{x}_{it}$ is correlated with both $\boldsymbol{\lambda}_i$ and $\boldsymbol{f}_t$. Remark 2 of Westerlund and Urbain (2013) questions the uncorrelated factor loadings assumption by arguing that a common shock that has a positive effect on savings, should have negative effects on interest rates. However, their discussion relates to the sign of the average effect of common shocks or the sign of the cross-section mean of loadings. Since the independence assumption does not restrict the sign of these means, the relevance of such a relaxation would be somewhat questionable.

Applying the pooled estimation to (8), we obtain the two-way FE estimator of β by

$$\hat{\boldsymbol{\beta}}_{FE} = \left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i\right)^{-1} \sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{y}}_i \tag{9}$$

where $\dot{\boldsymbol{X}}_i = (\dot{\boldsymbol{x}}_{i1}, ..., \dot{\boldsymbol{x}}_{iT})'$ and $\dot{\boldsymbol{y}}_i = (\dot{y}_{i1}, ..., \dot{y}_{iT})'$. Further, we propose the two consistent versions of the variance estimator, which are shown to be robust to the heteroscedasticity and the serial-correlation as well as the slope heterogeneity. The first is the nonparametric variance estimator, similarly applied in deriving the variance of the CCE estimator by Pesaran (2006):

$$\begin{aligned} &\hat{\boldsymbol{V}}^{NON}\left(\hat{\boldsymbol{\beta}}_{FE}\right) \\ &= \left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i\right)^{-1} \left(\sum_{i=1}^{N} \left(\dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i\right)\left(\hat{\boldsymbol{\beta}}_{FE,i} - \hat{\boldsymbol{\beta}}_{FE}\right)\left(\hat{\boldsymbol{\beta}}_{FE,i} - \hat{\boldsymbol{\beta}}_{FE}\right)'\left(\dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i\right)\right)\left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i\right)^{-1} \end{aligned} \tag{10}$$

where $\hat{\boldsymbol{\beta}}_{FE,i} = \left(\dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i\right)^{-1} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{y}}_i$. Next, we consider the following heteroscedasticity, autocorrelation and slope heterogeneity robust variance estimator (see HNY):

$$\hat{\boldsymbol{V}}^{HAC}\left(\hat{\boldsymbol{\beta}}_{FE}\right) = \left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i\right)^{-1} \left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \hat{\boldsymbol{u}}_i \hat{\boldsymbol{u}}_i' \dot{\boldsymbol{X}}_i\right)\left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i\right)^{-1} \tag{11}$$

where $\hat{\boldsymbol{u}}_i = \dot{\boldsymbol{y}}_i - \dot{\boldsymbol{X}}_i \hat{\boldsymbol{\beta}}_{FE}$.

We show that $\hat{\boldsymbol{\beta}}_{FE}$ is $\sqrt{N}$-consistent and asymptotically normal.

Theorem 1 *Under Assumption A, as $N, T \to \infty$,*

$$\sqrt{N}\left(\hat{\boldsymbol{\beta}}_{FE} - \boldsymbol{\beta}\right) \to_d N\left(0_{k\times 1}, \boldsymbol{\Psi}_{FE}^{-1} \boldsymbol{R}_{FE} \boldsymbol{\Psi}_{FE}^{-1}\right) \tag{12}$$

where

$$\boldsymbol{\Psi}_{FE} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} E\left(\frac{\dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i}{T}\right) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} E\left(\tilde{\boldsymbol{\Gamma}}_i' \frac{\tilde{\boldsymbol{F}}' \tilde{\boldsymbol{F}}}{T} \tilde{\boldsymbol{\Gamma}}_i\right) + \boldsymbol{\Sigma}_u,$$

$\tilde{\boldsymbol{F}} = \left[(\boldsymbol{f}_1 - \bar{\boldsymbol{f}}), ..., (\boldsymbol{f}_T - \bar{\boldsymbol{f}})\right]'$ *with* $\bar{\boldsymbol{f}} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{f}_t$, $\tilde{\boldsymbol{\Gamma}}_i = \boldsymbol{\Gamma}_i - \bar{\boldsymbol{\Gamma}}$ *with* $\bar{\boldsymbol{\Gamma}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\Gamma}_i$, *and* $\boldsymbol{R}_{FE} = \boldsymbol{R}_{1,FE}$ *if* $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ *and* $\boldsymbol{R}_{FE} = \boldsymbol{R}_{1,FE} + \boldsymbol{R}_{2,FE}$, *if* $\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\eta}_i$, *where*

$$\boldsymbol{R}_{1,FE} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} E\left(\frac{\dot{\boldsymbol{X}}_i' \tilde{\boldsymbol{F}}}{T} \tilde{\boldsymbol{\gamma}}_i \tilde{\boldsymbol{\gamma}}_i' \frac{\tilde{\boldsymbol{F}}' \dot{\boldsymbol{X}}_i}{T}\right) \tag{13}$$

$$\boldsymbol{R}_{2,FE} = \lim_{N\to\infty} \sum_{i=1}^{N} E\left(\frac{\dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i}{T} \boldsymbol{\eta}_i \boldsymbol{\eta}_i' \frac{\dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i}{T}\right), \tag{14}$$

and $\tilde{\boldsymbol{\gamma}}_i = \boldsymbol{\gamma}_i - \bar{\boldsymbol{\gamma}}$ *with* $\bar{\boldsymbol{\gamma}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\gamma}_i$. *Furthermore,*

$$\hat{\boldsymbol{V}}^{NON}\left(\hat{\boldsymbol{\beta}}_{FE}\right)^{-1/2}\left(\hat{\boldsymbol{\beta}}_{FE} - \boldsymbol{\beta}\right) \to_d N\left(0, \boldsymbol{I}_k\right) \text{ and } \hat{\boldsymbol{V}}^{HAC}\left(\hat{\boldsymbol{\beta}}_{FE}\right)^{-1/2}\left(\hat{\boldsymbol{\beta}}_{FE} - \boldsymbol{\beta}\right) \to_d N\left(0, \boldsymbol{I}_k\right). \tag{15}$$

The main message is that two-way FE estimation can produce a consistent estimator of $\boldsymbol{\beta}$ under Assumption A(v), irrespective of whether the rank condition holds or not. This simple fact that the FE estimator is not always inconsistent even in the presence of unobserved factor structure, has been almost neglected in the literature. Conversely, if the factor loadings, $\boldsymbol{\gamma}_i$ and $\boldsymbol{\Gamma}_i$, are correlated, the FE estimator becomes inconsistent. In this case the CCE estimator may also be inconsistent in general, but it can be consistent only in the special case with the full rank, namely $r = k + 1$. See Westerlund and Urbain (2013) for the simulation evidence showing that the CCE estimator performs poorly when the factor loadings are correlated.

Recently, Westerlund (2018) shows that if the true model is given by (1), one can obtain transformed regressors, say $\bar{\boldsymbol{x}}_{it}$, such that

$$\sum_i \bar{\boldsymbol{X}}_i = \mathbf{0}, \tag{16}$$

where $\bar{\boldsymbol{X}}_i = (\bar{\boldsymbol{x}}_{i1}, ..., \bar{\boldsymbol{x}}_{iT})'$. Then, the following pooled OLS estimator

$$\hat{\boldsymbol{\beta}}_{ZS} = \left(\sum_i \bar{\boldsymbol{X}}_i' \bar{\boldsymbol{X}}_i \right)^{-1} \sum_i \bar{\boldsymbol{X}}_i' \boldsymbol{y}_i, \tag{17}$$

will be consistent, where the subscript ZS stands for zero sum. However, Assumption A(v) is still maintained in this analysis, which is the crucial condition for consistency of the ZS estimator.[5] Furthermore, the use of both (16) and (17) raises some issues. First, it is not clear how the FE estimator belongs to the class of ZS estimators, $\hat{\boldsymbol{\beta}}_{ZS}$, since the FE estimator uses a transformed dependent variable while $\hat{\boldsymbol{\beta}}_{ZS}$ does not. Next, the following crucial restriction is imposed for consistency:

$$\left(\sum_i \bar{\boldsymbol{X}}_i' \bar{\boldsymbol{X}}_i \right)^{-1} \sum_i \bar{\boldsymbol{X}}_i' \boldsymbol{X}_i \boldsymbol{\beta} = \boldsymbol{\beta}, \tag{18}$$

where $\boldsymbol{X}_i = (\boldsymbol{x}_{i1}, ..., \boldsymbol{x}_{iT})'$. Though both FE and CCE estimators satisfy (18), it is unclear how to construct general $\bar{\boldsymbol{X}}_i$ satisfying this condition.

3 Testing for Correlated Factor Loadings

A number of specification tests have been proposed to test the validity of the cross section dependence or the presence of the multiplicative interactive effects in panels. The most popular test is the so-called CD test proposed by Pesaran (2015), who showed that the CD test can be applied to a wide variety of models, including heterogeneous dynamic panel data models, even with multiple breaks and non-stationary variables. However, the CD test fails to reject the null hypothesis of no error CSD when the factor loadings have zero means, implying that the CD test

[5] We notice that these results can be obtained only if the factor loadings are uncorrelated or the full rank condition is met. Surprisingly, Westerlund (2018) does not provide any simulation evidence for the cases with correlated factor loadings and rank deficiency.

will display very poor power when it is applied to cross-sectionally demeaned data. Sarafidis, Yamagata, and Robertson (2009) propose an alternative testing procedure for the homogeneous factor loadings after estimating a linear dynamic panel data model by GMM. This approach is valid only when N is large relative to T, but it can be applied to testing for any error CSD remaining after including time dummies. But, Sarafidis, Yamagata, and Robertson (2009) still maintain an assumption that the factor loadings are uncorrelated (see Assumption 5(b)). If the factor loadings are correlated under the alternative hypothesis, it is easily seen that the proposed GMM-based test will be invalid because the GMM estimator is no longer consistent under the alternative.

The PC estimator is consistent both under models with additive effects and with interactive effects, but less efficient than the FE estimator under the null model with additive effects only. On the other hand, the FE estimator is inconsistent under the alternative model with interactive effects and correlated factor loadings. Following this idea, Bai (2009), Section 9, advances a Hausman test for testing the null hypothesis of additive effects against the alternative of interactive effects. Focussing on the special cases, Castagnetti, Rossi, and Trapani (2015) propose two tests for the null of no factor structure: one for the null that factor loadings are cross sectionally homogeneous, and another for the null that common factors are homogeneous over time. Using extremes of the estimated loadings and common factors, they show that their statistics follow an asymptotic Gumbel distribution under the null.[6]

The conventional wisdom is that if the null hypothesis of no error CSD or additive effects is rejected, the use of the standard FE estimator would be invalid due to ignoring the potential endogeneity arising from the correlation between regressors and unobserved factors. We have shown that the presence of CSD or interactive effects does not always imply that the FE estimator is inconsistent even in the presence of unobserved factor structure. In particular, if the factor loadings are uncorrelated, we showed that the FE estimator is still consistent. More importantly, the FE estimator avoids any issue related to selecting the correct number of unobserved factors, which has been shown to significantly affect the performance of both CCE and PC estimators, e.g. Westerlund and Urbain (2015).

In this regard, it is rather surprising to find that the literature has been silent on investigating an important issue of testing the validity of uncorrelated factor loadings in panels with interactive effects. For example, Sarafidis, Yamagata, and Robertson (2009) maintains an assumption that factor loadings are uncorrelated under the alternative. If they are correlated, their proposed test becomes invalid. Notice that the Hausman test developed by Bai (2009) would be valid only if regressors are correlated with both factors and loadings. Further, it is easily seen that

[6] Castagnetti, Rossi, and Trapani (2015) do not consider the two-way FE estimator; they consider the one-way FE estimator when testing $\boldsymbol{f}_t = \boldsymbol{f}$ for all t whilst considering the between estimator when testing $\gamma_i = \gamma$ for all i. Furthermore, Castagnetti, Rossi, and Trapani (2015) show that the average-type statistics diverge under the null while the Hausman-type ones are inconsistent.

Hausman test has no power if the factor loadings are uncorrelated under the alternative model with interactive effects. This raises a potentially important research question. For large T, without loss of generality, we suppose that $\boldsymbol{f}_t$ represent the unobserved common policy or globalisation trend, and $\boldsymbol{\gamma}_i$ are the associated heterogeneous individual responses (parameters). In this context, it is natural to allow for $\boldsymbol{x}_{it}$ to be correlated with $\boldsymbol{f}_t$ to avoid the potential omitted variables bias. But, it still remains an important issue to test whether $\boldsymbol{x}_{it}$ are correlated with $\boldsymbol{\gamma}_i$ or not.

Given the pervasive evidence of cross sectionally dependent errors in panels (see Pesaran (2015)), as the main contribution of this paper, we proceed to develop a Hausman-type test that determines the presence of uncorrelated factor loadings in cross-sectionally correlated panels. In terms of the model, (1) and (2) with interactive effects, recall that both the two way FE estimator and the PC estimator are consistent under the null hypothesis that factor loadings, $\boldsymbol{\gamma}_i$ and Γ_i are uncorrelated. Only the latter is consistent under the alternative of correlated loadings. Following this idea, we propose the Hausman-type test based on the difference between the FE and PC estimators:

$$H = \left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{PC}\right)' \boldsymbol{V}^{-1} \left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{PC}\right) \tag{19}$$

where $\hat{\boldsymbol{\beta}}_{PC}$ is the PC estimator to be defined below, and $\boldsymbol{V} = Var\left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{PC}\right) = Var\left(\hat{\boldsymbol{\beta}}_{FE}\right) + Var\left(\hat{\boldsymbol{\beta}}_{PC}\right) - 2Cov\left(\hat{\boldsymbol{\beta}}_{FE}, \hat{\boldsymbol{\beta}}_{PC}\right)$. Notice that both estimators are consistent under the null hypothesis, but the PC estimator is more efficient than the FE estimator even under the null. This implies that

$$Var\left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{PC}\right) \neq Var\left(\hat{\boldsymbol{\beta}}_{FE}\right) - Var\left(\hat{\boldsymbol{\beta}}_{PC}\right)$$

in contrast to the well-established finding in Hausman (1978). We interpret (19) as a test for correlated factor loadings in panels with interactive effects.

Before developing the asymptotic theory for the Hausman-type statistic, we describe the asymptotic distribution of the bias-corrected PC estimator given by[7]

$$\hat{\boldsymbol{\beta}}_{PC} = \bar{\boldsymbol{\beta}}_{PC} - \frac{1}{N}\hat{\boldsymbol{c}}_{NT}$$

where

$$\bar{\boldsymbol{\beta}}_{PC} = \left(\sum_{i=1}^{N} \boldsymbol{X}_i' \boldsymbol{M}_{\hat{F}} \boldsymbol{X}_i\right)^{-1} \sum_{i=1}^{N} \boldsymbol{X}_i' \boldsymbol{M}_{\hat{F}} \boldsymbol{y}_i,$$

is the uncorrected PC estimator, $\boldsymbol{M}_{\hat{F}} = \boldsymbol{I}_T - \hat{\boldsymbol{F}}\left(\hat{\boldsymbol{F}}'\hat{\boldsymbol{F}}\right)^{-1}\hat{\boldsymbol{F}}'$, and $\hat{\boldsymbol{F}}$ is estimated by $\sqrt{T}$ times the eigenvectors corresponding to the r largest eigenvalues of $N^{-1}\sum_{i=1}^{N} \boldsymbol{Z}_i \boldsymbol{Z}_i'$, $\boldsymbol{Z}_i = (\boldsymbol{y}_i, \boldsymbol{X}_i)$,

[7]Notice that the model, (1) can be written as

$$\boldsymbol{y}_i = \alpha_i \boldsymbol{i}_T + \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{F}\gamma_i + \boldsymbol{\varepsilon}_i$$

and the one-way within transformation has already been applied such that

$$\tilde{\boldsymbol{y}}_i = \tilde{\boldsymbol{X}}_i \boldsymbol{\beta} + \tilde{\boldsymbol{F}}\gamma_i + \tilde{\boldsymbol{\varepsilon}}_i$$

where $\tilde{y}_{it} = y_{it} - \bar{y}_i$ with $\bar{y}_i = T^{-1}\sum_{i=1}^{T} y_{it}$. For notational simplicity we use $\boldsymbol{y}_i = \tilde{\boldsymbol{y}}_i$ and $\boldsymbol{X}_i = \tilde{\boldsymbol{X}}_i$.

and $\hat{\boldsymbol{c}}_{NT}$ is a bias correction term. We consider both versions of the bias-corrected PC estimators proposed by Bai (2009) and HNY (see Appendix 9 for the details).

Next, we propose two robust versions of the variance estimator as follows:

$$\hat{\boldsymbol{V}}^{NON}\left(\hat{\boldsymbol{\beta}}_{PC}\right) \tag{20}$$
$$= \left(\sum_{i=1}^{N} \boldsymbol{X}_i'\boldsymbol{M}_{\hat{F}}\boldsymbol{X}_i\right)^{-1} \left(\sum_{i=1}^{N} \left(\boldsymbol{X}_i'\boldsymbol{M}_{\hat{F}}\boldsymbol{X}_i\right)\left(\bar{\boldsymbol{\beta}}_{PC,i}-\bar{\boldsymbol{\beta}}_{PC}\right)\left(\bar{\boldsymbol{\beta}}_{PC,i}-\bar{\boldsymbol{\beta}}_{PC}\right)'\left(\boldsymbol{X}_i'\boldsymbol{M}_{\hat{F}}\boldsymbol{X}_i\right)\right) \left(\sum_{i=1}^{N} \boldsymbol{X}_i'\boldsymbol{M}_{\hat{F}}\boldsymbol{X}_i\right)^{-1}$$

where $\bar{\boldsymbol{\beta}}_{PC,i} = \left(\boldsymbol{X}_i'\boldsymbol{M}_{\hat{F}}\boldsymbol{X}_i\right)^{-1}\boldsymbol{X}_i'\boldsymbol{M}_{\hat{F}}\boldsymbol{y}_i$ and

$$\hat{\boldsymbol{V}}^{HAC}\left(\hat{\boldsymbol{\beta}}_{PC}\right) = \left(\sum_{i=1}^{N} \boldsymbol{X}_i'\boldsymbol{M}_{\hat{F}}\boldsymbol{X}_i\right)^{-1} \left(\sum_{i=1}^{N} \hat{\boldsymbol{X}}_i'\hat{\mathbf{V}}_i\hat{\mathbf{V}}_i'\hat{\boldsymbol{X}}_i\right) \left(\sum_{i=1}^{N} \boldsymbol{X}_i'\boldsymbol{M}_{\hat{F}}\boldsymbol{X}_i\right)^{-1} \tag{21}$$

where $\hat{\boldsymbol{V}}_i = \boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}_{PC}$.

We provide the asymptotic results for the $\hat{\boldsymbol{\beta}}_{PC}$ estimator in Theorem 2.

Theorem 2 *Suppose that Assumption A holds. In the homogeneous case with $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all i, as $N, T \to \infty$,and*

$$\sqrt{NT}\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right) \to_d N\left(0_{k\times 1}, \boldsymbol{\Psi}_{PC}^{-1}\boldsymbol{R}_{1,PC}\boldsymbol{\Psi}_{PC}^{-1}\right) \tag{22}$$

where

$$\boldsymbol{\Psi}_{PC} = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left(\frac{\mathbf{V}_i'\mathbf{V}_i}{T}\right),$$

$$\boldsymbol{R}_{1,PC} = \lim_{N\to\infty} N^{-1}\sum_{i=1}^{N} E\left(\frac{\mathbf{V}_i'\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{V}_i}{T}\right) \tag{23}$$

and $\mathbf{V}_i = (\boldsymbol{v}_{i1}, ..., \boldsymbol{v}_{iT})'$. In the heterogeneous case with $\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\eta}_i$, as $N, T \to \infty$,

$$\sqrt{N}\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right) \to_d N\left(0_{k\times 1}, \boldsymbol{\Psi}_{PC}^{-1}\boldsymbol{R}_{2,PC}\boldsymbol{\Psi}_{PC}^{-1}\right) \tag{24}$$

where

$$\boldsymbol{R}_{2,PC} = \lim_{N\to\infty} N^{-1}\sum_{i=1}^{N} E\left(\frac{\mathbf{V}_i'\mathbf{V}_i}{T}\boldsymbol{\eta}_i\boldsymbol{\eta}_i'\frac{\mathbf{V}_i'\mathbf{V}_i}{T}\right) \tag{25}$$

Furthermore,

$$\hat{\boldsymbol{V}}^{NON}\left(\hat{\boldsymbol{\beta}}_{PC}\right)^{-1/2}\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right) \to_d N\left(0, \boldsymbol{I}_k\right) \text{ and } \hat{\boldsymbol{V}}^{HAC}\left(\hat{\boldsymbol{\beta}}_{PC}\right)^{-1/2}\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right) \to_d N\left(0, \boldsymbol{I}_k\right). \tag{26}$$

Having established that the two versions of the robust estimator can consistently standardise the estimator, we propose to account for the covariance $Cov\left(\hat{\boldsymbol{\beta}}_{FE}, \hat{\boldsymbol{\beta}}_{PC}\right)$ by setting[8]

$$\hat{\boldsymbol{C}}^{NON}\left(\hat{\boldsymbol{\beta}}_{FE}, \hat{\boldsymbol{\beta}}_{PC}\right)$$
$$= \left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i\right)^{-1} \left(\sum_{i=1}^{N} \left(\dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i\right) \left(\hat{\boldsymbol{\beta}}_{FE,i} - \hat{\boldsymbol{\beta}}_{FE}\right) \left(\hat{\boldsymbol{\beta}}_{PC,i} - \hat{\boldsymbol{\beta}}_{PC}\right)' \left(\boldsymbol{X}_i' \boldsymbol{M}_{\hat{F}} \boldsymbol{X}_i\right)\right) \left(\sum_{i=1}^{N} \boldsymbol{X}_i' \boldsymbol{M}_{\hat{F}} \boldsymbol{X}_i\right)^{-1}$$

and

$$\hat{\boldsymbol{C}}^{HAC}\left(\hat{\boldsymbol{\beta}}_{FE}, \hat{\boldsymbol{\beta}}_{PC}\right) = \left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i\right)^{-1} \left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \hat{\boldsymbol{u}}_i \hat{\boldsymbol{V}}_i' \hat{\boldsymbol{X}}_i\right) \left(\sum_{i=1}^{N} \boldsymbol{X}_i' \boldsymbol{M}_{\hat{F}} \boldsymbol{X}_i\right)^{-1}.$$

Accordingly, we define two operating versions of the Hausman-type statistic by

$$H^{NON} = \left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{PC}\right)' \left(\hat{\boldsymbol{V}}^{NON}\right)^{-1} \left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{PC}\right) \tag{27}$$

$$H^{HAC} = \left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{PC}\right)' \left(\hat{\boldsymbol{V}}^{HAC}\right)^{-1} \left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{PC}\right) \tag{28}$$

where

$$\hat{\boldsymbol{V}}^{NON} = \hat{\boldsymbol{V}}^{NON}\left(\hat{\boldsymbol{\beta}}_{FE}\right) + \hat{\boldsymbol{V}}^{NON}\left(\hat{\boldsymbol{\beta}}_{PC}\right) - 2\hat{\boldsymbol{C}}^{NON}\left(\hat{\boldsymbol{\beta}}_{FE}, \hat{\boldsymbol{\beta}}_{PC}\right) \tag{29}$$

$$\hat{\boldsymbol{V}}^{HAC} = \hat{\boldsymbol{V}}^{HAC}\left(\hat{\boldsymbol{\beta}}_{FE}\right) + \hat{\boldsymbol{V}}^{HAC}\left(\hat{\boldsymbol{\beta}}_{PC}\right) - 2\hat{\boldsymbol{C}}^{HAC}\left(\hat{\boldsymbol{\beta}}_{FE}, \hat{\boldsymbol{\beta}}_{PC}\right) \tag{30}$$

We provide the main result in the following Theorem.

Theorem 3 *Under Assumption A, as* $N, T \to \infty$,

$$H^j \to_d \chi_k^2 \ \textit{for } j = NON, HAC$$

There is ample evidence, as will be seen in the Monte Carlo section, that the CCE and the FE estimators perform comparably to the PC estimator in small samples, if the null hypothesis is not rejected. Hence, the proposed Hausman-type test statistic has a clear operational rationale in practice.

4 Monte Carlo Simulations

4.1 Review of previous studies

Westerlund and Urbain (2013) examine the small sample performance of the CCE estimator, and find that the CCE does not perform very well in the presence of correlated factor loadings, especially if the rank condition is not satisfied. But, the CCE performs reasonably well under the full rank case. Karabiyik, Reese, and Westerlund (2017) discuss the role of the rank condition in

[8] Following Bai (2009), we have also employed the analytic (sandwich-form) variance estimator of $\boldsymbol{V}$, taking into account unknown form of heteroscedastic and autocorrelated errors. After conducting the preliminary simulations, we come to a conclusion that the two robust versions of the estimators perform more satisfactory even in the presence of slope parameter heterogeneity.

the CCE estimation, and show that the second moment matrix of the estimated factors becomes asymptotically singular if the number of factors is strictly less than the number of dependent and independent variables, invalidating the key arguments commonly applied to establish the asymptotic theory. Westerlund and Urbain (2015) provide a formal comparison between the cross-sectional average (CA) and principal component (PC) estimators by employing he same data generating process (DGP) in the framework of the factor augmented regressions.[9] They show that the two estimators are asymptotically the same only if $N/T \to 0$ whereas their asymptotic distributions are no longer equivalent, especially in terms of the asymptotic biases, if $N/T \to \tau > 0$.

Though a number of papers have examined the small sample performance of the CCE and PC estimators, we find that only two studies by Sarafidis and Wansbeek (2012) and Westerlund (2018), have explicitly analysed the performance of the FE estimator in the presence of CSD. Assuming the homogenous slope parameters with $N = 100$ and $T = 50$, Sarafidis and Wansbeek (2012) compare the performance of the CCE, the PC and the (two-way) FE estimators. If the factor loadings are uncorrelated and the rank condition is satisfied, they find that all these three estimators perform well in terms of bias and RMSE. When the factor loadings are correlated, the FE estimator is severely biased, irrespective of whether the rank condition is satisfied or not. The CCE estimator is substantially biased if the rank condition is violated. On the other hand, the performance of the PC estimator is not significantly affected by the presence of correlated factor loadings. Still, assuming that factor loadings are uncorrelated, Westerlund (2018) demonstrates that the performance of the FE and CEE estimators is similar and satisfactory. Surprisingly, however, he does no longer consider the role of the rank condition and/or the correlated factor loadings as in Westerlund and Urbain (2013).

We will examine the relative performance of the FE, CCE and PC estimators in the presence of correlated and uncorrelated factor loadings and/or rank deficiency. Furthermore, we investigate the small sample performance of the proposed Hausman-type statistic.

4.2 Monte Carlo design

Following the model, (1) and (2), we generate the data with $m_x = 1$, $m_f = 2$ and $\alpha_i = b_i = 0$:

$$y_{it} = \beta_i x_{it} + \gamma_{1i} f_{1t} + \gamma_{2i} f_{2t} + \varepsilon_{it}, \tag{31}$$

and

$$x_{it} = \Gamma_{1i} f_{1t} + \Gamma_{2i} f_{2t} + u_{it}, \tag{32}$$

where $(f_{1t}, f_{2t}, \varepsilon_{it}, u_{it})'$ are drawn from the multivariate normal distribution with zero means and covariance matrix, $\mathbf{\Sigma}_i = diag\left(\sigma^2_{f1}, \sigma^2_{f2}, \sigma^2_{\varepsilon_i}, \sigma^2_{u_i}\right)$. We follow Pesaran (2006) and Westerlund and

[9] Notice that the DGP and the estimators are not identical to what have proposed by Pesaran (2006) and Bai (2009).

Urbain (2013), and generate the factor loadings, $(\gamma_{1i}, \gamma_{2i})$ and $(\Gamma_{1i}, \Gamma_{2i})$ as follows:

- Experiment 1 with uncorrelated factor loadings and the full rank in which case $\gamma_{1i} \sim iidN(1,1)$, $\gamma_{2i} \sim iidN(0,1)$, $\Gamma_{1i} \sim iidN(0,1)$, $\Gamma_{2i} \sim iidN(1,1)$ such that $E\begin{pmatrix} \gamma_{1i} & \gamma_{2i} \\ \Gamma_{1i} & \Gamma_{2i} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

- Experiment 2 with uncorrelated factor loadings and the rank deficiency in which case $\gamma_{1i} \sim iidN(1,1)$, $\gamma_{2i} \sim iidN(0,1)$, $\Gamma_{1i} \sim iidN(1,1)$, $\Gamma_{2i} \sim iidN(0,1)$, such that $E\begin{pmatrix} \gamma_{1i} & \gamma_{2i} \\ \Gamma_{1i} & \Gamma_{2i} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$.

- Experiment 3 with correlated factor loadings and the full rank in which case: $\gamma_{1i} = \gamma_1 + \upsilon_{1i}$, $\gamma_{2i} = \gamma_2 + \upsilon_{2i}$, $\Gamma_{1i} = \Gamma_1 + \upsilon_{1i}$, and $\Gamma_{2i} = \Gamma_2 + \upsilon_{2i}$ with $\gamma_1 = 1$, $\gamma_2 = 0$, $\Gamma_1 = 2$, $\Gamma_2 = 1$ and $(\upsilon_{1i}, \upsilon_{2i}) \sim iidN(0, I_2)$, such that $E\begin{pmatrix} \gamma_{1i} & \gamma_{12} \\ \Gamma_{1i} & \Gamma_{12} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$

- Experiment 4 with correlated factor loadings and the rank deficiency in which case $\gamma_{1i} \sim iidN(1,1)$, $\gamma_{2i} \sim iidN(0,1)$, $\gamma_{1i} = \Gamma_{1i}$ and $\gamma_{2i} = \Gamma_{2i}$ such that $E\begin{pmatrix} \gamma_{1i} & \gamma_{2i} \\ \Gamma_{1i} & \Gamma_{2i} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$.

For the main slope parameter, β, we consider the homogeneous case with $\beta_i = 1$ for all i, and the heterogenous case with $\beta_i = 1 + \eta_i$ and $\eta_i \sim iidN(0, 0.04)$. We consider the following combination of $(N, T) = 20, 30, 50, 100, 200$, and set the number of replications at $R = 1,000$.

4.3 The small sample performance of FE, CCE and PC estimators

We examine the finite sample performance of the following estimators applied to the model (1) and (2): the two-way fixed effect (FE) estimator, $\hat{\beta}_{FE}$, the CCE estimator proposed by Pesaran (2006), $\hat{\beta}_{CCE}$, and the bias corrected principal component (PC) estimators proposed respectively by Bai (2009) and HNY, denoted $\hat{\beta}_{PCBai}$ and $\hat{\beta}_{PCHNY}$. We consider both pooled and mean group estimators except for $\hat{\beta}_{PCBai}$ (see Appendix 7 for details). Notice that consistency of the PC estimator depends crucially upon correctly selecting the number of unobserved factors (e.g. Moon and Weidner (2015)). In this regard, to address the potential uncertainty associated with the selection criteria, we initially consider the two information criteria, denoted IC_{p1} and AIC_1, proposed by (Bai and Ng, 2002, p 202). Overall, we find that the PC estimator using IC_{p1} outperforms that with AIC_1, and we only report the results based on IC_{p1}.

We report the following summary statistics:

- Bias: $\hat{\beta}_R - \beta_0$, where β_0 is a true parameter value and $\hat{\beta}_R = R^{-1}\sum_{r=1}^{R}\hat{\beta}_r$ is the mean coefficient of $\hat{\beta}_r$ across R replications.

- RMSE: the root mean square error estimated by $\sqrt{R^{-1}\sum_{r=1}^{R}\left(\hat{\beta}_r - \beta_0\right)^2}$

Table 1 shows the simulation results for Experiment 1 with the full rank, uncorrelated factor loadings and homogeneous β's. We find that the biases of all estimators are mostly negligible even for the relatively small samples with the FE performing slightly worse than other estimators when $N = 20$. The relative performance of the FE, CCE and PC estimators is generally in line with the simulation results reported in Chudik, Pesaran, and Tosetti (2011), Sarafidis and Wansbeek (2012) and HNY. The results for RMSEs display qualitatively similar patterns. RMSEs of CCE and PC estimators are lower than those of the FE estimators and decline as N or T grows.

An important exception is the poor performance of the PC estimator evaluated using the AIC_1 criterion.[10] In this case the biases are substantial and nonnegligible in small samples, especially for the $\hat{\beta}_{PCHNY}$ estimator. They decline only when both N and T become large. Further, their RMSEs are much larger than those of the other estimators and decrease only when both N and T are large. This clearly demonstrates the influence of the estimated number of factors for the PC estimator. Given that information criteria have very variable performance, this is a problematic issue for PC estimators but also highlights that the FE estimator can make an operational alternative.

On the other hand, the RMSE performance of the FE estimator improves only with N, consistent with the theoretical prediction that the FE estimator is $\sqrt{N}$-consistent in panels with interactive effects when factor loadings are uncorrelated. Finally, biases and RMSEs of the pooled and mean group estimators display almost identical patterns.

Table 2 presents simulation results for Experiment 1 with heterogeneous βs. The biases and the $RMSE$ of the four estimators are qualitatively similar to those reported in Table 1, confirming that the small sample performance of all three estimators are still reliable, as long as factor loadings are uncorrelated. Again, both the pooled and mean group estimators display similar performance.

Table 3 presents simulation results for Experiment 2 where factor loadings are uncorrelated but the rank condition is violated for the homogeneous βs. The performance of the CCE estimators tends to slightly deteriorate in small samples, but both bias and $RMSE$ declines with N and T. The $RMSE$s are higher than in the case with the full rank. This evidence is in line with Pesaran (2006). On the other hand, the bias and the RMSE of the PC and FE estimators do not appear to be affected by the rank deficiency.

Table 4 presents the results for Experiment 2 with heterogeneous β's. We observe qualitatively similar results to those reported in Table 3 with homogeneous β's. Bias and $RMSE$ of the CCE estimator are larger than those in Table 2. The performance of the CCE estimator improves slowly with N only, suggesting that the rank deficiency may slow down the performance of the CCE estimator. Again, the performance of the PC and FE estimators is similar to the previous cases. Finally, we find that the mean group estimator performs slightly better than the pooled estimator in small samples.

[10] For a complete comparison we report the simulation results based on AIC_1 in the Online Supplement.

Table 5 shows the results for Experiment 3 with correlated loadings and full rank for homogeneous β's. Now, only the FE estimator is severely biased. Next, the biases of the CCE estimator are not negligible for the small N, but its performance improves sharply with N, a consistent finding with Westerlund and Urbain (2013), who note that 'the problem with correlated loadings goes away if the rank condition is satisfied'. The overall performance of the PC estimator is qualitatively similar to the previous cases, confirming that it is still consistent with both N and T. We also observe the qualitatively similar results documented in Table 6 for Experiment 3 with heterogeneous β's.

Table 7 presents the simulation results for Experiment 4 with correlated loadings and the rank deficiency for homogeneous βs. Both CCE and FE estimators are severely biased, confirming our theoretical prediction that both estimators are inconsistent in the presence of correlated factors loadings as also discussed in Sarafidis and Wansbeek (2012) and Westerlund and Urbain (2013). On the other hand, the performance of the PC estimators are still qualitatively similar to those presented in Table 3. In Table 8 we report the simulation results for Experiment 4 with heterogeneous βs, which provide qualitatively similar results to those in Table 7.

Overall, our results show that, when the factor loadings are uncorrelated all the four estimators examined above show a similar and satisfactory performance, suggesting that the FE estimator can produce reliable results even in the presence of interactive effects. When factor loadings are correlated, however, the FE estimator becomes severely biased and the performance of the CCE estimator tends to worsen. Only under the full rank condition, the performance of the CCE improves with N. The performance of the bias-corrected PC estimator is qualitatively similar across all four experiments.

However, there is an important caveat. The estimated number of factors can make a considerable difference in the performance of the PC estimator and this issue needs to be handled carefully, in practice. Crucially, previous literature does not provide clear evidence on what is the best course of action to choose the number of factors. In this regard, we propose a pretest estimator which is constructed as follows. The pretest estimator, denoted $\hat{\beta}_{pretest}$, selects either the FE or the PC estimator depending on the Hausman-type test results. To be more specific, we first evaluate the H^{NON} and H^{HAC}statistics. If the null hypothesis of uncorrelated factor loadings is not rejected, then we select $\hat{\beta}_{pretest} = \hat{\beta}_{FE}$ while, if the null is rejected, we set $\hat{\beta}_{pretest} = \hat{\beta}_{PCBai}$ or $\hat{\beta}_{PCHNY}$. In the Online Supplement we have examined the finite sample performance of this pretest estimator under the same four experiments considered above. Its overall performance is satisfactory in terms of bias and RMSE, irrespective of whether factor loadings are correlated or not. This suggests that such an estimator has considerable potential as it alleviates the issue of selecting the number of factors, especially in the case where factor loadings are found to be uncorrelated in practice, see the prevasive evidence in the empirical applications below.

4.4 The performance of the Hausman-type statistic

In this section we examine the small sample performance of the H test statistics under the above four experiments. To construct the H statistic, we consider the difference between the FE estimator, $\boldsymbol{\beta}_{FE}$ and the bias corrected PC estimators, denoted $\boldsymbol{\beta}_{PCBai}$ and $\boldsymbol{\beta}_{PCHNY}$, standardised respectively by both versions of robust variance estimator, denoted NON and HAC.[11]

We examine size and the power of the H statistic, but we also report the coverage rates for the three estimators. We consider the cases with homogeneous $\beta_i = \beta$, and with heterogeneous $\beta_i = \beta + \eta_i$ and $\eta_i \sim N(0, 0.04)$. Further, we consider serially correlated errors given by

$$\varepsilon_{it} = \rho_\varepsilon \varepsilon_{i,t-1} + v_{\varepsilon it} \text{ and } u_{it} = \rho_u u_{i,t-1} + v_{uit} \text{ with } \rho_\varepsilon = \rho_u = 0 \text{ or } 0.5.$$

Hence, we examine the following two cases:[12]

Case 1: Homogeneous β's and no serial correlation; see Tables 9 and 10 for H^{NON} and H^{HAC} test results.

Case 2: Heterogeneous β's and serial correlation; see Tables 11 and 12 for H^{NON} and H^{HAC} test results.

Overall, the test performance reported in Tables (9)-(12), displays qualitatively satisfactory and similar results in terms of the empirical size and power of the H statistics. This confirms that all the estimators are consistent under the null for both homogeneous and heterogeneous βs. Furthermore, the satisfactory coverage rates revealed by the three estimators demonstrate that both nonparametric and HAC variance estimators are robust to serial correlation as well as the slope heterogeneity.

In Experiments 1 and 2, the sizes of both $H^{(NON)}$ and $H^{(HAC)}$ tests approach the nominal level (0.05) in most cases as the sample size rises. The power of the H test is always one under Experiments 3 and 4. In particular, when factor loadings are uncorrelated, $\boldsymbol{\beta}_{FE}$ is shown to be consistent and its coverage rate reaches the nominal 95% in Experiments 1 and 2, irrespective of the rank condition. In Experiments 3 and 4 when loadings are correlated, however, $\boldsymbol{\beta}_{FE}$ is significantly biased and displays a zero coverage rate. The coverage rates of both bias-corrected PC estimators tend to 95% under all four experiments.[13]

Finally, following Goncalves and Perron (2014), we have also developed a parametric bootstrap test statistic. The simulation results for the bootstrap-based statistics are qualitatively similar to

[11] In what follows, we apply the bias corrected PC estimators using the IC_{p1} criterion. We have also investigated the performance of the H statistics using the uncorrected PC estimators, and obtained qualitatively similar results.

[12] We have also considered the cases with homogeneous β's and serial correlation and with heterogeneous β's and no serial correlation, and obtained qualitatively similar results, which are reported in the Online Supplement.

[13] We have also examined the coverage rates for both estimators using the analytic variance estimators described in Bai (2009, Section 9). We find that, when errors are serially correlated and conditionally heteroscedastic and/or β's are heterogenous, coverage rates are inconsistent for the FE estimator and they are mostly well below the nominal level for the PC estimator. The latter evidence is also reported in Chudik, Pesaran, and Tosetti (2011) and Sarafidis and Wansbeek (2012). This clearly demonstrates an importance of using the robust variance estimators for a reliable inference.

those for the asymptotic counterparts. See the Online Supplement for details.

5 Empirical Applications

In this section, we aim to investigate the empirical relevance of the factor loadings correlation by applying our proposed statistics, H^{NON} and H^{HAC} defined in (27) and (28) to three empirical questions and six different datasets, which have been considered in existing literature. The details of the data definitions and the empirical specifications are provided in Appendix 8.

The Cobb-Douglas production function For the first dataset comprising four different cases - namely OECD members,[14] the EU27 countries, 20 Italian regions and 48 States in the U.S., we estimate a production function by the FE and PC estimators and then evaluate our proposed test for the presence of factor loadings correlation. Following Solow (1956) in the economic growth literature, we employ the classic Cobb-Douglas specification, detailed in Appendix 8. For OECD and EU27, the output is measured by the per capita GDP while the regressor is the capital-labour ratio. For the Italian regions, output is measured by per capita value added and in the fourth application for the US, the output is measured by per capita gross State product with the same regressor.

The Gravity model of bilateral trade flows Next, we consider the estimation of a gravity model of the bilateral trade flows for the EU14 countries. The data covers the period from 1960 to 2008. Here, we follow Serlenga and Shin (2007) and estimate the gravity panel data regression, in which the bilateral trade flow is set as a function of GDP, countries' similarity, relative factor endowment, the real exchange rate as well as the trade union and common currency dummies.

The gasoline demand function The final application aims at estimating the price and income elasticities of gasoline demand. In particular, we focus on estimating the demand function for gasoline using the data from Liu (2014), which contains quarterly data for the 50 States in the U.S. over the period 1994 to 2008.

In Table (13), we present the estimation and test results.[15] First of all, the test results for both H^{NON} and H^{HAC}, provide a surprisingly convincing finding that the null of uncorrelated factor loadings is rejected (at levels of significance greater than 1%) only in one out of six cases considered. We also report the results for the CD test proposed by Pesaran (2015), which tests the null of no (weak) CSD against the alternative of strong CSD, and the Hausman test proposed by Bai (2009), denoted H^{Bai}, which tests the null of additive-effects against the alternative of

[14] The same OECD sample has been used in Mastromarco, Serlenga, and Shin (2016).

[15] Table (13) shows estimation and test results obtained using the HNY PC estimator, which is proved to be more robust than the Bai PC estimator, see HNY. Results using the Bai PC are shown in the Online Supplement.

interactive-effects. The CD test strongly rejects the null hypothesis for all the datasets whilst the H^{Bai} test never rejects the null hypothesis of additive-effects model. These test results are rather in conflict, since the former suggests the presence of interactive effects while the latter suggests no such effects. As highlighted in Section 2, however, the rejection of CD test does not always imply that the FE estimator is biased even in panels with interactive effects. Further, in Section 3, we note that the H^{Bai} test has no power against the alternative model with interactive effects, if the factor loadings are uncorrelated. Indeed, such conflicting results could provide support that the panels are subject to interactive effects but factor loadings are uncorrelated.

Next, turning to the slope estimates provided by both FE and PC estimators, we find that they are mostly significant. Further, their magnitudes and signs are relatively similar to each other, and consistent with theoretical predictions, only with an exception, the gravity model of international trade.[16] These casual observations are formally confirmed by the Hausman test results. These findings suggest that if factor loadings are found to be uncorrelated, then standard FE estimation can still produce consistent estimators even in the presence of unobserved interactive effects, and can be more robust as it avoids the complex issue of selecting the correct number of unobserved factors, which would significantly affect PC estimators.

6 Conclusions

Over the last decades a large strand of the literature on panel data has focused on analysing cross sectional dependence, based on the (unobserved) factors or interactive effects models, which are implicitly understood to bias conventional estimators such as the two-way FE estimator, due to the potential endogeneity arising from the correlation between regressors and factors.

Two main approaches have been advocated to deal with this issue: the Pesaran (2006) CCE estimator and the Bai (2009) PC estimator. In this paper we have shown that the interactive effects model can be encompassed by the standard error components model under a mild restriction on the factor loadings. In particular, when the factor loadings are uncorrelated in panels with interactive effects, the two-way FE is estimator still consistent. By avoiding the nontrivial issue of applying the complex bias-corrections in conjunction with the reliable information criteria correctly selecting the number of unobserved factors, the FE estimator would provide a simple and well-established estimation strategy in practice.

Via Monte Carlo studies, we find that the FE and CCE estimators display a similar and satisfactory performance when factor loadings satisfy an uncorrelatedness condition. As expected, the performance of both CCE and FE estimators worsens significantly under correlated factor loadings whilst the performance of the PC estimator is not unduly affected.

[16] Notice that the FE estimation tends to produce substantially large coefficient on GDP, which has been widely reported in the literature.

This suggests that the presence of correlated factor loadings emerges as an influential but under-appreciated feature of the panel data model with interactive effects. As the main contribution of this paper, we propose a test for the validity of correlated factor loadings, based on the Hausman principle. Further Monte Carlo simulation results confirm that the size and the power of the proposed test for correlated factor loadings is quite satisfactory even in small samples.

Finally, we consider a number of panel dataset to explore empirical evidence related to factor loadings restrictions. We find strong evidence in favor of uncorrelated factor loadings, suggesting that the standard two-way FE estimator should be an important part of the toolkit of applied researchers, worried about the presence of cross sectional dependence in their dataset.

7 Appendix: Proofs

7.1 Preliminary Lemmata

We first provide two important Lemmas that extend the Law of large Numbers and Central Limit Theorem to cover the martingale difference case for the panel data.

Lemma 4 *Let* $W_{i,T}$ *and* $\mu_{i,T}$ *for* $i = 1, ..., N$ *and* $T = 1, ..., T$, *be arrays of matrices of random variables and constants such that* $W_{i,T} - \mu_{i,T}$ *is a martingale difference array where* $\sup_{i,T} E \|W_{i,T}\|^{1+\delta} < \infty$ *for some* $\delta > 0$. *Then, as* $(N, T) \to_j \infty$,

$$N^{-1} \sum_{i=1}^{N} W_{i,T} - \mu_{i,T} \to_p 0.$$

Proof. Note that by Theorem 12.11 of Davidson (1994), if $\sup_{i,T} E \|W_{i,T}\|^{1+\delta} < \infty$, then $\lim_{M\to\infty} \sup_{i,T} E \left(\|W_{i,T} - \mu_{i,T}\| I_{\{\|W_{i,T}-\mu_{i,T}\|>M\}} \right) = 0$, which is a generalisation of uniform integrability to arrays. Then, the result follows immediately by Corollary 19.9 of Davidson (1994). ■

Lemma 5 *Let* $w_{i,T}$ *and* $\mu_{i,T}$, *for* $i = 1, ..., N$ *and* $T = 1, ..., T$, *be arrays of vectors of random variables and constants such that* $w_{i,T} - \mu_{i,T}$ *is a martingale difference array where* $E \left[(w_{i,T} - \mu_{i,T}) (w_{i,T} - \mu_{i,T})' \right] = \Sigma_{i,T}$, *and* $\sup_{i,T} E \|w_{iT}\|^{2+\delta} < \infty$ *for some* $\delta > 0$. *Assume that* $\Sigma = \lim_{N,T\to\infty} N^{-1} \sum_{i=1}^{N} \Sigma_{i,T}$ *is positive definite and* $\sup_{N,T} N^{-1} \sum_{i=1}^{N} \Sigma_{i,T} < \infty$. *Then, as* $(N, T) \to_j \infty$,

$$N^{-1} \sum_{i=1}^{N} w_{iT} - \mu_{iT} \to_d N(0, \Sigma). \tag{33}$$

Proof. By Theorem 12.11 of Davidson (1994), if $\sup_{i,T} E \|w_{i,T}\|^{2+\delta} < \infty$, we then obtain the uniform integrability condition,

$$\lim_{M\to\infty} \sup_{i,T} E \left(\|W_{i,T} - \mu_{i,T}\| I_{\{\|W_{i,T}-\mu_{i,T}\|>M\}} \right) = 0.$$

Together with $\sup_{N,T} N^{-1} \sum_{i=1}^{N} \Sigma_{i,T} < \infty$, this implies that the Lindeberg condition holds by Theorem 23.18 of Davidson (1994). Then, by Theorem 23.16 of Davidson (1994), it follows that

$$\max_{i,T} N^{-1} (w_{iT} - \mu_{iT}) \to_p 0. \tag{34}$$

Finally, together with $\sup_{i,T} E \|w_{i,T}\|^{2+\delta} < \infty$, (34) implies (33) by Theorem 24.3 of Davidson (1994). ■

Next, we provide the proofs for Theorems 1, 2 and 3.

7.2 Proof of Theorem 1

Consider first the homogeneous case with $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all i in which case we have:

$$\hat{\boldsymbol{\beta}}_{FE} - \boldsymbol{\beta} = \left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i \right)^{-1} \sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{u}}_i = \left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i \right)^{-1} \sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \left(\tilde{\boldsymbol{F}} \tilde{\boldsymbol{\gamma}}_i + \dot{\boldsymbol{\varepsilon}}_i \right)$$

$$= \left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i \right)^{-1} \sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \left(\tilde{\boldsymbol{F}} \tilde{\boldsymbol{\gamma}}_i + \boldsymbol{\varepsilon}_i \right) + o_p(1), \tag{35}$$

where we use

$$\dot{\varepsilon}_{it} = \varepsilon_{it} - \varepsilon_{i.} - \varepsilon_{.t} + \varepsilon_{..} = \varepsilon_{it} + o_p(1).$$

Next, in the heterogeneous case with $\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\eta}_i$, we have

$$\hat{\boldsymbol{\beta}}_{FE} - \boldsymbol{\beta} = \left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i \right)^{-1} \sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \left(\dot{\boldsymbol{X}}_i \boldsymbol{\eta}_i + \tilde{\boldsymbol{F}} \tilde{\boldsymbol{\gamma}}_i + \dot{\boldsymbol{\varepsilon}}_i \right) \tag{36}$$

$$= \left(\sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i \right)^{-1} \sum_{i=1}^{N} \dot{\boldsymbol{X}}_i' \left(\dot{\boldsymbol{X}}_i \boldsymbol{\eta}_i + \tilde{\boldsymbol{F}} \tilde{\boldsymbol{\gamma}}_i + \dot{\boldsymbol{\varepsilon}}_i \right) + o_p(1).$$

Noting that $\boldsymbol{x}_{it} = \boldsymbol{\Gamma}_i' \boldsymbol{f}_t + \boldsymbol{u}_{it}$, and using Lemma 1, it is easily seen that as $(N, T) \to_j \infty$,

$$\left(\frac{1}{N} \sum_{i=1}^{N} \frac{\dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i}{T} \right)^{-1} \to_p \lim_{N,T \to \infty} \frac{1}{N} \sum_{i=1}^{N} E\left(\frac{\dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i}{T} \right) = \boldsymbol{\Psi}_{FE} \equiv \bar{\boldsymbol{\Gamma}} \boldsymbol{\Sigma}_{\boldsymbol{f}} \bar{\boldsymbol{\Gamma}}' + \boldsymbol{\Sigma}_u,$$

where $\bar{\boldsymbol{\Gamma}} = E(\boldsymbol{\Gamma}_i)$. Next, by the independence of $\tilde{\boldsymbol{\gamma}}_i$ and $\boldsymbol{\eta}_i$ each other and from $\dot{\boldsymbol{X}}_i$ and $\tilde{\boldsymbol{F}}$ across i, and using the fact that $E(\tilde{\boldsymbol{\gamma}}_i) = E(\boldsymbol{\eta}_i) = 0$, it follows that $\dot{\boldsymbol{X}}_i' \left(\tilde{\boldsymbol{F}} \tilde{\boldsymbol{\gamma}}_i + \dot{\boldsymbol{X}}_i \boldsymbol{\eta}_i \right)$ is a martingale difference sequence for any ordering across i. To see this, notice that for any ordering over i, we have:

$$\begin{aligned}
& E\left[\dot{\boldsymbol{X}}_i' \left(\dot{\boldsymbol{X}}_i \boldsymbol{\eta}_i + \tilde{\boldsymbol{F}} \tilde{\boldsymbol{\gamma}}_i + \boldsymbol{\varepsilon}_i \right) | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] \\
& = E\left[\dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i \boldsymbol{\eta}_i | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] + E\left[\dot{\boldsymbol{X}}_i' \tilde{\boldsymbol{F}} \tilde{\boldsymbol{\gamma}}_i | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] + E\left[\dot{\boldsymbol{X}}_i' \boldsymbol{\varepsilon}_i | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] \\
& = E\left[\dot{\boldsymbol{X}}_i' \tilde{\boldsymbol{F}} | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] E\left[\tilde{\boldsymbol{\gamma}}_i | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] \\
& + E\left[\dot{\boldsymbol{X}}_i' \dot{\boldsymbol{X}}_i | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] E\left[\boldsymbol{\eta}_i | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] \\
& + E\left[\dot{\boldsymbol{X}}_i | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] E\left[\boldsymbol{\varepsilon}_i | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] \text{ for } j \neq i
\end{aligned}$$

But, since

$$E\left[\tilde{\boldsymbol{\gamma}}_i | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] = E\left[\boldsymbol{\eta}_i | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] = E\left[\boldsymbol{\varepsilon}_i | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] = 0,\ j \neq i$$

hence

$$E\left[\dot{\boldsymbol{X}}_i' \left(\dot{\boldsymbol{X}}_i \boldsymbol{\eta}_i + \tilde{\boldsymbol{F}} \tilde{\boldsymbol{\gamma}}_i + \boldsymbol{\varepsilon}_i \right) | \dot{\boldsymbol{X}}_j, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{\gamma}}_j, \boldsymbol{\eta}_j \right] = 0,\ j \neq i$$

which proves the martingale difference property. Notice in our proofs that we repeatedly use the simple fact that the product of a stochastic process with a second process, that is independent over its index as well as of the first process, is a martingale difference process. Next, by Lemma 2, we find that $\left\{\frac{\dot{\mathbf{X}}_i'\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\}_{i=1}^N$, $\left\{\frac{\dot{\mathbf{X}}_i'\tilde{\boldsymbol{F}}\tilde{\boldsymbol{\gamma}}_i}{T}\right\}_{i=1}^N$ and $\left\{\frac{\dot{\mathbf{X}}_i'(\tilde{\boldsymbol{F}}\tilde{\boldsymbol{\gamma}}_i+\dot{\mathbf{X}}_i\boldsymbol{\eta}_i)}{T}\right\}_{i=1}^N$ are martingale difference series. Notice that $\left(\sum_{i=1}^N \dot{\mathbf{X}}_i'\dot{\mathbf{X}}_i\right)^{-1}\sum_{i=1}^N \dot{\mathbf{X}}_i'\boldsymbol{\varepsilon}_i = O_p\left(\frac{1}{\sqrt{NT}}\right)$, is of the smaller probability order of magnitude than the other two terms in the RHS of (35). Therefore, it follows as $(N,T)\rightarrow_j \infty$,

$$\frac{1}{\sqrt{N}}\sum_{i=1}^N \frac{\dot{\mathbf{X}}_i'\tilde{\boldsymbol{F}}\tilde{\boldsymbol{\gamma}}_i}{T} \rightarrow_d N\left(0, \boldsymbol{R}_{1,FE}\right)$$

and

$$\frac{1}{\sqrt{N}}\sum_{i=1}^N \frac{\dot{\mathbf{X}}_i'\left(\tilde{\boldsymbol{F}}\tilde{\boldsymbol{\gamma}}_i+\dot{\mathbf{X}}_i\boldsymbol{\eta}_i\right)}{T} \rightarrow^d N\left(0, \boldsymbol{R}_{1,FE}+\boldsymbol{R}_{2,FE}\right).$$

where $\boldsymbol{R}_{1,FE}$ and $\boldsymbol{R}_{2,FE}$ are defined in (13) and (14). This proves (12) in Theorem 1. We will prove (15) in the proof of Theorem 3.

7.3 Proof of Theorem 2

By Theorems 4 and 5 in Hayakawa, Nagata, and Yamagata (2018), we have for $\boldsymbol{\beta}_i = \boldsymbol{\beta}$:

$$\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta} = \left(\sum_{i=1}^N \boldsymbol{V}_i'\boldsymbol{V}_i\right)^{-1}\sum_{i=1}^N \boldsymbol{V}_i'\boldsymbol{\varepsilon}_i + R_{NT},$$

while for $\boldsymbol{\beta}_i = \boldsymbol{\beta}+\boldsymbol{\eta}_i$,

$$\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta} = \left(\sum_{i=1}^N \boldsymbol{V}_i'\boldsymbol{V}_i\right)^{-1}\sum_{i=1}^N \boldsymbol{V}_i'\boldsymbol{V}_i\boldsymbol{\eta}_i + R_{NT},$$

where R_{NT} denotes terms of the lower order of probability than the leading terms in the RHS of the above equations. Note that $\sum_{i=1}^N \boldsymbol{V}_i'\boldsymbol{\varepsilon}_i = O_p\left(\sqrt{NT}\right)$ and $\sum_{i=1}^N \boldsymbol{V}_i'\boldsymbol{V}_i\boldsymbol{\eta}_i = O_p\left(\sqrt{N}T\right)$. Using Lemmas 1 and 2, it follows that as $(N,T)\rightarrow_j \infty$,

$$\left(\frac{1}{N}\sum_{i=1}^N \frac{\boldsymbol{V}_i'\boldsymbol{V}_i}{T}\right)^{-1} \rightarrow_p \lim_{N,T\rightarrow\infty}\frac{1}{N}\sum_{i=1}^N E\left(\frac{\boldsymbol{V}_i'\boldsymbol{V}_i}{T}\right) = \boldsymbol{\Psi}_{PC}$$

$$\frac{1}{\sqrt{N}}\sum_{i=1}^N \frac{\boldsymbol{V}_i'\boldsymbol{\varepsilon}_i}{\sqrt{T}} \rightarrow_d N\left(0, \boldsymbol{R}_{1,PC}\right)$$

and

$$\frac{1}{\sqrt{N}}\sum_{i=1}^N \frac{\boldsymbol{V}_i'\boldsymbol{V}_i\boldsymbol{\eta}_i}{T} \rightarrow_d N\left(0, \boldsymbol{R}_{2,PC}\right)$$

where $\boldsymbol{R}_{1,PC}$ and $\boldsymbol{R}_{2,PC}$ are defined in (23) and (25). These prove (22) and (24). (26) follows by the proof of Theorem 3.

7.4 Proof of Theorem 3

Given Theorems 1 and 2, it suffices to derive the equivalence and consistency of the two robust covariance estimators for $\hat{\boldsymbol{\beta}}_{FE}$ and $\hat{\boldsymbol{\beta}}_{PC}$, which are given by (10), (11), (20) and (21), respectively. Rewrite them compactly as

$$\boldsymbol{V}^{NON}\left(\hat{\boldsymbol{\beta}}_{FE}\right)=\hat{\boldsymbol{\Psi}}_{FE}^{-1}\hat{\boldsymbol{R}}_{FE}^{NON}\hat{\boldsymbol{\Psi}}_{FE}^{-1} \text{ and } \boldsymbol{V}^{HAC}\left(\hat{\boldsymbol{\beta}}_{FE}\right)=\hat{\boldsymbol{\Psi}}_{FE}^{-1}\hat{\boldsymbol{R}}_{FE}^{HAC}\hat{\boldsymbol{\Psi}}_{FE}^{-1}$$

and

$$\boldsymbol{V}^{NON}\left(\hat{\boldsymbol{\beta}}_{PC}\right)=\hat{\boldsymbol{\Psi}}_{PC}^{-1}\hat{\boldsymbol{R}}_{PC}^{NON}\hat{\boldsymbol{\Psi}}_{PC}^{-1} \text{ and } \boldsymbol{V}^{HAC}\left(\hat{\boldsymbol{\beta}}_{PC}\right)=\hat{\boldsymbol{\Psi}}_{PC}^{-1}\hat{\boldsymbol{R}}_{PC}^{HAC}\hat{\boldsymbol{\Psi}}_{PC}^{-1}$$

where

$$\begin{aligned}
\hat{\boldsymbol{\Psi}}_{FE} &= \sum_{i=1}^{N}\dot{\boldsymbol{X}}_i'\dot{\boldsymbol{X}}_i,\ \hat{\boldsymbol{\Psi}}_{PC}=\sum_{i=1}^{N}\hat{\boldsymbol{X}}_i'\hat{\boldsymbol{X}}_i=\sum_{i=1}^{N}\boldsymbol{X}_i'\boldsymbol{M}_{\hat{F}}\boldsymbol{X}_i\\
\hat{\boldsymbol{R}}_{FE}^{NON} &= \sum_{i}^{N}\left(\dot{\boldsymbol{X}}_i'\dot{\boldsymbol{X}}_i\right)\left(\hat{\boldsymbol{\beta}}_{FE,i}-\hat{\boldsymbol{\beta}}_{FE}\right)\left(\hat{\boldsymbol{\beta}}_{FE,i}-\hat{\boldsymbol{\beta}}_{FE}\right)'\left(\dot{\boldsymbol{X}}_i'\dot{\boldsymbol{X}}_i\right)\\
\hat{\boldsymbol{R}}_{PC}^{NON} &= \sum_{i}^{N}\hat{\boldsymbol{X}}_i'\hat{\boldsymbol{X}}_i\left(\hat{\boldsymbol{\beta}}_{PC,i}-\hat{\boldsymbol{\beta}}_{PC}\right)\left(\hat{\boldsymbol{\beta}}_{PC,i}-\hat{\boldsymbol{\beta}}_{PC}\right)'\hat{\boldsymbol{X}}_i'\hat{\boldsymbol{X}}_i\\
\hat{\boldsymbol{R}}_{FE}^{HAC} &= \sum_{i}^{N}\dot{\boldsymbol{X}}_i'\hat{\boldsymbol{u}}_{FE,i}\hat{\boldsymbol{u}}_{FE,i}'\dot{\boldsymbol{X}}_i,\ \hat{\boldsymbol{R}}_{PC}^{HAC}=\sum_{i}^{N}\hat{\boldsymbol{X}}_i'\hat{\boldsymbol{u}}_{PC,i}\hat{\boldsymbol{u}}_{PC,i}'\hat{\boldsymbol{X}}_i
\end{aligned}$$

Finally, we define

$$\hat{\boldsymbol{C}}^{NON}\left(\hat{\boldsymbol{\beta}}_{FE},\hat{\boldsymbol{\beta}}_{PC}\right)=\hat{\boldsymbol{\Psi}}_{FE}^{-1}\hat{\boldsymbol{R}}_{FE,PC}^{NON}\hat{\boldsymbol{\Psi}}_{PC}^{-1} \text{ and } \hat{\boldsymbol{C}}^{HAC}\left(\hat{\boldsymbol{\beta}}_{FE},\hat{\boldsymbol{\beta}}_{PC}\right)=\hat{\boldsymbol{\Psi}}_{FE}^{-1}\hat{\boldsymbol{R}}_{FE,PC}^{HAC}\hat{\boldsymbol{\Psi}}_{PC}^{-1}$$

where

$$\begin{aligned}
\hat{\boldsymbol{R}}_{FE,PC}^{NON} &= \sum_{i}^{N}\dot{\boldsymbol{X}}_i'\dot{\boldsymbol{X}}_i\left(\hat{\boldsymbol{\beta}}_{FE,i}-\hat{\boldsymbol{\beta}}_{FE}\right)\left(\hat{\boldsymbol{\beta}}_{PC,i}-\hat{\boldsymbol{\beta}}_{PC}\right)'\hat{\boldsymbol{X}}_i'\hat{\boldsymbol{X}}_i,\\
\hat{\boldsymbol{R}}_{FE,PC}^{HAC} &= \sum_{i}^{N}\dot{\boldsymbol{X}}_i'\hat{\boldsymbol{u}}_{FE,i}\hat{\boldsymbol{u}}_{PC,i}'\hat{\boldsymbol{X}}_i
\end{aligned}$$

To establish that the two covariance estimators are (asymptotically) equivalent, we need to show that

$$\hat{\boldsymbol{R}}_{FE}^{NON}=\hat{\boldsymbol{R}}_{FE}^{HAC}+R_{NT} \tag{37}$$

$$\hat{\boldsymbol{R}}_{PC}^{NON}=\hat{\boldsymbol{R}}_{PC}^{HAC}+R_{NT} \tag{38}$$

$$\hat{\boldsymbol{R}}_{FE,PC}^{NON}=\hat{\boldsymbol{R}}_{FE,PC}^{HAC}+R_{NT} \tag{39}$$

where R_{NT} denotes terms of the lower order of probability than the leading terms in the RHS of the above equations. We focus on the PC estimator in (38). First, consider $\hat{\boldsymbol{R}}_{PC}^{HAC}$ and notice that

$$\hat{\boldsymbol{X}}_i'\hat{\boldsymbol{u}}_{PC,i}=\hat{\boldsymbol{X}}_i'\left(\boldsymbol{u}_{PC,i}+\hat{\boldsymbol{X}}_i\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\right).$$

By Theorem 6 of Hayakawa, Nagata, and Yamagata (2018), it follows that, as $\lim_{N,T\to\infty} \frac{T}{N} \to c \in (0, \Delta]$ with $\Delta < \infty$,

$$\sum_{i=1}^{N} \hat{\boldsymbol{X}}_i' \hat{\boldsymbol{u}}_{PC,i} \hat{\boldsymbol{u}}_{PC,i}' \hat{\boldsymbol{X}}_i = \sum_{i=1}^{N} \boldsymbol{V}_i' \boldsymbol{u}_{PC,i} \boldsymbol{u}_{PC,i}' \boldsymbol{V}_i + R_{NT}$$

Using $\boldsymbol{u}_{PC,i} = \boldsymbol{X}_i \boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i$, we have

$$\hat{\boldsymbol{R}}_{FE}^{HAC} = \sum_{i=1}^{N} \boldsymbol{V}_i' \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i' \boldsymbol{V}_i + \sum_{i=1}^{N} \boldsymbol{V}_i' \boldsymbol{V}_i \boldsymbol{\eta}_i \boldsymbol{\eta}_i' \boldsymbol{V}_i' \boldsymbol{V}_i + R_{NT}. \tag{40}$$

Next, it is easily seen that

$$\hat{\boldsymbol{\beta}}_{PC,i} - \boldsymbol{\beta} = \left(\hat{\boldsymbol{X}}_i' \hat{\boldsymbol{X}}_i \right)^{-1} \hat{\boldsymbol{X}}_i' \boldsymbol{\varepsilon}_i + \boldsymbol{\eta}_i$$

and

$$\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta} = \frac{1}{N} \sum_{i=1}^{N} \left[\left(\hat{\boldsymbol{X}}_i' \hat{\boldsymbol{X}}_i \right)^{-1} \hat{\boldsymbol{X}}_i' \boldsymbol{\varepsilon}_i + \boldsymbol{\eta}_i \right]$$

Then,

$$\begin{aligned}
\hat{\boldsymbol{X}}_i' \hat{\boldsymbol{X}}_i \left(\hat{\boldsymbol{\beta}}_{PC,i} - \hat{\boldsymbol{\beta}}_{PC} \right) &= \hat{\boldsymbol{X}}_i' \hat{\boldsymbol{X}}_i \left(\hat{\boldsymbol{\beta}}_{PC,i} - \boldsymbol{\beta}_i + \boldsymbol{\beta}_i - \boldsymbol{\beta} + \boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{PC} \right) \\
&= \hat{\boldsymbol{X}}_i' \hat{\boldsymbol{X}}_i \left(\hat{\boldsymbol{\beta}}_{PC,i} - \boldsymbol{\beta}_i \right) + \hat{\boldsymbol{X}}_i' \hat{\boldsymbol{X}}_i \boldsymbol{\eta}_i + \hat{\boldsymbol{X}}_i' \hat{\boldsymbol{X}}_i \left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{PC} \right) \\
&= \hat{\boldsymbol{X}}_i' \boldsymbol{\varepsilon}_i + \hat{\boldsymbol{X}}_i' \hat{\boldsymbol{X}}_i \boldsymbol{\eta}_i + \hat{\boldsymbol{X}}_i' \hat{\boldsymbol{X}}_i \left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{PC} \right)
\end{aligned}$$

By Theorem 6 of Hayakawa, Nagata, and Yamagata (2018), we obtain:

$$\hat{\boldsymbol{R}}_{PC}^{NON} = \sum_{i=1}^{N} \boldsymbol{V}_i' \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i' \boldsymbol{V}_i + \sum_{i=1}^{N} \boldsymbol{V}_i' \boldsymbol{V}_i \boldsymbol{\eta}_i \boldsymbol{\eta}_i' \boldsymbol{V}_i' \boldsymbol{V}_i + R_{N,T}. \tag{41}$$

This proves (38). Noticing that both $\frac{\boldsymbol{V}_i' \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i' \boldsymbol{V}_i}{T} - E\left(\frac{\boldsymbol{V}_i' \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i' \boldsymbol{V}_i}{T} \right)$ and $\frac{\boldsymbol{V}_i' \boldsymbol{V}_i}{T} \boldsymbol{\eta}_i \boldsymbol{\eta}_i' \frac{\boldsymbol{V}_i' \boldsymbol{V}_i}{T} - E\left(\frac{\boldsymbol{V}_i' \boldsymbol{V}_i}{T} \boldsymbol{\eta}_i \boldsymbol{\eta}_i' \frac{\boldsymbol{V}_i' \boldsymbol{V}_i}{T} \right)$ are *iid* and martingale difference processes over i and by Lemma 1, we have:

$$\frac{1}{N} \sum_{i=1}^{N} \frac{\boldsymbol{V}_i' \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i' \boldsymbol{V}_i}{T} \to_p \lim_{N,T\to\infty} \frac{1}{N} \sum_{i=1}^{N} E\left(\frac{\boldsymbol{V}_i' \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i' \boldsymbol{V}_i}{T} \right)$$

and

$$\frac{1}{N} \sum_{i=1}^{N} \frac{\boldsymbol{V}_i' \boldsymbol{V}_i}{T} \boldsymbol{\eta}_i \boldsymbol{\eta}_i' \frac{\boldsymbol{V}_i' \boldsymbol{V}_i}{T} \to_p \lim_{NT,\to\infty} \frac{1}{N} \sum_{i=1}^{N} E\left(\frac{\boldsymbol{V}_i' \boldsymbol{V}_i}{T} \boldsymbol{\eta}_i \boldsymbol{\eta}_i' \frac{\boldsymbol{V}_i' \boldsymbol{V}_i}{T} \right)$$

This provides consistency of both variance estimators. Along similar lines of derivation, it is straightforward to prove (37) and (39) Using the above results, it readily follows that

$$\left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{PC} \right)' \left(\hat{\boldsymbol{V}}^{NON} \right)^{-1} \left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{PC} \right) \sim \chi_k^2$$

and

$$\left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{PC} \right)' \left(\hat{\boldsymbol{V}}^{HAC} \right)^{-1} \left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{PC} \right) \sim \chi_k^2$$

irrespective of whether $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ or $\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\eta}_i$, where $\hat{\boldsymbol{V}}^{NON}$ and $\hat{\boldsymbol{V}}^{HAC}$ are defined in (29) and (30), respectively.

8 Appendix: The Data and Empirical Specifications

We describe the empirical specifications and the data in details. For the production function, we estimate the following panel data regression:

$$\ln\left(\frac{Y}{L}\right)_{it} = \beta \ln\left(\frac{K}{L}\right)_{it} + e_{it},\ e_{it} = \alpha_i + \boldsymbol{\gamma}_i'\boldsymbol{f}_t + \varepsilon_{it} \tag{42}$$

The first group consists of 26 OECD countries; Australia, Austria, Belgium, Canada, Chile, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Israel, Italy, Japan, Korea, Mexico, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Turkey, the U.K. and the U.S. The data is collected from PWT 7.0 and covers the period 1970-2010. Y is GDP measured in million U.S. \$ at the 2005 price, K the capital measured in millions U.S. \$, constructed using the perpetual inventory method (PIM), and L the labour measured as the total employment in thousands. The second group contains the EU27 countries; Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and the U.K. The data are extracted from PWT 9.0 over the period 1990-2015 and the definition of the variables, Y, K and L is the same as above. The third group includes 20 Italian regions over the period 1995-2016; Piemonte, Valle d'Aosta, Liguria, Lombardia, Trentino Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna, Toscana, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia and Sardegna. Due to the data availability, we construct Y by the value added measured in million Euros at the 2010 price, L by the total employment in thousands, and K by Gross Fixed Capital Formation in millions Euros. The data covers the period 1995 to 2000, gathered from ISTAT. The fourth group comprises 48 U.S. states. The data are taken from Munnell (1990), covering the period, 1970-1986. In this case, Y is the per capita gross state product, K is the private capital computed by apportioning Bureau of Economic Analysis (BEA) national stock estimates, and L is the number of employers in thousands in non-agricultural payrolls.

Next, we consider the gravity model specifications for the bilateral trade flows given by

$$\begin{aligned} \ln(trade_{it}) &= \beta_{gdp}\ln(gdp_{it}) + \beta_{rer}\ln(rer_{it}) + \beta_{sim}\ln(sim_{it}) + \beta_{rlf}\ln(rlf_{it}) \\ &+\beta_{cee}cee_{it} + \beta_{euro}euro_{it} + e_{it}, e_{it} = \alpha_i + \boldsymbol{\gamma}_i'\boldsymbol{f}_t + \varepsilon_{it} \end{aligned} \tag{43}$$

Here, $trade_{it}$ is the sum of bilateral import flows ($import_{odt}$) and export flows ($\exp ort_{odt}$) measured in million U.S. dollars at the 2000 price with o and d denoting the origin and the destination country, gdp_{it} is the sum of gdp_{ot} and gdp_{dt} both of which are measured as the gross domestic product at the 2000 dollar price, $rer_{it} = ner_{odt} \times xpi_{US}$ is the real exchange rate measured in the 2000 dollar price, where ner_{hft} is the bilateral nominal exchange rate normalised in terms of the

U.S. \$, sim is a measure of similarity in size constructed by

$$sim_{it} = \left[1 - \left(\frac{gdp_{ot}}{gdp_{ot} + gdp_{dt}} \right)^2 - \left(\frac{gdp_{dt}}{gdp_{ot} + gdp_{dt}} \right)^2 \right]$$

and $rlf_{it} = |pgdp_{ot} - pgdp_{dt}|$ measures countries' difference in relative factor endowment where $pgdp$ is per capita GDP. cee and $euro$ represent dummies equal to one when countries of origin and destination both belong to the European Economic Community and share the euro as common currency, respectively. The data are collected from the IMF Direction of Trade Statistics, and covers the period, 1960-2008. We consider a sample of 91 country-pairs amongst the EU14 member countries (Austria, Belgium-Luxemburg, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Portugal, Spain, Sweden and the U.K.).

Finally, we estimate the gasoline demand function by

$$\ln(q_{it}) = \beta_p \ln(p_{it}) + \beta_{inc} \ln(inc_{it}) + e_{it}, \; e_{it} = \alpha_i + \boldsymbol{\gamma}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{44}$$

where gasoline consumption, q_{it}, is approximated as monthly sales volumes of motor gasoline, per capita per day; p_{it} is the after tax gasoline prices computed by adding the state/federal tax rates to the motor gasoline sales to end user price and q_{it} represent the quarterly personal disposable income. Prices, income, and tax rates are converted to constant 2005 dollars using GDP implicit price deflators. The source of data is Liu (2014).

9 Appendix: The Bias Corrected PC Estimator

The bias corrected estimator proposed by Hayakawa, Nagata, and Yamagata (2018) is given by

$$\hat{\boldsymbol{\beta}}_{PCHNY} = \bar{\boldsymbol{\beta}}_{PC} - \frac{1}{N}\hat{\boldsymbol{c}}$$

where the factors are estimated by the eigenvectors corresponding to the largest r largest eigenvalues of the $T \times T$ matrix $\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{Z}_i \boldsymbol{Z}_i'$ being $\boldsymbol{Z}_i = (\boldsymbol{y}_i, \boldsymbol{X}_i)$, and the bias correction terms are given by

$$\hat{\boldsymbol{c}} = \left(\frac{1}{NT} \sum_{i=1}^{N} \hat{\boldsymbol{X}}_i' \hat{\boldsymbol{X}}_i \right)^{-1} \hat{\boldsymbol{\xi}}$$

where $\hat{\boldsymbol{X}}_i = \boldsymbol{M}_{\hat{F}} \boldsymbol{X}_i$, $\boldsymbol{M}_{\hat{F}} = \boldsymbol{I} - \hat{\boldsymbol{F}} \left(\hat{\boldsymbol{F}}' \hat{\boldsymbol{F}} \right)^{-1} \hat{\boldsymbol{F}}'$, and

$$\hat{\boldsymbol{\xi}} = -\frac{1}{N} \sum_{i=1}^{N} \hat{\boldsymbol{\Gamma}}_i \hat{\boldsymbol{\Upsilon}}_N^{-1} \hat{\boldsymbol{g}}_{1i} \hat{\sigma}_i^2 + \frac{1}{N} \sum_{i=1}^{N} \hat{\boldsymbol{\Gamma}}_i \hat{\boldsymbol{\Upsilon}}_N^{-1} \left(\frac{1}{N} \sum_{j=1}^{N} \hat{\boldsymbol{G}}_j \hat{\boldsymbol{\Omega}}_{EEj} \hat{\boldsymbol{G}}_j' \right) \hat{\boldsymbol{\Upsilon}}_N^{-1} \hat{\lambda}_i$$
$$- \frac{1}{N} \sum_{j=1}^{N} \hat{\boldsymbol{\Omega}}_{VEj} \hat{\boldsymbol{G}}_j' \hat{\boldsymbol{\Upsilon}}_N^{-1} \hat{\lambda}_i$$

with

$$\hat{\boldsymbol{\Gamma}}_i' = \frac{\hat{\boldsymbol{F}}'\boldsymbol{X}_i}{T};\ \hat{\boldsymbol{\Upsilon}}_N = \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{G}}_i\hat{\boldsymbol{G}}_i';\ \hat{\boldsymbol{G}}_i = \frac{\hat{\boldsymbol{F}}'\boldsymbol{Z}_i}{T};\ \hat{\boldsymbol{g}}_{1i} = \frac{\hat{\boldsymbol{F}}'\boldsymbol{y}_i}{T};\ \widehat{\sigma}_i^2 = \frac{\hat{\boldsymbol{u}}_i'\boldsymbol{M}_{\hat{F}}\hat{\boldsymbol{u}}_i}{T};\ \hat{\boldsymbol{u}}_i = \boldsymbol{y}_i - \boldsymbol{X}_i\,\hat{\boldsymbol{\beta}}_{PC}$$

$$\hat{\boldsymbol{\Omega}}_{EE,i} = \frac{\hat{\boldsymbol{E}}_i'\hat{\boldsymbol{E}}_i}{T};\ \hat{\boldsymbol{E}}_i = \boldsymbol{M}_{\hat{F}}\boldsymbol{Z}_i;\ \hat{\lambda}_i = \frac{\hat{\boldsymbol{F}}'\boldsymbol{u}_i}{T};\ \hat{\boldsymbol{\Omega}}_{VE,i} = \frac{\hat{\boldsymbol{V}}_i'\hat{\boldsymbol{E}}_i}{T};\ \hat{\boldsymbol{V}}_i = \boldsymbol{M}_{\hat{F}}\boldsymbol{X}_i$$

Tables

Table 1: Simulation results for Experiment 1 with uncorrelated factor loadings and the full rank for homogeneous β=1

T/N	20	30	50	100	200	20	30	50	100	200
			β_{CCEP}					β_{CCEMG}		
					Bias					
20	0.0011	-0.0013	0.0003	-0.0007	-0.0007	0.0018	-0.0014	0.0009	-0.0009	-0.0009
30	-0.0015	0.0009	-0.0007	-0.0004	0.0003	-0.0013	0.0003	-0.0007	-0.0003	0.0005
50	-0.0001	-0.0017	0.0003	0.0009	0.0007	-0.0003	-0.0018	0.0002	0.0008	0.0007
100	-0.0011	0.0000	-0.0004	-0.0002	-0.0002	-0.0010	-0.0001	-0.0004	-0.0002	-0.0002
200	-0.0005	-0.0003	-0.0003	-0.0002	-0.0002	-0.0003	-0.0003	-0.0003	-0.0002	-0.0002
					RMSE					
20	0.0578	0.0473	0.0351	0.0236	0.0163	0.0612	0.0494	0.0377	0.0252	0.0171
30	0.0492	0.0374	0.0274	0.0190	0.0134	0.0501	0.0389	0.0285	0.0195	0.0140
50	0.0391	0.0301	0.0213	0.0148	0.0101	0.0387	0.0307	0.0217	0.0152	0.0103
100	0.0308	0.0213	0.0154	0.0109	0.0074	0.0295	0.0212	0.0155	0.0110	0.0075
200	0.0261	0.0161	0.0114	0.0071	0.0049	0.0239	0.0158	0.0114	0.0072	0.0050
			β_{FEP}					β_{FEMG}		
					Bias					
20	0.0028	-0.0003	-0.0001	-0.0015	-0.0004	0.0039	-0.0015	0.0011	-0.0010	0.0000
30	-0.0012	-0.0028	-0.0002	-0.0011	-0.0004	0.0025	0.0001	0.0006	-0.0011	0.0000
50	0.0047	0.0019	0.0017	0.0000	0.0016	0.0032	0.0016	0.0008	-0.0005	0.0012
100	-0.0068	0.0017	-0.0007	-0.0002	-0.0016	-0.0066	0.0004	-0.0019	0.0003	-0.0004
200	-0.0002	-0.0010	-0.0007	-0.0021	-0.0001	-0.0009	-0.0006	-0.0004	-0.0018	-0.0003
					RMSE					
20	0.1224	0.0906	0.0752	0.0507	0.0355	0.1167	0.0914	0.0730	0.0509	0.0355
30	0.1124	0.0901	0.0723	0.0493	0.0345	0.1066	0.0868	0.0687	0.0486	0.0333
50	0.1109	0.0914	0.0713	0.0486	0.0342	0.1064	0.0881	0.0677	0.0452	0.0327
100	0.1111	0.0908	0.0681	0.0509	0.0325	0.1010	0.0848	0.0639	0.0474	0.0316
200	0.1097	0.0849	0.0694	0.0470	0.0350	0.1012	0.0789	0.0639	0.0439	0.0330
			β_{PCPBai}							
					Bias					
20	-0.0041	-0.0021	0.0002	-0.0006	0.0005					
30	0.0010	0.0007	0.0009	0.0009	-0.0008					
50	0.0016	0.0003	-0.0009	0.0009	0.0000					
100	0.0001	0.0004	0.0000	0.0001	-0.0004					
200	-0.0002	0.0005	0.0001	0.0001	-0.0002					
					RMSE					
20	0.0595	0.0477	0.0349	0.0244	0.0171					
30	0.0476	0.0369	0.0289	0.0199	0.0140					
50	0.0371	0.0278	0.0209	0.0145	0.0102					
100	0.0250	0.0201	0.0147	0.0104	0.0070					
200	0.0175	0.0137	0.0104	0.0074	0.0051					
			β_{PCPHNY}					$\beta_{PCMGHNY}$		
					Bias					
20	-0.0027	-0.0017	0.0005	-0.0002	0.0006	-0.0018	-0.0012	0.0003	-0.0005	0.0005
30	0.0011	0.0014	0.0008	0.0010	-0.0007	0.0014	0.0019	0.0007	0.0011	-0.0006
50	0.0004	0.0008	-0.0013	0.0007	0.0000	0.0003	0.0005	-0.0013	0.0006	0.0001
100	0.0005	0.0002	0.0000	0.0000	-0.0004	0.0006	0.0001	0.0000	0.0000	-0.0004
200	-0.0002	0.0002	0.0000	0.0001	-0.0002	-0.0003	0.0002	0.0000	0.0001	-0.0002
					RMSE					
20	0.0604	0.0482	0.0355	0.0242	0.0169	0.0638	0.0511	0.0381	0.0255	0.0176
30	0.0477	0.0377	0.0287	0.0199	0.0139	0.0498	0.0394	0.0295	0.0206	0.0144
50	0.0362	0.0280	0.0213	0.0146	0.0101	0.0376	0.0283	0.0219	0.0148	0.0103
100	0.0254	0.0203	0.0149	0.0103	0.0069	0.0258	0.0204	0.0151	0.0104	0.0070
200	0.0179	0.0135	0.0106	0.0075	0.0052	0.0181	0.0136	0.0107	0.0075	0.0052

Notes: CCE and CCEMG are the pooled and mean group common correlated estimators by Pesaran (2006); FEP and FEMG denote the pooled and mean group two-way fixed effects estimators. PCPBai is the iterative pooled principal component estimator by Bai (2009) whilst PCPHNY and PCMGHNY stand for the pooled and mean group principal component estimators proposed by Hayakawa, Nagata, and Yamagata (2018) The PC estimators are bias-corrected and evaluated using the IC_{p1} criterion by Bai and Ng (2002).

Table 2: Simulation results for Experiment 1 with uncorrelated factor loadings and the full rank for heterogeneous $\beta_i = 1 + \eta_i$, $\eta_i \sim iidN(0, 0.04)$

T/N	20	30	50	100	200	20	30	50	100	200
			β_{CCEP}					β_{CCEMG}		
					Bias					
20	-0.0038	-0.0013	0.0015	-0.0001	-0.0003	-0.0041	-0.0026	0.0023	0.0000	-0.0002
30	-0.0037	0.0019	0.0021	-0.0003	-0.0007	-0.0027	0.0017	0.0017	-0.0008	-0.0008
50	0.0010	0.0008	0.0001	0.0014	0.0007	0.0006	0.0016	0.0002	0.0015	0.0006
100	0.0019	0.0002	-0.0008	-0.0004	0.0005	0.0015	0.0003	-0.0007	-0.0004	0.0005
200	-0.0011	0.0008	-0.0007	-0.0003	0.0003	-0.0006	0.0007	-0.0006	-0.0005	0.0003
					RMSE					
20	0.0730	0.0609	0.0444	0.0319	0.0229	0.0741	0.0613	0.0454	0.0322	0.0234
30	0.0670	0.0546	0.0407	0.0281	0.0195	0.0660	0.0546	0.0410	0.0282	0.0194
50	0.0598	0.0461	0.0347	0.0255	0.0181	0.0582	0.0458	0.0343	0.0254	0.0179
100	0.0538	0.0416	0.0311	0.0224	0.0163	0.0525	0.0407	0.0312	0.0222	0.0162
200	0.0545	0.0401	0.0294	0.0213	0.0151	0.0529	0.0397	0.0293	0.0212	0.0150
			β_{FEP}					β_{FEMG}		
					Bias					
20	0.0008	0.0042	0.0006	0.0012	-0.0004	-0.0012	0.0002	-0.0004	0.0014	-0.0016
30	-0.0073	-0.0022	0.0010	0.0009	0.0004	-0.0017	-0.0023	0.0034	-0.0007	0.0004
50	-0.0011	-0.0001	-0.0017	0.0006	0.0029	0.0022	0.0053	-0.0002	0.0003	0.0031
100	0.0052	-0.0001	-0.0021	-0.0004	0.0015	0.0073	0.0010	-0.0017	-0.0007	0.0001
200	-0.0037	0.0010	0.0011	0.0029	-0.0003	-0.0014	0.0009	-0.0001	0.0022	0.0005
					RMSE					
20	0.1285	0.1041	0.0849	0.0578	0.0418	0.1230	0.1035	0.0808	0.0558	0.0398
30	0.1253	0.0995	0.0816	0.0579	0.0405	0.1198	0.0992	0.0760	0.0538	0.0387
50	0.1251	0.1001	0.0798	0.0559	0.0390	0.1152	0.0938	0.0749	0.0530	0.0369
100	0.1230	0.0968	0.0790	0.0544	0.0380	0.1114	0.0887	0.0715	0.0489	0.0354
200	0.1255	0.0996	0.0778	0.0533	0.0378	0.1135	0.0913	0.0717	0.0485	0.0339
			β_{PCPBai}							
					Bias					
20	0.0029	-0.0024	-0.0002	-0.0008	0.0000					
30	0.0045	-0.0033	-0.0023	-0.0004	0.0004					
50	-0.0012	0.0023	0.0002	0.0017	0.0000					
100	-0.0019	-0.0018	-0.0005	-0.0008	-0.0002					
200	-0.0006	0.0019	-0.0006	-0.0006	-0.0004					
					RMSE					
20	0.0773	0.0577	0.0476	0.0327	0.0229					
30	0.0685	0.0541	0.0405	0.0277	0.0193					
50	0.0604	0.0469	0.0346	0.0247	0.0174					
100	0.0515	0.0433	0.0320	0.0229	0.0162					
200	0.0491	0.0388	0.0306	0.0210	0.0146					
			β_{PCPHNY}					$\beta_{PCMGHNY}$		
					Bias					
20	0.0016	-0.0018	0.0002	-0.0005	-0.0001	0.0000	-0.0008	0.0004	-0.0008	-0.0005
30	0.0042	-0.0023	-0.0020	-0.0004	0.0004	0.0028	-0.0019	-0.0020	0.0001	0.0002
50	-0.0011	0.0021	0.0002	0.0015	0.0001	-0.0008	0.0024	0.0005	0.0015	0.0002
100	-0.0020	-0.0013	-0.0007	-0.0008	-0.0002	-0.0019	-0.0010	-0.0005	-0.0010	-0.0003
200	-0.0002	0.0021	-0.0005	-0.0007	-0.0004	-0.0004	0.0021	-0.0005	-0.0007	-0.0004
					RMSE					
20	0.0767	0.0583	0.0468	0.0327	0.0229	0.0779	0.0599	0.0478	0.0332	0.0231
30	0.0675	0.0548	0.0402	0.0280	0.0194	0.0671	0.0548	0.0403	0.0283	0.0195
50	0.0599	0.0468	0.0346	0.0248	0.0173	0.0591	0.0467	0.0346	0.0244	0.0172
100	0.0518	0.0432	0.0320	0.0228	0.0162	0.0512	0.0428	0.0320	0.0229	0.0161
200	0.0487	0.0390	0.0308	0.0209	0.0146	0.0487	0.0388	0.0307	0.0208	0.0145

Notes: See notes to Table 1.

Table 3: Simulation results for Experiment 2 with uncorrelated factor loadings and the rank deficiency for homogeneous β=1

T/N	20	30	50	100	200	20	30	50	100	200
			β_{CCEP}					β_{CCEMG}		
					Bias					
20	-0.0019	0.0014	0.0049	0.0016	0.0004	0.0010	0.0021	0.0033	0.0002	0.0010
30	0.0015	-0.0012	0.0010	-0.0008	-0.0004	0.0023	-0.0010	0.0012	-0.0001	-0.0004
50	0.0019	0.0019	0.0018	0.0012	0.0012	0.0014	0.0007	0.0002	-0.0002	0.0010
100	0.0059	0.0032	0.0008	0.0006	0.0011	0.0033	0.0034	0.0015	0.0010	0.0011
200	0.0052	0.0012	-0.0011	-0.0001	0.0005	0.0039	0.0015	0.0002	0.0001	0.0001
					RMSE					
20	0.1067	0.0855	0.0612	0.0457	0.0327	0.0982	0.0796	0.0571	0.0417	0.0297
30	0.0986	0.0809	0.0619	0.0419	0.0322	0.0861	0.0700	0.0532	0.0379	0.0284
50	0.0967	0.0766	0.0574	0.0420	0.0288	0.0820	0.0623	0.0482	0.0356	0.0243
100	0.0939	0.0791	0.0545	0.0421	0.0293	0.0768	0.0653	0.0445	0.0340	0.0235
200	0.0902	0.0722	0.0583	0.0383	0.0295	0.0725	0.0559	0.0460	0.0314	0.0233
			β_{FEP}					β_{FEMG}		
					Bias					
20	-0.0032	-0.0005	0.0018	0.0016	0.0009	0.0010	0.0016	0.0020	0.0018	0.0013
30	-0.0006	-0.0005	-0.0001	-0.0005	-0.0002	-0.0003	-0.0006	0.0009	-0.0006	0.0000
50	0.0003	-0.0020	0.0020	0.0014	0.0006	0.0001	-0.0021	0.0015	0.0005	0.0011
100	-0.0002	0.0049	-0.0005	0.0006	0.0001	-0.0007	0.0056	0.0002	0.0016	0.0003
200	0.0069	0.0053	-0.0014	-0.0023	0.0007	0.0053	0.0047	-0.0010	-0.0019	0.0000
					RMSE					
20	0.1235	0.0943	0.0737	0.0518	0.0366	0.1237	0.0926	0.0728	0.0505	0.0362
30	0.1169	0.0954	0.0715	0.0513	0.0355	0.1124	0.0888	0.0663	0.0500	0.0346
50	0.1160	0.0883	0.0674	0.0504	0.0331	0.1065	0.0833	0.0657	0.0488	0.0318
100	0.1112	0.0939	0.0679	0.0500	0.0347	0.1030	0.0866	0.0649	0.0469	0.0320
200	0.1099	0.0856	0.0700	0.0499	0.0345	0.1005	0.0806	0.0653	0.0459	0.0327
			β_{PCPBai}							
					Bias					
20	-0.0021	-0.0026	0.0020	0.0010	-0.0005					
30	-0.0009	-0.0004	-0.0006	0.0001	0.0004					
50	0.0010	-0.0014	0.0001	0.0004	-0.0002					
100	0.0014	0.0005	0.0007	0.0003	0.0001					
200	0.0009	-0.0002	-0.0003	0.0001	0.0001					
					RMSE					
20	0.0607	0.0464	0.0345	0.0240	0.0171					
30	0.0484	0.0377	0.0289	0.0196	0.0141					
50	0.0344	0.0281	0.0211	0.0153	0.0105					
100	0.0263	0.0198	0.0150	0.0103	0.0073					
200	0.0177	0.0133	0.0108	0.0074	0.0052					
			β_{PCPHNY}					$\beta_{PCMGHNY}$		
					Bias					
20	-0.0029	-0.0016	0.0020	0.0011	-0.0005	-0.0038	-0.0011	0.0021	0.0015	-0.0004
30	-0.0003	0.0002	-0.0007	0.0001	0.0005	-0.0005	0.0003	-0.0008	0.0001	0.0004
50	0.0007	-0.0010	0.0002	0.0005	-0.0002	0.0007	-0.0010	0.0000	0.0005	-0.0002
100	0.0016	0.0002	0.0006	0.0003	0.0001	0.0017	0.0003	0.0006	0.0003	0.0001
200	0.0002	-0.0002	-0.0005	0.0001	0.0001	0.0002	-0.0003	-0.0005	0.0000	0.0001
					RMSE					
20	0.0606	0.0468	0.0351	0.0240	0.0173	0.0641	0.0502	0.0367	0.0255	0.0185
30	0.0503	0.0382	0.0287	0.0197	0.0141	0.0525	0.0394	0.0296	0.0205	0.0147
50	0.0359	0.0279	0.0209	0.0153	0.0104	0.0370	0.0286	0.0212	0.0155	0.0107
100	0.0263	0.0197	0.0152	0.0101	0.0073	0.0266	0.0199	0.0153	0.0103	0.0074
200	0.0176	0.0132	0.0107	0.0074	0.0051	0.0176	0.0133	0.0108	0.0075	0.0051

Notes: See notes to Table 1.

Table 4: Simulation results for Experiment 2 with uncorrelated factor loadings and the rank deficiency for heterogeneous $\beta_i = 1 + \eta_i$, $\eta_i \sim iidN(0, 0.04)$

T/N	20	30	50	100	200	20	30	50	100	200
	β_{CCEP}					β_{CCEMG}				
	Bias									
20	-0.0025	-0.0035	0.0001	-0.0015	0.0013	0.0011	-0.0054	-0.0008	-0.0012	0.0020
30	0.0052	0.0011	0.0015	-0.0015	0.0018	0.0018	0.0010	0.0013	-0.0016	0.0013
50	0.0010	0.0015	-0.0010	0.0001	0.0025	-0.0013	0.0014	0.0001	0.0001	0.0023
100	-0.0024	0.0010	-0.0029	-0.0001	0.0016	-0.0018	-0.0001	-0.0019	0.0005	0.0010
200	0.0010	-0.0054	0.0005	0.0005	-0.0011	-0.0010	-0.0048	0.0001	0.0010	-0.0006
	RMSE									
20	0.1168	0.0906	0.0703	0.0495	0.0351	0.1033	0.0819	0.0648	0.0443	0.0322
30	0.1107	0.0956	0.0691	0.0492	0.0347	0.0947	0.0802	0.0591	0.0428	0.0299
50	0.1071	0.0830	0.0668	0.0481	0.0339	0.0900	0.0717	0.0555	0.0402	0.0276
100	0.1022	0.0833	0.0657	0.0474	0.0320	0.0843	0.0700	0.0535	0.0384	0.0271
200	0.0980	0.0840	0.0665	0.0452	0.0333	0.0833	0.0694	0.0538	0.0367	0.0276
	β_{FEP}					β_{FEMG}				
	Bias									
20	-0.0009	-0.0052	-0.0012	-0.0006	0.0006	-0.0001	-0.0032	-0.0008	0.0004	0.0005
30	0.0052	-0.0019	0.0017	0.0005	0.0006	0.0027	-0.0003	0.0022	0.0006	0.0013
50	0.0037	0.0056	-0.0018	0.0020	0.0014	0.0034	0.0032	-0.0036	0.0011	0.0009
100	-0.0021	0.0004	0.0012	-0.0010	0.0011	-0.0005	0.0017	0.0011	0.0009	0.0003
200	-0.0018	0.0011	0.0031	0.0002	-0.0007	-0.0033	-0.0009	0.0025	0.0003	0.0000
	RMSE									
220	0.1370	0.1054	0.0801	0.0593	0.0403	0.1301	0.0974	0.0767	0.0550	0.0398
30	0.1339	0.1106	0.0791	0.0573	0.0396	0.1214	0.1031	0.0735	0.0551	0.0366
50	0.1207	0.0994	0.0778	0.0571	0.0398	0.1140	0.0927	0.0716	0.0521	0.0377
100	0.1191	0.0985	0.0770	0.0537	0.0377	0.1082	0.0901	0.0687	0.0502	0.0346
200	0.1175	0.1015	0.0785	0.0541	0.0391	0.1111	0.0932	0.0716	0.0502	0.0362
	β_{PCPBai}									
	Bias									
20	-0.0026	-0.0003	-0.0035	-0.0017	0.0004					
30	-0.0009	-0.0005	-0.0029	0.0009	-0.0009					
50	-0.0031	-0.0041	0.0008	-0.0008	0.0001					
100	-0.0016	-0.0017	-0.0014	0.0008	0.0000					
200	-0.0014	0.0003	-0.0025	-0.0009	-0.0007					
	RMSE									
20	0.0814	0.0630	0.0460	0.0323	0.0228					
30	0.0674	0.0541	0.0397	0.0292	0.0203					
50	0.0599	0.0473	0.0370	0.0245	0.0182					
100	0.0521	0.0425	0.0320	0.0231	0.0157					
200	0.0484	0.0399	0.0309	0.0218	0.0149					
	β_{PCPHNY}					$\beta_{PCMGHNY}$				
	Bias									
20	0.0014	0.0012	-0.0017	-0.0013	0.0007	0.0009	0.0013	-0.0010	-0.0008	0.0009
30	0.0013	0.0012	-0.0018	0.0011	-0.0005	0.0011	0.0020	-0.0022	0.0011	-0.0003
50	-0.0022	-0.0031	0.0014	-0.0004	0.0002	-0.0018	-0.0031	0.0016	0.0002	0.0004
100	0.0007	-0.0004	-0.0004	0.0013	0.0002	0.0005	-0.0007	-0.0006	0.0013	0.0001
200	0.0007	0.0017	-0.0013	-0.0005	-0.0005	0.0009	0.0018	-0.0012	-0.0004	-0.0005
	RMSE									
20	0.0788	0.0623	0.0458	0.0329	0.0229	0.0816	0.0632	0.0475	0.0332	0.0233
30	0.0671	0.0539	0.0405	0.0293	0.0202	0.0672	0.0536	0.0403	0.0294	0.0202
50	0.0593	0.0465	0.0365	0.0247	0.0181	0.0585	0.0462	0.0365	0.0244	0.0181
100	0.0525	0.0419	0.0320	0.0231	0.0158	0.0520	0.0419	0.0320	0.0231	0.0157
200	0.0489	0.0398	0.0309	0.0217	0.0149	0.0486	0.0396	0.0307	0.0216	0.0148

Notes: See notes to Table 1.

Table 5: Simulation results for Experiment 3 with correlated factor loadings and the full rank for homogeneous β=1

T/N	20	30	50	100	200	20	30	50	100	200
			β_{CCEP}					β_{CCEMG}		
					Bias					
20	0.0749	0.0496	0.0297	0.0131	0.0071	0.0689	0.0462	0.0282	0.0124	0.0069
30	0.0736	0.0510	0.0312	0.0145	0.0079	0.0677	0.0476	0.0304	0.0141	0.0078
50	0.0743	0.0501	0.0304	0.0151	0.0071	0.0682	0.0471	0.0292	0.0148	0.007
100	0.0740	0.0482	0.0308	0.0149	0.0074	0.0670	0.0451	0.0296	0.0145	0.0073
200	0.0727	0.0493	0.0300	0.0151	0.0075	0.0662	0.0460	0.0287	0.0148	0.0074
					RMSE					
20	0.1020	0.0727	0.0472	0.0276	0.0184	0.0966	0.0716	0.0479	0.0287	0.0192
30	0.0910	0.0673	0.0428	0.0243	0.0159	0.0848	0.0641	0.0426	0.0249	0.0161
50	0.0889	0.0611	0.0381	0.0215	0.0125	0.0820	0.0582	0.0373	0.0214	0.0126
100	0.0865	0.0552	0.0352	0.0182	0.0104	0.0774	0.0517	0.034	0.0179	0.0104
200	0.0819	0.0548	0.0327	0.0171	0.0091	0.0735	0.0508	0.0314	0.0168	0.0091
			β_{FEP}					β_{FEMG}		
					Bias					
20	0.6510	0.6523	0.6556	0.6577	0.6578	0.54	0.5371	0.5383	0.5376	0.5379
30	0.6566	0.6605	0.6624	0.6596	0.6621	0.5402	0.5425	0.5428	0.5381	0.5398
50	0.6605	0.6584	0.6603	0.6633	0.6615	0.5416	0.5385	0.5376	0.5401	0.5373
100	0.6574	0.662	0.6609	0.6642	0.6648	0.5377	0.5405	0.538	0.5396	0.5385
200	0.6568	0.6617	0.6624	0.6647	0.6647	0.5373	0.539	0.5383	0.5389	0.5377
					RMSE					
20	0.6562	0.6564	0.6591	0.6604	0.6601	0.5462	0.5419	0.5422	0.5404	0.5401
30	0.6605	0.6636	0.6647	0.6615	0.6636	0.5449	0.5463	0.5456	0.5402	0.5414
50	0.6637	0.661	0.662	0.6644	0.6625	0.5456	0.5415	0.5396	0.5414	0.5384
100	0.6600	0.6639	0.6621	0.665	0.6654	0.541	0.5428	0.5394	0.5405	0.5391
200	0.6592	0.6633	0.6634	0.6652	0.6651	0.5402	0.5412	0.5395	0.5396	0.5381
			β_{PCPBai}							
					Bias					
20	0.0068	0.0031	-0.0007	0.0008	0.0005					
30	0.0023	0.0010	0.0008	0.0001	0.0006					
50	0.0005	-0.0001	0.0004	0.0001	0.0003					
100	0.0007	-0.0003	-0.0004	0.0003	0.0002					
200	0.0006	0.0001	0.0001	-0.0003	0.0000					
					RMSE					
20	0.0596	0.0461	0.0356	0.0238	0.0173					
30	0.0465	0.0369	0.0282	0.0198	0.0138					
50	0.0352	0.0279	0.0208	0.0149	0.0102					
100	0.0235	0.0190	0.0147	0.0100	0.0073					
200	0.0168	0.0134	0.0105	0.0072	0.0050					
			β_{PCPHNY}					$\beta_{PCMGHNY}$		
					Bias					
20	0.0063	0.0025	-0.0008	0.0008	0.0004	0.0069	0.0029	-0.0013	0.0009	0.0004
30	0.0052	0.0009	0.0008	0.0001	0.0006	0.0053	0.0010	0.0009	-0.0001	0.0005
50	0.0017	0.0012	0.0004	0.0001	0.0003	0.0017	0.0011	0.0005	0.0001	0.0002
100	0.0010	0.0000	0.0000	0.0003	0.0002	0.0009	-0.0001	-0.0001	0.0004	0.0001
200	0.0007	0.0002	0.0002	-0.0002	0.0000	0.0007	0.0002	0.0002	-0.0001	0.0000
					RMSE					
20	0.0579	0.0452	0.0353	0.0238	0.0173	0.0626	0.0481	0.0371	0.0255	0.0186
30	0.0456	0.0365	0.0280	0.0198	0.0138	0.0477	0.0383	0.0291	0.0204	0.0145
50	0.0349	0.0277	0.0208	0.0149	0.0102	0.0359	0.0284	0.0213	0.0152	0.0104
100	0.0234	0.0190	0.0147	0.0100	0.0073	0.0238	0.0193	0.0149	0.0100	0.0074
200	0.0168	0.0133	0.0105	0.0072	0.0050	0.0170	0.0135	0.0105	0.0073	0.0050

Notes: See notes to Table 1.

Table 6: Simulation results for Experiment 3 with correlated factor loadings and the full rank for heterogeneous $\beta_i = 1 + \eta_i$, $\eta_i \sim iidN(0, 0.04)$

T/N	20	30	50	100	200	20	30	50	100	200
	β_{CCEP}					β_{CCEMG}				
	Bias									
20	0.0728	0.0484	0.0301	0.0158	0.0086	0.0678	0.0459	0.0282	0.0155	0.0089
30	0.0727	0.0468	0.0297	0.0130	0.0088	0.0675	0.0440	0.0283	0.0130	0.0083
50	0.0750	0.0502	0.0287	0.0150	0.0074	0.0683	0.0471	0.0274	0.0145	0.0073
100	0.0746	0.0474	0.0304	0.0152	0.0073	0.0682	0.0444	0.0291	0.0149	0.0071
200	0.0767	0.0513	0.0298	0.0144	0.0075	0.0700	0.0482	0.0287	0.0141	0.0074
	RMSE									
20	0.1121	0.0808	0.0566	0.0346	0.0249	0.1060	0.0791	0.0570	0.0352	0.0250
30	0.1007	0.0738	0.0500	0.0330	0.0227	0.0955	0.0710	0.0493	0.0326	0.0225
50	0.1011	0.0711	0.0465	0.0291	0.0185	0.0940	0.0678	0.0450	0.0288	0.0184
100	0.0959	0.0644	0.0443	0.0273	0.0175	0.0886	0.0613	0.0432	0.0269	0.0175
200	0.0984	0.0675	0.0439	0.0261	0.0168	0.0901	0.0644	0.0428	0.0259	0.0167
	β_{FEP}					β_{FEMG}				
	Bias									
20	0.6517	0.6528	0.6545	0.6564	0.6574	0.5413	0.5372	0.5364	0.5378	0.5389
30	0.6528	0.6562	0.6591	0.6606	0.6602	0.5371	0.5381	0.5383	0.5378	0.5383
50	0.6586	0.6568	0.6571	0.6629	0.6637	0.5384	0.5360	0.5344	0.5393	0.5388
100	0.6589	0.6596	0.6619	0.6625	0.6637	0.5403	0.5386	0.5393	0.5379	0.5385
200	0.6589	0.6623	0.6626	0.6624	0.6651	0.5396	0.5411	0.5387	0.5365	0.5382
	RMSE									
20	0.6595	0.6586	0.6588	0.6595	0.6602	0.5492	0.5435	0.5410	0.5408	0.5415
30	0.6587	0.6612	0.6623	0.6629	0.6621	0.5438	0.5434	0.5416	0.5401	0.5402
50	0.6643	0.6607	0.6599	0.6647	0.6650	0.5445	0.5403	0.5373	0.5411	0.5400
100	0.6639	0.6628	0.6642	0.6639	0.6644	0.5456	0.5422	0.5418	0.5393	0.5393
200	0.6635	0.6657	0.6648	0.6634	0.6657	0.5446	0.5446	0.5409	0.5377	0.5389
	β_{PCPBai}									
	Bias									
20	0.0012	-0.0028	-0.0014	-0.0018	-0.0010					
30	-0.0056	-0.0048	0.0000	-0.0007	-0.0012					
50	-0.0055	-0.0040	-0.0019	-0.0013	-0.0004					
100	-0.0043	-0.0041	-0.0024	-0.0016	-0.0005					
200	-0.0042	-0.0035	-0.0016	-0.0021	-0.0005					
	RMSE									
20	0.0759	0.0613	0.0453	0.0329	0.0234					
30	0.0653	0.0518	0.0403	0.0286	0.0193					
50	0.0572	0.0453	0.0362	0.0254	0.0180					
100	0.0528	0.0421	0.0333	0.0233	0.0160					
200	0.0500	0.0391	0.0312	0.0214	0.0150					
	β_{PCPHNY}					$\beta_{PCMGHNY}$				
	Bias									
20	0.0052	0.0013	0.0013	-0.0005	-0.0004	0.0046	0.0018	0.0007	-0.0005	-0.0004
30	0.0028	-0.0016	0.0025	0.0006	-0.0005	0.0025	-0.0012	0.0027	0.0009	-0.0004
50	0.0013	0.0012	0.0007	0.0001	0.0004	0.0007	0.0012	0.0008	0.0003	0.0004
100	0.0017	-0.0016	0.0006	-0.0002	0.0003	0.0016	-0.0010	0.0003	0.0000	0.0003
200	0.0006	0.0008	0.0011	-0.0006	0.0002	0.0007	0.0007	0.0010	-0.0006	0.0002
	RMSE									
20	0.0748	0.0610	0.0453	0.0328	0.0233	0.0762	0.0622	0.0462	0.0336	0.0238
30	0.0648	0.0516	0.0405	0.0286	0.0193	0.0651	0.0515	0.0409	0.0287	0.0194
50	0.0569	0.0450	0.0361	0.0254	0.0180	0.0571	0.0449	0.0360	0.0251	0.0178
100	0.0527	0.0419	0.0332	0.0232	0.0160	0.0525	0.0417	0.0330	0.0231	0.0158
200	0.0498	0.0390	0.0312	0.0213	0.0150	0.0495	0.0387	0.0310	0.0211	0.0150

Notes: See notes to Table 1.

Table 7: Simulation results for Experiment 4 with correlated factor loadings and the rank deficiency for homogeneous β=1

T/N	20	30	50	100	200	20	30	50	100	200
			β_{CCEP}					β_{CCEMG}		
					Bias					
20	0.4771	0.4831	0.4848	0.4858	0.4863	0.3502	0.3504	0.3479	0.3472	0.3450
30	0.4809	0.4843	0.4876	0.4912	0.4906	0.3522	0.3463	0.3469	0.3473	0.3451
50	0.4865	0.4880	0.4934	0.4905	0.4966	0.3520	0.3475	0.3488	0.3437	0.3468
100	0.4891	0.4906	0.4953	0.4947	0.4987	0.3513	0.3483	0.3484	0.3444	0.3458
200	0.4856	0.4929	0.4972	0.4957	0.4977	0.3471	0.3468	0.3480	0.3448	0.3450
					RMSE					
20	0.4922	0.4960	0.4957	0.4945	0.4940	0.3628	0.3603	0.3558	0.3529	0.3498
30	0.4925	0.4935	0.4952	0.4971	0.4964	0.3615	0.3536	0.3524	0.3512	0.3487
50	0.4966	0.4954	0.4990	0.4945	0.5000	0.3602	0.3532	0.3531	0.3465	0.3490
100	0.4971	0.4963	0.4992	0.4972	0.5005	0.3574	0.3527	0.3513	0.3462	0.3470
200	0.4926	0.4974	0.5004	0.4976	0.4990	0.3528	0.3504	0.3505	0.3462	0.3459
			β_{FEP}					β_{FEMG}		
					Bias					
20	0.6515	0.6542	0.6573	0.6614	0.6588	0.5401	0.5387	0.5397	0.5421	0.5388
30	0.6548	0.6583	0.6598	0.6614	0.6612	0.5418	0.5390	0.5398	0.5398	0.5389
50	0.6568	0.6581	0.6595	0.6611	0.6648	0.5390	0.5384	0.5378	0.5375	0.5407
100	0.6589	0.6605	0.6618	0.6633	0.6645	0.5384	0.5391	0.5377	0.5376	0.5379
200	0.6555	0.6613	0.6630	0.6641	0.6655	0.5366	0.5385	0.5384	0.5380	0.5387
					RMSE					
20	0.6572	0.6584	0.6606	0.6640	0.6610	0.5465	0.5435	0.5435	0.5448	0.5411
30	0.6589	0.6614	0.6622	0.6631	0.6628	0.5467	0.5426	0.5426	0.5417	0.5406
50	0.6600	0.6604	0.6613	0.6623	0.6658	0.5432	0.5411	0.5398	0.5389	0.5417
100	0.6615	0.6625	0.6630	0.6641	0.6650	0.5416	0.5415	0.5392	0.5385	0.5385
200	0.6581	0.6628	0.6641	0.6647	0.6659	0.5399	0.5405	0.5398	0.5387	0.5391
			β_{PCPBai}							
					Bias					
20	0.0064	0.0013	0.0016	-0.0006	0.0006					
30	-0.0019	0.0025	0.0003	0.0002	0.0004					
50	0.0008	-0.0003	0.0005	0.0008	0.0001					
100	0.0001	0.0006	0.0003	0.0005	-0.0001					
200	0.0002	0.0001	-0.0002	0.0001	0.0001					
					RMSE					
20	0.0593	0.0460	0.0345	0.0245	0.0171					
30	0.0453	0.0371	0.0278	0.0192	0.0133					
50	0.0346	0.0276	0.0214	0.0149	0.0106					
100	0.0247	0.0189	0.0150	0.0101	0.0072					
200	0.0168	0.0134	0.0103	0.0072	0.0050					
			β_{PCPHNY}					$\beta_{PCMGHNY}$		
					Bias					
20	0.0072	0.0027	0.0029	-0.0003	0.0008	0.0072	0.0026	0.0028	-0.0002	0.0008
30	0.0019	0.0033	0.0007	0.0005	0.0005	0.0016	0.0028	0.0006	0.0002	0.0005
50	0.0017	0.0009	0.0007	0.0009	0.0001	0.0014	0.0004	0.0006	0.001	0.0001
100	0.0004	0.0007	0.0005	0.0006	-0.0001	0.0004	0.0006	0.0006	0.0005	-0.0001
200	0.0003	0.0004	-0.0001	0.0002	0.0001	0.0003	0.0005	-0.0001	0.0002	0.0001
					RMSE					
20	0.0618	0.0478	0.0352	0.0247	0.0171	0.0669	0.0516	0.0375	0.0261	0.0182
30	0.0472	0.0382	0.0282	0.0193	0.0134	0.0500	0.0393	0.0293	0.02	0.014
50	0.0359	0.0282	0.0216	0.0151	0.0106	0.0366	0.0288	0.0221	0.0154	0.0108
100	0.0257	0.0196	0.0154	0.0103	0.0073	0.0260	0.0199	0.0156	0.0103	0.0073
200	0.0175	0.0138	0.0106	0.0073	0.005	0.0175	0.0138	0.0106	0.0073	0.0051

Notes: See notes to Table 1.

Table 8: Simulation results for Experiment 4 with correlated factor loadings and the rank deficiency for heterogeneous $\beta_i = 1 + \eta_i$, $\eta_i \sim iidN(0, 0.04)$

T/N	20	30	50	100	200	20	30	50	100	200
	β_{CCEP}					β_{CCEMG}				
	Bias									
20	0.4731	0.4771	0.4785	0.4777	0.4758	0.3478	0.3478	0.3434	0.3405	0.3384
30	0.4750	0.4766	0.4839	0.4857	0.4879	0.3469	0.3412	0.3453	0.3447	0.3429
50	0.4843	0.4839	0.4880	0.4895	0.4867	0.3472	0.3450	0.3458	0.3447	0.3395
100	0.4882	0.4939	0.4882	0.4873	0.4923	0.3518	0.3475	0.3429	0.3396	0.3425
200	0.4842	0.4887	0.4909	0.4925	0.4947	0.3466	0.3454	0.3444	0.3426	0.3427
	RMSE									
20	0.4926	0.4925	0.4902	0.4873	0.4856	0.3647	0.3607	0.3528	0.3470	0.3445
30	0.4915	0.4875	0.4927	0.4919	0.4934	0.3602	0.3499	0.3521	0.3491	0.3465
50	0.4960	0.4932	0.4949	0.4941	0.4906	0.3563	0.3526	0.3511	0.3482	0.3422
100	0.4990	0.5008	0.4933	0.4906	0.4947	0.3609	0.3535	0.3468	0.3421	0.3441
200	0.4946	0.4955	0.4950	0.4951	0.4963	0.3549	0.3512	0.3477	0.3447	0.3439
	β_{FEP}					β_{FEMG}				
	Bias									
20	0.6495	0.6548	0.6516	0.6571	0.6558	0.5387	0.5403	0.5353	0.5375	0.5357
30	0.6504	0.6564	0.6600	0.6605	0.6617	0.5358	0.5382	0.5412	0.5399	0.5383
50	0.6563	0.6556	0.6608	0.6643	0.6628	0.5359	0.5361	0.5398	0.5418	0.5376
100	0.6601	0.6632	0.6591	0.6618	0.6654	0.5416	0.5399	0.5365	0.5377	0.5400
200	0.6576	0.6603	0.6618	0.6653	0.6665	0.5383	0.5379	0.5388	0.5388	0.5389
	RMSE									
20	0.6583	0.6608	0.6561	0.6603	0.6584	0.5483	0.5470	0.5402	0.5407	0.5383
30	0.6579	0.6612	0.6635	0.6627	0.6634	0.5435	0.5434	0.5447	0.5422	0.5400
50	0.6618	0.6599	0.6636	0.6660	0.6640	0.5417	0.5407	0.5427	0.5436	0.5389
100	0.6651	0.6665	0.6615	0.6632	0.6663	0.5471	0.5436	0.5389	0.5392	0.5408
200	0.6629	0.6635	0.6639	0.6664	0.6672	0.5434	0.5414	0.5409	0.5400	0.5396
	β_{PCPBai}									
	Bias									
20	0.0013	-0.0012	-0.0011	-0.0005	0.0002					
30	-0.0028	0.0002	-0.0001	-0.0007	-0.0009					
50	-0.0031	-0.0026	-0.0008	-0.0005	-0.0003					
100	-0.0045	-0.0020	-0.0009	-0.0009	0.0004					
200	-0.0019	-0.0018	-0.0013	-0.0008	-0.0005					
	RMSE									
20	0.0738	0.0597	0.0460	0.0323	0.0227					
30	0.0638	0.0530	0.0404	0.0277	0.0205					
50	0.0575	0.0473	0.0350	0.0254	0.0175					
100	0.0518	0.0403	0.0307	0.0232	0.0161					
200	0.0500	0.0384	0.0306	0.0209	0.0154					
	β_{PCPHNY}					$\beta_{PCMGHNY}$				
	Bias									
20	0.0064	0.0022	0.0011	0.0007	0.0008	0.0066	0.0024	0.0014	0.0007	0.0007
30	0.0034	0.0034	0.0019	0.0005	-0.0003	0.0028	0.0032	0.0018	0.0006	-0.0001
50	0.0016	0.0009	0.0012	0.0003	0.0002	0.0020	0.0007	0.0009	0.0004	0.0002
100	-0.0016	0.0010	0.0010	0.0000	0.0009	-0.0018	0.0009	0.0011	0.0000	0.0008
200	0.0018	0.0012	0.0006	0.0001	0.0000	0.0019	0.0011	0.0006	0.0002	0.0000
	RMSE									
20	0.0769	0.0617	0.0462	0.0324	0.0228	0.0803	0.0638	0.0471	0.0327	0.0230
30	0.0652	0.0545	0.0409	0.0279	0.0205	0.0650	0.0543	0.0409	0.0280	0.0204
50	0.0586	0.0482	0.0352	0.0255	0.0175	0.0586	0.0480	0.0352	0.0254	0.0174
100	0.0523	0.0406	0.0308	0.0232	0.0161	0.0519	0.0403	0.0305	0.0231	0.0161
200	0.0507	0.0386	0.0305	0.0209	0.0154	0.0503	0.0383	0.0304	0.0209	0.0154

Notes: See notes to Table 1.

Table 9: Size and power of the H^{NON} statistic and coverage rates at 95 % level.

	Experiment 1						Experiment 3				
	Size of the H^{NON} with β_{PCBai}						Power of the H^{NON} with β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.072	0.059	0.059	0.057	0.047	50	1	1	1	1	1
100	0.058	0.055	0.058	0.054	0.051	100	1	1	1	1	1
150	0.073	0.070	0.053	0.047	0.047	150	1	1	1	1	1
200	0.061	0.050	0.049	0.056	0.048	200	1	1	1	1	1
500	0.071	0.054	0.048	0.046	0.060	500	1	1	1	1	1
	Size of the H^{NON} with β_{PCHNY}						Power of the H^{NON} with β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.082	0.069	0.073	0.062	0.053	50	1	1	1	1	1
100	0.059	0.056	0.063	0.060	0.052	100	1	1	1	1	1
150	0.072	0.069	0.053	0.048	0.048	150	1	1	1	1	1
200	0.059	0.052	0.050	0.056	0.050	200	1	1	1	1	1
500	0.070	0.054	0.048	0.046	0.060	500	1	1	1	1	1
	Coverage rates β_{FE}						Coverage rates β_{FE}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.919	0.939	0.950	0.941	0.942	50	0	0	0	0	0
100	0.936	0.950	0.942	0.952	0.944	100	0	0	0	0	0
150	0.939	0.931	0.947	0.950	0.947	150	0	0	0	0	0
200	0.939	0.955	0.949	0.939	0.939	200	0	0	0	0	0
500	0.932	0.944	0.948	0.954	0.945	500	0	0	0	0	0
	Coverage rates β_{PCBai}						Coverage rates β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.92	0.939	0.943	0.936	0.95	50	0.922	0.942	0.937	0.941	0.947
100	0.917	0.923	0.926	0.953	0.955	100	0.925	0.948	0.936	0.942	0.953
150	0.91	0.936	0.938	0.957	0.937	150	0.931	0.921	0.933	0.942	0.942
200	0.916	0.934	0.933	0.951	0.951	200	0.918	0.938	0.939	0.932	0.935
500	0.911	0.929	0.924	0.934	0.948	500	0.927	0.937	0.943	0.945	0.948
	Coverage rates β_{PCHNY}						Coverage rates β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.909	0.927	0.928	0.929	0.940	50	0.928	0.943	0.938	0.942	0.949
100	0.914	0.923	0.919	0.947	0.952	100	0.932	0.948	0.935	0.940	0.954
150	0.914	0.929	0.937	0.956	0.933	150	0.919	0.941	0.933	0.942	0.942
200	0.913	0.935	0.935	0.950	0.947	200	0.924	0.934	0.946	0.932	0.935
500	0.914	0.928	0.923	0.934	0.948	500	0.924	0.936	0.942	0.946	0.948
	Experiment 2						Experiment 4				
	Size of the H^{NON} with β_{PCBai}						Power of the H^{NON} with β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.090	0.059	0.061	0.044	0.038	50	1	1	1	1	1
100	0.073	0.050	0.052	0.059	0.054	100	1	1	1	1	1
150	0.051	0.053	0.037	0.056	0.051	150	1	1	1	1	1
200	0.076	0.051	0.058	0.049	0.047	200	1	1	1	1	1
500	0.072	0.057	0.042	0.069	0.048	500	1	1	1	1	1
	Size of the H^{NON} with β_{PCHNY}						Power of the H^{NON} with β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.090	0.059	0.061	0.044	0.038	50	1	1	1	1	1
100	0.074	0.050	0.052	0.059	0.054	100	1	1	1	1	1
150	0.054	0.052	0.037	0.056	0.051	150	1	1	1	1	1
200	0.076	0.050	0.057	0.049	0.047	200	1	1	1	1	1
500	0.073	0.056	0.041	0.069	0.048	500	1	1	1	1	1
	Coverage rates β_{FE}						Coverage rates β_{FE}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.925	0.943	0.939	0.944	0.961	50	0	0	0	0	0
100	0.927	0.956	0.945	0.936	0.941	100	0	0	0	0	0
150	0.951	0.955	0.962	0.95	0.941	150	0	0	0	0	0
200	0.928	0.947	0.947	0.948	0.951	200	0	0	0	0	0
500	0.935	0.942	0.956	0.933	0.953	500	0	0	0	0	0
	Coverage rates β_{PCBai}						Coverage rates β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.912	0.937	0.942	0.951	0.956	50	0.915	0.935	0.938	0.947	0.951
100	0.891	0.930	0.941	0.936	0.961	100	0.776	0.953	0.945	0.947	0.949
150	0.916	0.910	0.930	0.944	0.943	150	0.754	0.652	0.936	0.939	0.936
200	0.913	0.911	0.919	0.934	0.945	200	0.829	0.783	0.605	0.948	0.960
500	0.912	0.938	0.917	0.924	0.938	500	0.884	0.855	0.826	0.768	0.951
	Coverage rates β_{PCHNY}						Coverage rates β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.912	0.936	0.942	0.951	0.956	50	0.906	0.926	0.934	0.939	0.939
100	0.922	0.930	0.941	0.936	0.961	100	0.918	0.948	0.939	0.943	0.949
150	0.917	0.931	0.930	0.944	0.943	150	0.909	0.939	0.934	0.939	0.933
200	0.924	0.942	0.956	0.934	0.945	200	0.915	0.927	0.951	0.947	0.956
500	0.904	0.939	0.928	0.935	0.938	500	0.922	0.935	0.948	0.950	0.950

Notes: FE denotes the two-way fixed effect estimators; PCBai is the iterative pooled principal component estimator by Bai (2009) while PCHNY is the pooled principal component estimator by Hayakawa, Nagata, and Yamagata (2018). The PC estimators are bias-corrected and evaluated using the IC_{p1} criterion by Bai and Ng (2002). H^{NON} is the H-statistic defined in (27).

Table 10: Size and power of the H^{HAC} statistic and coverage rates at 95 % level.

	Experiment 1						Experiment 3				
	Size of the H^{HAC} with β_{PCBai}						Power of the H^{HAC} with β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.091	0.062	0.049	0.061	0.050	50	1	1	1	1	1
100	0.072	0.057	0.052	0.052	0.054	100	1	1	1	1	1
150	0.058	0.059	0.054	0.063	0.047	150	1	1	1	1	1
200	0.059	0.056	0.060	0.060	0.041	200	1	1	1	1	1
500	0.079	0.063	0.048	0.050	0.055	500	1	1	1	1	1
	Size of the H^{HAC} with β_{PCHNY}						Power of the H^{HAC} with β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.102	0.067	0.061	0.071	0.058	50	1	1	1	1	1
100	0.072	0.06	0.054	0.052	0.059	100	1	1	1	1	1
150	0.055	0.059	0.055	0.067	0.048	150	1	1	1	1	1
200	0.057	0.056	0.061	0.06	0.043	200	1	1	1	1	1
500	0.077	0.063	0.048	0.049	0.055	500	1	1	1	1	1
	Coverage rates β_{FE}						Coverage rates β_{FE}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.921	0.935	0.946	0.953	0.958	50	0	0	0	0	0
100	0.938	0.957	0.946	0.943	0.948	100	0	0	0	0	0
150	0.940	0.944	0.941	0.936	0.951	150	0	0	0	0	0
200	0.949	0.951	0.942	0.940	0.956	200	0	0	0	0	0
500	0.923	0.941	0.953	0.952	0.945	500	0	0	0	0	0
	Coverage rates β_{PCBai}						Coverage rates β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.926	0.937	0.942	0.941	0.936	50	0.922	0.935	0.934	0.945	0.937
100	0.903	0.927	0.953	0.949	0.953	100	0.929	0.941	0.932	0.941	0.947
150	0.908	0.946	0.935	0.938	0.940	150	0.926	0.913	0.951	0.939	0.952
200	0.926	0.932	0.943	0.944	0.945	200	0.928	0.926	0.950	0.947	0.950
500	0.904	0.939	0.940	0.931	0.935	500	0.921	0.942	0.939	0.950	0.950
	Coverage rates β_{PCHNY}						Coverage rates β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.913	0.931	0.927	0.930	0.925	50	0.923	0.936	0.935	0.943	0.936
100	0.898	0.925	0.952	0.947	0.948	100	0.923	0.940	0.934	0.941	0.950
150	0.911	0.948	0.934	0.934	0.937	150	0.928	0.917	0.951	0.939	0.953
200	0.926	0.934	0.945	0.940	0.943	200	0.924	0.928	0.948	0.947	0.950
500	0.904	0.939	0.940	0.935	0.935	500	0.920	0.941	0.938	0.956	0.950
	Experiment 2						Experiment 4				
	Size of the H^{HAC} with β_{PCBai}						Power of the H^{HAC} with β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.070	0.070	0.056	0.060	0.054	50	1	1	1	1	1
100	0.076	0.050	0.065	0.055	0.049	100	1	1	1	1	1
150	0.070	0.073	0.050	0.051	0.045	150	1	1	1	1	1
200	0.064	0.061	0.061	0.061	0.054	200	1	1	1	1	1
500	0.059	0.065	0.060	0.052	0.048	500	1	1	1	1	1
	Size of the H^{HAC} with β_{PCHNY}						Power of the H^{HAC} with β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.086	0.076	0.076	0.072	0.072	50	1	1	1	1	1
100	0.077	0.060	0.071	0.066	0.058	100	1	1	1	1	1
150	0.069	0.076	0.054	0.055	0.045	150	1	1	1	1	1
200	0.068	0.069	0.057	0.062	0.059	200	1	1	1	1	1
500	0.060	0.064	0.061	0.050	0.049	500	1	1	1	1	1
	Coverage rates β_{FE}						Coverage rates β_{FE}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.935	0.938	0.949	0.947	0.949	50	0	0	0	0	0
100	0.914	0.952	0.942	0.946	0.942	100	0	0	0	0	0
150	0.931	0.926	0.954	0.948	0.960	150	0	0	0	0	0
200	0.933	0.936	0.937	0.943	0.945	200	0	0	0	0	0
500	0.941	0.937	0.944	0.951	0.954	500	0	0	0	0	0
	Coverage rates β_{PCBai}						Coverage rates β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.934	0.931	0.943	0.935	0.946	50	0.930	0.952	0.943	0.959	0.957
100	0.899	0.945	0.929	0.945	0.956	100	0.675	0.943	0.952	0.956	0.961
150	0.900	0.924	0.940	0.941	0.949	150	0.754	0.661	0.941	0.952	0.955
200	0.916	0.914	0.918	0.946	0.957	200	0.793	0.786	0.584	0.942	0.938
500	0.913	0.930	0.921	0.929	0.952	500	0.885	0.847	0.835	0.797	0.948
	Coverage rates β_{PCHNY}						Coverage rates β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.909	0.921	0.916	0.915	0.923	50	0.922	0.945	0.926	0.947	0.942
100	0.901	0.932	0.921	0.938	0.948	100	0.907	0.936	0.951	0.954	0.956
150	0.916	0.926	0.935	0.935	0.948	150	0.922	0.921	0.939	0.949	0.952
200	0.913	0.933	0.934	0.944	0.952	200	0.911	0.940	0.953	0.937	0.934
500	0.925	0.926	0.922	0.934	0.950	500	0.920	0.925	0.940	0.938	0.948

Notes: FE denotes the two-way fixed effect estimators; PCBai is the iterative pooled principal component estimator by Bai (2009) while PCHNY is the pooled principal component estimator by Hayakawa, Nagata, and Yamagata (2018). The PC estimators are bias-corrected and evaluated using the IC_{p1} criterion by Bai and Ng (2002). H^{HAC} is the H-statistic defined in (28).

Table 11: Size and power of the H^{NON} statistic and coverage rates at 95 % level for serirally correlated errors and heterogeneous βs.

	Experiment 1						Experiment 3				
	Size of the H^{NON} with β_{PCBai}						Power of the H^{NON} with β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.068	0.057	0.054	0.043	0.054	50	1	1	1	1	1
100	0.061	0.049	0.065	0.062	0.046	100	1	1	1	1	1
150	0.073	0.057	0.052	0.049	0.042	150	1	1	1	1	1
200	0.059	0.055	0.066	0.048	0.051	200	1	1	1	1	1
500	0.060	0.062	0.054	0.061	0.057	500	1	1	1	1	1
	Size of the H^{NON} with β_{PCHNY}						Power of the H^{NON} with β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.072	0.064	0.064	0.057	0.065	50	1	1	1	1	1
100	0.060	0.055	0.072	0.063	0.046	100	1	1	1	1	1
150	0.073	0.056	0.053	0.051	0.045	150	1	1	1	1	1
200	0.059	0.055	0.066	0.050	0.054	200	1	1	1	1	1
500	0.059	0.062	0.053	0.060	0.058	500	1	1	1	1	1
	Coverage rates β_{FE}						Coverage rates β_{FE}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.944	0.948	0.945	0.951	0.948	50	0	0	0	0	0
100	0.942	0.952	0.938	0.949	0.955	100	0	0	0	0	0
150	0.932	0.943	0.945	0.957	0.949	150	0	0	0	0	0
200	0.941	0.941	0.931	0.947	0.952	200	0	0	0	0	0
500	0.926	0.937	0.952	0.934	0.939	500	0	0	0	0	0
	Coverage rates β_{PCBai}						Coverage rates β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.926	0.914	0.942	0.941	0.941	50	0.924	0.930	0.938	0.946	0.943
100	0.922	0.948	0.951	0.943	0.937	100	0.834	0.936	0.943	0.946	0.952
150	0.919	0.924	0.936	0.945	0.941	150	0.809	0.777	0.938	0.944	0.947
200	0.919	0.929	0.944	0.945	0.935	200	0.884	0.856	0.815	0.947	0.942
500	0.916	0.945	0.938	0.914	0.952	500	0.896	0.913	0.901	0.907	0.952
	Coverage rates β_{PCHNY}						Coverage rates β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.921	0.909	0.930	0.926	0.930	50	0.838	0.870	0.845	0.851	0.851
100	0.916	0.942	0.943	0.942	0.937	100	0.860	0.889	0.901	0.900	0.896
150	0.919	0.924	0.934	0.943	0.938	150	0.875	0.893	0.909	0.917	0.937
200	0.916	0.929	0.940	0.943	0.932	200	0.904	0.897	0.911	0.931	0.928
500	0.917	0.943	0.940	0.914	0.951	500	0.906	0.924	0.933	0.943	0.947
	Experiment 2						Experiment 4				
	Size of the H^{NON} with β_{PCBai}						Power of the H^{NON} with β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.064	0.054	0.062	0.057	0.055	50	1	1	1	1	1
100	0.056	0.068	0.055	0.050	0.044	100	1	1	1	1	1
150	0.061	0.053	0.051	0.052	0.051	150	1	1	1	1	1
200	0.067	0.050	0.064	0.048	0.054	200	1	1	1	1	1
500	0.072	0.062	0.065	0.062	0.053	500	1	1	1	1	1
	Size of the H^{NON} with β_{PCHNY}						Power of the H^{NON} with β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.080	0.061	0.074	0.075	0.064	50	1	1	1	1	1
100	0.063	0.071	0.058	0.056	0.049	100	1	1	1	1	1
150	0.065	0.051	0.056	0.054	0.054	150	1	1	1	1	1
200	0.072	0.046	0.066	0.048	0.057	200	1	1	1	1	1
500	0.070	0.058	0.058	0.060	0.053	500	1	1	1	1	1
	Coverage rates β_{FE}						Coverage rates β_{FE}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.938	0.933	0.942	0.946	0.953	50	0	0	0	0	0
100	0.943	0.943	0.933	0.937	0.955	100	0	0	0	0	0
150	0.943	0.945	0.956	0.949	0.946	150	0	0	0	0	0
200	0.941	0.951	0.935	0.957	0.943	200	0	0	0	0	0
500	0.924	0.944	0.949	0.938	0.950	500	0	0	0	0	0
	Coverage rates β_{PCBai}						Coverage rates β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.915	0.931	0.945	0.950	0.938	50	0.925	0.949	0.940	0.958	0.939
100	0.918	0.938	0.935	0.933	0.960	100	0.871	0.940	0.938	0.943	0.951
150	0.916	0.945	0.943	0.930	0.935	150	0.908	0.882	0.952	0.947	0.941
200	0.903	0.931	0.936	0.945	0.957	200	0.904	0.891	0.857	0.945	0.947
500	0.927	0.950	0.954	0.944	0.952	500	0.903	0.927	0.934	0.945	0.952
	Coverage rates β_{PCHNY}						Coverage rates β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.899	0.921	0.930	0.928	0.930	50	0.920	0.941	0.932	0.952	0.927
100	0.913	0.933	0.931	0.926	0.954	100	0.901	0.936	0.935	0.941	0.951
150	0.925	0.939	0.939	0.927	0.930	150	0.920	0.933	0.951	0.946	0.940
200	0.915	0.931	0.946	0.944	0.954	200	0.925	0.934	0.934	0.943	0.946
500	0.925	0.948	0.953	0.940	0.952	500	0.908	0.940	0.940	0.950	0.952

Notes: see notes to Table 9 .

Table 12: Size and power of the H^{HAC} statistic and coverage rates at 95 % level for serirally correlated errors and heterogeneous βs.

	Experiment 1						Experiment 3				
	Size of the H^{HAC} with β_{PCBai}						Power of the H^{HAC} with β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.068	0.074	0.045	0.054	0.053	50	1	1	1	1	1
100	0.068	0.053	0.054	0.052	0.045	100	1	1	1	1	1
150	0.070	0.064	0.056	0.053	0.047	150	1	1	1	1	1
200	0.053	0.060	0.058	0.059	0.043	200	1	1	1	1	1
500	0.068	0.055	0.062	0.059	0.053	500	1	1	1	1	1
	Size of the H^{HAC} with β_{PCHNY}						Power of the H^{HAC} with β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.074	0.076	0.054	0.068	0.064	50	1	1	1	1	1
100	0.066	0.057	0.057	0.054	0.048	100	1	1	1	1	1
150	0.070	0.062	0.056	0.058	0.051	150	1	1	1	1	1
200	0.054	0.059	0.056	0.061	0.044	200	1	1	1	1	1
500	0.068	0.055	0.061	0.059	0.053	500	1	1	1	1	1
	Coverage rates β_{FE}						Coverage rates β_{FE}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.937	0.941	0.940	0.947	0.950	50	0	0	0	0	0
100	0.929	0.938	0.947	0.943	0.950	100	0	0	0	0	0
150	0.924	0.933	0.940	0.947	0.954	150	0	0	0	0	0
200	0.943	0.943	0.943	0.928	0.956	200	0	0	0	0	0
500	0.924	0.944	0.952	0.931	0.946	500	0	0	0	0	0
	Coverage rates β_{PCBai}						Coverage rates β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.923	0.941	0.951	0.938	0.939	50	0.927	0.920	0.939	0.948	0.930
100	0.909	0.937	0.939	0.944	0.960	100	0.755	0.926	0.943	0.932	0.955
150	0.916	0.911	0.936	0.937	0.953	150	0.748	0.711	0.943	0.935	0.937
200	0.911	0.931	0.938	0.950	0.959	200	0.836	0.790	0.663	0.944	0.963
500	0.906	0.937	0.937	0.941	0.932	500	0.895	0.902	0.885	0.868	0.938
	Coverage rates β_{PCHNY}						Coverage rates β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.917	0.937	0.941	0.924	0.929	50	0.857	0.858	0.853	0.904	0.870
100	0.910	0.933	0.936	0.942	0.955	100	0.891	0.894	0.917	0.891	0.925
150	0.913	0.908	0.936	0.932	0.950	150	0.842	0.902	0.923	0.908	0.921
200	0.911	0.932	0.940	0.947	0.957	200	0.900	0.897	0.909	0.931	0.941
500	0.905	0.937	0.937	0.940	0.931	500	0.906	0.927	0.930	0.915	0.934
	Experiment 2						Experiment 4				
	Size of the H^{HAC} with β_{PCBai}						Power of the H^{HAC} with β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.078	0.054	0.057	0.051	0.050	50	1	1	1	1	1
100	0.075	0.051	0.052	0.059	0.063	100	1	1	1	1	1
150	0.064	0.064	0.042	0.054	0.053	150	1	1	1	1	1
200	0.091	0.044	0.060	0.051	0.036	200	1	1	1	1	1
500	0.086	0.055	0.066	0.040	0.049	500	1	1	1	1	1
	Size of the H^{HAC} with β_{PCHNY}						Power of the H^{HAC} with β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.086	0.064	0.072	0.060	0.067	50	1	1	1	1	1
100	0.074	0.056	0.056	0.066	0.071	100	1	1	1	1	1
150	0.064	0.067	0.045	0.057	0.053	150	1	1	1	1	1
200	0.091	0.044	0.064	0.052	0.040	200	1	1	1	1	1
500	0.084	0.056	0.064	0.037	0.049	500	1	1	1	1	1
	Coverage rates β_{FE}						Coverage rates β_{FE}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.929	0.948	0.937	0.936	0.952	50	0	0	0	0	0
100	0.918	0.942	0.951	0.942	0.925	100	0	0	0	0	0
150	0.943	0.942	0.959	0.937	0.952	150	0	0	0	0	0
200	0.914	0.954	0.942	0.946	0.957	200	0	0	0	0	0
500	0.925	0.941	0.938	0.954	0.952	500	0	0	0	0	0
	Coverage rates β_{PCBai}						Coverage rates β_{PCBai}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.918	0.942	0.937	0.928	0.934	50	0.930	0.934	0.928	0.941	0.937
100	0.910	0.936	0.925	0.936	0.941	100	0.862	0.937	0.935	0.941	0.939
150	0.920	0.930	0.938	0.942	0.947	150	0.882	0.899	0.947	0.949	0.948
200	0.907	0.929	0.949	0.945	0.940	200	0.895	0.889	0.885	0.935	0.945
500	0.922	0.932	0.941	0.944	0.952	500	0.925	0.934	0.937	0.940	0.946
	Coverage rates β_{PCHNY}						Coverage rates β_{PCHNY}				
T/N	50	100	150	200	500	T/N	50	100	150	200	500
50	0.910	0.925	0.919	0.917	0.918	50	0.921	0.924	0.920	0.933	0.933
100	0.910	0.927	0.918	0.928	0.934	100	0.916	0.931	0.933	0.935	0.934
150	0.921	0.933	0.934	0.938	0.947	150	0.910	0.948	0.947	0.945	0.945
200	0.909	0.934	0.947	0.944	0.936	200	0.911	0.922	0.944	0.934	0.943
500	0.918	0.932	0.936	0.947	0.951	500	0.933	0.934	0.948	0.961	0.946

Notes: see notes to Table 10 .

Table 13: Empirical applications to six different dataset.

		β_{FE}	se_{NON}	se_{HAC}	β_{PCHNY}	se_{NON}	se_{HAC}	H^{NON}	H^{HAC}	H^{Bai}	CD
Production function as in (42)											
OECD	$\beta_{\frac{k}{l}}$	0.589	0.015	0.015	0.596	0.012	0.012	0.0001	0.0001	0.010	24.27
								(0.998)	(0.998)	(0.919)	(0.013)
EU27	$\beta_{\frac{k}{l}}$	0.654	0.042	0.038	0.554	0.019	0.019	0.151	0.169	0.001	10.28
								(0.697)	(0.681)	(0.968)	(0.031)
ITA	$\beta_{\frac{k}{l}}$	0.537	0.042	0.035	0.311	0.016	0.016	0.885	1.195	0.016	17.77
								(0.346)	(0.274)	(0.897)	(0.017)
US	$\beta_{\frac{k}{l}}$	0.165	0.011	0.011	0.149	0.042	0.042	0.098	0.094	0.006	14.35
								(0.753)	(0.758)	(0.934)	(0.022)
Gravity model as in (43)											
EU14	β_{gdp}	3.299	0.050	0.039	1.959	0.011	0.011				
$k=6$	β_{rer}	0.037	0.006	0.004	0.024	0.007	0.007				
	β_{sim}	1.362	0.036	0.022	1.402	0.067	0.067				
	β_{rlf}	0.033	0.006	0.002	0.079	0.011	0.008				
	β_{cee}	0.291	0.006	0.004	0.354	0.010	0.011	13.86	14.09	0.124	97.30
	β_{emu}	0.254	0.008	0.006	0.113	0.013	0.013	(0.031)	(0.028)	(0.999)	(0.000)
Gasoline demand function as in (44)											
	β_p	-0.161	0.008	0.007	-0.121	0.002	0.002	0.747	0.590	0.001	85.59
	β_i	0.436	0.022	0.020	0.398	0.010	0.010	(0.688)	(0.744)	(0.999)	(0.000)

Notes: FE denotes the two-way fixed effecs estimator. PCHNY is the pooled principal component estimator by Hayakawa, Nagata, and Yamagata (2018). The PC estimators are bias-corrected. Applying IC_{p1} criterion by Bai and Ng (2002), we extract the four factors in the gravity dataset while we estimate two factors in all other dataset. se_{NON} and se_{HAC} are the standard errors evaluated respectively by the non parametric NON and HAC variance-covariance estimators. H^{NON} and H^{HAC} are the H-statistic defined in (27) and (28) with the p-values inside (.). H^{Bai} is the Hausman test statistic proposed by Bai (2009) while CD is the Pesaran (2015)'s CD test applied to the one way fixed effects residuals.

References

BAI, J. (2009): "Panel Data Models with Interactive Fixed Effects," *Econometrica*, 77(4), 1229–1279.

BAI, J., AND S. NG (2002): "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70(1), 191–221.

CASTAGNETTI, C., E. ROSSI, AND L. TRAPANI (2015): "Testing for no factor structures: On the use of Hausman-type statistics," *Economics Letters*, 130, 66 – 68.

CHARBONNEAU, K. B. (2017): "Multiple fixed effects in binary response panel data models," *Econometrics Journal*, 20(3), S1–S13.

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

CHUDIK, A., M. H. PESARAN, AND E. TOSETTI (2011): "Weak and strong cross section dependence and estimation of large panels," *The Econometrics Journal*, 14(1), C45–C90.

COAKELY, J., A. M. FUERTES, AND R. SMITH (2006): "Unobserved heterogeneity in panel time series models," *Computational Statistics and Data Analysis*, 50(1), 2361–2380.

DAVIDSON, J. (1994): *Stochastic Limit Theory*. Oxford University Press.

FERNANDEZ-VAL, I., AND M. WEIDNER (2016): "Individual and time effects in nonlinear panel models with large N, T," *Journal of Econometrics*, 192(1), 291 – 312.

GONCALVES, S., AND B. PERRON (2014): "Bootstrapping factor-augmented regression models," *Journal of Econometrics*, 182(1), 156–173.

HAUSMAN, J. A. (1978): "Specification Tests in Econometrics," *Econometrica*, 46(6), 1251–1271.

HAYAKAWA, K., S. NAGATA, AND T. YAMAGATA (2018): "A Robust Approach to Heteroskedasticity, Error Serial Correlation and Slope Heterogeneity for Large Linear Panel Data Models with Interactive Effects," *SSRN Electronic Journal*.

KAPETANIOS, G., AND M. H. PESARAN (2005): "Alternative Approaches to Estimation and Inference in Large Multifactor Panels: Small Sample Results with an Application to Modelling of Asset Returns," CESifo Working Paper Series 1416, CESifo Group Munich.

KAPETANIOS, G., M. H. PESARAN, AND T. YAMAGATA (2011): "Panels with non-stationary multifactor error structures," *Journal of Econometrics*, 160(2), 326 – 348.

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

Liu, W. (2014): "Modeling gasoline demand in the United States: A flexible semiparametric approach," *Energy Economics*, 45, 244 – 253.

Mastromarco, C., L. Serlenga, and Y. Shin (2016): "Modelling Technical Efficiency in Cross Sectionally Dependent Stochastic Frontier Panels," *Journal of Applied Econometrics*, 31(1), 281–297.

Moon, H. R., and M. Weidner (2015): "Linear Regression for Panel With Unknown Number of Factors as Interactive Fixed Effects," *Econometrica*, 83(4), 1543–1579.

Munnell, A. (1990): "How does public infrastructure affect regional economic performance?," *Conference Series ; [Proceedings]*, 34, 69–112.

Pesaran, M. H. (2006): "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica*, 74(4), 967–1012.

——— (2015): "Testing Weak Cross-Sectional Dependence in Large Panels," *Econometric Reviews*, 34(6), 1089–1117.

Sarafidis, V., and T. Wansbeek (2012): "Cross-Sectional Dependence in Panel Data Analysis," *Econometric Reviews*, 31(5), 483–531.

Sarafidis, V., T. Yamagata, and D. Robertson (2009): "A test of cross section dependence for a linear dynamic panel model with regressors," *Journal of Econometrics*, 148(2), 149 – 161.

Serlenga, L., and Y. Shin (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.

Solow, R. (1956): "A Contribution to The Theory of Economic Growth," *The Quarterly Journal of Economics*, 70, 65 âĂŞ 94.

Westerlund, J. (2018): "On Estimation and Inference in Heterogeneous Panel Regressions with Interactive Effects," *Journal of Time Series Analysis*, Forthcoming.

Westerlund, J., and J.-P. Urbain (2013): "On the estimation and inference in factor-augmented panel regressions with correlated loadings," *Economics Letters*, 119(3), 247 – 250.

——— (2015): "Cross-sectional averages versus principal components," *Journal of Econometrics*, 185(2), 372 – 377.