Journal of Business & Economic Statistics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ubes20

A Robust Approach to Heteroscedasticity, Error Serial Correlation and Slope Heterogeneity in Linear Models with Interactive Effects for Large Panel Data

Guowei Cui, Kazuhiko Hayakawa, Shuichi Nagata & Takashi Yamagata

To cite this article: Guowei Cui, Kazuhiko Hayakawa, Shuichi Nagata & Takashi Yamagata (2022): A Robust Approach to Heteroscedasticity, Error Serial Correlation and Slope Heterogeneity in Linear Models with Interactive Effects for Large Panel Data, Journal of Business & Economic Statistics, DOI: 10.1080/07350015.2022.2077349

To link to this article: https://doi.org/10.1080/07350015.2022.2077349

JOURNAL OF BUSINESS & ECONOMIC STATISTICS
2022, VOL. 00, NO. 0, 1–14
https://doi.org/10.1080/07350015.2022.2077349

A Robust Approach to Heteroscedasticity, Error Serial Correlation and Slope Heterogeneity in Linear Models with Interactive Effects for Large Panel Data

Guowei Cui[a], Kazuhiko Hayakawa[b], Shuichi Nagata[c], and Takashi Yamagata[d,e]

[a]School of Economics, Huazhong University of Science and Technology, Wuhan, China; [b]Graduate School of Social Sciences, Department of Economics, Higashi-Hiroshima City, Hiroshima, Japan; [c]School of Business Administration, Kwansei Gakuin University, Nishinomiya, Hyogo, Japan; [d]Department of Economics and Related Studies, University of York, York, UK; [e]Institute of Social and Economic Research, Osaka University, Ibaraki, Osaka, Japan

ABSTRACT
In this article, we propose a robust approach against heteroscedasticity, error serial correlation and slope heterogeneity in linear models with interactive effects for large panel data. First, consistency and asymptotic normality of the pooled iterated principal component (IPC) estimator for random coefficient and homogeneous slope models are established. Then, we prove the asymptotic validity of the associated Wald test for slope parameter restrictions based on the panel heteroscedasticity and autocorrelation consistent (PHAC) variance matrix estimator for both random coefficient and homogeneous slope models, which does not require the Newey-West type time-series parameter truncation. These results asymptotically justify the use of the same pooled IPC estimator and the PHAC standard error for both homogeneous-slope and heterogeneous-slope models. This robust approach can significantly reduce the model selection uncertainty for applied researchers. In addition, we propose a Lagrange Multiplier (LM) test for correlated random coefficients with covariates. This test has nontrivial power against correlated random coefficients, but not for random coefficients and homogeneous slopes. The LM test is important because the IPC estimator becomes inconsistent with correlated random coefficients. The finite sample evidence and an empirical application support the reliability and the usefulness of our robust approach.

ARTICLE HISTORY
Received December 2020
Accepted May 2022

KEYWORDS
Correlated random coefficients; Interactive effects; Panel data; Slope heterogeneity

1. Introduction

In recent years, the increasing availability of panel data with both a large cross-sectional dimension N and a large time series dimension T has necessitated the development of statistical methods to exploit the rich information they contain, but it has also created technical challenges. In particular, the control of cross-sectional dependence, parameter heterogeneity and serial dependence has been the main focus of the literature. Among other things, the modeling of the cross-sectional dependence by interactive effects has been widely adopted in recent years.

In particular, the past decade has seen significant development in estimation and inferential methods for linear models with interactive effects for large panel data. Two estimation approaches are particularly popular. The first involves eliminating the interactive effects from both the error term and the regressors. Representative methods for static models include the Common Correlated Effects (CCE) approach proposed by Pesaran (2006), the principal component (PC) approach investigated by Westerlund and Urbain (2015), the maximum likelihood (ML) approach of Bai and Li (2014), and the two-step instrumental variable (2SIV) approach proposed by Cui et al. (2021), among others. The second approach asymptotically eliminates the interactive effects from the error term only. The representative method is the iterative principal components (IPC) estimator of Bai (2009), further developed by Moon and Weidner (2015), and Bai and Liao (2017), among many others.

Concerning the $\sqrt{NT}$-consistent estimation and the associated inference against error serial correlation and heteroscedasticity, the CCE, PC, 2SIV, and IPC approaches typically permit idiosyncratic errors to be serially correlated and heteroscedastic (see, Remark 1 in Westerlund and Urbain 2015, Assumptions 2.1 and 2.2 in Cui et al. 2021, Remark 8 in Bai 2009), whereas the ML approach generally does not allow such a degree of robustness (see Assumption B in Bai and Li 2014). The CCE, PC, and IPC estimators, however, will have asymptotic biases, essentially due to the endogeneity arising from estimating the interactive effects. Note that the properties of the asymptotic biases in the CCE and PC estimators and those of the IPC estimator are different, because of the difference in their approaches for eliminating the interactive effects. These asymptotic biases should be controlled, particularly for inference, either analytically or numerically by, for example, jackknife subsampling. For the latter, see Fernández and Weidner (2016) and Westerlund (2018).

CONTACT Takashi Yamagata takashi.yamagata@york.ac.uk Department of Economics and Related Studies, University of York, York YO10 5DD, UK.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/UBES.

For testing linear restrictions on the slope coefficients, most existing work routinely employs the Newey and West (1987) type heteroscedasticity and autocorrelation consistent (HAC) variance-covariance estimator, which requires time-series parameter truncation. This includes the CCE and the PC estimators (see Westerlund 2018), as well as the IPC estimator (see Remark 8 in Bai 2009). The 2SIV estimator of Cui et al. (2021) and other related papers, including Greenaway-McGrevy, Han, and Sul (2012), Bonhomme and Manresa (2015), and Fernández and Weidner (2016), employ the Hansen (2007) type panel HAC (PHAC) variance-covariance estimator *without* the parameter truncation. In this article we employ this PHAC estimator.[1]

One of the attractive features of the CCE, PC, and 2SIV approaches is that they permit estimation and inference for the models with (cross-sectionally) heterogeneous slopes. For such models, the parameter of interest is often the population average of slope coefficients, which is typically estimated as an average of estimators of cross-section specific slopes. This is called a mean-group (MG) estimator. For the MG estimators based on the CCE, PC, and 2SIV approaches, see Pesaran (2006) and Cui et al. (2021). Interestingly, as shown by Pesaran (2006) and Reese and Westerlund (2018), if the cross-sectional variation of slopes is independent of the regressors (i.e., if slopes are random coefficients), the aforementioned pooled CCE and PC estimators are $\sqrt{N}$-consistent to the population average of slope coefficients. Pesaran (2006) proposes a nonparametric estimator of the variance-covariance matrix of the $\sqrt{N}$-consistent pooled CCE estimator. This estimator uses the sample variance-covariance matrix of cross-section specific slope estimates. The evidence therein has shown that this estimator behaves very well in finite samples.[2]

On the other hand, the IPC approaches typically do not permit models with heterogeneous slopes.[3] In this article we investigate asymptotic properties of the pooled IPC estimator for the models with heterogeneous slopes assuming the regressors are subject to a factor structure. To the best of our knowledge, this is new to the literature. In this article, we show that the pooled IPC estimator is $\sqrt{N}$-consistent for the models with heterogeneous slopes (and confirm it is $\sqrt{NT}$-consistent for the models with homogeneous slopes), when the regressors have the factor structure. We also prove that the use of the PHAC variance-covariance matrix estimator for the pooled IPC estimator is asymptotically justified for the models with heterogeneous slopes, as well as for the models with homogeneous slopes. Our findings essentially imply that the inference based on the pooled IPC estimator and the associated PHAC variance estimator is asymptotically valid for both homogeneous slope *and* heterogeneous slope models.

Building upon these novel asymptotic results, we propose a robust approach against heteroscedasticity, error serial correlation and slope heterogeneity in linear models with interactive effects for large panel data. The basic idea behind this approach is classical. For the models with homogeneous slopes and homoscedastic and uncorrelated idiosyncratic errors, the pooled estimator is expected to be most efficient. Otherwise, an alternative estimator, such as a suitable generalized least squares (GLS) estimator, can be more efficient; see Bai and Liao (2017) for the extension of the IPC estimator in this direction. This potential efficiency gain hinges on the correct specification of, in our context, error serial correlation, heteroscedasticity and slope heterogeneity, and achieving such specifications simultaneously does not seem easy in practice. In this kind of situation, the use of the pooled estimator, together with a robust variance-covariance matrix estimator for inference, is very likely to have practical appeal.[4]

Another main contribution of this article is a novel test for correlation and dependency between the random coefficients and covariates (hereafter called correlated random coefficients). For the proposed robust approach to work, the slope heterogeneity, if any, has to be essentially independent of the covariates (i.e., random coefficients). The proposed test is designed to detect departures from this independence and slope homogeneity. Note that the existing slope homogeneity tests for large panel data, such as Pesaran and Yamagata (2008) and Su and Chen (2013), among others, have power against random coefficients as well as correlated random coefficients, and these are not suitable for our purpose.

Here, we briefly discuss a related study, Galvao and Kato (2014), which considers estimation and inference of pseudo-true parameters under some model misspecifications, such as inconsistent estimators in the fixed effect models with correlated random coefficients (see sec. 5.1.2 in Galvao and Kato 2014). Their theoretical results are closely related to ours in that the convergence rate of the pooled estimator changes from $\sqrt{NT}$ to $\sqrt{N}$ under correlated random coefficients. It is important to emphasize that this article deals with consistent estimation of true parameters of interest and the associated inference in the models with interactive effects, which is not explicitly considered in Galvao and Kato (2014). In addition, in contrast to Galvao and Kato (2014), our approach advocates a method that enables empirical researchers to statistically identify misspecified models (i.e., correlated random coefficient models) that can result in inconsistent estimation.

We have examined the finite sample performance of the estimators, the tests of parameter restrictions using the PHAC variance estimator, and the LM test for correlated random coefficients. The results show that the size of the proposed robust Wald test with the bias-corrected IPC estimator is sufficiently close to the nominal level in both slope homogeneity and slope heterogeneity, and that the LM test for correlated random coefficients has correct size under both slope homogeneity and random coefficients while exhibiting high power for correlated random coefficients. The practical usefulness of our robust approach is illustrated by applying the proposed methods to analyze the Feldstein–Horioka puzzle, presented in Feldstein and Horioka (1980).

[1] In the fixed T panel data analysis, the PHAC estimator is proposed by Arellano (1987) and widely used in the related literature.

[2] In this article, we propose to use the PHAC estimator for the $\sqrt{N}$-consistent estimator, which is different from Pesaran (2006) and Reese and Westerlund (2018). See Section 2.

[3] One important exception is Ando and Bai (2017), in which shrinkage estimators for sparse heterogeneous slopes in the models with interactive effects are considered.

[4] In this article, we do not consider the model with the group-wise homogeneous slope model, which is considered by Su and Ju (2018).

The article is organized as follows. The asymptotic properties of the IPC estimator and the associated robust Wald test statistic based on the PHAC variance-covariance estimator are investigated in Section 2. In Section 3, we introduce the LM test statistic for correlated random coefficients and study its limiting distribution. The finite sample performance of the IPC estimators, the robust Wald test and the LM test for correlated random coefficients is investigated using the Monte Carlo method in Section 4. An empirical application is provided in Section 5. Section 6 contains some concluding remarks. Proofs of the main results, additional discussions and experimental results are relegated to the supplementary materials.

Notations: we denote the largest eigenvalues of the $N \times N$ matrix $\mathbf{A} = (a_{ij})$ by $\mu_{\max}(\mathbf{A})$, its trace by $\text{tr}(\mathbf{A}) = \sum_{i=1}^{N} a_{ii}$, its Frobenius norm by $\|\mathbf{A}\| = \sqrt{\text{tr}(\mathbf{A}'\mathbf{A})}$. The projection matrix on $\mathbf{A}$ is $\mathbf{P_A} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ and $\mathbf{M_A} = \mathbf{I} - \mathbf{P_A}$. Δ is a generic positive constant large enough, $\delta_{NT}^2 = \min\{N, T\}$. We use $N, T \to \infty$ to denote that N and T pass to infinity jointly. We use $\lesssim (\gtrsim)$ to represent $\leq (\geq)$ up to a positive constant factor. For any positive sequences a_n and b_n that converge to some points or diverge as $n \to \infty$, we write $a_n \asymp b_n$ if $a_n \lesssim b_n$ and $a_n \gtrsim b_n$.

2. Panel Data Models with Interactive Effects

We consider the following panel data model with possibly heterogeneous coefficients,

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta}_i + \boldsymbol{\lambda}_i^{0\prime}\mathbf{f}_t^0 + \varepsilon_{it}, \quad (i = 1, 2, \ldots, N; t = 1, 2, \ldots, T), \tag{1}$$

$$\mathbf{x}_{it} = \boldsymbol{\Gamma}_i^{0\prime}\mathbf{g}_t^0 + \mathbf{v}_{it}, \tag{2}$$

where $\mathbf{x}_{it}$ is a $k \times 1$ vector of regressors, $\mathbf{f}_t^0$ and $\mathbf{g}_t^0$ denote $r_1 \times 1$ and $r_2 \times 1$ vectors of latent factors, respectively. Correspondingly, their factor loadings are $\boldsymbol{\lambda}_i^0$ and $\boldsymbol{\Gamma}_i^0 = (\boldsymbol{\gamma}_{1i}^0, \ldots, \boldsymbol{\gamma}_{ki}^0)$. ε_{it} and $\mathbf{v}_{it}$ are the idiosyncratic disturbance terms.

If we stack the Equations (1) and (2) over t, we have

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{F}^0\boldsymbol{\lambda}_i^0 + \boldsymbol{\varepsilon}_i, \tag{3}$$

$$\mathbf{X}_i = \mathbf{G}^0\boldsymbol{\Gamma}_i^0 + \mathbf{V}_i, \tag{4}$$

where $\mathbf{y}_i = (y_{i1}, \ldots, y_{iT})'$, $\mathbf{X}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iT})'$, $\mathbf{F}^0 = (\mathbf{f}_1^0, \ldots, \mathbf{f}_T^0)'$, $\mathbf{G}^0 = (\mathbf{g}_1^0, \ldots, \mathbf{g}_T^0)'$, $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iT})'$ and $\mathbf{V}_i = (\mathbf{v}_{i1}, \ldots, \mathbf{v}_{iT})'$. Supposing random coefficients

$$\boldsymbol{\beta}_i = \boldsymbol{\beta}^0 + \boldsymbol{\eta}_i, \tag{5}$$

we can rewrite the model (3) using the process (4) as

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}^0 + \mathbf{W}^0\boldsymbol{\vartheta}_i^0 + \mathbf{V}_i\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i, \tag{6}$$

where $\mathbf{W}^0 = (\mathbf{G}^0, \mathbf{F}^0)$ and $\boldsymbol{\vartheta}_i^0 = (\boldsymbol{\eta}_i'\boldsymbol{\Gamma}_i^{0\prime}, \boldsymbol{\lambda}_i^{0\prime})'$. The results in this article will hold when $\mathbf{F}^0$ and $\mathbf{G}^0$ share common factors (say, $\mathbf{W}^0 = \mathbf{G}^0 \cup \mathbf{F}^0$ and $\mathbf{F}^0 \cap \mathbf{G}^0 \neq \mathbf{0}$). However, for simplicity, we suppose $\mathbf{F}^0 \cap \mathbf{G}^0 = \mathbf{0}$ and $\text{rank}(\mathbf{W}^0) = r_1 + r_2$.

Noting that for the IPC estimator the interactive effects $\mathbf{F}^0\boldsymbol{\lambda}_i^0$ for the models with homogeneous slopes and $\mathbf{W}^0\boldsymbol{\vartheta}_i^0$ for the models with heterogeneous slopes are extracted from the error vector $\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i^0$, we employ the following common expression to describe both homogeneous-slope and heterogeneous-slope models:

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}^0 + \mathbf{u}_i, \text{ with } \mathbf{u}_i = \mathbf{H}^0\boldsymbol{\phi}_i^0 + \mathbf{e}_i, \tag{7}$$

where $\{\mathbf{H}^0, \boldsymbol{\phi}_i^0, \mathbf{e}_i\} = \{\mathbf{F}^0, \boldsymbol{\lambda}_i^0, \boldsymbol{\varepsilon}_i\}$ and $r = r_1$ for the models with homogeneous slopes and $\{\mathbf{H}^0, \boldsymbol{\phi}_i^0, \mathbf{e}_i\} = \{\mathbf{W}^0, \boldsymbol{\vartheta}_i^0, \mathbf{V}_i\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i\}$ and $r = r_1 + r_2$ for the models with heterogeneous slopes. Hereafter, we use Equation (7) to describe both homogeneous-slope and homogeneous-slope models.

Controlling the interactive effects, $\mathbf{H}^0\boldsymbol{\phi}_i^0$, typically results in asymptotic biases of the $\sqrt{NT}$-consistent CCE, PC, and IPC estimators, as shown by Pesaran (2006), Bai (2009), and Westerlund and Urbain (2015), among others.

As discussed in the introduction section, in this article we focus on the iterative principal component (IPC) estimator proposed by Bai (2009), which estimates the interactive effects in $\mathbf{u}_i$ by the PC method and the slope coefficient $\boldsymbol{\beta}^0$ by the least-squares method, iteratively.

In particular, define the least squares objective function

$$\text{SSR}(\boldsymbol{\beta}, \mathbf{H}) = \frac{1}{NT}\sum_{i=1}^{N}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{H}\boldsymbol{\phi}_i)'(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{H}\boldsymbol{\phi}_i)$$

subject to the constraints $\mathbf{H}'\mathbf{H}/T = \mathbf{I}_r$ and $\sum_{i=1}^{N}\boldsymbol{\phi}_i\boldsymbol{\phi}_i'$ being diagonal.

The least squares estimator for $(\boldsymbol{\beta}, \mathbf{H})$, denoted by $(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{H}})$, is the IPC estimator, which is the solution to the following nonlinear equations:

$$\widehat{\boldsymbol{\beta}} = \left(\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{H}}}\mathbf{X}_i\right)^{-1}\left(\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{H}}}\mathbf{y}_i\right)$$

$$\left[\frac{1}{NT}\sum_{i=1}^{N}\left(\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}\right)\left(\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}\right)'\right]\widehat{\mathbf{H}} = \widehat{\mathbf{H}}\boldsymbol{\mathcal{V}}_{NT} \tag{8}$$

where $\boldsymbol{\mathcal{V}}_{NT}$ is a diagonal matrix that contains the r largest eigenvalues of the matrix in the square brackets in decreasing order. Given $(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{H}})$, we can estimate $\boldsymbol{\phi}_i$ by

$$\widehat{\boldsymbol{\phi}}_i = \frac{1}{T}\widehat{\mathbf{H}}'(\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}).$$

The number of factors, r, and an initial estimator of $\boldsymbol{\beta}^0$, should be chosen to obtain the IPC estimator. In this section and the next, we suppose that r and the initial estimator of $\boldsymbol{\beta}^0$ are given, and the procedure to choose them is discussed in Section 4.1.

Remark 1. Permitting different sets of interactive effects in $\mathbf{X}_i$ and $\mathbf{u}_i$ is important, because the IPC estimator extracts the interactive effects in $\mathbf{u}_i$ only, which does not necessarily project out the factors in $\mathbf{X}_i$. Nevertheless, our results in this article hold when $\mathbf{X}_i$ and $\mathbf{u}_i$ share the same set of interactive effects. See Norkutė et al. (2021) and Cui et al. (2021) for further discussions.

Remark 2. The factor model for the regressors specified in (2) has been widely employed in the literature: see Pesaran (2006), Bai and Li (2014), Westerlund and Urbain (2015), Li, Cui, and Lu (2020), Norkutė et al. (2021), among many others. For the IPC estimators, typically no factor structure in the regressors is imposed. However, only the model with homogeneous slopes is considered; see Bai (2009) and Moon and Weidner (2015), among others. Importantly, our new results shown below will reveal that, the pooled IPC estimator remains consistent for the random coefficient models with interactive effects if the regressors have a factor structure as in (2).

For the asymptotic analysis of the estimator, we impose the following assumptions.

Assumption 1 (idiosyncratic error in y). (i) ε_{it} is distributed independently across i; (ii) $\mathbb{E}(\varepsilon_{it}) = 0$ and $\mathbb{E}|\varepsilon_{it}|^8 \leq \Delta$; (iii) $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}\mathbb{E}|\varepsilon_{is}\varepsilon_{it}| \leq \Delta$; (iv) $\mathbb{E}|N^{-1/2}\sum_{i=1}^{N}[\varepsilon_{is}\varepsilon_{it} - \mathbb{E}(\varepsilon_{is}\varepsilon_{it})]|^4 \leq \Delta$ for each (s,t); (v) $N^{-1}T^{-2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{w=1}^{T}|\text{cov}(\varepsilon_{is}\varepsilon_{it},\varepsilon_{ir}\varepsilon_{iw})| \leq \Delta$; (vi) $\boldsymbol{\Omega}_{\varepsilon,i} = \mathbb{E}(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i')$ is positive definite and its largest eigenvalue is bounded, uniformly every i and T.

Assumption 2 (idiosyncratic error in x). Let $v_{\ell it}$ be the ℓth element of $\mathbf{v}_{it}$ and $\mathbf{v}_{\ell,i} = (v_{\ell i1},\ldots,v_{\ell iT})'$. Then we assume that (i) $\mathbf{v}_{it}$ is independently distributed across i and independent of $\{\varepsilon_{js}\}$ for $1 \leq j \leq N$ and $1 \leq s \leq T$; (ii) $\mathbb{E}(\mathbf{v}_{it}) = \mathbf{0}$ and $\mathbb{E}\|\mathbf{v}_{it}\|^8 \leq \Delta$; (iii) $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}\mathbb{E}|v_{\ell is}v_{\ell it}| \leq \Delta$; (iv) $\mathbb{E}|N^{-1/2}\sum_{i=1}^{N}[v_{\ell is}v_{\ell it} - \mathbb{E}(v_{\ell is}v_{\ell it})]|^4 \leq \Delta$ for every ℓ, s, and t; (v) $N^{-1}T^{-2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{w=1}^{T}|\text{cov}(v_{\ell is}v_{\ell it}, v_{\ell ir}v_{\ell iw})| \leq \Delta$; (vi) the largest eigenvalue of $\mathbb{E}(\mathbf{v}_{\ell,i}\mathbf{v}_{\ell,i}')$ is bounded uniformly for every ℓ, i, and T.

Assumption 3 (random coefficients). $\boldsymbol{\eta}_i$ is independent across i, and is independent of ε_{it}, $\mathbf{v}_{it}$, $\boldsymbol{\lambda}_i^0$ and $\boldsymbol{\Gamma}_i^0$; $\mathbb{E}(\boldsymbol{\eta}_i) = \mathbf{0}$ and $\mathbb{E}\|\boldsymbol{\eta}_i\|^4 \leq \Delta$; $\mathbb{E}(\boldsymbol{\eta}_i\boldsymbol{\eta}_i') = \boldsymbol{\Omega}_{\eta,i}$ is a fixed positive definite matrix uniformly over i.

Assumption 4 (factor components). (i) Let $\mathbf{w}_t^0 = (\mathbf{f}_t^{0\prime},\mathbf{g}_t^{0\prime})'$. $\mathbb{E}\|\mathbf{w}_t^0\|^4 \leq \Delta$ and $T^{-1}\sum_{t=1}^{T}\mathbf{w}_t^0\mathbf{w}_t^{0\prime} \xrightarrow{p} \boldsymbol{\Sigma}_w = \begin{bmatrix} \boldsymbol{\Sigma}_f & \boldsymbol{\Sigma}_{fg} \\ \boldsymbol{\Sigma}_{gf} & \boldsymbol{\Sigma}_g \end{bmatrix}$, which is positive definite and $\mathbf{w}_t^0$ is independent of $\mathbf{v}_{it}$ and ε_{it}; (ii) $\mathbb{E}\|\boldsymbol{\Gamma}_i^0\|^4 \leq \Delta$, $\mathbb{E}\|\boldsymbol{\lambda}_i^0\|^4 \leq \Delta$, $N^{-1}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\boldsymbol{\Gamma}_i^{0\prime} \xrightarrow{p} \boldsymbol{\Sigma}_\Gamma$, $N^{-1}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\boldsymbol{\eta}_i\boldsymbol{\eta}_i'\boldsymbol{\Gamma}_i^{0\prime} \xrightarrow{p} \boldsymbol{\Sigma}_{\gamma\eta}$ and $N^{-1}\sum_{i=1}^{N}\boldsymbol{\lambda}_i^0\boldsymbol{\lambda}_i^{0\prime} \xrightarrow{p} \boldsymbol{\Sigma}_\lambda$, where $\boldsymbol{\Sigma}_\Gamma$, $\boldsymbol{\Sigma}_{\gamma\eta}$, and $\boldsymbol{\Sigma}_\lambda$ are positive definite matrices. $\boldsymbol{\Gamma}_i^0$ and $\boldsymbol{\lambda}_i^0$ are independent of $\mathbf{v}_{it}$ and ε_{it}.

Assumption 5 (identification and variance matrices). We assume that the four matrices

$$\mathbf{A}_0 = \lim_{N,T\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\mathbb{E}(\mathbf{V}_i'\mathbf{V}_i),$$

$$\mathbf{A}_1 = \operatorname*{plim}_{N,T\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{Z}_i,$$

$$\mathbf{C}_0 = \lim_{N,T\to\infty}\frac{1}{NT^2}\sum_{i=1}^{N}\mathbb{E}(\mathbf{V}_i'\mathbf{V}_i\boldsymbol{\eta}_i\boldsymbol{\eta}_i'\mathbf{V}_i'\mathbf{V}_i),$$

$$\mathbf{B}_1 = \operatorname*{plim}_{N,T\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\boldsymbol{\Omega}_{\varepsilon,i}\mathbf{M}_{\mathbf{H}^0}\mathbf{Z}_i$$

are fixed and positive definite, where $\mathbf{Z}_i = \mathbf{X}_i - N^{-1}\sum_{j=1}^{N}\boldsymbol{\phi}_i^{0\prime}\boldsymbol{\Upsilon}_{\phi^0}^{-1}\boldsymbol{\phi}_j^0\mathbf{X}_j$, $\boldsymbol{\Upsilon}_{\phi^0} = N^{-1}\sum_{i=1}^{N}\boldsymbol{\phi}_i^0\boldsymbol{\phi}_i^{0\prime}$.

Assumption 6 (Central Limit Theorem). When $N,T\to\infty$ such that $T/N \to \rho \in (0,\Delta]$, $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\boldsymbol{\varepsilon}_i \xrightarrow{d} N(\mathbf{0},\mathbf{B}_1)$.

Assumptions 1 and 2 are very similar to Assumption in Bai (2009), though, ε_{it} and $\mathbf{v}_{it}$ are independent of each other and independent over i. Assumption 3 is a typical assumption for random coefficient models with interactive effects (see, e.g., Assumption HET in Reese and Westerlund 2018). Assumption 4 implies that the factor loadings $\boldsymbol{\Gamma}_i^0$ and $\boldsymbol{\lambda}_i^0$ are allowed to be correlated with each other. Assumption 5 is for identification of the parameters and for estimation of the variance-covariance matrices. Assumption 6 is imposed following Bai (2009) for the asymptotic normality of the pooled IPC estimator for the models with homogeneous slopes.

In what follows, we investigate asymptotic properties of the IPC estimator for the models with homogeneous and heterogeneous slopes, assuming that the regressors have a factor structure. As discussed in the introduction section, the results for the models with heterogeneous slopes are new additions to the literature. We first derive asymptotic expressions of the pooled IPC estimator, $\widehat{\boldsymbol{\beta}}$, for the models with homogeneous and heterogeneous slopes in Proposition 1. These confirm the presence of asymptotic bias in $\widehat{\boldsymbol{\beta}}$ for the models with homogeneous slopes, as shown by Bai (2009) and Cui et al. (2021), among others. After introducing an asymptotic expression for the bias-corrected IPC estimator, which is common to the models with homogeneous and heterogeneous slopes in Proposition 2, we present asymptotic normality of the bias-corrected IPC estimator for both models in Theorem 1. Finally, in Theorem 2, we show the asymptotic validity of the robust Wald test for linear constraints on $\boldsymbol{\beta}^0$, which is based on the same statistic for both homogeneous-slope and heterogeneous-slope models.

A similar line of discussion in Bai (2009) and Cui et al. (2021) for the model with homogeneous slopes, and in Pesaran (2006) for the case of heterogeneous slopes provides the following asymptotic expressions:

Proposition 1. Suppose that Assumptions 1–5 hold.

(a) For the model with homogeneous slopes, we have

$$\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0 = \left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{Z}_i\right)^{-1}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\boldsymbol{\varepsilon}_i + \frac{1}{N}\boldsymbol{\xi}_{NT} + \frac{1}{T}\boldsymbol{\zeta}_{NT} + O_p(\delta_{NT}^{-3}), \tag{9}$$

where the bias terms are given by

$$\boldsymbol{\xi}_{NT} = -\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{Z}_i\right)^{-1}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{Z}_i'\mathbf{H}^0(\mathbf{H}^{0\prime}\mathbf{H}^0)^{-1}\boldsymbol{\Upsilon}_{\phi^0}^{-1}\boldsymbol{\phi}_i^0\mathbb{E}(e_{it}^2),$$

$$\boldsymbol{\zeta}_{NT} = -\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{Z}_i\right)^{-1}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{X}_i'\mathbf{M}_{\mathbf{H}^0}\mathbb{E}(\mathbf{e}_j\mathbf{e}_j')\mathbf{H}^0(\mathbf{H}^{0\prime}\mathbf{H}^0)^{-1}\boldsymbol{\Upsilon}_{\phi^0}^{-1}\boldsymbol{\phi}_i^0.$$

(b) For the model with heterogeneous slopes, we have

$$\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0 = \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{V}_i' \mathbf{V}_i \right)^{-1} \frac{1}{NT} \sum_{i=1}^{N} \mathbf{V}_i' \mathbf{V}_i \boldsymbol{\eta}_i + O_p(\delta_{NT}^{-2}). \tag{10}$$

A few remarks regarding Proposition 1 are warranted. Proposition 1(a) is essentially the same result for the models with the homogeneous slopes shown in Bai (2009) and Cui et al. (2021). Because of the estimation of the interactive effects, $\mathbf{X}_i$ has to be replaced by $\mathbf{Z}_i$, and the bias terms $\boldsymbol{\xi}_{NT}$ and $\boldsymbol{\zeta}_{NT}$ are present. Under our assumptions $\boldsymbol{\xi}_{NT}$ and $\boldsymbol{\zeta}_{NT}$ are $O_p(1)$, because $\mathbf{Z}_i$ can be correlated with $\boldsymbol{\phi}_i^0$ and $\mathbb{E}(\mathbf{e}_j \mathbf{e}_j')$ is not necessarily proportional to $\mathbf{I}_T$. Proposition 1(b) reveals that the pooled IPC estimator is $\sqrt{N}$-consistent and asymptotically free from the effects of interactive effects. See a similar result for pooled CCE and PC estimators in Pesaran (2006) and Reese and Westerlund (2018).

To derive an asymptotically unbiased estimator of $\boldsymbol{\beta}^0$ in the case of homogeneous slope, we should estimate the bias terms. Following Bai (2009) and Moon and Weidner (2015), we propose an analytically bias corrected estimator, which is defined by

$$\widetilde{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}} - \frac{1}{N} \widehat{\boldsymbol{\xi}}_{NT} - \frac{1}{T} \widehat{\boldsymbol{\zeta}}_{NT}, \tag{11}$$

where

$$\widehat{\boldsymbol{\xi}}_{NT} = -\left(\frac{1}{NT} \sum_{i=1}^{N} \widehat{\mathbf{Z}}_i' \mathbf{M}_{\widehat{\mathbf{H}}} \widehat{\mathbf{Z}}_i \right)^{-1} \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} \widehat{\mathbf{Z}}_i' \widehat{\mathbf{H}} \widehat{\boldsymbol{\Upsilon}}_{\phi}^{-1} \widehat{\boldsymbol{\phi}}_i \widehat{e}_{it}^2$$

$$\widehat{\boldsymbol{\zeta}}_{NT} = -\left(\frac{1}{NT} \sum_{i=1}^{N} \widehat{\mathbf{Z}}_i' \mathbf{M}_{\widehat{\mathbf{H}}} \widehat{\mathbf{Z}}_i \right) \frac{1}{N} \sum_{i=1}^{N} \frac{\widehat{\mathbf{X}_i' \mathbf{M}_{\mathbf{H}} \boldsymbol{\Omega} \mathbf{H}}}{T} \widehat{\boldsymbol{\Upsilon}}_{\phi}^{-1} \widehat{\boldsymbol{\phi}}_i$$

with $\widehat{\mathbf{Z}}_i = \mathbf{X}_i - N^{-1} \sum_{j=1}^{N} \widehat{\boldsymbol{\phi}}_i' \widehat{\boldsymbol{\Upsilon}}_{\phi}^{-1} \widehat{\boldsymbol{\phi}}_j \mathbf{X}_j$, $\widehat{\boldsymbol{\Upsilon}}_{\phi} = N^{-1} \sum_{i=1}^{N} \widehat{\boldsymbol{\phi}}_i \widehat{\boldsymbol{\phi}}_i'$,

$$\frac{\widehat{\mathbf{X}_i' \mathbf{M}_{\mathbf{H}} \boldsymbol{\Omega} \mathbf{H}}}{T} = \frac{1}{TN} \sum_{j=1}^{N} \left[\sum_{t=1}^{T} \widehat{e}_{jt}^2 \widehat{\mathbf{x}}_{it} \widehat{\mathbf{h}}_t' + \sum_{s=1}^{S} \sum_{t=s+1}^{T} \left(1 - \frac{s}{S+1} \right) \widehat{e}_{jt} \widehat{e}_{j,t-s} \left(\widehat{\mathbf{x}}_{it} \widehat{\mathbf{h}}_{t-s}' + \widehat{\mathbf{x}}_{i,t-s} \widehat{\mathbf{h}}_t' \right) \right] \tag{12}$$

and $\widehat{e}_{it}$ being the tth element of $\widehat{\mathbf{e}}_i = \mathbf{y}_i - \mathbf{X}_i \widehat{\boldsymbol{\beta}} - \widehat{\mathbf{H}} \widehat{\boldsymbol{\phi}}_i$.

Following Bai's (2009) suggestion we employ the Newey and West (1987) type estimator of $\boldsymbol{\zeta}_{NT}$ as shown in Equation (12). In the Monte Carlo experiment in Section 4, we set $S = \lfloor T^{1/4} \rfloor$.

Remark 3. Recently, jackknife bias-correction, a numerical way of asymptotically eliminating the asymptotic biases, was proposed and applied in Fernández and Weidner (2016) and Westerlund (2018), among others. The bias terms in Equation (11) are $O_p(N^{-1})$ and $O_p(T^{-1})$, which indicates that the jackknife bias-correction of Dhaene and Jochmans (2015) by splitting the time-series and cross-section dimensions would be applicable for $\widehat{\boldsymbol{\beta}}$. We provide the procedure for this jackknife bias-correction and examine its finite sample performance in Section 4.

Now consider the robust approach against slope heterogeneity. Such robustness may be achieved when the leading terms in Equations (9) and (10) share a common expression, and the estimated asymptotic bias terms are bounded for the models with heterogeneous slopes. Indeed, we prove that for the models with heterogeneous slopes $\frac{1}{NT} \sum_{i=1}^{N} \mathbf{Z}_i' \mathbf{M}_{\mathbf{H}^0} \mathbf{Z}_i = \frac{1}{NT} \sum_{i=1}^{N} \mathbf{V}_i' \mathbf{V}_i + o_p(1)$, $\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{Z}_i' \mathbf{M}_{\mathbf{H}^0} \mathbf{u}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i' \mathbf{V}_i \boldsymbol{\eta}_i + o_p(1)$, $\widehat{\boldsymbol{\zeta}}_{NT} = O_p(1)$ and $\widehat{\boldsymbol{\zeta}}_{NT} = O_p(1)$, and for the models with homogeneous slopes $\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{Z}_i' \mathbf{M}_{\mathbf{H}^0} \mathbf{u}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{Z}_i' \mathbf{M}_{\mathbf{H}^0} \boldsymbol{\varepsilon}_i$ under Assumptions 1–5; see the supplementary materials.

The following proposition shows a common stochastic representation of the bias-corrected IPC estimator for the models with homogeneous and heterogeneous slopes.

Proposition 2. Suppose that Assumptions 1–5 hold. Then, we have

$$\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^0 = \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{Z}_i' \mathbf{M}_{\mathbf{H}^0} \mathbf{Z}_i \right)^{-1} \times \frac{1}{NT} \sum_{i=1}^{N} \mathbf{Z}_i' \mathbf{M}_{\mathbf{H}^0} \mathbf{u}_i + O_p(\delta_{NT}^{-\epsilon}) \tag{13}$$

where $\epsilon = 3$ for the models with homogeneous slopes and $\epsilon = 2$ for the models with heterogeneous slopes, as $(N, T) \to \infty$ and $T/N \to \rho \in (0, \Delta]$.

Equipped with Assumption 6, we are ready to present the asymptotic normality of the bias-corrected estimator $\widetilde{\boldsymbol{\beta}}$.

Theorem 1. Suppose that Assumptions 1–6 hold and $T/N \to \rho \in (0, \Delta]$.

(a) For the model with homogeneous slopes, we have

$$\sqrt{NT}(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^0) \xrightarrow{d} N(\mathbf{0}, \mathbf{A}_1^{-1} \mathbf{B}_1 \mathbf{A}_1^{-1}).$$

(b) For the model with heterogeneous slopes, we have

$$\sqrt{N}(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^0) \xrightarrow{d} N(\mathbf{0}, \mathbf{A}_0^{-1} \mathbf{C}_0 \mathbf{A}_0^{-1}),$$

where $\mathbf{A}_0$, $\mathbf{A}_1$, $\mathbf{B}_1$, and $\mathbf{C}_0$ are defined in Assumption 5.

Let us now turn our attention to the estimation of the asymptotic variance-covariance matrix of $\widetilde{\boldsymbol{\beta}}$. Define the panel heteroscedasticity and autocorrelation consistent (PHAC) variance-covariance matrix estimator by

$$\widehat{\boldsymbol{\Sigma}}_{\widetilde{\beta}} = \left(\sum_{i=1}^{N} \widehat{\mathbf{Z}}_i' \mathbf{M}_{\widehat{\mathbf{H}}} \widehat{\mathbf{Z}}_i \right)^{-1} \left(\sum_{i=1}^{N} \widehat{\mathbf{Z}}_i' \mathbf{M}_{\widehat{\mathbf{H}}} \widehat{\mathbf{u}}_i \widehat{\mathbf{u}}_i' \mathbf{M}_{\widehat{\mathbf{H}}} \widehat{\mathbf{Z}}_i \right) \times \left(\sum_{i=1}^{N} \widehat{\mathbf{Z}}_i' \mathbf{M}_{\widehat{\mathbf{H}}} \widehat{\mathbf{Z}}_i \right)^{-1}, \tag{14}$$

where $\widehat{\mathbf{u}}_i = \mathbf{y}_i - \mathbf{X}_i \widehat{\boldsymbol{\beta}}$. Note that this PHAC estimator does not require Newey and West (1987) type time-series parameter truncation.

One of the highlights of this article is the revelation of the asymptotic validity of the PHAC variance estimator for both homogeneous-slope and heterogeneous-slope models with interactive effects.[5] This may be intuitively understood by the following discussion, with the help of Propositions 1 and 2 and Theorem 1. For the models with homogeneous slopes, Equations (9) and (13) yield

$$\begin{aligned}\sqrt{NT}(\widetilde{\boldsymbol{\beta}}-\boldsymbol{\beta}^0) &= \left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{Z}_i\right)^{-1} \\ &\quad\times\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{u}_i+o_p(1) \\ &= \left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{Z}_i\right)^{-1} \\ &\quad\times\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\boldsymbol{\varepsilon}_i+o_p(1).\end{aligned}$$

From this, we can easily infer that $NT\widehat{\boldsymbol{\Sigma}}_{\widetilde{\beta}}\xrightarrow{p}\mathbf{A}_1^{-1}\mathbf{B}_1\mathbf{A}_1^{-1}$. For the models with heterogeneous slopes, Equations (10) and (13) give

$$\begin{aligned}\sqrt{N}(\widetilde{\boldsymbol{\beta}}-\boldsymbol{\beta}^0) &= \left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{Z}_i\right)^{-1} \\ &\quad\times\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{u}_i+o_p(1) \\ &= \left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i\right)^{-1} \\ &\quad\times\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i\boldsymbol{\eta}_i+o_p(1), \qquad (15)\end{aligned}$$

from which, we can conjecture that $N\widehat{\boldsymbol{\Sigma}}_{\widetilde{\beta}}\xrightarrow{p}\mathbf{A}_0^{-1}\mathbf{C}_0\mathbf{A}_0^{-1}$.

In the following theorem, we formally establish the limit distribution of the Wald test statistic for linear restrictions on $\boldsymbol{\beta}^0$, based on the bias-corrected IPC estimator $\widetilde{\boldsymbol{\beta}}$ and the PHAC variance estimator $\widehat{\boldsymbol{\Sigma}}_{\widetilde{\beta}}$.

Theorem 2. Consider testing m linearly independent restrictions on $\boldsymbol{\beta}^0$, $H_0:\boldsymbol{\mathcal{R}}\boldsymbol{\beta}^0=\boldsymbol{r}$ against $H_1:\boldsymbol{\mathcal{R}}\boldsymbol{\beta}^0\neq\boldsymbol{r}$, where $\boldsymbol{\mathcal{R}}$ is an $m\times k$ fixed matrix of full row rank. Consider the model (7) and the Wald statistic

$$\widetilde{W}_{NT}=(\boldsymbol{\mathcal{R}}\widetilde{\boldsymbol{\beta}}-\boldsymbol{r})'\left(\boldsymbol{\mathcal{R}}\widehat{\boldsymbol{\Sigma}}_{\widetilde{\beta}}\boldsymbol{\mathcal{R}}'\right)^{-1}(\boldsymbol{\mathcal{R}}\widetilde{\boldsymbol{\beta}}-\boldsymbol{r}) \qquad (16)$$

where $\widetilde{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\Sigma}}_{\widetilde{\beta}}$ are defined by (11) and (14), respectively. Suppose that Assumptions 1–6 hold. Then, under the null hypothesis, for the models with homogeneous slopes ($\boldsymbol{\eta}_i=\mathbf{0}$ for all i) and for the models with heterogeneous slopes, $\widetilde{W}_{NT}\xrightarrow{d}\chi_m^2$, as $(N,T)\to\infty$ and $T/N\to\rho\in(0,\Delta]$.

To our knowledge, this is the first article which shows the consistency of the PHAC estimator in the random coefficients models with interactive effects for large panel data. As discussed in the introduction section, the use of the PHAC variance estimator for the models with homogeneous slopes with large N and T is proposed by Hansen (2007) for the fixed effects models, and considered by Greenaway-McGrevy, Han, and Sul (2012), Vogelsang (2012), Fernández and Weidner (2016), and Cui et al. (2021) for the models with homogeneous slopes, among others.

Remark 4. Our approach is also robust against mixtures of homogeneous and heterogeneous slopes.[6] To illustrate this, consider the model without common components and the case in which the k slopes are partitioned in such a way that $k=k_1+k_2$, where $\boldsymbol{\beta}_i=(\boldsymbol{\beta}_{1i}',\boldsymbol{\beta}_2^{0\prime})'$, $\boldsymbol{\beta}_{1i}=\boldsymbol{\beta}_1^0+\boldsymbol{\eta}_{1i}$, $\mathbb{E}(\boldsymbol{\eta}_{1i})=\mathbf{0}$ and $\text{var}(\boldsymbol{\eta}_{1i})=\boldsymbol{\Omega}_{1i}$, with $\boldsymbol{\beta}^0=(\boldsymbol{\beta}_1^{0\prime},\boldsymbol{\beta}_2^{0\prime})'$. The expansion of the pooled ordinary least squares (OLS) estimator $\widehat{\boldsymbol{\beta}}=\left(\widehat{\boldsymbol{\beta}}_1',\widehat{\boldsymbol{\beta}}_2'\right)'$ gives

$$\begin{aligned}\sqrt{N}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right) &= \left(\frac{\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i}{NT}\right)^{-1}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{u}_i \\ &= \left(\frac{\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i}{NT}\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\mathbf{X}_{1i}}{T}\right)\boldsymbol{\eta}_{1i} \\ &\quad+O_p\left(1/\sqrt{T}\right) \\ &\xrightarrow{d} N\left(\mathbf{0},\mathbf{A}^{-1}\mathbf{C}\mathbf{A}^{-1}\right)\end{aligned}$$

where $\mathbf{A}=\operatorname*{plim}_{N,T\to\infty}(NT)^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i$, which is accordingly partitioned as

$$\mathbf{A}=\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix}, \quad\text{and}$$

$$\mathbf{C}=\operatorname*{plim}_{N,T\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\mathbf{X}_{1i}}{T}\right)\boldsymbol{\eta}_{1i}\boldsymbol{\eta}_{1i}'\left(\frac{\mathbf{X}_{1i}'\mathbf{X}_i}{T}\right),$$

which are assumed to be fixed and positive definite. Also note that the convergence rate of ($\widehat{\boldsymbol{\beta}}_1$ and) $\widehat{\boldsymbol{\beta}}_2$ is $\sqrt{N}$, as the variation of $\widehat{\boldsymbol{\beta}}$ is dominated by that of $\boldsymbol{\eta}_{1i}$.

Now consider a special case in which $\text{plim}_{T\to\infty}\mathbf{X}_{1i}'\mathbf{X}_{2i}=\mathbf{0}$. Then, a similar discussion gives $\sqrt{N}\left(\hat{\boldsymbol{\beta}}_1-\boldsymbol{\beta}_1^0\right)\xrightarrow{d}N\left(\mathbf{0},\mathbf{A}_{11}^{-1}\mathbf{C}_{11}\mathbf{A}_{11}^{-1}\right)$ and $\sqrt{NT}\left(\hat{\boldsymbol{\beta}}_2-\boldsymbol{\beta}_2^0\right)\xrightarrow{d}N\left(\mathbf{0},\mathbf{A}_{22}^{-1}\mathbf{B}_{22}\mathbf{A}_{22}^{-1}\right)$, where $\mathbf{C}_{11}=\text{plim}N,T\to\infty(NT^2)^{-1}\sum_{i=1}^{N}\mathbf{X}_{1i}'\mathbf{X}_{1i}\boldsymbol{\eta}_{1i}\boldsymbol{\eta}_{1i}'\mathbf{X}_{1i}'\mathbf{X}_{1i}$ and $\mathbf{B}_{22}=\text{plim}_{N,T\to\infty}(NT)^{-1}\sum_{i=1}^{N}\mathbf{X}_{2i}'\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{X}_{2i}$. Therefore, the asymptotic normality of $\widehat{\boldsymbol{\beta}}$, the consistency of the PHAC estimator and the asymptotic validity of Wald test hold with mixtures of homogeneous and heterogeneous slopes.

[5] See Galvao and Kato (2014) for a similar discussion in misspecified dynamic panel data models.

[6] We do not consider cross-sectional and/or time-series structural breaks in $\boldsymbol{\beta}_i$ which is beyond the scope of this article.

3. LM Test for Correlated Random Coefficients with Covariates

As discussed earlier, the proposed robust approach is asymptotically justified for the models with homogeneous slopes and for random coefficient models, in which the slope heterogeneity is generally independent of the variations of the regressors (see Assumption 3). In particular, when the heterogeneous coefficients are correlated with the defactored regressor matrix, $\mathbf{M}_{\mathbf{H}^0}\mathbf{X}_i$, the approach may not work. To see this, consider the bias-corrected IPC estimator $\widetilde{\boldsymbol{\beta}}$ defined by Equation (11) for the models with heterogeneous slopes. Equation (15) gives $\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^0 \approx \left(\sum_{i=1}^N \mathbf{V}_i'\mathbf{V}_i\right)^{-1} \sum_{i=1}^N \mathbf{V}_i'\mathbf{V}_i\boldsymbol{\eta}_i$. If $\mathbb{E}\left(\boldsymbol{\eta}_i|\mathbf{V}_i\right) \neq \mathbf{0}$, which violates the random coefficient assumption (in Assumption 3), $\mathbb{E}\left[\mathbf{V}_i'\mathbf{V}_i\mathbb{E}\left(\boldsymbol{\eta}_i|\mathbf{V}_i\right)\right]$ is not necessarily a zero vector, and this may render $\widetilde{\boldsymbol{\beta}}$ inconsistent, in general. Note that even if we had nonzero correlations between $\boldsymbol{\eta}_i$ and the factor loadings, the slope estimator would remain consistent so long as the regressor matrix is defactored, because the leading term in Equation (15) does not contain the factor loadings. Therefore, we focus on the correlated random coefficients caused by $\mathbb{E}\left(\boldsymbol{\eta}_i|\mathbf{V}_i\right) \neq \mathbf{0}$.

There are studies which propose tests to detect slope heterogeneity for large panel data. However, most of them have focused only on testing the slope heterogeneity, not the dependence of the slope heterogeneity on the regressors. Such examples are found in Pesaran and Yamagata (2008) and Su and Chen (2013), among others. These existing tests are not suitable for our robust approach, because they have power against the random coefficient models as well as the correlated random coefficient models.

In view of this, we propose a novel Lagrange Multiplier (LM) or Score test of dependence of random coefficients on covariates. The proposed test is designed to detect departures from the assumption of independence between slope heterogeneity and the regressors, or departures from the assumption of slope homogeneity toward slope heterogeneity dependent on the regressors. Hereafter, for simplicity, we call it the LM test for correlated random coefficients (CRC). Specifically, we extend the test of correlated random effects proposed in Wooldridge (2010, chap. 11.7.4) for short panels in the fixed effects models. Therein, using a similar approach taken by Mundlak (1978) and Chamberlain (1980), the correlated heterogeneity is modeled as a linear function of within average of the regressors, $\boldsymbol{\beta}_i = \boldsymbol{\mu} + \theta\bar{\mathbf{x}}_i + \boldsymbol{\omega}_i$, where $\bar{\mathbf{x}}_i = T^{-1}\sum_{t=1}^T \mathbf{x}_{it}$ and $\boldsymbol{\omega}_i$ is a mean-zero random vector which is uncorrelated with $\bar{\mathbf{x}}_i$. From equations (4) and (5), it is easily seen that the error term $\mathbf{u}_i$ for the model of omitted slope heterogeneity will contain the term $\mathbf{X}_i\boldsymbol{\beta}_i$ and this motivates his test of correlated random coefficients, which examines the significance of $\mathbf{X}_i\bar{\mathbf{x}}_i$. As pointed out by Wooldridge (2010, p. 386), this test does not have power against uncorrelated slope heterogeneity, which is the desirable property for our approach. We extend the test by considering the models with interactive effects for large N and T and by permitting the alternatives, $\mathbb{E}\left(\boldsymbol{\eta}_i|\mathbf{V}_i\right)$, to be a nonlinear function of $\mathbf{V}_i$.

To motivate the LM test for correlated random coefficients, suppose that the random part of the coefficients, $\boldsymbol{\eta}_i$ (see Equation (5)), is decomposed into correlated and uncorrelated parts:

$$\boldsymbol{\eta}_i = \boldsymbol{\psi}\left(\mathbf{V}_i\right) + \boldsymbol{\omega}_i, \tag{17}$$

where $\mathbb{E}[\boldsymbol{\psi}\left(\mathbf{V}_i\right)|\mathbf{V}_i] \neq \mathbf{0}$ but $\mathbb{E}[\boldsymbol{\omega}_i|\mathbf{V}_i] = \mathbf{0}$. Various functional forms of $\boldsymbol{\psi}(.)$ can be entertained. As an illustration, suppose the ℓth element of $\boldsymbol{\psi}(\mathbf{V}_i)$ to be $\psi_{\ell i} = \frac{1}{\sqrt{T}}\sum_{t=1}^T\{f(\nu_{\ell it}) - \mathbb{E}[f(\nu_{\ell it})]\}$, where the function $f(.)$ satisfies certain regularity conditions. Then, the qth-order Maclaurin series expansion of $f(\nu_{\ell it}) - \mathbb{E}[f(\nu_{\ell it})]$ gives $\psi_{\ell i} = \sum_{p=1}^q \theta_p \frac{1}{\sqrt{T}}\sum_{t=1}^T[\nu_{\ell it}^p - \mathbb{E}(\nu_{\ell it}^p)] + R_q$, where $\theta_p = \frac{1}{p!}f^{(p)}(0)$, $f^{(p)}(.)$ is the pth-order derivative of $f(.)$, and R_q is the reminder term. For testing purposes, examining whether the first few terms of the expansion, say $g(\leq q)$ terms, are zero or not may suffice. This motivates our test of correlated random coefficients (defined below) for the null hypothesis,

$$H_0 : \boldsymbol{\theta} = \mathbf{0} \text{ against } H_1 : \boldsymbol{\theta} \neq \mathbf{0}, \tag{18}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_g)'$.[7]

Having discussed a process of the correlated random coefficients, let us turn our attention to the test statistic. Following on from the above discussion, we define the infeasible score of the LM test for the models with homogeneous (respectively, heterogeneous) slopes as $\frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{L}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{u}_i$ (resp. $\frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{L}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{u}_i$), where $\mathbf{L}_i = \mathbf{X}_i\sqrt{T}\left(\boldsymbol{\Xi}_i - \boldsymbol{\Xi}\right)$,

$$\boldsymbol{\Xi}_i = \left(\bar{\mathbf{v}}_i^{(1)}, \bar{\mathbf{v}}_i^{(2)}, \ldots, \bar{\mathbf{v}}_i^{(g)}\right), \tag{19}$$

$\boldsymbol{\Xi} = \lim_{N\to\infty} N^{-1}\sum_{i=1}^N \mathbb{E}[\boldsymbol{\Xi}_i]$, with $\bar{\mathbf{v}}_i^{(p)} = \left(\bar{\nu}_{1i}^{(p)}, \bar{\nu}_{2i}^{(p)}, \ldots, \bar{\nu}_{ki}^{(p)}\right)'$, $\bar{\nu}_{\ell i}^{(p)} = T^{-1}\sum_{t=1}^T \nu_{\ell it}^p$, $\nu_{\ell it}^p$ is the pth power of the (t, ℓ) element of $\mathbf{V}_i$, $p = 1, \ldots, g$. Observe that the direct alternative model of $\boldsymbol{\psi}(\mathbf{V}_i)$ for this score is $\sqrt{T}\left(\boldsymbol{\Xi}_i - \boldsymbol{\Xi}\right)\boldsymbol{\theta}$, where $\boldsymbol{\theta}$ is the $g \times 1$ constant vector.

To construct the associated feasible score, $\mathbf{L}_i$, $\mathbf{H}^0$, and $\mathbf{u}_i$ should be replaced by their sample counterparts, which are $\widehat{\mathbf{L}}_i = \mathbf{X}_i\sqrt{T}\left(\widehat{\boldsymbol{\Xi}}_i - \overline{\boldsymbol{\Xi}}\right)$, $\widehat{\mathbf{H}}$, and $\widetilde{\mathbf{u}} = \mathbf{y}_i - \mathbf{X}_i\widetilde{\boldsymbol{\beta}}$, respectively, where $\widehat{\boldsymbol{\Xi}}_i = \left(\widehat{\bar{\mathbf{v}}}_i^{(1)}, \widehat{\bar{\mathbf{v}}}_i^{(2)}, \ldots, \widehat{\bar{\mathbf{v}}}_i^{(g)}\right)$, $\widehat{\bar{\mathbf{v}}}_i^{(p)} = \left(\widehat{\bar{\nu}}_{1i}^{(p)}, \widehat{\bar{\nu}}_{2i}^{(p)}, \ldots, \widehat{\bar{\nu}}_{ki}^{(p)}\right)'$, $\widehat{\bar{\nu}}_{\ell i}^{(p)} = T^{-1}\sum_{t=1}^T \widehat{\nu}_{\ell it}^p$ with $\widehat{\mathbf{V}}_i = \mathbf{M}_{\widehat{\mathbf{H}}}\mathbf{X}_i = (\widehat{\nu}_{\ell it})$, $p = 1, \ldots, g$. The resulting feasible score for the models with homogeneous (resp. heterogeneous) slopes is $\frac{1}{\sqrt{NT}}\sum_{i=1}^N \widehat{\mathbf{L}}_i'\mathbf{M}_{\widehat{\mathbf{H}}}\widetilde{\mathbf{u}}_i$ (respectively, $\frac{1}{\sqrt{NT}}\sum_{i=1}^N \widehat{\mathbf{L}}_i'\mathbf{M}_{\widehat{\mathbf{H}}}\widetilde{\mathbf{u}}_i$). Let us sketch the derivation of the LM test statistic next. For the models with homogeneous slopes, a familiar expansion gives $\widetilde{\mathbf{u}}_i = \mathbf{u}_i - \mathbf{X}_i\left(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^0\right)$, which yields

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^N \widehat{\mathbf{L}}_i'\mathbf{M}_{\widehat{\mathbf{H}}}\widetilde{\mathbf{u}}_i = \frac{1}{\sqrt{NT}}\sum_{i=1}^N \widehat{\mathbf{L}}_i'\mathbf{M}_{\widehat{\mathbf{H}}}\mathbf{u}_i + \frac{1}{NT}\sum_{i=1}^N \widehat{\mathbf{L}}_i'\mathbf{M}_{\widehat{\mathbf{H}}}\mathbf{X}_i\sqrt{NT}\left(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^0\right).$$

For a reason similar to why $\widehat{\boldsymbol{\beta}}$ has asymptotic bias terms, the first term of the right hand side of the above equation will have asymptotic bias terms. For the second term, demeaning $\mathbf{L}_i$ and $\mathbf{X}_i$ is required for a reason similar to why the asymptotic variance of $\widehat{\boldsymbol{\beta}}$ involves $\mathbf{Z}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{Z}_i$ rather than $\mathbf{X}_i'\mathbf{M}_{\mathbf{H}^0}\mathbf{X}_i$. Define such a demeaned $\mathbf{L}_i$, $\mathfrak{L}_i = \mathbf{L}_i - N^{-1}\sum_{j=1}^N \boldsymbol{\phi}_i^{0\prime}\boldsymbol{\Upsilon}_{\phi^0}^{-1}\boldsymbol{\phi}_j^0\mathbf{L}_j$. For the LM test, we impose the following assumption, which is similar to Assumption 6:

[7] For testing $H_0 : \theta_p = 0$ for $p = 1, \ldots, \infty$, a nonparametric regression approach (e.g., sieve or spline regressions) could be employed.

Assumption 7. When $N, T \to \infty$ such that $T/N \to \rho \in (0, \Delta]$, $\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathfrak{L}_i' \mathbf{M}_{\mathbf{H}^0} \boldsymbol{\varepsilon}_i \xrightarrow{d} N(\mathbf{0}, \mathbf{B}_2)$, where $\mathbf{B}_2 = \text{plim}_{N,T\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \mathfrak{L}_i' \mathbf{M}_{\mathbf{H}^0} \boldsymbol{\Omega}_{\varepsilon,i} \mathbf{M}_{\mathbf{H}^0} \mathfrak{L}_i$, which is positive definite.

For the slope heterogeneous case, the use of the same score scaled by $\frac{1}{\sqrt{T}}$ will be justified. For more details, see proof of Theorem 3 in the supplementary materials.

Consequently, the LM test statistic for correlated random coefficients, which is common to the model with homogeneous slopes and the random coefficient model, is defined by

$$\text{LM}_{\text{CRC}}^{(g)} = \widehat{\mathbf{s}}^{\dagger\prime} \widehat{\boldsymbol{\Sigma}}_{ss}^{-1} \widehat{\mathbf{s}}^{\dagger} \tag{20}$$

where $\widehat{\mathbf{s}}^{\dagger} = \widehat{\mathbf{s}} - N^{-1} \widehat{\boldsymbol{\xi}}_{NT}^{\dagger} - T^{-1} \widehat{\boldsymbol{\zeta}}_{NT}^{\dagger}$ with $\widehat{\mathbf{s}} = \sum_{i=1}^{N} \widehat{\mathbf{L}}_i' \mathbf{M}_{\widehat{\mathbf{H}}} \widetilde{\mathbf{u}}_i$,

$$\boldsymbol{\xi}_{NT}^{\dagger} = -\frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} \widehat{\mathfrak{L}}_i' \widehat{\mathbf{H}} \widehat{\boldsymbol{\Upsilon}}_{\phi}^{-1} \widehat{\boldsymbol{\phi}}_i \widehat{e}_{it}^2, \ \boldsymbol{\zeta}_{NT}^{\dagger}$$
$$= -\frac{1}{N} \sum_{i=1}^{N} \frac{\widehat{\mathfrak{L}_i' \mathbf{M}_{\mathbf{H}} \boldsymbol{\Omega} \mathbf{H}}}{T} \widehat{\boldsymbol{\Upsilon}}_{\phi}^{-1} \widehat{\boldsymbol{\phi}}_i,$$

$\widehat{\mathfrak{L}}_i = \widehat{\mathbf{L}}_i - N^{-1} \sum_{j=1}^{N} \widehat{\boldsymbol{\phi}}_i' \widehat{\boldsymbol{\Upsilon}}_{\phi}^{-1} \widehat{\boldsymbol{\phi}}_j \widehat{\mathbf{L}}_j$, $\frac{\widehat{\mathfrak{L}_i' \mathbf{M}_{\mathbf{H}} \boldsymbol{\Omega} \mathbf{H}}}{T}$ is computed analogously to the computation of $\frac{\widehat{\mathbf{X}_i' \mathbf{M}_{\mathbf{H}} \boldsymbol{\Omega} \mathbf{H}}}{T}$ (see Equation (12)), and $\widehat{\boldsymbol{\Sigma}}_{ss} = \sum_{i=1}^{N} \widehat{\mathbf{K}}_i' \mathbf{M}_{\widehat{\mathbf{H}}} \widetilde{\mathbf{u}}_i \widetilde{\mathbf{u}}_i' \mathbf{M}_{\widehat{\mathbf{H}}} \widehat{\mathbf{K}}_i$ with

$$\widehat{\mathbf{K}}_i = \widehat{\mathfrak{L}}_i - \widehat{\mathbf{Z}}_i \left(\sum_{i=1}^{N} \widehat{\mathbf{Z}}_i' \mathbf{M}_{\widehat{\mathbf{H}}} \widehat{\mathbf{Z}}_i \right)^{-1} \sum_{i=1}^{N} \widehat{\mathbf{Z}}_i' \mathbf{M}_{\widehat{\mathbf{H}}} \widehat{\mathfrak{L}}_i.$$

Now we are ready to present the results of the limiting distribution of the LM test statistic.

Theorem 3. Consider the model (7). Suppose that Assumptions 1–7 hold. Then, under the null hypothesis specified in (18), $\text{LM}_{\text{CRC}}^{(g)} \xrightarrow{d} \chi_g^2$, as $N, T \to \infty$ and $T/N \to \rho \in (0, \Delta]$, for both homogeneous-slope and heterogeneous-slope models.

In practice, the degree of polynomial, g, for the statistic $\text{LM}_{\text{CRC}}^{(g)}$, should be chosen. In order to approximate the function $\boldsymbol{\psi}(\mathbf{V}_i)$ to the extent that the LM test has sufficient power, a small value of g, such as two, may be sufficient. The experimental results reported below support the choice of $g = 2$. In addition, the asymptotic local power analysis, which is provided in the supplementary materials, suggests that the LM test has nontrivial power when $||\boldsymbol{\theta}|| \asymp \frac{1}{\sqrt{NT}}$ and that the LM test associated with the pth column of the test matrix $\mathbf{L}_i$ has nontrivial power only against the pth term in the expansion of $\boldsymbol{\psi}(\mathbf{V}_i)$, that is, $H_1 : \theta_p \neq 0$. The experimental results shown below support these analytical results.

When the null hypothesis of random coefficient models is rejected in favor of the alternatives, it is preferable to employ estimators which are consistent to $\boldsymbol{\beta}^0$. For the estimation of the models with interactive effects for large panel data, the CCE and 2SIV mean group estimators and the ML estimators proposed by Pesaran (2006), Cui et al. (2021), and Li, Cui, and Lu (2020), among others, would be possible choices.

4. Monte Carlo Experiments

In this section we investigate the finite sample performance of our robust approach against slope heterogeneity, error serial correlation and heteroscedasticity. We consider the performance of the (analytically) bias-corrected IPC estimator defined by (11), which is denoted as BC-IPC. In addition, as discussed in Remark 3, we introduce a jackknife bias-corrected IPC estimator, which is denoted as JK-IPC. To evaluate accuracy of the bias-correction, the IPC estimator without bias-corrections is considered as well. In particular, we examine biases, standard deviations and root mean square errors (RMSEs) of these estimators, and empirical size of the (Wald) test for linear restrictions of slope coefficients. Furthermore, performance of the LM test for correlated random coefficients is examined.

Before introducing our experimental design, we discuss how to choose the initial estimator and the number of factors r for the IPC estimation in Section 4.1, then the procedure for computing the jackknife bias-corrected IPC estimator in Section 4.2.

4.1. Initial Value for the IPC Estimator and the Number of Factors

In practice, to compute the IPC estimator, an initial (consistent) estimator of $\boldsymbol{\beta}^0$ and the number of factors in $\mathbf{u}_i$ must be given. Exploiting the factor structure in $\mathbf{X}_i$ and $\mathbf{u}_i$, we propose to employ the consistent PC estimator of Westerlund and Urbain (2015) and Reese and Westerlund (2018) as the initial estimator of $\boldsymbol{\beta}^0$, which is defined as $\widehat{\boldsymbol{\beta}}_{\text{PC}} = \left(\sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{W}}} \mathbf{X}_i \right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{W}}} \mathbf{y}_i$, where $\widehat{\mathbf{W}}$ is a $T \times \widehat{r_1 + r_2}$ PC estimator of a span of $\mathbf{W}^0 = (\mathbf{F}^0, \mathbf{G}^0)$. $\widehat{r_1 + r_2}$ and $\widehat{\mathbf{W}}$ are estimated using $\{\mathbf{X}_i, \mathbf{y}_i\}_{i=1}^{N}$, applying the Eigenvalue Ratio (ER) or the Growth Ratio (GR) estimators proposed by Ahn and Horenstein (2013), and the PC method, respectively. The estimate of the number of factors in $\mathbf{u}_i$, $\widehat{r}$, is obtained using the residual, $\widehat{\mathbf{u}}_{\text{PC}i} = \mathbf{y}_i - \mathbf{X}_i \widehat{\boldsymbol{\beta}}_{\text{PC}}$, applying the ER or GR methods.[8] In the experiment below, we set the maximum number of factors to six to estimate $r_1 + r_2$ and r. Alternatively, the CCE or 2SIV estimator of Pesaran (2006) and Cui et al. (2021) can be used as an initial estimator of $\boldsymbol{\beta}^0$, which is not pursued in this article.

4.2. Jackknife Bias Correction

In Sections 2 and 3, the analytical bias correction of the IPC estimator $\widehat{\boldsymbol{\beta}}$ is considered. Instead, as discussed in Remark 3, subsampling methods can be employed for bias corrections. Specifically, we consider the split panel jackknife in both the cross-section and time dimensions, which is proposed in Fernández and Weidner (2016).

To define the jackknife bias-corrected IPC estimator, assuming that N and T are even numbers for simplicity, consider the index set, $\mathcal{N} = \{1, \ldots, N\}$, which is divided into two subgroups, $\mathcal{N}_1 = \{1, \ldots, N/2\}$ and $\mathcal{N}_2 = \{(N/2) + 1, \ldots, N\}$. Similarly define $\mathcal{T} = \{1, \ldots, T\}$, $\mathcal{T}_1 = \{1, \ldots, T/2\}$ and $\mathcal{T}_2 = \{(T/2) + 1, \ldots, T\}$. Let us denote the IPC estimator as $\widehat{\boldsymbol{\beta}}_{\mathcal{N},\mathcal{T}} (= \widehat{\boldsymbol{\beta}})$, in order to explicitly express the dependence on

[8] Alternatively, we may use the criteria proposed by Bai and Ng (2002).

the (sub)sample, $(i,t) \in (\mathcal{N},\mathcal{T})$. Then, we define the jackknife bias-corrected IPC estimator as $\widehat{\boldsymbol{\beta}}_{JK} = \left(\widehat{\boldsymbol{\beta}}_{TS} + \widehat{\boldsymbol{\beta}}_{CS}\right) - \widehat{\boldsymbol{\beta}}_{\mathcal{N},\mathcal{T}}$, where $\widehat{\boldsymbol{\beta}}_{TS} = 2\widehat{\boldsymbol{\beta}}_{\mathcal{N},\mathcal{T}} - \frac{\widehat{\boldsymbol{\beta}}_{\mathcal{N},\mathcal{T}_1}+\widehat{\boldsymbol{\beta}}_{\mathcal{N},\mathcal{T}_2}}{2}$ and $\widehat{\boldsymbol{\beta}}_{CS} = 2\widehat{\boldsymbol{\beta}}_{\mathcal{N},\mathcal{T}} - \frac{\widehat{\boldsymbol{\beta}}_{\mathcal{N}_1,\mathcal{T}}+\widehat{\boldsymbol{\beta}}_{\mathcal{N}_2,\mathcal{T}}}{2}$. For more details, see the supplementary materials and Section 3.2 in Fernández and Weidner (2016).

4.3. Experimental Design

Consider the following data generating process:

$$y_{it} = \sum_{\ell=1}^{k} x_{\ell it}\beta^0_{\ell i} + f^0_{1t}\lambda^0_{1i} + f^0_{2t}\lambda^0_{2i} + \sigma_{\varepsilon,it}\varepsilon_{it},$$
$$i = 1,2,\dots,N; t = 1,2,\dots,T \tag{21}$$

where $\lambda^0_{si} \sim \text{iid}N(0,1)$, $f^0_{st} = \rho_f f^0_{s,t-1} + \sqrt{1-\rho_f^2}\epsilon_{f,st}$, $\epsilon_{f,st} \sim \text{iid}N(0,1)$ with $f^0_{s0} \sim \text{iid}N(0,1)$ for $s = 1,2,3$, $\varepsilon_{it} = \rho_\varepsilon \varepsilon_{it-1} + \sqrt{1-\rho_\varepsilon^2}\epsilon_{\varepsilon,it}$, $\epsilon_{\varepsilon,it} \sim \text{iid}N(0,1)$ with $\varepsilon_{i0} \sim \text{iid}N(0,1)$, and

$$\sigma_{\varepsilon,it} = \left(\kappa_{\varepsilon,i}\kappa_{\varepsilon,t}\right)^{1/2}, \kappa_{\varepsilon,i} \sim \text{iid}U(0.5, 1.5) \text{ and } \kappa_{\varepsilon,t} = 0.5 + t/T. \tag{22}$$

The regressors $x_{\ell it}$, $\ell = 1,2,\dots,k$, are generated as

$$x_{\ell it} = f^0_{1t}\gamma^0_{\ell 1i} + f^0_{3t}\gamma^0_{\ell 3i} + 0.3\sigma_{v,it}v_{\ell it}, \tag{23}$$

where $v_{\ell it} = \rho_v v_{\ell it-1} + \sqrt{1-\rho_v^2}\varpi_{\ell it}$. Note that f^0_{1t} is common to both y_{it} and $x_{\ell it}$, whereas f^0_{2t} and f^0_{3t} are unique to y_{it} and $x_{\ell it}$, respectively. $\varpi_{\ell it}$ is generated as $\varpi_{\ell it} = \left(\varpi^*_{\ell it} - c\right)/\sqrt{2c}$, $\varpi^*_{\ell it} \sim \text{iid}\chi^2_c$ and $v_{\ell i0} = \left(v^*_{\ell i0} - c\right)/\sqrt{2c}$, $v^*_{\ell i0} \sim \text{iid}\chi^2_c$ with $c = 6$.[9] The factor loadings in $x_{\ell it}$ are generated as

$$\gamma^0_{\ell si} = 0.7\lambda^0_{si} + \left(1 - 0.7^2\right)^{1/2}\varphi_{\ell si}, \tag{24}$$

$\varphi_{\ell si} \sim \text{iid}N(0,1)$ for $\ell = 1,\dots,k$ and $s = 1,3$, so that they are correlated with factor loadings in y_{it}.

$$\sigma_{v,it} = \left(\kappa_{v,i}\kappa_{v,t}\right)^{1/2}, \kappa_{v,i} \sim \text{iid}U(0.5, 1.5) \text{ and } \kappa_{v,t} = 4.5 + t/T. \tag{25}$$

Finally we have

$$\beta_{\ell i} = \beta^0_\ell + \sigma_\eta \eta_{\ell i}, \tag{26}$$
$$\eta_{\ell i} = \rho_{x\eta}\psi_{\ell i} + \sqrt{1-\rho_{x\eta}^2}\omega_{\ell i} \tag{27}$$

$\omega_{\ell i} \sim \text{iid}N(0,1)$ for $\ell = 1,\dots,k$ and $\psi^{(\mathcal{P})}_{\ell i} = \sum_{p=1}^{q}\theta_p\widetilde{z}_{\ell i,p}/\sqrt{\sum_{p=1}^{q}\theta_p^2}$, where $\mathcal{P} = \{p : \theta_p \neq 0, p \in I\}$ with $I = \{1,2,\dots,q\}$ and $\theta_p \in \{0,1\}$,

$$\widetilde{z}_{\ell i,p} = \frac{z_{\ell i,p} - \bar{z}_{\ell,p}}{s_{z\ell,p}}$$

with $\bar{z}_{\ell,p} = N^{-1}\sum_{i=1}^{N} z_{\ell i,p}$, $s^2_{z\ell,p} = (N-1)^{-1}\sum_{i=1}^{N}\left(z_{\ell i,p} - \bar{z}_{\ell,p}\right)^2$ and $z_{\ell i,p} = T^{-1}\sum_{t=1}^{T} v^p_{\ell it}$. Observe that the variance of $\beta_{\ell i}$ is controlled to be σ^2_η. θ_p selects the term $\widetilde{z}_{\ell i,p}$. Specifically, in the experiment, we consider $\psi^{(1,2,3,4)}_{\ell i} = \sum_{p=1}^{4}\widetilde{z}_{\ell i,p}/\sqrt{4}$ and $\psi^{(p)}_{\ell i} = \widetilde{z}_{\ell i,p}$, $p = 1,\dots,4$. It is interesting to examine the impact of such linear and highly nonlinear relationships on the estimators and the power of the LM tests.

[9] We also considered the case in which $\varpi_{\ell it}$ is generated from normal distribution. The results are not reported because they were qualitatively very similar.

Table 1. Experimental designs.

Design	Model	r	σ_η	$\rho_{x\eta}$	$\psi^{(\mathcal{P})}_{\ell i}$
1	Homogeneous slopes	2	0	0	–
2	Heterogeneous slopes uncorrelated with $x_{\ell it}$	3	0.50	0	–
3	Heterogeneous slopes correlated with $x_{\ell it}$	3	0.50	0.50	$\psi^{(1,2,3,4)}_{\ell i}$
4	Heterogeneous slopes correlated with $x_{\ell it}$	3	0.50	0.50	$\psi^{(1)}_{\ell i}$
5	Heterogeneous slopes correlated with $x_{\ell it}$	3	0.50	0.50	$\psi^{(2)}_{\ell i}$
6	Heterogeneous slopes correlated with $x_{\ell it}$	3	0.50	0.50	$\psi^{(3)}_{\ell i}$
7	Heterogeneous slopes correlated with $x_{\ell it}$	3	0.50	0.50	$\psi^{(4)}_{\ell i}$

NOTE: r is the number of factors in the error term. σ_η is standard deviation of $\beta_{\ell i} - \beta^0_\ell (= \sigma_\eta\eta_{\ell i})$, and $\rho_{x\eta}$ is the correlation between $\eta_{\ell i}$ and a function of the defactored regressor $\psi^{(\mathcal{P})}_{\ell i}$, where $\psi^{(1,2,3,4)}_{\ell i} = \sum_{p=1}^{4}\widetilde{z}_{\ell i,p}/\sqrt{4}$, $\psi^{(p)}_{\ell i} = \widetilde{z}_{\ell i,p}$ for $p = 1,\dots,4$, and $\widetilde{z}_{\ell i,p}$ is the cross-sectionally standardized within average of the defactored regressor, $T^{-1}\sum_{t=1}^{T} v^p_{\ell it}$, for $\ell = 1,\dots,k$. In the article we report the results of designs 1–3, with the rest in the supplementary materials.

We set $k = 2$ (two regressors) for all experiments. Before the estimation, the data is within-transformed and cross-sectionally demeaned, to make the results invariant to the inclusion of (additive) individual effects and time effects. For parameter values, we set $\left(\beta^0_1, \beta^0_2\right) = (1,1)$.

We consider seven experimental designs as outlined in Table 1. Designs 1–7 are considered to investigate the finite sample behavior of the slope estimator, the associated robust Wald test and the LM test for correlated random coefficients.

Design 1 corresponds to the model with homogeneous slopes ($\sigma_\eta = 0$) and Design 2 to the model with random coefficients ($\sigma_\eta = 0.5$ and $\rho_{x\eta} = 0$). Designs 3-7 correspond to models with heterogeneous slopes that depend on the regressors ($\sigma_\eta = 0.5$ and $\rho_{x\eta} = 0.5$) in different ways. These designs are considered in order to investigate the impact of such dependence on the estimators and the power properties of the LM test. In Design 3, the slope heterogeneity contains a fourth-degree polynomial of the defactored regressors ($\psi^{(1,2,3,4)}_{\ell i}$), which is highly nonlinear. In Designs 4–7, the slope heterogeneity contains $\psi^{(1)}_{\ell i}$, $\psi^{(2)}_{\ell i}$, $\psi^{(3)}_{\ell i}$, and $\psi^{(4)}_{\ell i}$, respectively. In the article we present the results for Designs 1 to 3, with the rest in the supplementary materials.

We consider all combinations of $N = 50, 100, 200$ and $T = 25, 50, 100, 200$. Throughout the experiment, we set $\rho_f = 0.5$, $\rho_\varepsilon = 0.5$ and $\rho_v = 0.5$. All tests are conducted at 5% significance level. All experimental results are based on 2000 replications.

4.4. Results

In order to compute the IPC estimator, it is necessary to estimate the number of factors in the error term, $\mathbf{u}_i$. The frequencies of the number of factors estimated by the ER and GR methods proposed by Ahn and Horenstein (2013) are reported in Table 2; see Section 4.1 for the estimation procedure. Note that, as summarized in Table 1, r is two in the models with homogeneous slopes (Design 1) and three in the models with random coefficients (Design 2). The results in Table 2 suggest that ER

Table 2. Frequency of the estimated number of factors in the error term by the ER and GR methods.

	Design 1: homogeneous slopes ($r = 2$)						Design 2: random coefficients ($r = 3$)					
Freq	ER			GR			ER			GR		
T/N	50	100	200	50	100	200	50	100	200	50	100	200
						$\widehat{r} = 1$						
25	1.0	0.3	0.2	3.6	1.5	1.0	3.2	1.4	0.7	13.5	7.9	5.0
50	0.0	0.0	0.0	0.2	0.0	0.0	0.4	0.0	0.0	3.1	0.6	0.1
100	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.6	0.1	0.0
200	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.3	0.0	0.0
						$\widehat{r} = 2$						
25	98.9	99.7	99.8	96.5	98.5	99.1	16.6	10.3	6.7	23.4	17.9	14.4
50	100.0	100.0	100.0	99.9	100.0	100.0	6.5	1.4	0.5	11.7	4.3	1.7
100	100.0	100.0	100.0	100.0	100.0	100.0	2.3	0.1	0.0	6.4	0.6	0.0
200	100.0	100.0	100.0	100.0	100.0	100.0	1.7	0.0	0.0	4.2	0.1	0.0
						$\widehat{r} = 3$						
25	0.1	0.0	0.0	0.0	0.0	0.0	80.1	88.3	92.7	63.1	74.3	80.7
50	0.0	0.0	0.0	0.0	0.0	0.0	93.2	98.7	99.6	85.2	95.2	98.3
100	0.0	0.0	0.0	0.0	0.0	0.0	97.7	100.0	100.0	93.1	99.4	100.0
200	0.0	0.0	0.0	0.0	0.0	0.0	98.3	100.0	100.0	95.5	99.9	100.0
						$\widehat{r} = 4$						
25	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0
50	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
200	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

NOTE: We set the maximum number of factors to six to estimate r (and $r_1 + r_2$). The frequencies of $\widehat{r} = 5$ and 6 were zero. ER and GR stands for the Eigenvalue Ratio and Growth Ratio methods proposed by Ahn and Horenstein (2013). See Table 1 for a summary of experimental designs.

and GR estimate r precisely enough for all designs, though the ER estimator is slightly more accurate than the GR, especially when T is not large. Based on this result, the experimental results reported hereafter are based on the r estimated by the ER method.

Given the estimated number of factors, we can proceed to the estimation of $\boldsymbol{\beta}^0$ and associated inferences based on the robust PHAC standard errors. In this experiment, we consider three different estimators of $\boldsymbol{\beta}^0$: the IPC estimator without bias correction (IPC), the analytically bias-corrected IPC estimator (BC-IPC), and the jackknife bias-corrected IPC estimator (JK-IPC). Table 3 summarizes the performance of these IPC estimators for Designs 1–3. The "ER" signifies the dependence of the estimators on the ER estimation of r. In the table, the bias (Bias), standard deviation (SD), root mean square error (RMSE) of the estimates for β_1^0, and the size (in percent) of the robust Wald test for $H_0 : \beta_1^0 = 1$ (Size), are reported. Bias, SD, and RMSE are all multiplied by 100. The results for β_2^0 are not reported because they are qualitatively very similar.

First, we discuss the results for Design 1 (the model with homogeneous slopes). As can be seen in the table, the IPC estimator (without bias-correction) suffers from a large bias. As a result, the size of the robust Wald test deviates significantly from the nominal level. In contrast, the BC-IPC estimator with the analytical bias correction and the JK-IPC estimator with the jackknife bias correction successfully reduce the bias. The BC-IPC estimator and the JK-IPC estimator perform comparably. However, when the sample size is small (e.g., $T = 25$ or $N = 50$), the jackknife method can reduce the bias more effectively than the analytical bias correction. Meanwhile, the jackknife bias correction is likely to lead to greater sampling variability than the analytical bias correction. The standard deviation of the JK-IPC estimator is uniformly larger than the standard deviation of the BC-IPC estimator. As a result, the BC-IPC estimator uniformly outperforms the JK-IPC estimator in terms of the mean squared error criterion.[10] Reflecting this small sample bias, there is a mild size distortion in the robust Wald test for small T or small N, which tends to disappear as N and T increase.

The results for Design 2 (random coefficient models) in Table 3 are qualitatively similar to those for Design 1. However, reflecting the fact that the IPC estimators are less efficient for random coefficient models, there is a moderate small sample bias in the BC-IPC and JK-IPC estimators, which dissipates as N increases. Due to this bias, there is a moderate size distortion in the robust Wald test for small N, but it quickly decreases as the sample size increases.

Next, let us look at the results for Design 3, in which the heterogeneity of the slopes depends on the regressors. The bias of the IPC, BC-IPC, and JK-IPC estimators is large and persistent for any sample size. As a consequence, the size distortions of the robust Wald tests are large and increase as the sample size rises. This implies the importance of the LM test for correlated random coefficients in our robust approach.

Table 4 reports the size and power of the LM test for the dependence of the random coefficients on the regressors. We note that in our alternative model (Design 3), the slope heterogeneity contains a fourth-degree polynomial of the defactored regressors (i.e., $\psi_{\ell i}^{(1,2,3,4)}$), which is highly nonlinear. In this experiment, we consider $\mathrm{LM}_{\mathrm{CRC}}^{(1)}$ and $\mathrm{LM}_{\mathrm{CRC}}^{(2)}$ tests, which are defined in equation (20). Recall that we recommend the test using the $\mathrm{LM}_{\mathrm{CRC}}^{(2)}$ statistic, because it is expected to have nontrivial power against a sufficiently broad class of alternative models, while the $\mathrm{LM}_{\mathrm{CRC}}^{(1)}$ test may not be sufficiently powerful against certain alternatives (see discussions in Section 3). The results reported in Table 4 show that the size of the $\mathrm{LM}_{\mathrm{CRC}}^{(1)}$ and

[10] A similar superior performance of analytical bias-correction over jackknife bias-correction is found in Fernández and Weidner (2016).

Table 3. Finite sample properties of the IPC estimators of β_1^0 and the associated robust Wald test with interactive effects, heteroscedastic and serially correlated idiosyncratic errors.

Design 1: homogeneous slopes $\left(\sigma_{\eta=0},\ \rho_{x\eta}=0\right)$

	Bias(×100)			SD(×100)			RMSE(×100)			Size(%)		
T/N	50	100	200	50	100	200	50	100	200	50	100	200
						IPC(ER)						
25	−3.599	−1.674	−0.739	6.556	4.418	3.074	7.478	4.725	3.161	49.4	44.8	39.0
50	−4.069	−1.978	−0.992	4.485	2.878	1.974	6.056	3.492	2.209	57.5	46.6	39.6
100	−4.315	−2.085	−1.034	3.122	1.887	1.250	5.326	2.812	1.622	71.6	54.4	41.2
200	−4.721	−2.274	−1.113	2.323	1.424	0.996	5.262	2.683	1.494	87.9	68.5	51.7
						BC-IPC(ER)						
25	0.228	0.160	0.099	3.042	2.083	1.441	3.050	2.089	1.445	8.2	7.6	6.0
50	0.047	0.033	0.016	1.958	1.367	0.939	1.959	1.368	0.939	7.4	6.9	5.6
100	−0.003	−0.006	0.008	1.497	1.010	0.710	1.497	1.010	0.710	7.7	6.7	5.7
200	0.006	0.000	0.001	1.218	0.825	0.573	1.218	0.825	0.573	7.0	5.9	5.4
						JK-IPC(ER)						
25	0.096	0.093	0.038	3.506	2.274	1.580	3.507	2.276	1.580	11.8	11.0	9.2
50	0.015	0.007	0.002	2.156	1.480	1.005	2.156	1.480	1.005	9.3	9.6	7.6
100	−0.033	−0.014	0.004	1.590	1.061	0.746	1.590	1.061	0.746	9.4	7.3	6.4
200	−0.022	−0.004	−0.002	1.292	0.856	0.590	1.292	0.856	0.590	9.5	7.2	5.9

Design 2: heterogeneous slopes uncorrelated with $x_{\ell it}$ $\left(\sigma_\eta=0.5,\ \rho_{x\eta}=0\right)$

	Bias(×100)			SD(×100)			RMSE(×100)			Size(%)		
T/N	50	100	200	50	100	200	50	100	200	50	100	200
						IPC(ER)						
25	−4.294	−2.282	−0.963	10.680	7.490	5.247	11.511	7.830	5.334	16.4	12.3	9.7
50	−4.693	−2.506	−1.143	9.069	6.257	4.322	10.212	6.741	4.471	17.6	11.7	8.7
100	−4.901	−2.611	−1.205	8.114	5.763	3.984	9.479	6.327	4.162	19.2	12.7	8.4
200	−5.371	−2.770	−1.311	7.666	5.487	3.824	9.360	6.147	4.042	20.4	13.1	7.9
						BC-IPC(ER)						
25	0.163	0.099	0.155	10.455	7.240	4.939	10.457	7.241	4.942	12.6	9.3	6.2
50	−0.202	−0.174	−0.017	9.073	6.150	4.219	9.076	6.153	4.219	12.8	8.6	7.1
100	−0.233	−0.197	−0.049	8.328	5.742	4.004	8.331	5.746	4.005	12.9	8.7	7.3
200	−0.263	−0.171	−0.050	7.933	5.555	3.860	7.937	5.557	3.861	12.8	9.1	6.5
						JK-IPC(ER)						
25	−0.414	−0.199	−0.077	12.167	7.994	5.314	12.174	7.997	5.315	18.0	11.9	9.0
50	−0.184	−0.204	−0.082	9.763	6.342	4.373	9.765	6.345	4.373	15.1	9.5	8.1
100	−0.100	−0.121	−0.060	8.726	5.865	4.062	8.726	5.867	4.062	14.5	9.7	7.4
200	−0.093	−0.083	−0.027	8.145	5.617	3.874	8.145	5.617	3.874	13.9	9.0	6.4

Design 3: heterogeneous slopes correlated with $x_{\ell it}$ $\left(\sigma_\eta=0.5,\ \rho_{x\eta}=0.5,\ \psi_{\ell i}^{(1,2,3,4)}\right)$

	Bias(×100)			SD(×100)			RMSE(×100)			Size(%)		
T/N	50	100	200	50	100	200	50	100	200	50	100	200
						IPC(ER)						
25	1.070	3.570	4.964	11.075	7.527	5.137	11.127	8.331	7.144	15.3	12.8	18.6
50	1.066	3.746	5.217	9.357	6.377	4.164	9.418	7.396	6.675	13.5	13.3	22.6
100	0.882	3.735	5.256	8.228	5.596	3.568	8.275	6.728	6.353	11.5	12.5	23.1
200	0.778	3.876	5.493	7.879	5.248	3.320	7.917	6.524	6.418	11.7	13.5	26.9
						BC-IPC(ER)						
25	4.567	5.160	5.592	10.452	7.129	4.913	11.406	8.801	7.444	15.3	16.3	22.0
50	5.065	5.823	6.250	9.064	6.149	3.964	10.383	8.469	7.401	16.9	19.6	29.6
100	5.221	5.956	6.356	8.318	5.555	3.565	9.821	8.145	7.287	18.5	20.2	34.4
200	5.540	6.291	6.662	8.050	5.322	3.368	9.772	8.240	7.465	21.4	23.7	40.0
						JK-IPC(ER)						
25	4.798	5.451	5.789	12.302	8.062	5.462	13.205	9.732	7.959	22.3	20.5	26.0
50	5.937	6.444	6.572	9.981	6.522	4.141	11.613	9.168	7.768	21.7	24.1	33.7
100	6.118	6.548	6.656	8.793	5.734	3.634	10.712	8.704	7.583	23.4	25.1	38.2
200	6.560	6.909	6.981	8.521	5.404	3.398	10.754	8.772	7.764	26.4	27.7	43.3

$\mathrm{LM}_{\mathrm{CRC}}^{(2)}$ tests is very close to the nominal level for Design 1 (the model with homogeneous slopes) and Design 2 (random coefficient models). In the results for Design 3 (correlated random coefficients), we find that the power of the $\mathrm{LM}_{\mathrm{CRC}}^{(2)}$ test is high and rises as N and T increase, while the $\mathrm{LM}_{\mathrm{CRC}}^{(1)}$ test is much less powerful. Table C.3 in the supplementary materials reports the power of the $\mathrm{LM}_{\mathrm{CRC}}^{(1)}$ and $\mathrm{LM}_{\mathrm{CRC}}^{(2)}$ tests for Designs 4–7. Therein it is found that the $\mathrm{LM}_{\mathrm{CRC}}^{(2)}$ test remains as powerful across Designs 4–7 as in Design 3, whilst the $\mathrm{LM}_{\mathrm{CRC}}^{(1)}$ test is substantially less powerful in Design 5 ($\psi_{\ell i}^{(2)}$) and Design 7 ($\psi_{\ell i}^{(4)}$). These findings are in line with the results of our asymptotic local power analysis and they support our claim that the $\mathrm{LM}_{\mathrm{CRC}}^{(2)}$ test is powerful against a broad class of alternative models of correlated coefficients.

Table 4. Size and power of the $LM^{(1)}_{CRC}$ and $LM^{(2)}_{CRC}$ tests of correlated random coefficient.

	Size (Design 1) $\sigma_\eta = 0, \rho_{x\eta} = 0$			Size (Design 2) $\sigma_\eta = 0.5, \rho_{x\eta} = 0$			Power (Design 3) $\sigma_\eta = 0.5. \rho_{x\eta} = 0.5$		
T,N	50	100	200	50	100	200	50	100	200
				$LM^{(1)}_{CRC}$					
25	2.6	2.6	3.1	2.7	2.4	2.4	5.9	7.1	9.7
50	3.4	3.3	2.3	2.5	3.0	3.3	8.9	12.0	16.8
100	3.3	2.0	2.6	2.7	2.3	2.3	12.1	16.9	28.8
200	3.4	2.9	2.3	2.7	3.0	2.4	18.8	32.2	51.4
				$LM^{(2)}_{CRC}$					
25	2.9	3.5	3.9	3.4	3.7	3.6	9.9	18.3	34.2
50	2.9	2.7	3.2	4.4	3.9	3.8	22.7	48.8	77.8
100	2.9	2.8	3.4	4.4	3.2	3.0	45.3	83.2	97.2
200	2.9	3.0	2.9	4.5	3.7	3.7	68.4	97.7	100.0

NOTE: For Designs 1–3, see Table 1. The $LM^{(g)}_{CRC}$ is the proposed LM test of correlated random effects defined by (20), based on the BC-IPC estimator. The test statistics are referred to the 95% quantile of χ^2_g distribution. All results are based on 2000 replications. The number of factors is estimated by the ER method.

Table 5. Estimation results for the Feldstein–Horioka puzzle based on OECD data covering 1968–1996, ($N = 24, T = 29$).

	Estimated number of factors	
$\widehat{r}$	$\widehat{r}_{,ER} = 1$	$\widehat{r}_{,GR} = 1$
	LM test for correlated random coefficient	
	test statistic	p-value
$LM^{(2)}_{CRC}$	0.5101	0.7749
	Estimation results with $\widehat{r} = 1$	
	BC-IPC	JK-CCEP
Coef.	0.4238	0.3518
S.E.	0.0891	0.1193

NOTE: BC-IPC is the bias-corrected IPC estimator with the PHAC standard error. JK-CCEP is the jackknife bias-corrected CCEP estimator of Westerlund (2018) with the nonparametric standard error of Pesaran (2006) for random coefficients. For the estimation of the number of factors and computation of BC-IPC, data are demeaned in both dimensions while undemeaned data is used for JK-CCEP.

5. An Empirical Application: Feldstein–Horioka Puzzle

In this section, we present an empirical illustration of the proposed method. Specifically, we analyze the Feldstein–Horioka (F–H) puzzle, which has attracted recent attention in the literature.

Feldstein and Horioka (1980) found that in OECD countries, the long-run average of the national savings rate is significantly correlated with the long-run average of the domestic investment rate. This empirical finding seems to conflict with macroeconomic theory, which predicts that the determinants of saving and investment are different and that, assuming perfect international capital mobility, the investment decisions in a country should not be constrained by domestic saving. This is known as the F–H puzzle, and is included as one of the six major puzzles in international macroeconomics by Obstfeld and Rogoff (2000).

Several econometric methods have been applied to investigate the F–H puzzle, including instrumental variable methods, cointegration regressions and the standard fixed effect models.[11] Recently, Giannone and Lenza (2010) and Ginama, Hayakawa, and Kanemi (2018) analyzed the F–H puzzle by considering a panel regression model with interactive effects.

In this section, we apply our approach to analyze the F–H puzzle by using the OECD panel dataset covering 1968–1996 with $N = 24$ and $T = 29$, as used in Ginama, Hayakawa, and Kanemi (2018). Specifically, we estimate the following model:

$$\left(\frac{I_{it}}{Y_{it}}\right) = \beta_i \left(\frac{S_{it}}{Y_{it}}\right) + \boldsymbol{\lambda}_i^{0\prime}\mathbf{f}_t^0 + \varepsilon_{it},$$

where I_{it}, S_{it}, and Y_{it} denote investment, gross savings, and GDP for country i and year t, respectively. This specification is more general than considered in the literature in that the coefficient is allowed to be heterogeneous and interactive effects are included.

We now describe how we proceed when using our approach. Our interest lies in estimating the population mean of β_i. First, we need to determine the number of factors in the error term for the model in which the slope heterogeneity (if any) is ignored. Here, we follow the procedure described in Section 4.1. Having estimated the number of factors in the error term, we conduct the LM test for correlated random coefficients, which is proposed in Section 3. If the null hypothesis is not rejected, we proceed to statistical inference using the robust PHAC standard errors; otherwise we need to consider an alternative approach, as described at the end of Section 3.

Table 5 provides the estimation results. As a comparison, we computed a jackknife bias-corrected CCEP estimator proposed by Westerlund (2018) and the associated nonparametric standard error of Pesaran (2006, eq. (69)) for random coefficient models.[12] With regard to the number of factors, both $\widehat{r}_{GR}$ and $\widehat{r}_{ER}$ suggest that the number of factors is one. The LM test for correlated random coefficients, $LM^{(2)}_{CRC}$, does not reject the null hypothesis of random coefficients or homogeneous slopes, which validates our robust approach. Note that the PHAC standard error is asymptotically valid for the models with homogeneous slopes or random coefficients. Meanwhile, the nonparametric standard error used for the CCEP estimator is developed primarily for the models with random coefficients, and a different standard error may have to be employed for the models with homogeneous slopes; see Theorem 4 in Pesaran (2006). Nevertheless, the BC-IPC and JK-CCEP estimation results show that the savings-GDP ratio is highly significant in the regression of the investment-GDP ratio, indicating the existence of the F–H puzzle.

6. Concluding Remarks

In this article, we have proposed a robust approach against heteroscedasticity, error serial correlation and slope heterogeneity in linear models with interactive effects for large panel data. First, consistency and asymptotic normality of the pooled iterated principal component (IPC) estimator for the models

[11] See sec. 1 of Ginama, Hayakawa, and Kanemi (2018) for a brief literature survey.

[12] For the IPC estimation, the data are within-transformed and cross-sectionally demeaned, while for the CCEP estimation the data are within transformed only.

with random coefficients and homogeneous slopes have been established. Then, we have proved the asymptotic validity of the associated Wald test for slope parameter restrictions based on the panel heteroscedasticity and autocorrelation consistent (PHAC) variance matrix estimator for the models with random coefficients and homogeneous slopes, which does not require the Newey-West type time-series parameter truncation. These results asymptotically justify the use of the same pooled IPC estimator and the PHAC standard error for both homogeneous-slope and heterogeneous-slope models. This robust approach can significantly reduce the model selection uncertainty for applied researchers. In addition, we have proposed a Lagrange Multiplier (LM) test for correlated random coefficients with covariates. This test has nontrivial power against correlated random coefficients, but not for random coefficients and homogeneous slopes. The LM test is important because the IPC estimator becomes inconsistent with correlated random coefficients. The finite sample evidence and an empirical application support the usefulness and reliability of our robust approach.

We have examined finite sample performance of the estimators, the tests of linear restrictions on the slope parameters, and the LM tests for correlated random coefficients. We have examined the finite sample performance of the estimators, the tests of parameter restrictions using the PHAC variance estimator, and the LM test for correlated random coefficients. The results show that the size of the proposed robust Wald test with the bias-corrected IPC estimator is sufficiently close to the nominal level in both slope homogeneity and slope heterogeneity, and that the LM test for correlated random coefficients has correct size under both slope homogeneity and random coefficients while exhibiting high power for correlated random coefficients. The finite sample evidence, together with the empirical application for the Feldstein–Horioka puzzle, support the reliability and the usefulness of the proposed robust approach.

It is well recognized that bootstrapping in general can provide more accurate inference; see, for example, Hall (1992). Recently, there has been a rapid growth in the study of bootstrap methods for factor models; see Gonçalves and Perron (2014, 2020), among others. This is an interesting research direction, as applying bootstrap methods to panel data models with interactive effects can lead to more precise inferences.

The proposed approach investigated for the IPC estimator is most likely to be applicable for pooled estimators in other approaches, including the Common Correlated Effects (CCE) estimator of Pesaran (2006), the principal component (PC) estimator investigated by Westerlund and Urbain (2015), and the two-step instrumental variable (2SIV) estimator proposed by Cui et al. (2021), among others. In particular, for these estimators the factors in the regressors and in the error term are extracted, whilst the IPC estimator only exploits the information of the factors in the error term. Studying such applications seems to be an interesting extension.

As emphasized in the article, when the null hypothesis of random coefficient models is rejected in favor of the alternatives, it is preferable to employ estimators which are consistent for the models with correlated random coefficients. In this situation, the mean group estimators based on the CCE, 2SIV and ML approaches proposed by Pesaran (2006), Cui et al. (2021), and Li, Cui, and Lu (2020), among others, seem to be possible choices. Investigating inferential methods for such estimators under correlated random coefficients seems to be an intriguing future research theme.

Supplementary Materials

The supplementary material consists of online appendices that discuss the experimental results with proofs of key results and additional discussion, and computational codes that replicate the experimental and empirical results.

Acknowledgments

The authors thank Christian Hansen (the editor), the associate editor, and two anonymous referees for their valuable comments and suggestions. We are grateful to Milda Norkutė, Joakim Westerlund and the participants of the 2019 Asia Meeting of the Econometric Society for their helpful comments on the earlier version of this article.

Funding

This work was supported by JSPS KAKENHI (grant numbers 20H01484, 20H05631, 20K20760, 21H00700, and 21H04397).

References

Ahn, S., and Horenstein, A. (2013), "Eigenvalue Ratio Test for the Number of Factors," *Econometrica*, 81, 1203–1227. [8,9,10]

Ando, T., and Bai, J. (2017), "Clustering Huge Number of Financial Time Series: A Panel Data Approach with High-Dimensional Predictors and Factor Structures," *Journal of the American Statistical Association*, 112, 1182–1198. [2]

Arellano, M. (1987), "Computing Robust Standard Errors for Within Group Estimators," *Oxford Bulletin of Economics and Statistics*, 49, 431–434. [2]

Bai, J. (2009), "Panel Data Models with Interactive Fixed Effects," *Econometrica*, 77, 1229–1279. [1,2,3,4,5]

Bai, J., and Li, K., (2014), "Theory and Methods of Panel Data Models with Interactive Effects," *Annals of Statistics*, 42, 142–170. [1,3]

Bai, J., and Liao, Y. (2017), "Inferences in Panel Data with Interactive Effects Using Large Covariance Matrices," *Journal of Econometrics*, 200, 59–78. [1,2]

Bai, J., and Ng, S. (2002), "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70, 191–221. [8]

Bonhomme, S., and Manresa, E. (2015), "Grouped Patterns of Heterogeneity in Panel Data," *Econometrica*, 83, 1147–1184. [2]

Chamberlain, G. (1980), "Analysis of Covariance with Qualitative Data," *Review of Economic Studies*, 47, 225–238. [7]

Cui, G., Norkute, M., Sarafidis, V., and Yamagata, T. (2021), "Two-Stage Instrumental Variable Estimation of Linear Panel Data Models with Interactive Effects," *Econometrics Journal*, 25, 340–361. [1,2,3,4,5,6,8,13]

Dhaene, G., and Jochmans, K. (2015), "Split-Panel Jackknife Estimation of Fixed-Effect Models," *Review of Economic Studies*, 82, 991–1030. [5]

Feldstein, M. and Horioka, C. Y. (1980), "Domestic Saving and International Capital Flows," *Economic Journal*, 90, 314–329. [2,12]

Fernández, V., and Weidner, M. (2016), "Individual and Time Effects in Nonlinear Panel Models with Large *N*, *T*," *Journal of Econometrics*, 192, 291–312. [1,2,5,6,8,9,10]

Galvao, A. F., and Kato, K. (2014), "Estimation and Inference for Linear Panel Data Models Under Misspecification When both *n* and *T* are Large," *Journal of Business and Economic Statistics*, 32, 285–309. [2,6]

Giannone, D., and Lenza, M. (2010), "The Feldstein–Horioka Fact," in *NBER International Seminar on Macroeconomics* (Vol. 6), eds. R. Clarida and F. Giavazzi, pp. 103–117, Chicago, IL: University of Chicago Press. [12]

Ginama, I., Hayakawa, K., and Kanemi, T. (2018), "Examining the Feldstein–Horioka Puzzle Using Common Factor Panels and Interval Estimation," *Japan and the World Economy*, 48, 11–21. [12]

Gonçalves, S., and Perron, B. (2014), "Bootstrapping Factor-Augmented Regression Models," *Journal of Econometrics*, 182, 156–173. [13]

——— (2020), "Bootstrapping Factor Models with Cross Sectional Dependence," *Journal of Econometrics*, 218, 476–495. [13]

Greenaway-McGrevy, R., Han, C., and Sul, D. (2012), "Asymptotic Distribution of Factor Augmented Estimators for Panel Regression," *Journal of Econometrics*, 169, 48–45. [2,6]

Hall, P. (1992), *The Bootstrap and Edgeworth Expansion*, New York: Springer. [13]

Hansen, C. (2007), "Asymptotic Properties of a Robust Variance Matrix Estimator for Panel Data When t is Large," *Journal of Econometrics*, 141, 597–620. [2,6]

Li, K., Cui, G., and Lu, L. (2020), "Efficient Estimation of Heterogeneous Coefficients in Panel Data Models with Common Shocks," *Journal of Econometrics*, 216, 327–353. [3,8,13]

Moon, H. R., and Weidner, M. (2015), "Linear Regression for Panel with Unknown Number of Factors as Interactive Fixed Effects," *Econometrica*, 83, 1543–1579. [1,3,5]

Mundlak, Y. (1978), "On the Pooling of Time Series and Cross Section Data," *Econometrica*, 46, 69–85. [7]

Newey, W. K., and West, K. D. (1987), "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708. [2,5]

Norkutė, M., Sarafidis, V., Yamagata, T., and Cui, G. (2021), "Instrumental Variable Estimation of Dynamic Linear Panel Data Models with Defactored Regressors and a Multifactor Error Structure," *Journal of Econometrics*, 220, 416–446. [3]

Obstfeld, M., and Rogoff, K. (2000), "The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?," *NBER Macroeconomics Annual*, 15, 339–390. [12]

Pesaran, M. H. (2006), "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica*, 74, 967–1012. [1,2,3,4,5,8,12,13]

Pesaran, M. H., and Yamagata, T. (2008), "Testing Slope Homogeneity in Large Panels," *Journal of Econometrics*, 142, 50–93. [2,7]

Reese, S., and Westerlund, J. (2018), "Estimation of Factor-Augmented Panel Regressions with Weakly Influential Factors," *Econometric Reviews*, 37, 401–465. [2,4,5,8]

Su, L. and Chen, Q. (2013), "Testing Homogeneity in Panel Data Models with Interactive Fixed Effects," *Econometric Theory*, 29, 1079–1135. [2,7]

Su, L., and Ju, G. (2018), "Identifying Latent Grouped Patterns in Panel Data Models with Interactive Fixed Effects," *Journal of Econometrics*, 206, 554–573. [2]

Vogelsang, T. J. (2012), "Heteroskedasticity, Autocorrelation, and Spatial Correlation Robust Inference in Linear Panel Models with Fixed-Effects," *Journal of Econometrics*, 166, 303–319. [6]

Westerlund, J. (2018), "CCE in Panels with General Unknown Factors," *Econometrics Journal*, 21, 264–276. [1,2,5,12]

Westerlund, J., and Urbain, J. P. (2015), "Cross-Sectional Averages versus Principal Components," *Journal of Econometrics*, 185, 372–377. [1,3,8,13]

Wooldridge, J. (2010), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press. [7]