An LM Test for the Conditional Independence between Regressors and Factor Loadings in Panel Data Models with Interactive Effects

George Kapetanios, Laura Serlenga & Yongcheol Shin

View supplementary material

Published online: 04 Aug 2023.

Submit your article to this journal

Article views: 174

View related articles

View Crossmark data

Taylor & Francis
Taylor & Francis Group

🔓 OPEN ACCESS | Check for updates

An LM Test for the Conditional Independence between Regressors and Factor Loadings in Panel Data Models with Interactive Effects

George Kapetanios[a], Laura Serlenga*[b], and Yongcheol Shin[c]

[a]King's College London, London, UK; [b]University of Bari, Bari, Italy; [c]University of York, York, UK

ABSTRACT

A huge literature on modeling cross-sectional dependence in panels has been developed using interactive effects (IE). One area of contention is the hypothesis concerned with whether the regressors and factor loadings are correlated or not. Under the null hypothesis that they are conditionally independent, we can still apply the consistent and robust two-way fixed effects estimator. As an important specification test we develop an LM test for both static and dynamic panels with IE. Simulation results confirm the satisfactory performance of the LM test in small samples. We demonstrate its usefulness with an application to a total of 22 datasets, including static panels with a small T and dynamic panels with serially correlated factors, providing convincing evidence that the null hypothesis is not rejected in many datasets.

1. Introduction

The pervasive evidence in favor of the presence of strong cross-sectional dependence (CSD) in panels (e.g., Pesaran 2015) has prompted the development of rigorous econometric methodology for explicitly modeling CSD, mainly through multiplicative interactive effects (IE). Two leading approaches have received considerable attention. The first approach proposed by Bai (2009), on the basis of the principal component (PC) estimation, estimates the factors and the main parameters in an iterative manner. This has been extended by Moon and Weidner (2015), Fernandez-Val and Weidner (2016), and Charbonneau (2017). The second approach, referred to as the common correlated effects (CCE) estimator advanced by Pesaran (2006), treats factors as nuisance terms by removing their effects through proxying them by the cross-section averages of the dependent and independent variables. A few extensions have been developed by Kapetanios, Pesaran, and Yamagata (2011), Chudik and Pesaran (2015), Westerlund and Urbain (2015), and Juodis (2022). See Chudik and Pesaran (2015) for an extended survey.

The conventional wisdom is that the two-way fixed effects (FE) estimator would be inconsistent in the presence of IE, due to ignoring the correlation between the regressors and factors/loadings (e.g., Bai 2009). However, if the regressors are uncorrelated with the factor loadings (even if still correlated with factors), then the two-way FE estimator is still unbiased. This has been noted earlier by Coakley, Fuertes, and Smith (2006), Sarafidis and Wansbeek (2012), and Westerlund (2019a).

Kapetanios, Serlenga, and Shin (2023) formally establish that the FE estimator is consistent and asymptotically normally distributed under this situation while proposing a robust heteroscedasticity and autocorrelation consistent (HAC) variance estimator to deal with the presence of IE in error components. In what follows, we refer to this estimator as the "robust FE estimator".

A number of specification tests have been proposed for testing the presence of CSD or IE in panels, see Pesaran (2015), Sarafidis, Yamagata, and Robertson (2009), Bai (2009), Castagnetti, Rossi, and Trapani (2015), and Westerlund (2019b). As discussed above, however, the rejection of the null hypothesis by these tests does not always imply that the FE estimator is inconsistent. Surprisingly, the literature has been, in general, silent on investigating whether regressors are uncorrelated with factor loadings. This is the crucial hypothesis to be tested because if it is not rejected, then we can still apply the robust FE estimator as well as the PC estimator, though the consistency of the latter requires that the estimated number of factors is equal to or larger than the true one. Furthermore, under this situation, the FE estimator can be easily applied to static panels with a small number of time periods or dynamic panels with serially correlated factors.

In order to fill this gap, Kapetanios, Serlenga, and Shin (2023) developed a Hausman test that determines whether the regressors and factor loadings are correlated or not. It focuses on testing the null hypothesis that the difference between the FE and PC estimators is zero in population. This procedure is intuitive since we care about the cost of using the potentially inaccurate

---

CONTACT George Kapetanios ✉ george.kapetanios@kcl.ac.uk 🏛 King's College London, London, UK.
*Current affiliation: European Commission, JRC.

FE estimator. Within the context of the factor augmented panel data model, where both the dependent variable and regressors share overlapping sets of factors, this is equivalent to testing that regressors and loadings are conditionally independent. However, such a Hausman test requires the use of a well-specified PC estimator for its implementation (e.g., the Stata commands by Kripfganz and Sarafidis 2021), suggesting that it may still be cumbersome from a practical perspective.

As a main contribution in this article we propose an LM test that does not require the use of the PC estimator at all. After applying the two-way within transformation to the panel data model with IE, we only use the FE estimator to obtain the residuals and construct an LM test statistic for determining the validity of the orthogonal moment conditions between transformed regressors and residuals under the null hypothesis. Further, by a portmanteau characterization of the LM test, we propose the use of the first principal component of the regressors, in constructing the LM test statistic. We show that this approach can provide a valid test that is well behaved under the null and the alternative hypotheses (see Theorem 1 and the supporting simulation evidence in Section S4.4 in the online supplement). Importantly, we also establish that the validity of the LM test does not require consistent estimation of individual slope parameters even for small T.

Further, Kapetanios, Serlenga, and Shin (2023) consider only the static panel data model with large T, while the current paper extends the LM test to the dynamic panel data model as well as the static panel data model with small T. In addition, we note that the performance of the Hausman test proposed by Kapetanios, Serlenga, and Shin (2023) is very sensitive to model misspecification.

Next, for a dynamic panel data model with IE and serially correlated factors, we suggest the use of an autoregressive distributed lag (ARDL) approximation in constructing the LM test statistic. We establish that the corresponding LM test follows the χ^2 distribution under the null hypothesis whereas it diverges under the alternative hypothesis.

It is important to emphasize that our approach is designed to develop a crucial specification test that is missing in the literature rather than provide a model building strategy, analogous to conducting a serial correlation test. The proposed LM test then complements the popular CD test by Pesaran (2015) and the refined CD* test recently advanced by Pesaran and Xie (2022). Given the pervasive evidence in favor of strong CSD, we suggest applying the LM test to determine whether the form of IE invalidates the consistency of the FE estimator or not. However, it is worth emphasizing that both CD and CD* tests are not always satisfactory. The CD test fails to reject the null hypothesis of weak error CSD when loadings have zero means, implying that it displays very poor power when applied to cross-sectionally demeaned data. Furthermore, Juodis and Reese (2021) show that the application of the CD test to regression residuals obtained from models with IE, involves a bias, resulting in erroneous rejections under the null hypothesis (e.g., Mastromarco, Serlenga, and Shin 2016). Thus, the CD* test is developed to correct the asymptotic bias of the CD test using the estimates of factor loadings and error variances, though its performance crucially relies upon a tuning parameter (the optimal number of factors). On the contrary, we can establish the validity of the LM test under weaker conditions, irrespective of the presence of CSD or selecting the correct number of factors, as a tool for testing the consistency of the FE estimator for the panel data models with IE.

Via Monte Carlo experiments, we find that the finite sample performance of the LM test is satisfactory even under pronounced parameter heterogeneity, serial and weak cross-sectional error correlation. Furthermore, its performance remains satisfactory for static panels with a small number of time periods and for dynamic panels with serially correlated factors. The LM test is also robust to model misspecification, unlike the Hausman test proposed by Kapetanios, Serlenga, and Shin (2023) that tends to display severe size distortions in the presence of neglected regressors.

We demonstrate the usefulness of the LM test with an application to a total of 22 datasets employed in the literature. At the 5% (10%) significance the LM test does not reject the null hypothesis in 12 (10) out of 14 datasets under the static panels whilst the null is marginally rejected only once out of five datasets under the dynamic framework. Furthermore, the LM test does not reject the null hypothesis for all three static panels with a small T. These results provide convincing evidence that the null hypothesis of conditional independence between factor loadings and the regressors is not rejected for many datasets.

The article proceeds as follows. Section 2 develops the LM test for the null hypothesis of the conditional independence between the regressors and factor loadings for static and dynamic panel data models with IE and presents the asymptotic theory. Section 3 investigates the finite sample performance of the LM test. Section 4 presents empirical evidence. Section 5 offers concluding remarks. Mathematical proofs and additional analytic/simulation results are relegated in the online supplement.

2. The LM Test

Consider a heterogeneous panel data model with interactive effects:

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

where y_{it} is the dependent variable of the ith cross-sectional unit in period t, $\boldsymbol{x}_{it} = (x_{i,it}, \ldots, x_{k,it})'$ is the $k \times 1$ vector of covariates with $\boldsymbol{\beta}_i$ the $k \times 1$ vector of parameters, and ε_{it} is an idiosyncratic error. $\boldsymbol{f}_t$ is an $r \times 1$ vector of unobserved factors and $\boldsymbol{\gamma}_i$ is an $r \times 1$ vector of heterogeneous loadings.

For consistent estimation of $\boldsymbol{\beta}$, two main approaches have been proposed to account for unobserved factors. The CCE estimator advanced by Pesaran (2006), imposes that $\boldsymbol{x}_{it}$ share the same factors, $\boldsymbol{f}_t$:

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

where $\boldsymbol{\Lambda}_i$ an $r \times k$ matrix of heterogeneous loadings and $\boldsymbol{v}_{it} = (v_{1it}, \ldots, v_{kit})'$ are idiosyncratic errors. CCE approximates $\boldsymbol{f}_t$ by the cross-section averages of the dependent and independent variables. The DGP condition in (2) simplifies the analysis considerably, but can be viewed as restrictive (see Remark 4). Next, Bai (2009) allows $\boldsymbol{x}_{it}$ to be arbitrarily correlated with both $\boldsymbol{\gamma}_i$ and $\boldsymbol{f}_t$, and proposes an iterative PC approach that estimates the factors/loadings together with $\boldsymbol{\beta}$.

A number of specification tests have been proposed to test the presence of CSD or IE in panels. The most popular test is the cross-section dependence (CD) test proposed by Pesaran (2015, 2021) that is increasingly used as an ex-post diagnostic tool. The CD test may be used as a model-selection tool, with a reduction in the absolute value of the CD statistic typically being interpreted as an indication of an improved model specification. Bai (2009) advances a Hausman test for testing the null hypothesis of (two-way) additive fixed effects (i.e., $\gamma_i' f_t = \alpha_i + \theta_t$) against the alternative of the multiplicative IE as

$$H_B = \left(\hat{\boldsymbol{\beta}}_{\text{FE}} - \hat{\boldsymbol{\beta}}_{\text{PC}} \right)' \boldsymbol{V}_B^{-1} \left(\hat{\boldsymbol{\beta}}_{\text{FE}} - \hat{\boldsymbol{\beta}}_{\text{PC}} \right). \tag{3}$$

Under the null hypothesis, Bai (2009) derives that $\boldsymbol{V}_B = \text{var}\left(\hat{\boldsymbol{\beta}}_{\text{PC}} \right) - \text{var}\left(\hat{\boldsymbol{\beta}}_{\text{FE}} \right)$ and $H_B \to_d \chi_k^2$ where $\text{var}\left(\hat{\boldsymbol{\beta}}_{\text{FE}} \right)$ is the variance estimator provided by the two-way FE estimation and $\text{var}\left(\hat{\boldsymbol{\beta}}_{\text{PC}} \right)$ is the variance estimator that accommodates unknown forms of heteroscedasticity and autocorrelation in errors. Westerlund (2019b) proposes an alternative Hausman test, denoted H_W, by replacing the PC estimator with the CCE estimator, see Section S2 in the Online Supplement for details.

The conventional wisdom is that if the CD test rejects the null hypothesis of weak error CSD, then the FE estimator is biased/inconsistent due to the endogeneity arising from the correlation between the regressors and unobserved factors/loadings. Notice, however, that if the regressors are conditionally independent from loadings, the two-way FE estimator is still consistent for the panel data model with IE. This has been noted earlier by Coakley, Fuertes, and Smith (2006), Sarafidis and Wansbeek (2012), and Westerlund (2019a). Under this situation, Kapetanios, Serlenga, and Shin (2023) formally establish that the FE estimator is consistent and follows an asymptotic normal distribution and propose the robust FE estimator using the HAC variance estimator to deal with the presence of IE in error components, that is shown to be robust to the presence of heteroscedastic and serially-correlated disturbances as well as parameter heterogeneity.

In empirical applications below we apply both CD test and Bai's Hausman test to several datasets employed in the literature. Surprisingly, we find the conflicting result that the CD test strongly rejects the null hypothesis of weak error CSD while the H_B test rarely rejects the null of (FE) additive-effects in most datasets. A careful inspection reveals that if factor loadings and the regressors are conditionally independent, then the H_B test becomes inconsistent. To investigate this issue, we examine the powers of both CD and H_B tests for the heterogeneous panel data with IE in Section S2 in the Online Supplement. We consider the two experiments. Under Experiment 1 we generate the loadings to be independent from the regressors while they are correlated under Experiment 2. In both Experiments we maintain that the regressors are correlated with factors. As expected, the CD test results display that the null of weak residual CSD is strongly rejected for all the data generating process (DGP) under both Experiments. Under Experiment 1, however, the H_B test does not display any power and its rejection probability is close to and sometimes lower than the nominal size, especially in the presence of serially correlated errors. This is because the FE estimator is still consistent under Experiment 1. The H_B test is

consistent only under Experiment 2. This confirms the limitation of applying the H_B test in practice.

This suggests that the null hypothesis of the conditional independence between the regressors and factor loadings emerges as an influential but underappreciated feature of the panel data model with IE. Surprisingly, the literature has been silent on investigating this important issue. Pesaran (2006) implicitly assumes that the factor loadings γ_i in (1) and Λ_i in (2), are uncorrelated, see also Sarafidis, Yamagata, and Robertson (2009). Westerlund and Urbain (2013) questions the assumption of uncorrelated factor loadings while Bai (2009) shows via simulations that the CCE estimator is biased when x_{it} is correlated with both γ_i and f_t. In this regard, we consider the null and alternative hypotheses as follows:

$$H_0 : E\left(\gamma_i - E(\gamma_i) \middle| x_{it}, f_t \right) = 0 \ \forall i \tag{4}$$

$$H_1 : E\left(\gamma_i - E(\gamma_i) \middle| x_{it}, f_t \right) \neq 0 \ \forall i. \tag{5}$$

This can be regarded as the important misspecification test because if the null hypothesis (4) is not rejected, then the (robust) two-way FE estimator is still valid to employ in the panel data with IE.

Notice that $\gamma_i' f_t$ in (1) can be expressed equally as

$$\gamma_i' f_t = \mu + \alpha_i + \theta_t + \mathring{\gamma}_i' \mathring{f}_t \tag{6}$$

where $\mu = \bar{\gamma}' \bar{f}, \alpha_i = \gamma_i' \bar{f}, \theta_t = \bar{\gamma}' f_t, \mathring{\gamma}_i = \gamma_i - \bar{\gamma}$ and $\mathring{f}_t = f_t - \bar{f}$ with $\bar{\gamma} = N^{-1} \sum_{i=1}^{N} \gamma_i$ and $\bar{f} = T^{-1} \sum_{t=1}^{T} f_t$. Using (6), we rewrite (1) as

$$y_{it} = \boldsymbol{\beta}_i' x_{it} + \mu + \alpha_i + \theta_t + \mathring{\gamma}_i' \mathring{f}_t + \varepsilon_{it} \tag{7}$$

Applying the two-way within transformation to (7), we obtain:

$$\ddot{y}_{it} = \boldsymbol{\beta}_i' \ddot{x}_{it} + \ddot{u}_{it}, \ \ddot{u}_{it} = \mathring{\gamma}_i' \mathring{f}_t + \ddot{\varepsilon}_{it} \tag{8}$$

where $\ddot{y}_{it} = y_{it} - \bar{y}_{i.} - \bar{y}_{.t} + \bar{y}_{..}, y_{i.} = T^{-1} \sum_{t=1}^{T} y_{it}, y_{.t} = N^{-1} \sum_{i=1}^{N} y_{it}$ and $\bar{y}_{..} = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it}$. Similarly for $\ddot{x}_{it}$ and $\ddot{\varepsilon}_{it}$. Under the null hypothesis, (4), it is easily seen by $E(\mathring{\gamma}_i) = E(\gamma_i - \bar{\gamma}) = 0$ and assuming independence of γ_i from all other random quantities in the model that $\ddot{x}_{it}$ is uncorrelated with the transformed composite error, $\ddot{u}_{it} = \mathring{\gamma}_i' \mathring{f}_t + \ddot{\varepsilon}_{it}$, provided that x_{it} are strictly exogenous with respect to ε_{it} because $E\left(\ddot{x}_{it} \mathring{\gamma}_i' \mathring{f}_t \right) = E\left\{ \ddot{x}_{it} \mathring{f}_t' E\left(\mathring{\gamma}_i \middle| \ddot{x}_{it}, \mathring{f}_t \right) \right\} = 0$ (see also Section 5 in Hsiao 2018). Therefore, under the null hypothesis (4), we obtain the following $k \times 1$ vector of moment conditions:

$$E\left(\ddot{u}_{it} \middle| \ddot{x}_{it} \right) = \boldsymbol{0}. \tag{9}$$

The conventional panel data without IE satisfies this condition under the maintained assumption that the regressors are exogenous. Now, our main contribution lies in showing that this moment condition can also be satisfied for the panel data model with IE under our null hypothesis.

We obtain an individual two-way FE estimator of $\boldsymbol{\beta}_i$ and the mean group (MG) estimator by

$$\hat{\boldsymbol{\beta}}_{\text{FE},i} = \left(\ddot{X}_i' \ddot{X}_i \right)^{-1} \ddot{X}_i' \ddot{y}_i \text{ and } \hat{\boldsymbol{\beta}}_{\text{FE,MG}} = \frac{1}{N} \sum_{i=1}^{N} \hat{\boldsymbol{\beta}}_{\text{FE},i} \tag{10}$$

Notice that $\hat{\boldsymbol{\beta}}_{\text{FE},i}$ is unbiased but inconsistent for $\boldsymbol{\beta}_i$ whereas $\hat{\boldsymbol{\beta}}_{\text{FE,MG}}$ is $\sqrt{N}$-consistent for $\boldsymbol{\beta} = E(\boldsymbol{\beta}_i)$. We have: $\hat{\boldsymbol{\beta}}_{\text{FE},i} - \boldsymbol{\beta}_i = \left(\frac{\ddot{X}_i'\ddot{X}_i}{T}\right)^{-1}\left(\frac{\ddot{X}_i'\dot{F}}{T}\right)\hat{\boldsymbol{\gamma}}_i + \left(\frac{\ddot{X}_i'\ddot{X}_i}{T}\right)^{-1}\left(\frac{\ddot{X}_i'\boldsymbol{\varepsilon}_i}{T}\right)$. Clearly, $\left(\frac{\ddot{X}_i'\ddot{X}_i}{T}\right)^{-1} = O_p(1)$, $\left(\frac{\ddot{X}_i'\dot{F}}{T}\right) = O_p(1)$ and $\frac{\ddot{X}_i'\boldsymbol{\varepsilon}_i}{T} = O_p(T^{-1/2})$. Although $E\left(\frac{\ddot{X}_i'\dot{F}}{T}\right)\hat{\boldsymbol{\gamma}}_i = 0$, we still have $\left(\frac{\ddot{X}_i'\dot{F}}{T}\right)\hat{\boldsymbol{\gamma}}_i = O_p(1)$. Hence, $\hat{\boldsymbol{\beta}}_{\text{FE},i}$ is unbiased but inconsistent. Next, we have: $\hat{\boldsymbol{\beta}}_{\text{FE,MG}} - \boldsymbol{\beta} = O_p(N^{-1/2})$ such that the MG estimator is $\sqrt{N}$ consistent (see Kapetanios, Serlenga, and Shin 2023). We conducted (unreported) simulation exercises, finding that the biases of the individual and MG estimators are mostly negligible. The variance of the MG estimator declines with N, but that of the individual estimator does not decline with T. Conversely, if x_{it} and $\boldsymbol{\gamma}_i$ are correlated, both estimators are biased and inconsistent due to $E(\ddot{x}_{it}\ddot{u}_{it}) \neq 0$.

Stacking (8) over t, we have:

$$\ddot{\boldsymbol{y}}_i = \ddot{X}_i\boldsymbol{\beta}_i + \ddot{\boldsymbol{u}}_i, \; \ddot{\boldsymbol{u}}_i = \dot{F}\hat{\boldsymbol{\gamma}}_i + \ddot{\boldsymbol{\varepsilon}}_i \tag{11}$$

where $\ddot{\boldsymbol{y}}_i = (\ddot{y}_{i1},\ldots,\ddot{y}_{iT})'$, $\ddot{\boldsymbol{u}}_i = (\ddot{u}_{i1},\ldots,\ddot{u}_{iT})'$, $\ddot{X}_i = (\ddot{x}_{i1},\ldots,\ddot{x}_{iT})'$, $\dot{F} = (\dot{f}_1,\ldots,\dot{f}_T)'$ and $\ddot{\boldsymbol{\varepsilon}}_i = (\ddot{\varepsilon}_{i1},\ldots,\ddot{\varepsilon}_{iT})'$. Under the moment conditions in (9), it is easily seen that

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\ddot{X}_i'\ddot{\boldsymbol{u}}_i}{T} \to_d N(0, \boldsymbol{V}) \text{ under } H_0 \tag{12}$$

where $\boldsymbol{V} = \lim_{N,T\to\infty}\frac{1}{N}\sum_{i=1}^{N}\frac{\ddot{X}_i'\ddot{\boldsymbol{u}}_i\ddot{\boldsymbol{u}}_i'\ddot{X}_i}{T^2}$. Under H_1 we have $E(\ddot{x}_{it}\ddot{u}_{it}) \neq 0$, from which we have:

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\ddot{X}_i'\ddot{\boldsymbol{u}}_i}{T} = O_p\left(\sqrt{N}\right) \tag{13}$$

Using the difference between (12) and (13), we can construct the LM test statistic as follows:

$$\text{LM} = \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\ddot{X}_i'\ddot{\boldsymbol{u}}_i}{T}\right)'\boldsymbol{V}^{-1}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\ddot{X}_i'\ddot{\boldsymbol{u}}_i}{T}\right)$$
$$\to_d \chi_k^2 \text{ under } H_0 \tag{14}$$

whereas the test diverges under H_1.

To develop an operational version of the LM test, we replace $\ddot{\boldsymbol{u}}_i$ by the FE residuals given by

$$\widehat{\ddot{\boldsymbol{u}}}_i = \ddot{\boldsymbol{y}}_i - \ddot{X}_i\hat{\boldsymbol{\beta}}_{\text{FE},i} \tag{15}$$

The corresponding statistic is identical to zero since $\ddot{X}_i$ are orthogonal to $\widehat{\ddot{\boldsymbol{u}}}_i$, that is, $\sum_{i=1}^{N}\ddot{X}_i'\widehat{\ddot{\boldsymbol{u}}}_i = 0$. We propose replacing $\ddot{X}_i$ by a fitted value, denoted $\widehat{\ddot{X}}_i$, and obtaining $\widehat{\ddot{X}}_i$ in a simple manner. We first estimate factors, denoted $\widehat{F}_X$, obtained by $\sqrt{T}$ times the $T \times r$ eigenvectors corresponding to the r largest eigenvalues of the covariance matrix of the regressors, $\frac{1}{NT}\sum_{i=1}^{N}\ddot{X}_i\ddot{X}_i'$. Next, by a portmanteau characterization of the LM test, we propose to employ only the first estimated factor corresponding to the largest eigenvalue, though the true number of factors, r is larger than one (see the proof of Theorem 1 and the supporting simulation evidence in Section S4.4 in the Online supplement.) Then, the final version of the LM test is constructed as follows:

$$\text{LM}_X = \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\widehat{\ddot{X}}_i'\widehat{\ddot{\boldsymbol{u}}}_i}{T}\right)'\hat{\boldsymbol{V}}^{-1}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\widehat{\ddot{X}}_i'\widehat{\ddot{\boldsymbol{u}}}_i}{T}\right) \tag{16}$$

where $\widehat{\ddot{X}}_i = \widehat{F}_X^{(1)}\left(\widehat{F}_X^{(1)\prime}\widehat{F}_X^{(1)}\right)^{-1}\widehat{F}_X^{(1)\prime}\ddot{X}_i$ with $\widehat{F}_X = \left(\widehat{F}_X^{(1)},\ldots,\widehat{F}_X^{(r)}\right)$, and $\hat{\boldsymbol{V}} = \frac{1}{N}\sum_{i=1}^{N}\frac{\widehat{\ddot{X}}_i'\widehat{\ddot{\boldsymbol{u}}}_i\widehat{\ddot{\boldsymbol{u}}}_i'\widehat{\ddot{X}}_i}{T^2}$.

We present Assumption A containing the regularity conditions.

Assumption A. (i) $E(\varepsilon_{it}) = 0$, $E\left(\varepsilon_{it}^2\right) = \sigma_{\varepsilon_i}^2$ and $E\left(\varepsilon_{it}^{2+\delta}\right) < \infty$ for $i = 1,\ldots,N$; $t = 1,\ldots,T$ and for some $\delta > 0$. Let $\mathcal{F}_{-\varepsilon_{it}}$ be the σ-field of all stochastic elements in the panel data model, apart from ε_{it}. Then, $E\left(\varepsilon_{it}\mid \mathcal{F}_{-\varepsilon_{it}}\right) = 0$.

(ii) f_t has finite $2 + \delta$ moments for some $\delta > 0$.

(iii) $\boldsymbol{\gamma}_i$ has finite mean, $\bar{\boldsymbol{\gamma}}$ and positive definite variance, $\boldsymbol{\Sigma}_{\gamma}$. Let $\mathcal{F}_{-\gamma_i}$ be the σ-field of all stochastic elements apart from $\boldsymbol{\gamma}_i$. Then, under the null hypothesis (4), we have:

$$E\left(\boldsymbol{\gamma}_i - \bar{\boldsymbol{\gamma}}\mid \mathcal{F}_{-\gamma_i}\right) = 0. \tag{17}$$

Under the alternative hypothesis (5), $\sum_{i=1}^{N}E\left(\frac{\ddot{X}_i'\dot{F}\hat{\boldsymbol{\gamma}}_i}{T}\right) = O(N)$, and $\frac{\ddot{X}_i'\dot{F}\hat{\boldsymbol{\gamma}}_i}{T} - E\left(\frac{\ddot{X}_i'\dot{F}\hat{\boldsymbol{\gamma}}_i}{T}\right)$ is a spatial martingale difference process (see Definition 1 in the online supplement).

(iv) x_{it} have finite $2+\delta$ moments. $\left(\frac{\ddot{X}_i'\ddot{X}_i}{T}\right)^{-1}$ and $\left(\frac{\widehat{F}_X^{(1)\prime}\widehat{F}_X^{(1)}}{T}\right)^{-1}$ exist for all $T \geq T_0$. Furthermore, the elements of $\left(\frac{\ddot{X}_i'\ddot{X}_i}{T}\right)^{-1}$, $\left(\frac{\widehat{F}_X^{(1)\prime}\widehat{F}_X^{(1)}}{T}\right)^{-1}$, $\frac{\widehat{F}_X^{(1)\prime}\ddot{X}_i}{T}$, $\frac{\ddot{X}_i'\dot{F}}{T}$ and $\frac{\widehat{F}_X^{(1)\prime}\dot{F}}{T}$ have finite $2 + \delta$ moments for all $T \geq T_0$ and for some $\delta > 0$ uniformly over i. Finally, $\sup_{\ell,i,j,t,i\neq j}\text{corr}(x_{\ell,it}, x_{\ell,jt}) < 1$.

Assumption A is much weaker than what is usually found in the factor literature (e.g., Bai 2009; Karabiyik, Reese, and Westerlund 2017; Cui et al. 2023; Cui, Sarafidis, and Yamagata 2022), mainly because the LM test requires the PC extraction from regressors x_{it} only, not the residuals from the main regression, (1). Further, we do not require that the unobserved factors are consistently estimated by PC, thus, reducing the need for restrictive assumptions. This is a substantial relaxation since we only require existence of $2 + \delta$ moments of ε_{it} instead of eighth moments, usually imposed in this literature. In Section S.4.3 in the Online Supplement, we show via further simulations that if the higher moments (greater than 3) do not exist, then the FE estimator outperforms the PC estimator in all sample sizes. For simplicity we make a zero conditional expectation assumption for the idiosyncratic errors ε_{it}, though it can be easily extended to the case with a serial and (weak) cross-sectional correlation at the expense of slightly more convoluted proofs (see Assumption A' used for the dynamic panel data model in the Online Appendix and the simulation evidence in Section 3). We impose a zero conditional expectation condition in (17) under the null

hypothesis, that is weaker than independence of $\boldsymbol{\gamma}_i$ across i. It is slightly stronger than (4) but makes the theoretical derivations much simpler. The first three parts of Assumption A(iv) are standard identification conditions on the existence of moments and their relevant inverses, and needed to derive the asymptotic distribution of the LM test for fixed T, though it also holds as $T \rightarrow \infty$. The final part of A(iv) is a regularity condition ensuring no pathological cases of perfect multicollinearity of the regressors. This condition is simply satisfied for the case where (2) holds, if $E(\boldsymbol{v}_{it}\boldsymbol{v}'_{it})$ is positive definite uniformly over i and t, though we do not need to impose this condition for the static model. Finally, we slightly restrict our definition of the null hypothesis in line with the technical proofs. Under the alternative hypothesis, we require that the sum of the expectations of the product of the regressors and the factor components, over panel units, explodes.

We have the main theoretical result in Theorem 1.

Theorem 1. Consider the static heterogeneous panel data model with IE given by (1). Under Assumption A and under the null hypothesis, (4), as $N \rightarrow \infty$, the LM_X test statistic in (16) asymptotically follows the χ^2_k distribution, where k is the number of regressors. Under the alternative hypothesis, (5), the LM_X test is consistent.

Remark 1. Theorem 1 holds for finite T as well as for $T \rightarrow \infty$. Initially, we place our work within a macroeconometric literature that views an effective estimation of $\boldsymbol{\beta}$ as a main aim. For large T, it is natural to maintain an assumption that the regressors are correlated with factors, $\boldsymbol{f}_t$ (representing the common policy or globalization trend) to avoid any omitted variables bias while it remains an important issue to test whether the regressors are correlated with the heterogeneous loadings. Moreover, there is a growing literature in microeconometrics that focuses on estimating causal effects by expanding models using factor components. Examples include estimating causal/counterfactual effects using synthetic controls (e.g., Arkhangelsky et al. 2021) or cautioning about the causal interpretation of the two-way FE estimator in difference-in-difference setups (e.g., Athey and Imbens 2022). In this regard our specification test can be regarded as a significant contribution in this literature with large N and small T, as it can provide support for the use of consistent and robust FE estimator under the null hypothesis.

Remark 2. Notice that the LM_X test remains valid in the absence of a factor structure in y_{it}. In this case the null hypothesis, (4), makes no sense since there are no loadings. But, we can still entertain the moment condition in (9) under which it is easily seen that the LM test, based on extracting a principal component that is well defined in the absence of a factor structure, follows a χ^2 distribution (see the simulation evidence in Section 3). Hence, we can establish the validity of the LM test under weak conditions, as a tool for testing the consistency of the two-way FE estimation, irrespective of the presence of IE in the model.

Remark 3. We may construct the FE residuals using the MG estimator such as $\widehat{\widehat{\boldsymbol{u}}}_i = \ddot{\boldsymbol{y}}_i - \ddot{\boldsymbol{X}}_i \hat{\boldsymbol{\beta}}_{\mathrm{FE,MG}}$. We can still show that the corresponding LM statistic follows a limiting χ^2 distribution under the null hypothesis (4), though there are complex techni-

cal issues related to establishing consistency of the HAC variance estimator used in constructing the LM statistic. Thus, we focus on the LM_X test based on the individual FE estimator given by $\widehat{\boldsymbol{u}}_i = \ddot{\boldsymbol{y}}_i - \ddot{\boldsymbol{X}}_i \hat{\boldsymbol{\beta}}_{\mathrm{FE},i}$. The main advantage of this approach lies in that the validity of the LM test does not require consistent estimation of $\boldsymbol{\beta}_i$ for small T. Furthermore, as we employ the individual FE estimator from the transformed regression (8), not the MG estimator, our approach does not need to make a random coefficient assumption about $\boldsymbol{\beta}_i$ (say, the $k \times 1$ vector of $\boldsymbol{\beta}_i$ generated as $\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\eta}_i$, where $\boldsymbol{\eta}_i$ is independent across i with $E(\boldsymbol{\eta}_i) = 0$ and $E(\boldsymbol{\eta}_i \boldsymbol{\eta}'_i) = \Omega_{\eta\eta,i}$), implying that our main results hold irrespective of whether slope parameters are homogeneous or heterogeneous (see the proof of Theorem 1 in the Online Supplement).

Remark 4. It is important to emphasize that we do not have to impose the DGP condition for $\boldsymbol{x}_{it}$ in (2) in Theorem 1. $\boldsymbol{x}_{it}$ may have no factor or can contain different factors from those in y_{it}. For example, we can consider the general case where y_{it} and $\boldsymbol{x}_{it}$ share a subset of common factors $\boldsymbol{g}_t$ while they are subject to further specific factors, $\boldsymbol{f}_{1t}$ and $\boldsymbol{f}_{2t}$ such that $\boldsymbol{f}_{yt} = (\boldsymbol{g}'_t, \boldsymbol{f}'_{1t})'$ and $\boldsymbol{f}_{xt} = (\boldsymbol{g}'_t, \boldsymbol{f}'_{2t})'$ (see Monte Carlo evidence in Section 3). The only condition required is A(iii). However, if we consider the dynamic panel data, we need to impose a more specific structure such as (2) (see Assumption A' in the Online Appendix).

Remark 5. We can generalize our testing approach through providing links with the monograph of Godfrey (1989) on the LM testing approach. Consider the model,

$$y_{it} = \boldsymbol{\beta}'_i \boldsymbol{x}_{it} + \boldsymbol{\gamma}'_i \boldsymbol{z}_{it} + \varepsilon_{it} \qquad (18)$$

where $\boldsymbol{z}_{it}$ can be any set of regressors that satisfy the usual regularity conditions. Suppose that the estimated model is given by

$$y_{it} = \boldsymbol{\beta}'_i \boldsymbol{x}_{it} + u_{it} \qquad (19)$$

which is misspecified by missing $\boldsymbol{z}_{it}$ such that $u_{it} = \boldsymbol{\gamma}'_i \boldsymbol{z}_{it} + \varepsilon_{it} = q_{it} + \varepsilon_{it}$ with $q_{it} = \boldsymbol{\gamma}'_i \boldsymbol{z}_{it}$. By applying a two-way within transformation, we obtain the estimated model by $\ddot{y}_{it} = \boldsymbol{\beta}'_i \ddot{\boldsymbol{x}}_{it} + \ddot{u}_{it}$ and $\ddot{u}_{it} = \ddot{q}_{it} + \ddot{\varepsilon}_{it}$ where $\ddot{q}_{it} = \mathring{\boldsymbol{\gamma}}'_i \ddot{\boldsymbol{z}}_{it}$. Let $\boldsymbol{s}_{it}$ be a set of variables, designed to capture some model misspecification. $\boldsymbol{s}_{it}$ can overlap partly or wholly with $\boldsymbol{z}_{it}$, or be a proxy for $\boldsymbol{z}_{it}$. For example, we have: $\boldsymbol{s}_{it} = \hat{\boldsymbol{f}}_t$ and $\boldsymbol{z}_{it} = \boldsymbol{f}_t$. Define $\boldsymbol{\phi} = \mathrm{plim}_{N,T \rightarrow \infty} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \boldsymbol{s}_{it} \ddot{q}_{it}$ that is a composite parameter depending on a number of model parameters such as regression coefficients and correlations between $\boldsymbol{s}_{it}$, $\boldsymbol{x}_{it}$ and $\boldsymbol{z}_{it}$. Notice that $\sum_{i=1}^{N} \sum_{t=1}^{T} \boldsymbol{s}_{it} \ddot{q}_{it}$ is part of the score function of the regression model,

$$y_{it} = \boldsymbol{\beta}'_i \boldsymbol{x}_{it} + \boldsymbol{\delta}'_i \boldsymbol{s}_{it} + \varepsilon_{it} \qquad (20)$$

evaluated at $(\boldsymbol{\beta}'_i, \boldsymbol{\delta}'_i)' = (\boldsymbol{\beta}'_i, \boldsymbol{0}')'$. If $\boldsymbol{\phi} = 0$, we can use (19) to carry out a consistent estimation of $\boldsymbol{\beta}$, rather than using the more complex model, (20) that requires the construction of $\boldsymbol{s}_{it}$, as a means of approximating (18). It would be cumbersome to construct $\boldsymbol{s}_{it}$ in some instances, as is the case for $\boldsymbol{s}_{it} = \hat{\boldsymbol{f}}_t$ and $\boldsymbol{z}_{it} = \boldsymbol{f}_t$. So we formally test the null hypothesis, $\boldsymbol{\phi} = 0$, using the test statistic given by

$$\left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\ddot{\boldsymbol{S}}'_i \widehat{\widehat{\boldsymbol{u}}}_i}{T} \right)' \boldsymbol{V}^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\ddot{\boldsymbol{S}}'_i \widehat{\widehat{\boldsymbol{u}}}_i}{T} \right)$$

where $\ddot{\boldsymbol{S}}_i = (\ddot{\boldsymbol{s}}_{i1}, \ldots, \ddot{\boldsymbol{s}}_{iT})'$ and $\widehat{\ddot{\boldsymbol{u}}}_i = \ddot{\boldsymbol{y}}_i - \ddot{\boldsymbol{X}}_i \hat{\boldsymbol{\beta}}_{\text{FE}}$. This is not a test of whether adding $\boldsymbol{s}_{it}$ to (19) improves the model fit. Such a test would be based on a panel version of the F-test, and would reject the null, $\boldsymbol{\phi} = 0$, as long as $\boldsymbol{\gamma}_i \neq 0$. What we test here is whether we can consistently estimate $\boldsymbol{\beta}$ by (19), coupled with the FE estimation. Notice that the two null hypotheses ($H_0^{(1)} : \boldsymbol{\phi} = 0$ and $H_0^{(2)} : \boldsymbol{\gamma}_i = 0, \forall i$) overlap under many circumstances. For example, $\boldsymbol{\phi} = 0$ requires $\boldsymbol{\gamma}_i = 0$ under (1) and (2), if $N = 1$. The fact that they do not overlap in our setting is closely linked to our null hypothesis that $\boldsymbol{x}_{it}$ is conditionally independent from $\boldsymbol{\gamma}_i, \forall i$.

Remark 6. Another useful property of the LM test is its relative robustness to model misspecification under the null hypothesis. To formalize this we augment the model (18) to $y_{it} = \boldsymbol{\beta}'_i \boldsymbol{x}_{it} + \boldsymbol{\gamma}'_i \boldsymbol{f}_t + \boldsymbol{\delta}'_i \boldsymbol{z}_{it} + \varepsilon_{it}$ and complete this specification by assuming that $\boldsymbol{x}_{it}$ follows (2) and setting $\boldsymbol{z}_{it} = \boldsymbol{\Delta}'_i \boldsymbol{f}_t + \boldsymbol{\xi}_{it}$ where $\boldsymbol{\Delta}_i$ is iid across i with finite mean, $\bar{\boldsymbol{\Delta}}$ and variance, $\boldsymbol{\Sigma}_\Delta$. $\boldsymbol{\Delta}_i$ are independent of ε_{jt} and $\boldsymbol{f}_t$ for all i, j and t. Suppose that we retain (19) as the estimated model. Then, we examine if

$$\text{plim}_{N,T\to\infty} \frac{\widehat{\ddot{\boldsymbol{X}}}'_i \widehat{\ddot{\boldsymbol{u}}}_i}{T} = 0, \tag{21}$$

which determines whether the LM test retains its validity as a test of the null hypothesis, (4) under the model misspecification. If $\boldsymbol{z}_{it}$ and $\boldsymbol{x}_{it}$ are uncorrelated (i.e., $\boldsymbol{v}_{it}$ and $\boldsymbol{\xi}_{it}$ are uncorrelated), it is straightforward to show that $\widehat{\ddot{\boldsymbol{u}}}_{it} = \left(\hat{\boldsymbol{\beta}}'_{\text{FE}} - \boldsymbol{\beta}' \right) \ddot{\boldsymbol{x}}_{it} + \ddot{\boldsymbol{\gamma}}'_i \dot{\boldsymbol{f}}_t + \boldsymbol{\delta}'_i \ddot{\boldsymbol{z}}_{it} + \ddot{\varepsilon}_{it}$ such that (21) holds since $p\lim_{N,T\to\infty} \hat{\boldsymbol{\beta}}_{\text{FE}} - \boldsymbol{\beta} = 0$. On the other hand, if $\boldsymbol{z}_{it}$ and $\boldsymbol{x}_{it}$ are correlated, then $p\lim_{N,T\to\infty} \hat{\boldsymbol{\beta}}_{\text{FE}} - \boldsymbol{\beta} = \boldsymbol{\beta}_b \neq 0$. Hence,

$$\boldsymbol{\beta}'_b \ddot{\boldsymbol{x}}_{it} + \ddot{\boldsymbol{\gamma}}'_i \dot{\boldsymbol{f}}_t + \ddot{q}_{it} + \ddot{\varepsilon}_{it} = \left(\boldsymbol{\beta}'_b \mathring{\boldsymbol{\Lambda}}'_i + \mathring{\boldsymbol{\gamma}}'_i + \boldsymbol{\delta}'_i \mathring{\boldsymbol{\Delta}}'_i \right) \dot{\boldsymbol{f}}_t$$
$$+ \boldsymbol{\beta}'_b \ddot{\boldsymbol{v}}_{it} + \boldsymbol{\delta}'_i \ddot{\boldsymbol{\xi}}_{it} + \ddot{\varepsilon}_{it}.$$

As $\widehat{\ddot{\boldsymbol{X}}}'_i \widehat{\ddot{\boldsymbol{u}}}_i$ contains the term $\mathring{\boldsymbol{\Lambda}}'_i \dot{\boldsymbol{F}}' \dot{\boldsymbol{F}} \mathring{\boldsymbol{\Lambda}}_i \boldsymbol{\beta}_b$, (21) does not hold. Then, the LM test diverges with probability approaching one even under the null hypothesis.

Remark 7. An extension of the model (1) to cover nonstationary factor models is feasible. For the case of $I(1)$ factors (with $I(0)$ errors, ε_{it} and $\boldsymbol{v}_{it}$), it is easily seen that Theorem 1 holds under Assumption A and the null hypothesis, (4). We explored this issue in a limited Monte Carlo study. The results, not reported for brevity, are available upon request.

Remark 8. Consider another modification, arising from weak loading correlations. Define two sets of units. The first is the set of unit indices, denoted $\mathcal{I}$, for which (17) holds while the complement is the set of unit indices, denoted $\mathcal{I}^c$, for which (17) does not hold. Denote the cardinality of $\mathcal{I}^c$ by $N_1 = O(N^\varpi)$. We refer to this setting as a weak factor case with exponent ϖ. We then have:

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T} = \frac{1}{\sqrt{N}} \sum_{i\in\mathcal{I}} \frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T} + \frac{1}{\sqrt{N}} \sum_{i\in\mathcal{I}^C} \frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T}$$

As long as $\varpi < 1/2$, it follows that

$$\frac{1}{\sqrt{N}} \sum_{i\in\mathcal{I}} \frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T} = \frac{\sqrt{N - N_1}}{\sqrt{N}} \frac{1}{\sqrt{N - N_1}} \sum_{i\in\mathcal{I}} \frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T}$$
$$\to_d N(0, \boldsymbol{R}),$$

where

$$\boldsymbol{R} = \lim_{N,T\to\infty} \frac{1}{N} \sum_{i=1}^N E\left(\frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}}}{T} \mathring{\boldsymbol{\gamma}}_i \mathring{\boldsymbol{\gamma}}'_i \frac{\dot{\boldsymbol{F}}' \ddot{\boldsymbol{X}}_i}{T} \right),$$
$$\frac{1}{\sqrt{N}} \sum_{i\in\mathcal{I}^C} \frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T} = o_p(1).$$

This implies that the presence of weak loading correlation ($\varpi < 1/2$) does not affect the LM test inference. On the other hand, if $\varpi \geq 1/2$,

$$\frac{1}{\sqrt{N}} \sum_{i\in\mathcal{I}^C} \frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T} = \frac{1}{\sqrt{N}} \sum_{i\in\mathcal{I}^C} \left[\frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T} - E\left(\frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T} \right) \right]$$
$$+ \frac{1}{\sqrt{N}} \sum_{i\in\mathcal{I}^C} E\left(\frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T} \right).$$

But,

$$\frac{1}{\sqrt{N}} \sum_{i\in\mathcal{I}^C} \left[\frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T} - E\left(\frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T} \right) \right]$$
$$= O_p(1), \quad \frac{1}{\sqrt{N}} \sum_{i\in\mathcal{I}^C} E\left(\frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T} \right) = O_p\left(N^{\varpi-1/2}\right),$$

implying $\frac{1}{\sqrt{N}} \sum_{i\in\mathcal{I}^C} \frac{\ddot{\boldsymbol{X}}'_i \dot{\boldsymbol{F}} \mathring{\boldsymbol{\gamma}}_i}{T} = O_p\left(N^{\varpi-1/2}\right)$. Corollary 1 summarizes the above discussions.

Corollary 1. Consider the panel data model with IE given by (1) under a weak factor setting with exponent ϖ. Under Assumption A, if $\varpi < 1/2$, then the LM_X test statistic in (16) follows the χ_k^2 distribution where k is the number of the regressors. If $\varpi > 1/2$, the LM_X test is consistent.

2.1. Extension to Dynamic Panels with Large T

The homogeneous panel data model with lagged dependent variables and IE has been analyzed by Moon and Weidner (2015, 2017), who propose a quasi maximum likelihood (QML) estimator. Song (2013) extends the iterative PC analysis of Bai (2009) for dynamic panels under parameter heterogeneity, but provides an asymptotic theory for the individual estimator only. Chudik and Pesaran (2015) extend the CCE approach to heterogeneous dynamic panels with IE, and propose a CCEMG estimator by augmenting the model with a sufficient number of lagged cross-sectional averages. Via Monte Carlo simulations, they demonstrate that the CCEMG estimator and the MG estimator based on Song's (2013) individual PC estimator perform better than Bai's IPC estimator and the QML estimator by Moon and Weidner (2015). The (bias-corrected) CCEMG estimator of the autoregressive parameters on the lagged dependent variable still exhibits significant bias, mainly due to the Nickel bias of order

$O(T^{-1})$. Recently, Norkute et al. (2021) develop two instrumental variable (IV) estimators for dynamic panel data models with exogenous covariates and a multifactor error structure, when both N and T are large, by projecting out the common factors from the exogenous covariates and constructing instruments based on defactored covariates. For a homogeneous model they propose a two-step IV estimator, which is not subject to either the small T Nickel bias or asymptotic biases, unlike the QML estimator by Moon and Weidner (2015). For a heterogeneous model they propose a mean group IV estimator, which is shown to outperform the CCEMG estimator by Chudik and Pesaran (2015).

Consider the heterogeneous dynamic panel data model with IE:

$$\rho_i(L)y_{it} = \boldsymbol{\beta}_i' \boldsymbol{x}_{it} + u_{it}, \; u_{it} = \boldsymbol{\gamma}_i' \boldsymbol{f}_t + \varepsilon_{it} \qquad (22)$$

where $\rho_i(L) = 1 - \sum_{j=1}^p \rho_{ij} L^j$. If the null hypothesis, (4) is not rejected, then the (modified) two-way FE estimator will still be valid to apply to the dynamic panel data with IE. Notice that only the *MG* estimator is consistent for dynamic panels if the parameters are heterogeneous (Pesaran and Smith 1995).

We first consider the case with serially uncorrelated factors, $\boldsymbol{f}_t$. To show that the dynamic two-way FE estimator of ρ_{ij} and $\boldsymbol{\beta}_i$ from (22) is unbiased under the null, we should ensure that $\lim_{T\to\infty} E(\ddot{y}_{i,t-j}\ddot{u}_{it}) = 0$ for $j = 1, \ldots, p$ and $E(\ddot{\boldsymbol{x}}_{it}\ddot{u}_{it}) = 0$. It is clear by (9) that $E(\ddot{\boldsymbol{x}}_{it}\ddot{u}_{it}) = E(\ddot{\boldsymbol{x}}_{it}\boldsymbol{\gamma}_i'\dot{\boldsymbol{f}}_t) + E(\ddot{\boldsymbol{x}}_{it}\ddot{\varepsilon}_{it}) = 0$. Consider $E(\ddot{y}_{i,t-1}\ddot{u}_{it}) = E(\ddot{y}_{i,t-1}\boldsymbol{\gamma}_i'\dot{\boldsymbol{f}}_t) + E(\ddot{y}_{i,t-1}\ddot{\varepsilon}_{it})$. Under Assumption A, $\lim_{T\to\infty} E(\ddot{y}_{i,t-1}\ddot{\varepsilon}_{it}) = 0$. Further, it is easily seen that $\lim_{T\to\infty} E(\ddot{y}_{i,t-1}\boldsymbol{\gamma}_i'\dot{\boldsymbol{f}}_t) = 0$ if factors, $\boldsymbol{f}_t$ are serially uncorrelated. Hence, under the null, it is valid to apply the FE estimator to the dynamic panel data model with IE. We then construct the LM_X test using the dynamic FE residuals given by

$$\widehat{\boldsymbol{u}}_i = \ddot{\boldsymbol{y}}_i - \sum_{j=1}^p \hat{\rho}_{ij}\ddot{\boldsymbol{y}}_{i,-j} - \ddot{\boldsymbol{X}}_i\hat{\boldsymbol{\beta}}_i \qquad (23)$$

where $\hat{\rho}_{ij}$ and $\hat{\boldsymbol{\beta}}_i$ are the two-way dynamic FE estimator obtained from (22). By Theorem 1, it is easily seen that $LM_X \to_d \chi_k^2$ under the null hypothesis, (4).

Next, to deal with the challenging case where $\boldsymbol{f}_t$ are serially correlated, we suppose that $\boldsymbol{f}_t$ follow an infinite order vector autoregressive (VAR) process: $\Phi(L)\boldsymbol{f}_t = \boldsymbol{\varepsilon}_{ft}$ where $\Phi(L) = \sum_{j=1}^\infty \Phi_j L^j$ and $\boldsymbol{\varepsilon}_{ft}$ is an $r \times 1$ vector of iid errors. Now, the dynamic FE estimator of ρ_{ij} and $\boldsymbol{\beta}_i$ from (22) is biased even under the null, (4) because $\lim_{T\to\infty} E(\ddot{y}_{i,t-j}\ddot{u}_{it}) \neq 0$ for $j \geq 1$ due to $E(\dot{\boldsymbol{f}}_t\dot{\boldsymbol{f}}_{t-j}') \neq 0$. This suggests that the LM_X test constructed using the FE residuals in (23), suffer from size distortions.

In this regard, for a valid inference, we propose the use of the following autoregressive distributed lag (ARDL) approximation of (22):

$$y_{it} = \sum_{j=1}^\infty \delta_{ij}y_{it-j} + \sum_{j=0}^\infty \boldsymbol{\psi}_{ij}' \boldsymbol{x}_{it-j} + u_{it}, \; u_{it} = \boldsymbol{\gamma}_i' \boldsymbol{\varepsilon}_{ft} + \varepsilon_{it} \quad (24)$$

This representation follows since y_{it} and $\boldsymbol{x}_{it}$ are generated by a state space model where $\boldsymbol{f}_t$ is the unobserved state. By

Theorem 1.2.1 of Hannan and Deistler (1989), the ARDL representation formalizes the usual projection argument: $y_{it} - E(y_{it}|y_{it-1}, \ldots, \boldsymbol{x}_{it}, \boldsymbol{x}_{it-1}, \ldots) = \boldsymbol{\gamma}_i'\boldsymbol{\varepsilon}_{ft} + \varepsilon_{it}$. Then, it is easily seen that $E(\ddot{\boldsymbol{x}}_{i,t-j}u_{it}) = E(\ddot{\boldsymbol{x}}_{i,t-j}(\boldsymbol{\gamma}_i'\dot{\boldsymbol{\varepsilon}}_{ft} + \ddot{\varepsilon}_{it})) = 0$ for $j \geq 0$ and $\lim_{T\to\infty} E(\ddot{y}_{i,t-j}\ddot{u}_{it}) = \lim_{T\to\infty} E(\ddot{y}_{i,t-j}(\boldsymbol{\gamma}_i'\dot{\boldsymbol{\varepsilon}}_{ft} + \ddot{\varepsilon}_{it})) = 0$ for $j \geq 1$. Next, we propose the finite order ARDL approximation of (24):

$$y_{it} = \sum_{j=1}^{p_T} \delta_{ij}y_{it-j} + \sum_{j=0}^{p_T} \boldsymbol{\psi}_{ij}' \boldsymbol{x}_{it-j} + u_{it} + o_p(1) \qquad (25)$$

where we suggest using $p_T = O(\ln T)$ or $p_T = O(T^{1/3})$ (e.g., Lewis and Reinsel 1988; Hannan and Deistler 1989; Chudik and Pesaran 2015). Under the null hypothesis, (4), the two-way dynamic ARDL estimator of δ_{ij} and $\boldsymbol{\psi}_{ij}$ from (25), denoted $\hat{\delta}_{ij}$ and $\hat{\boldsymbol{\psi}}_{ij}$, will be unbiased. Hence, when constructing the proper LM_X test, we propose the use of the following ARDL residuals:

$$\widehat{\boldsymbol{u}}_i = \ddot{\boldsymbol{y}}_i - \sum_{j=1}^{p_T} \hat{\delta}_{ij}\ddot{\boldsymbol{y}}_{i,-j} - \sum_{j=0}^{p_T} \ddot{\boldsymbol{X}}_{i,-j}\hat{\boldsymbol{\psi}}_{ij} \qquad (26)$$

Assumption B. The lag polynomial, $\rho_i(L) = 1 - \sum_{j=1}^{p_i} \rho_{ij} L^j$ in (22) has the roots outside the unit circle such that $|\rho_{ij}| \leq c^j$ for some $0 < c < 1$. The VAR lag polynomial, $\Phi(L) = \sum_{j=1}^\infty \Phi_j L^j$ satisfy $\|\Phi_j\|_s \leq c^j$ where $\|.\|_s$ denotes spectral norm, and $0 < c < 1$. The idiosyncratic errors, ε_{it} and $\boldsymbol{\varepsilon}_{ft}$ are serially uncorrelated and mutually independent.

To show that the LM_X test constructed, using the ARDL residuals in (26), follows the χ^2 distribution asymptotically under the null hypothesis, (4), we need to strengthen Assumption A, replaced by Assumption A' in the Online Supplement. We also consider a setting with $T \to \infty$ and impose the DGP condition in (2). These refinements are needed to accommodate the case with serially correlated factors.

Theorem 2. Consider the dynamic heterogeneous panel data model with IE given by (22) and (2), the ARDL representation, (24) and the ARDL approximation, (25). Then, under Assumption A' and B and under the null hypothesis, (4), as $N, T \to \infty$, the LM_X test statistic in (16), constructed using the ARDL residuals in (26), follows a χ_k^2 distribution where k is the dimension of $\boldsymbol{X}_i$. Under the alternative hypothesis, (5), the LM_X test is consistent.

Remark 9. Theorems 1 and 2 do not require any particular relative rates for N and T for the validity of the LM test. For large N and T, we only need to show that the dominant term of the LM test statistic tends to a normal variate while the remaining terms tend to zero. This is shown to be theoretically less demanding than deriving convergence rates for the pooled or MG estimators that would involve handling asymptotic biases.

Remark 10. In practice we suggest to apply the LM_X test with the ARDL residuals to the dynamic panel data with IE, irrespective of whether factors are serially correlated or not. See Section S4.5 in the Online Supplement for the comprehensive simulation results. In the presence of serially correlated factors, the dynamics given by $\rho_i(L)$ in (22), do not capture the full dynamics of

the panel data model due to neglecting the dynamics embedded in f_t. In this regard, if the LM_X test does not reject the null hypothesis, then we suggest the ARDL model, (25) as a proper dynamic specification.

Remark 11. It is not straightforward to find the one-to-one relationship between the parameters in (22) and (25). If we are still interested in consistently estimating the parameters of the model, (22) under the null hypothesis, then we may consider applying the IV estimation, similarly to Norkute et al. (2021). For convenience we focus on the simple dynamic panel model with IE:

$$y_{it} = \rho_i y_{i,t-1} + \beta_i' x_{it} + u_{it}, \ u_{it} = \gamma_i' f_t + \varepsilon_{it}$$

or in the matrix notation,

$$y_i = \phi_i y_{i,-1} + X_i \beta_i + F\lambda_i + e_i = W_i \theta_i + u_i$$

where $W_i = (y_{i,-1}, X_i,)$ and $\theta_i = (\phi_i, \beta_i')'$. We consider the (internal) IV matrix given by $Z_i = (X_{i,-1}, X_i)$, which satisfies the orthogonality condition, $E[\ddot{Z}_i' \ddot{u}_i] = 0$ under the null. Given $E[\ddot{Z}_i' \ddot{W}_i] \neq 0$, θ_i can be consistently estimated by $\hat{\theta}_{IV,i} = \left[\left(\ddot{W}_i' \ddot{Z}_i \right) \left(\ddot{Z}_i' \ddot{Z}_i \right)^{-1} \left(\ddot{Z}_i' \ddot{W}_i \right) \right]^{-1} \left[\left(\ddot{W}_i' \ddot{Z}_i \right) \left(\ddot{Z}_i' \ddot{Z}_i \right)^{-1} \left(\ddot{Z}_i' \ddot{y}_i \right) \right]$, which is referred to as the IVFE estimator. We find via simulations that $\hat{\rho}_{IV,i}$ and $\hat{\beta}_{IV,i}$ and their MG counterparts show negligible biases though the bias of ρ_i is higher than that of β_i in small samples. But, the LM_X test constructed using the IVFE residuals given by (S.50), tends to over-reject the null for large N, especially if the serial correlation of the factors is pronounced. See Section S4.6 in the Online Supplement.

3. Monte Carlo Simulations

W examine the size and power performance of the LM_X statistic in (16) under both static and dynamic panel data frameworks.

3.1. Monte Carlo Design

As a benchmark, we generate the data with two regressors ($k = 2$) and two factors ($r = 2$):

$$y_{it} = \beta_{i1} x_{it1} + \beta_{i2} x_{it2} + \gamma_{i1} f_{t1} + \gamma_{i2} f_{t2} + \varepsilon_{it}, \quad (27)$$

$$x_{it1} = \Lambda_{i11} f_{t1} + \Lambda_{i12} f_{t2} + v_{it1} \text{ and } x_{it2} = \Lambda_{i21} f_{t1} + \Lambda_{i22} f_{t2} + v_{it2}. \quad (28)$$

We allow serial correlation and weak cross-sectional correlation in ε_{it} by generating $\varepsilon_{it} = \rho_\varepsilon \varepsilon_{i,t-1} + v_{it} + \theta \sum_{|h| \leq 8} v_{i-h,t}$ with $v_{it} \sim$ iidN(0, 1), $\rho_\varepsilon = 0.5$ and $\theta = 0.2$. We generate $(v_{it1}, v_{it2})' \sim$ iidN(0, I_2) and $(f_{t1}, f_{t2})' \sim$ iidN(0.5, I_2) (the simulation results for serially correlated factors are qualitatively similar and available upon request). To test the validity of the LM test, we generate the factor loadings, $(\gamma_{i1}, \gamma_{i2})$, $(\Lambda_{i11}, \Lambda_{i12})$ and $(\Lambda_{i21}, \Lambda_{i22})$ under the following two experiments:

- Experiment 1 (independent factor loadings): $\gamma_{i1} \sim$ iidU(0, 1), $\gamma_{i2} \sim$ iidU(0, 1), $\Lambda_{i11} \sim$ iidU(0, 1), $\Lambda_{i12} \sim$ iidU(0, 1), $\Lambda_{i21} \sim$ iidU(0, 1), and $\Lambda_{i22} \sim$ iidU(0, 2) such that

$$E \begin{pmatrix} \gamma_{i1} & \gamma_{i2} \\ \Lambda_{i11} & \Lambda_{i12} \\ \Lambda_{i21} & \Lambda_{i22} \end{pmatrix} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

- Experiment 2 (correlated factor loadings): $\gamma_{i1} \sim$ iidU(0, 1), $\gamma_{i2} \sim$ iidU(0, 2), $\Lambda_{i11} = \gamma_{i1}$, $\Lambda_{i12} \sim$ iidU(0, 1), $\Lambda_{i21} \sim$ iidU(0, 1), and $\Lambda_{i22} = \gamma_{i2}$.

We explore the performance of the LM_X test under the parameter heterogeneity by generating $\beta_{ik} = 1 + \eta_{ik}$ for $k = \{1, 2\}$. We have also constructed DGPs with the different number of regressors and factors; namely for $k = \{1, 2, 3\}$ and $r = \{0, 1, 2, 3, 4\}$. See Section S3 in the Online Supplement for a full description. We examine the following three cases for $r = \{0, 1, 2, 3, 4\}$: Case 1: Weak heterogeneity with $\eta_{ik} \sim$ iidN(0, 0.04); Case 2: Medium heterogeneity with $\eta_{ik} \sim$ iidN(0, 0.25); Case 3: Strong heterogeneity with $\eta_{ik} \sim$ iidN(0, 1). We consider the following combinations of $(N, T) = \{30, 50, 100, 200\}$. We set the number of replications at $R = 1000$.

3.2. The Performance of the LM_X Statistic

We evaluate the size of the LM_X test under Experiment 1 and the power under Experiment 2. If $r = 0$, we only report the size of the test. Table 1 reports the size performance under Experiment 1 with $k = \{1, 2, 3\}$, respectively. The size of the LM_X test is close to the nominal level (5%) for all sample sizes and its performance is shown to be invariant to the number of regressors, the number of factors and the different strengths of parameter heterogeneity.

In Table 2 we present the power performance. Overall, the LM_X test is reasonably powerful even for $T = 30$, but it becomes consistent as the sample size increases. Even as the parameter heterogeneity gets stronger and/or the number of regressors and factors increases, the power of the LM_X test remains satisfactory in most cases.

We find that the finite sample performance of the LM_X test is satisfactory in all cases considered, even under strong heterogeneity. Overall simulation results demonstrate that the LM_X test is robust to serial correlation and weak cross-sectional correlation in errors as well as the slope heterogeneity.

Section S.4.1 in the Online Supplement explores the performance of the LM test when both factors and loadings are generated by a zero mean process. The performances of the LM test reported in Tables S.5 and S.6 are qualitatively similar to those in Tables 1 and 2. Furthermore, in Section S4.2, we report additional simulation results for the following cases: Case A with homogeneous parameters, $\beta = 1$ for all i and iid errors; Case B with homogeneous parameters and serially and weakly cross-sectionally correlated errors; Case C with heterogeneous parameters, $\beta_{ik} = 1 + \eta_{ik}$ and iid errors; Case D with heterogeneous parameters, and serially and weakly cross-sectionally correlated errors. The simulation results for Cases A–C are qualitatively similar to those reported here for Case D.

Further, we evaluate the size and power performance the LM_X test in panels with small T ($T = 3, 5$) in conjunction with ($N = 30, 50, 100, 200$), using DGPs with one regressor and two factors (see Section S.3.1.2 for details) and set $\beta_1 = 1$. We consider two cases:

- Case A with iid error, $\varepsilon_{it} \sim$ iidN(0, 1).

Table 1. The size of the LM_X test under Experiment 1.

	Experiment 1											
	Case 1				Case 2				Case 3			
T/N	30	50	100	200	30	50	100	200	30	50	100	200
						$k=1$						
						$r=0$						
30	0.034	0.048	0.047	0.049	0.050	0.053	0.054	0.061	0.036	0.039	0.048	0.040
50	0.033	0.039	0.053	0.053	0.038	0.036	0.045	0.064	0.039	0.050	0.035	0.054
100	0.047	0.050	0.051	0.048	0.052	0.044	0.052	0.045	0.028	0.037	0.051	0.051
200	0.038	0.042	0.048	0.039	0.043	0.034	0.051	0.043	0.046	0.027	0.049	0.040
						$r=1$						
30	0.037	0.050	0.057	0.048	0.045	0.047	0.075	0.044	0.044	0.044	0.053	0.044
50	0.059	0.045	0.054	0.046	0.056	0.048	0.042	0.044	0.061	0.046	0.046	0.040
100	0.043	0.059	0.038	0.049	0.043	0.047	0.055	0.059	0.053	0.050	0.043	0.058
200	0.059	0.054	0.041	0.048	0.049	0.040	0.054	0.053	0.045	0.053	0.054	0.050
						$r=2$						
30	0.046	0.054	0.047	0.063	0.053	0.049	0.053	0.042	0.057	0.034	0.061	0.035
50	0.052	0.050	0.047	0.043	0.051	0.046	0.066	0.054	0.040	0.054	0.045	0.062
100	0.050	0.037	0.059	0.049	0.047	0.060	0.054	0.050	0.045	0.041	0.048	0.045
200	0.052	0.055	0.045	0.051	0.047	0.063	0.057	0.052	0.051	0.045	0.053	0.044
						$r=3$						
30	0.050	0.040	0.046	0.055	0.049	0.051	0.046	0.036	0.044	0.055	0.049	0.051
50	0.041	0.045	0.047	0.036	0.044	0.041	0.059	0.048	0.048	0.038	0.042	0.048
100	0.052	0.049	0.056	0.048	0.043	0.041	0.054	0.053	0.054	0.044	0.043	0.049
200	0.042	0.044	0.060	0.051	0.045	0.057	0.051	0.052	0.040	0.051	0.042	0.046
						$r=4$						
30	0.053	0.055	0.040	0.064	0.053	0.045	0.041	0.058	0.056	0.053	0.056	0.056
50	0.050	0.046	0.066	0.041	0.045	0.039	0.047	0.052	0.056	0.046	0.043	0.053
100	0.044	0.061	0.043	0.042	0.058	0.050	0.050	0.055	0.052	0.051	0.050	0.045
200	0.046	0.048	0.047	0.043	0.044	0.041	0.058	0.053	0.043	0.051	0.036	0.048
						$k=2$						
						$r=0$						
30	0.035	0.052	0.050	0.045	0.039	0.053	0.037	0.046	0.041	0.045	0.046	0.049
50	0.044	0.053	0.040	0.058	0.029	0.042	0.056	0.052	0.052	0.033	0.050	0.061
100	0.038	0.046	0.050	0.048	0.043	0.048	0.040	0.054	0.029	0.048	0.050	0.045
200	0.041	0.044	0.051	0.031	0.031	0.040	0.054	0.041	0.042	0.036	0.044	0.045
						$r=1$						
30	0.035	0.051	0.053	0.057	0.040	0.051	0.042	0.046	0.053	0.044	0.048	0.044
50	0.040	0.039	0.053	0.046	0.028	0.045	0.032	0.045	0.044	0.039	0.052	0.058
100	0.045	0.045	0.040	0.048	0.041	0.047	0.052	0.048	0.045	0.039	0.050	0.047
200	0.041	0.054	0.050	0.051	0.049	0.043	0.046	0.055	0.037	0.037	0.041	0.055
						$r=2$						
30	0.046	0.037	0.049	0.047	0.039	0.044	0.043	0.049	0.048	0.059	0.057	0.045
50	0.042	0.041	0.045	0.044	0.049	0.042	0.048	0.055	0.045	0.046	0.059	0.033
100	0.032	0.048	0.048	0.058	0.043	0.044	0.045	0.050	0.039	0.045	0.045	0.051
200	0.038	0.058	0.046	0.055	0.035	0.039	0.049	0.044	0.035	0.041	0.045	0.056
						$,r=3$						
30	0.050	0.051	0.043	0.050	0.033	0.051	0.059	0.045	0.030	0.037	0.044	0.050
50	0.038	0.055	0.058	0.040	0.045	0.041	0.049	0.059	0.035	0.041	0.050	0.056
100	0.042	0.034	0.045	0.044	0.035	0.037	0.037	0.050	0.028	0.040	0.047	0.047
200	0.041	0.048	0.047	0.053	0.039	0.040	0.054	0.051	0.039	0.040	0.050	0.043
						$r=4$						
30	0.052	0.028	0.079	0.051	0.046	0.039	0.046	0.062	0.049	0.048	0.041	0.042
50	0.039	0.058	0.049	0.043	0.033	0.050	0.045	0.046	0.036	0.038	0.038	0.049
100	0.038	0.043	0.050	0.053	0.043	0.042	0.038	0.036	0.034	0.036	0.042	0.054
200	0.046	0.040	0.046	0.052	0.036	0.053	0.052	0.052	0.040	0.043	0.052	0.049
						$k=3$						
						$r=0$						
30	0.036	0.043	0.037	0.043	0.034	0.042	0.046	0.056	0.038	0.047	0.027	0.044
50	0.042	0.047	0.052	0.047	0.036	0.044	0.050	0.043	0.032	0.034	0.048	0.034
100	0.030	0.042	0.049	0.058	0.032	0.036	0.041	0.045	0.032	0.044	0.043	0.048
200	0.030	0.044	0.060	0.044	0.030	0.049	0.038	0.037	0.032	0.047	0.059	0.057
						$r=1$						
30	0.032	0.048	0.045	0.035	0.033	0.038	0.035	0.043	0.031	0.038	0.046	0.047
50	0.029	0.038	0.043	0.047	0.035	0.051	0.046	0.046	0.024	0.034	0.041	0.051
100	0.032	0.040	0.050	0.049	0.032	0.047	0.051	0.048	0.026	0.033	0.035	0.037
200	0.050	0.052	0.057	0.048	0.032	0.044	0.047	0.055	0.033	0.042	0.048	0.054

(*Continued*)

Table 1. (*Continued*)

	Experiment 1											
	Case 1				Case 2				Case 3			
T/N	30	50	100	200	30	50	100	200	30	50	100	200
							$r = 2$					
30	0.041	0.039	0.054	0.047	0.026	0.051	0.037	0.049	0.030	0.036	0.045	0.044
50	0.034	0.034	0.043	0.047	0.034	0.034	0.047	0.034	0.039	0.044	0.027	0.043
100	0.033	0.047	0.041	0.052	0.027	0.044	0.045	0.054	0.028	0.038	0.035	0.054
200	0.039	0.047	0.047	0.043	0.035	0.047	0.046	0.046	0.028	0.045	0.045	0.048
							$r = 3$					
30	0.036	0.050	0.046	0.041	0.033	0.039	0.044	0.058	0.033	0.032	0.039	0.045
50	0.037	0.036	0.044	0.041	0.026	0.046	0.049	0.031	0.030	0.040	0.045	0.043
100	0.034	0.048	0.057	0.045	0.047	0.041	0.036	0.051	0.039	0.030	0.039	0.037
200	0.035	0.032	0.060	0.047	0.030	0.044	0.050	0.053	0.035	0.037	0.040	0.046
							$r = 4$					
30	0.023	0.056	0.045	0.053	0.036	0.037	0.041	0.047	0.019	0.044	0.040	0.036
50	0.036	0.046	0.036	0.052	0.029	0.040	0.037	0.044	0.029	0.041	0.041	0.051
100	0.028	0.033	0.056	0.047	0.035	0.039	0.045	0.063	0.019	0.046	0.035	0.038
200	0.039	0.035	0.040	0.046	0.026	0.041	0.054	0.055	0.039	0.035	0.045	0.044

NOTE: LM_X denotes the LM test in (16), where we construct the residuals $\ddot{u}_{it}$ in (15) using the individual FE estimator, $\hat{\beta}_{FE,i}$. We assume parameter heterogeneity, $\beta_{ik} = 1 + \eta_{ik}$ and examine three cases: Case 1 (weak heterogeneity) with $\eta_{ik} \sim iidN(0, 0.04)$; Case 2 (medium heterogeneity) with $\eta_{ik} \sim iidN(0, 0.25)$; Case 3 (strong heterogeneity) with $\eta_{ik} \sim iidN(0, 1)$. k is the number of regressors and r the number of factors.

Table 2. The power of the LM_X test under Experiment 2.

	Experiment 2											
	Case 1				Case 2				Case 3			
T/N	30	50	100	200	30	50	100	200	30	50	100	200
							$k = 1$					
							$r = 1$					
30	0.998	1	1	1	0.973	0.998	1	1	0.699	0.894	0.999	1
50	1	1	1	1	0.993	1	1	1	0.730	0.916	0.999	1
100	1	1	1	1	0.998	1	1	1	0.774	0.941	1	1
200	1	1	1	1	1	1	1	1	1	0.956	1	1
							$r = 2$					
30	0.996	1	1	1	0.940	0.996	1	1	0.638	0.863	0.990	1
50	1	1	1	1	0.977	0.998	1	1	0.686	0.882	0.991	1
100	1	1	1	1	0.986	1	1	1	0.701	0.886	0.999	1
200	1	1	1	1	0.995	1	1	1	0.709	0.922	0.999	1
							$r = 3$					
30	0.992	0.999	1	1	0.929	0.991	1	1	0.588	0.807	0.968	0.998
50	1	1	1	1	0.970	0.997	1	1	0.616	0.843	0.983	1
100		1	1	1	0.977	1	1	1	0.664	0.856	0.993	1
200	1	1	1	1	0.991	1	1	1	0.660	0.899	0.993	1
							$r = 4$					
30	0.992	1	1	1	0.880	0.984	1	1	0.539	0.746	0.948	0.997
50	0.999	1	1	1	0.935	0.997	1	1	0.608	0.786	0.967	1
100	1	1	1	1	0.962	0.998	1	1	0.622	0.828	0.982	1
200	1	1	1	1	0.969	1	1	1	0.610	0.842	0.988	1
							$k = 2$					
							$r = 1$					
30	0.993	1	1	1	0.802	0.969	1	1	0.308	0.569	0.873	0.996
50	0.998	1	1	1	0.831	0.978	0.999	1	0.349	0.588	0.911	0.998
100	1	1	1	1	0.885	0.991	1	1	0.364	0.620	0.926	0.998
200	1	1	1	1	0.880	0.992	1	1	0.382	0.609	0.916	1
							$r = 2$					
30	0.973	0.998	1	1	0.711	0.956	0.998	0.999	0.307	0.539	0.861	0.992
50	0.995	0.999	1	1	0.810	0.965	1	1	0.327	0.570	0.898	0.999
100	1	1	1	1	0.860	0.988	1	1	0.357	0.599	0.917	0.997
200	1	1	1	1	0.882	0.988	1	1	0.336	0.592	0.923	0.998
							$r = 3$					
30	0.956	0.993	1	0.999	0.695	0.927	0.996	0.999	0.306	0.511	0.833	0.978
50	0.995	1	1	1	0.804	0.975	1	1	0.330	0.593	0.887	0.992
100	0.999	1	1	1	0.854	0.981	1	1	0.341	0.600	0.899	0.997
200	1	1	1	1	0.896	0.995	1	1	0.352	0.609	0.927	0.999

Table 2. (Continued)

	Experiment 2											
	Case 1				Case 2				Case 3			
T/N	30	50	100	200	30	50	100	200	30	50	100	200
						$r = 4$						
30	0.935	0.991	0.999	1	0.665	0.901	0.988	0.999	0.278	0.474	0.786	0.962
50	0.991	0.999	1	1	0.744	0.954	0.999	1	0.289	0.523	0.839	0.982
100	0.996	1	1	1	0.810	0.970	1	1	0.293	0.537	0.853	0.988
200	0.999	1	1	1	0.822	0.979	1	1	0.320	0.564	0.871	0.997
						$k = 3$						
						$r = 1$						
30	0.946	0.999	1	1	0.518	0.831	0.993	1	0.167	0.314	0.646	0.935
50	0.987	1	1	1	0.561	0.853	0.997	1	0.194	0.334	0.686	0.945
100	0.993	1	1	1	0.621	0.897	0.997	1	0.177	0.356	0.691	0.954
200	1	1	1	1	0.642	0.899	0.999	1	0.179	0.347	0.693	0.956
						$r = 2$						
30	0.900	0.988	1	1	0.459	0.779	0.988	1	0.144	0.280	0.601	0.918
50	0.967	1	1	1	0.528	0.876	0.997	1	0.182	0.320	0.652	0.946
100	0.995	1	1	1	0.588	0.900	0.996	1	0.159	0.347	0.699	0.958
200	0.992	1	1	1	0.590	0.911	0.999	1	0.159	0.331	0.700	0.967
						$r = 3$						
30	0.863	0.992	0.999	1	0.460	0.754	0.977	0.999	0.143	0.318	0.620	0.918
50	0.942	1	1	1	0.497	0.847	0.988	1	0.157	0.331	0.658	0.930
100	0.989	1	1	1	0.588	0.886	0.998	1	0.180	0.349	0.717	0.955
200	0.997	1	1	1	0.596	0.901	1	1	0.167	0.359	0.712	0.967
						$r = 4$						
30	0.828	0.978	1	1	0.404	0.713	0.963	0.998	0.138	0.287	0.586	0.865
50	0.940	0.996	1	1	0.465	0.805	0.987	0.999	0.147	0.298	0.612	0.897
100	0.975	1	1	1	0.502	0.809	0.992	1	0.155	0.314	0.630	0.927
200	0.995	1	1	1	0.553	0.855	0.998	1	0.151	0.306	0.665	0.945

NOTE: see notes to Table 1.

- Case B with serially and weakly cross-sectionally correlated error generated by $\varepsilon_{it} = \rho_\varepsilon \varepsilon_{i,t-1} + \upsilon_{it} + \theta \sum_{1 \leq h \leq 8} \upsilon_{i-h,t}$ with $\upsilon_{it} \sim \text{iidN}(0,1)$, $\rho_\varepsilon = 0.5$ and $\theta = 0.2$.

Table 3 reports the size and power performance of the LM_X test for $T = 3, 5$. The size of the LM_X test is close to the nominal level (5%) for all cases considered while its performance is shown to be invariant to the number of factors. The size performance of the LM_X test is also robust to serial and weak cross-sectional error correlation. Given the fixed T (as small as $T = 3$), the power of the LM_X test rises monotonically with N. Overall, the power of the LM_X test remains satisfactory.

3.3. Robustness to Misspecification

To investigate the size performance of the LM_X test in the presence of model misspecification, we generate the DGP using three regressors ($k = 3$) and two factors ($r = 2$), but apply the LM_X test as if $k = 1$:

$$y_{it} = \beta_{i1}x_{it1} + \beta_{i2}x_{it2} + \beta_{i3}x_{it3} + \gamma_{i1}f_{1t} + \gamma_{i2}f_{2t} + \varepsilon_{it}, \quad (29)$$

$$x_{it1} = \Lambda_{i11}f_{1t} + \Lambda_{i12}f_{2t} + \upsilon_{it1}, x_{it2} = \Lambda_{i21}f_{1t} + \Lambda_{i22}f_{2t} + \upsilon_{it2},$$
$$x_{it3} = \Lambda_{i31}f_{1t} + \Lambda_{i32}f_{2t} + \upsilon_{it3}.$$

We allow serial correlation and weak cross-sectional correlation in ε_{it} by generating $\varepsilon_{it} = \rho_\varepsilon \varepsilon_{i,t-1} + \upsilon_{it} + \theta \sum_{1 \leq h \leq 8} \upsilon_{i-h,t}$ with $\upsilon_{it} \sim$ iidN$(0,1)$, $\rho_\varepsilon = 0.5$ and $\theta = 0.2$. We generate $(\upsilon_{it1}, \upsilon_{it2}, \upsilon_{it3})' \sim$ iidN$(0, \boldsymbol{I}_3)$ and $(f_{t1}, f_{t2})' \sim$ iidN$(0.5, \boldsymbol{I}_2)$. We generate the factor loadings only under Experiment 1 (independent factor loadings)

as follows: $\gamma_{i1} \sim \text{iidU}(0,1)$, $\gamma_{i2} \sim \text{iidU}(0,1)$; $\Lambda_{i11} \sim \text{iidU}(0,1)$, $\Lambda_{i12} \sim \text{iidU}(0,1)$; $\Lambda_{i21} \sim \text{iidU}(0,1)$, $\Lambda_{i22} \sim \text{iidU}(0,1)$, $\Lambda_{i31} \sim \text{iidU}(0,1)$, $\Lambda_{i32} \sim \text{iidU}(0,2)$ such that

$$E \begin{pmatrix} \gamma_{i1} & \gamma_{i2} \\ \Lambda_{i11} & \Lambda_{i12} \\ \Lambda_{i21} & \Lambda_{i22} \\ \Lambda_{i31} & \Lambda_{i32} \end{pmatrix} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \\ 0.5 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

We explore the size performance of the LM_X test by generating $\beta_{ik} = 1 + \eta_{ik}$ for $k = 1, 2, 3$, and examine the following three cases for $r = 0, 1, 2, 3, 4$: Case 1: weak heterogeneity with $\eta_{ik} \sim \text{iidN}(0, 0.04)$; Case 2: medium heterogeneity with $\eta_{ik} \sim \text{iidN}(0, 0.25)$; Case 3: strong heterogeneity with $\eta_{ik} \sim \text{iidN}(0, 1)$. We consider the combinations of $(N, T) = \{30, 50, 100, 200\}$, and set the number of replications at $R = 1000$.

The results in Table 4 display that the size of the LM_X test is mostly close to the nominal level (5%) in all cases considered, confirming that it is robust to model misspecication. For comparison, we report the size performance of the Hausman test proposed by Kapetanios, Serlenga, and Shin (2023), which is severely oversized in the presence of neglected regressors, leading to an incorrect rejection of the null hypothesis in all cases.

3.4. Extension to Heterogeneous Dynamic Panel Data Models with Large T

We generate the heterogeneous dynamic panel data with a single regressor ($k = 1$) and two factors ($r = 2$):

$$y_{it} = \rho_i y_{i,t-1} + \beta_i x_{it} + \gamma_{i1} f_{1t} + \gamma_{i2} f_{2t} + \varepsilon_{it}, \ x_{it}$$
$$= \Lambda_{i11} f_{1t} + \Lambda_{i12} f_{2t} + \upsilon_{it} \quad (30)$$

Table 3. The size and power of the LM_X test with small T and $k = 1$.

	Experiment 1							
	Case A				Case B			
T/N	30	50	100	200	30	50	100	200
				$r = 1$				
3	0.047	0.041	0.071	0.046	0.049	0.042	0.047	0.050
5	0.045	0.043	0.048	0.058	0.048	0.050	0.054	0.045
				$r = 2$				
3	0.054	0.055	0.054	0.066	0.049	0.042	0.047	0.050
5	0.042	0.045	0.044	0.051	0.048	0.050	0.054	0.045
				$r = 3$				
3	0.044	0.052	0.056	0.044	0.042	0.039	0.046	0.047
5	0.039	0.056	0.054	0.045	0.042	0.038	0.046	0.055
				$r = 4$				
3	0.033	0.047	0.048	0.061	0.048	0.057	0.041	0.057
5	0.053	0.046	0.053	0.051	0.044	0.064	0.050	0.041
	Experiment 2							
				$r = 1$				
3	0.141	0.262	0.413	0.564	0.153	0.248	0.381	0.540
5	0.456	0.640	0.829	0.912	0.408	0.576	0.776	0.876
				$r = 2$				
3	0.138	0.261	0.378	0.567	0.138	0.247	0.358	0.545
5	0.404	0.615	0.809	0.920	0.372	0.526	0.743	0.887
				$r = 3$				
3	0.146	0.242	0.394	0.530	0.140	0.213	0.358	0.546
5	0.376	0.601	0.781	0.919	0.358	0.544	0.754	0.896
				$r = 4$				
3	0.148	0.210	0.384	0.552	0.140	0.202	0.322	0.541
5	0.329	0.549	0.775	0.912	0.304	0.463	0.724	0.881

NOTE: LM_X denotes the LM test defined in (16), where we construct the residuals $\ddot{u}_{it}$ in (15) using the individual FE estimator, $\hat{\beta}_{FE,i}$. The performance of the LM_X test is explored under parameter homogeneity, $\beta = 1$ and $k = 1$. We examine two cases. Case A with iid error, $\varepsilon_{it} \sim iidN(0,1)$ and Case B with serially and weakly cross-sectionally correlated error generated by $\varepsilon_{it} = \rho_\varepsilon \varepsilon_{i,t-1} + \upsilon_{it} + \theta \sum_{1 \leq h \leq 8} \upsilon_{i-h,t}$ with $\upsilon_{it} \sim iidN(0,1)$, $\rho_\varepsilon = 0.5$ and $\theta = 0.2$.

Table 4. The size of the LM_X and H_{KSS} tests in the presence of the model misspecification.

	Case 1				Case 2				Case 3			
T/N	30	50	100	200	30	50	100	200	30	50	100	200
						$r = 0$						
						LM_X						
30	0.044	0.041	0.050	0.047	0.046	0.053	0.059	0.059	0.042	0.036	0.048	0.046
50	0.046	0.044	0.067	0.048	0.046	0.048	0.056	0.055	0.045	0.046	0.057	0.047
100	0.041	0.039	0.054	0.052	0.050	0.045	0.040	0.040	0.035	0.037	0.041	0.048
200	0.044	0.041	0.053	0.037	0.036	0.048	0.052	0.053	0.035	0.046	0.038	0.038
						H_{KSS}						
30	0.911	0.917	0.945	0.946	0.944	0.948	0.964	0.978	0.953	0.970	0.977	0.987
50	0.934	0.939	0.963	0.975	0.971	0.975	0.989	0.995	0.967	0.980	0.984	0.996
100	0.967	0.968	0.974	0.991	0.980	0.983	0.992	0.995	0.965	0.975	0.997	0.996
200	0.984	0.986	0.992	0.995	0.995	0.993	0.996	0.999	0.972	0.979	0.993	0.998
						$r = 1$						
						LM_X						
30	0.046	0.042	0.048	0.047	0.055	0.048	0.046	0.048	0.048	0.047	0.053	0.049
50	0.061	0.047	0.047	0.040	0.037	0.052	0.043	0.046	0.051	0.042	0.056	0.046
100	0.045	0.048	0.051	0.055	0.050	0.042	0.046	0.042	0.048	0.061	0.037	0.044
200	0.065	0.054	0.060	0.052	0.055	0.054	0.048	0.055	0.040	0.044	0.061	0.058
						H_{KSS}						
30	0.841	0.814	0.809	0.820	0.886	0.872	0.898	0.870	0.909	0.911	0.906	0.910
50	0.846	0.839	0.810	0.823	0.908	0.919	0.897	0.875	0.925	0.933	0.913	0.912
100	0.864	0.861	0.830	0.814	0.935	0.918	0.897	0.906	0.934	0.936	0.941	0.932
200	0.893	0.850	0.824	0.791	0.942	0.945	0.916	0.913	0.948	0.943	0.936	0.940
						$r = 2$						
						LM_X						
30	0.046	0.047	0.050	0.041	0.050	0.038	0.046	0.061	0.050	0.054	0.055	0.045
50	0.054	0.053	0.047	0.054	0.056	0.057	0.059	0.061	0.039	0.043	0.058	0.052
100	0.050	0.048	0.046	0.055	0.044	0.038	0.040	0.054	0.044	0.035	0.048	0.047
200	0.063	0.052	0.033	0.039	0.053	0.045	0.044	0.042	0.035	0.043	0.046	0.054

(Continued)

Table 4. (*Continued*)

	Case 1				Case 2				Case 3			
T/N	30	50	100	200	30	50	100	200	30	50	100	200
							H_{KSS}					
30	0.823	0.824	0.794	0.781	0.839	0.838	0.820	0.817	0.833	0.862	0.818	0.844
50	0.867	0.835	0.778	0.767	0.854	0.860	0.830	0.824	0.871	0.858	0.846	0.876
100	0.890	0.857	0.775	0.713	0.894	0.871	0.867	0.828	0.868	0.868	0.885	0.860
200	0.890	0.849	0.769	0.690	0.891	0.877	0.867	0.844	0.884	0.886	0.875	0.884
							$r = 3$					
							LM_X					
30	0.033	0.036	0.051	0.056	0.053	0.066	0.060	0.051	0.036	0.051	0.051	0.051
50	0.048	0.049	0.054	0.058	0.049	0.041	0.048	0.051	0.047	0.042	0.047	0.057
100	0.050	0.052	0.035	0.043	0.052	0.045	0.041	0.042	0.040	0.044	0.049	0.050
200	0.044	0.041	0.046	0.048	0.052	0.054	0.054	0.060	0.053	0.050	0.038	0.041
							H_{KSS}					
30	0.866	0.863	0.835	0.798	0.864	0.828	0.794	0.818	0.799	0.815	0.801	0.798
50	0.898	0.899	0.858	0.756	0.866	0.868	0.813	0.788	0.810	0.837	0.826	0.828
100	0.914	0.902	0.826	0.728	0.873	0.864	0.849	0.805	0.851	0.849	0.849	0.842
200	0.953	0.924	0.824	0.762	0.912	0.878	0.832	0.804	0.866	0.852	0.844	0.852
							$r = 4$					
							LM_X					
30	0.039	0.045	0.050	0.044	0.050	0.051	0.050	0.050	0.045	0.054	0.040	0.064
50	0.041	0.058	0.063	0.053	0.049	0.050	0.051	0.041	0.034	0.041	0.053	0.053
100	0.058	0.050	0.049	0.060	0.046	0.053	0.056	0.048	0.047	0.054	0.052	0.056
200	0.059	0.044	0.047	0.058	0.056	0.044	0.054	0.053	0.041	0.050	0.046	0.042
							H_{KSS}					
30	0.915	0.887	0.839	0.800	0.796	0.795	0.764	0.677	0.716	0.714	0.705	0.637
50	0.951	0.926	0.866	0.787	0.846	0.821	0.777	0.710	0.728	0.730	0.726	0.713
100	0.970	0.944	0.877	0.777	0.862	0.815	0.774	0.747	0.749	0.723	0.734	0.709
200	0.979	0.956	0.909	0.800	0.863	0.846	0.777	0.725	0.776	0.736	0.723	0.704

NOTE: H_{KSS} is the Hausman test statistic proposed by Kapetanios, Serlenga, and Shin (2023), see Section S6 in the Online Supplement. See also notes to Table 1.

We allow weak error cross-sectional correlation by generating $\varepsilon_{it} = \upsilon_{it} + \theta \sum_{1 \leq h \leq 8} \upsilon_{i-h,t}$ with $\upsilon_{it} \sim$ iidN$(0, 1)$ and $\theta = 0.2$. We generate $\upsilon_{it} \sim$ iidN$(0, 1)$, and serially correlated factors as $f_{1t} = \rho_{f1} f_{1,t-1} + \nu_{1t}$ and $f_{2t} = \rho_{f2} f_{2,t-1} + \nu_{2t}$ where $(\nu_{1t}, \nu_{2t})' \sim$ iidN$(0, \boldsymbol{I}_2)$ and $\rho_{f1} = \rho_{f2} = 0.8$. We generate the factor loadings under the following two settings:

- Experiment 1 (independent factor loadings): $\gamma_{i1} \sim$ iidU$(0, 1)$, $\gamma_{i2} \sim$ iidU$(0, 1)$, $\Lambda_{i11} \sim$ iidU$(0, 1)$, $\Lambda_{i12} \sim$ iidU$(0, 2)$ such that $E \begin{pmatrix} \gamma_{i1} & \gamma_{i2} \\ \Lambda_{i11} & \Lambda_{i12} \end{pmatrix} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 1 \end{pmatrix}$.
- Experiment 2 (correlated factor loadings): $\gamma_{i1} = \Lambda_{i11} \sim$ iidU$(0, 1)$ and $\gamma_{i2} = \Lambda_{i12} \sim$ iidU$(0, 1)$.

We explore the performance of the LM_X test under parameter heterogeneity by generating $\rho_i \sim U(\rho_L, \rho_U)$ and $\beta_i = 1 + \eta_i$. We examine the following three cases for $r = 0, 1, 2, 3, 4$: Case 1: weak heterogeneity with $\rho_i \sim U(0.4, 0.6)$ and $\eta_i \sim N(0, 0.04)$; Case 2: medium heterogeneity with $\rho_i \sim U(0.25, 0.75)$ and $\eta_i \sim N(0, 0.25)$; Case 3: strong heterogeneity with $\rho_i \sim U(0.1, 0.9)$ and $\eta_i \sim N(0, 1)$. To explicitly take into account the presence of serially correlated factors, we use the ARDL approximation, (25) and construct the LM_X test using the ARDL residuals given by (26). We consider the combinations of $(N, T) = \{30, 50, 100, 200\}$, and set the number of replications at $R = 1,000$.

Table 5 reports the size and power performance under Experiments 1 and 2, respectively. The size of the LM_X test is close to the nominal level (5%) for all sample sizes and its performance

is invariant to the number of factors and the different strengths of parameter heterogeneity. Under Experiment 2 the LM_X test is reasonably powerful in small samples, but it becomes consistent as the sample size increases. Even if the parameter heterogeneity gets stronger and/or the number of factors increases, the power of the LM_X test remains satisfactory in most cases. This suggests that the size and power performance of the LM_X tests, using the ARDL residuals, is satisfactory for the heterogeneous dynamic panel data with IE, irrespective of whether the factors are serially correlated or not.

In Section S4.5 in the Online Supplement we have conducted comprehensive simulation exercises, covering the following 4 cases: Case A with serially uncorrelated factors and the dynamic FE residuals given by (23); Case B with serially uncorrelated factors and the ARDL residuals given by (26); Case C with serially correlated factors with $\rho_{f1} = \rho_{f2} = 0.8$ and the dynamic FE residuals; Case D with serially correlated factors and the ARDL residuals. The performance of the LM_X test under Case 1 is satisfactory while the simulation results for Case 2 are qualitatively similar to those for Case 1. This suggests that the performance of the LM_X test using either dynamic FE residuals or ARDL residuals is satisfactory, only if the factors are serially uncorrelated. Under Case 3, however, the LM_X test suffers from severe size distortions that worsen as the number of factors rises. This clearly suggests that we need to deal with serially correlated factors through the use of the ARDL approximation for correct inference. The simulation results under Case 4 are qualitatively similar to those for Case 2. Based on this evidence we propose the use of the ARDL residuals when constructing the LM_X test in practice.

Table 5. The size and power of the LM_X test for the heterogeneous dynamic panel data.

	Experiment 1											
	Case 1				Case 2				Case 3			
T/N	30	50	100	200	30	50	100	200	30	50	100	200
						$r = 0$						
30	0.051	0.055	0.051	0.051	0.055	0.049	0.043	0.055	0.050	0.048	0.042	0.049
50	0.057	0.047	0.050	0.054	0.053	0.038	0.040	0.056	0.047	0.053	0.041	0.045
100	0.028	0.045	0.050	0.055	0.046	0.051	0.050	0.045	0.036	0.039	0.065	0.046
200	0.044	0.040	0.056	0.053	0.032	0.054	0.051	0.056	0.036	0.045	0.051	0.058
						$r = 1$						
30	0.065	0.047	0.060	0.053	0.054	0.048	0.054	0.047	0.044	0.048	0.050	0.041
50	0.052	0.057	0.049	0.040	0.037	0.051	0.060	0.055	0.037	0.045	0.044	0.061
100	0.052	0.037	0.052	0.054	0.044	0.056	0.054	0.044	0.042	0.049	0.048	0.050
200	0.041	0.057	0.045	0.050	0.047	0.054	0.043	0.050	0.034	0.041	0.062	0.044
						$r = 2$						
30	0.058	0.051	0.045	0.052	0.044	0.059	0.048	0.059	0.032	0.053	0.047	0.053
50	0.045	0.042	0.046	0.047	0.045	0.047	0.045	0.048	0.047	0.042	0.054	0.053
100	0.058	0.055	0.044	0.048	0.060	0.041	0.039	0.048	0.044	0.045	0.043	0.048
200	0.037	0.057	0.048	0.055	0.046	0.055	0.048	0.056	0.038	0.039	0.056	0.043
						$r = 3$						
30	0.058	0.051	0.045	0.052	0.044	0.059	0.048	0.059	0.032	0.053	0.047	0.053
50	0.045	0.042	0.046	0.047	0.045	0.047	0.045	0.048	0.047	0.042	0.054	0.053
100	0.058	0.055	0.044	0.048	0.060	0.041	0.039	0.048	0.044	0.045	0.043	0.048
200	0.037	0.057	0.048	0.055	0.046	0.055	0.048	0.058	0.038	0.039	0.056	0.043
						$r = 4$						
30	0.050	0.054	0.047	0.058	0.039	0.064	0.054	0.048	0.050	0.048	0.041	0.056
50	0.046	0.056	0.062	0.046	0.040	0.057	0.060	0.045	0.050	0.056	0.045	0.047
100	0.052	0.053	0.043	0.065	0.052	0.058	0.055	0.043	0.047	0.051	0.049	0.041
200	0.059	0.046	0.051	0.056	0.049	0.056	0.048	0.047	0.041	0.045	0.055	0.055
	Experiment 2											
	Case 1				Case 2				Case 3			
T/N	30	50	100	200	30	50	100	200	30	50	100	200
						$r = 1$						
30	0.999	1	1	1	0.962	1	1	1	0.698	0.901	0.999	1
50	1	1	1	1	0.989	1	1	1	0.730	0.919	0.999	1
100	1	1	1	1	0.995	1	1	1	0.748	0.947	0.998	1
200	1	1	1	1	0.994	1	1	1	0.738	0.919	1	1
						$r = 2$						
30	0.99	1	1	1	0.921	0.984	1	1	0.595	0.841	0.975	1
50	1	1	1	1	0.952	0.997	1	1	0.648	0.866	0.989	1
100	1	1	1	1	0.969	0.999	1	1	0.645	0.872	0.990	1
200	1	1	1	1	0.970	1	1	1	0.661	0.871	0.995	1
						$r = 3$						
30	0.946	0.997	1	1	0.777	0.922	0.996	1	0.491	0.698	0.909	0.986
50	0.990	1	1	1	0.845	0.955	0.997	1	0.541	0.719	0.938	0.996
100	0.997	1	1	1	0.885	0.976	1	1	0.595	0.781	0.948	0.998
200	0.999	1	1	1	0.923	0.989	1	1	0.609	0.794	0.960	0.999
						$r = 4$						
30	0.855	0.982	0.999	1	0.650	0.836	0.988	0.999	0.410	0.624	0.836	0.967
50	0.948	0.997	1	1	0.761	0.911	0.989	0.999	0.491	0.705	0.879	0.968
100	0.989	1	1	1	0.844	0.946	0.990	0.999	0.579	0.751	0.888	0.982
200	0.994	1	1	1	0.876	0.949	0.995	1	0.585	0.786	0.916	0.978

NOTE: LM_X denotes the LM test defined in (16), where we construct $\ddot{u}_{it}$ by the individual ARDL residuals in (26). The performance of the LM_X test is explored under parameter heterogeneities—$\rho_i \sim U(\rho_L, \rho_U)$ and $\beta_{ik} = 1 + \eta_{ik}$—and serially correlated factors. We examine three cases of heterogeneities. Case 1 (weak heterogeneity) with $\rho_i \sim U(0.4, 0.6)$ and $\eta_i \sim N(0, 0.04)$; Case 2 (medium heterogeneity) with $\rho_i \sim U(0.25, 0.75)$ and $\eta_i \sim N(0, 0.25)$; Case 3 (strong heterogeneity) with $\rho_i \sim U(0.1, 0.9)$ and $\eta_i \sim N(0, 1)$. Factors are generated as $f_{1t} = \rho_{f1} f_{1,t-1} + v_{1t}$ and $f_{2t} = \rho_{f2} f_{2,t-1} + v_{2t}$ where $(v_{1t}, v_{2t})' \sim iidN(0, I_2)$ and $\rho_{f1} = \rho_{f2} = 0.8$. r is the number of factors.

4. Empirical Applications

To investigate the empirical relevance of the null hypothesis of (conditional) independence between regressors and factor loadings in the heterogeneous panel data model with IE, we apply the LM_X test to a total of 22 datasets (17 under a static panel data framework including 3 panels with small T and 5 under a dynamic panel data framework). In what follows we briefly describe the datasets used in the applications (see Section S5 in the Online Supplement for a full description).

4.1. Static Panel Data Framework

4.1.1. Cobb-Douglas Production Function
We estimate the production function in five different cases: the OECD members ($N = 26$ and $T = 41$, Mastromarco, Serlenga,

and Shin 2016), the Italian regions ($N = 20$ and $T = 21$), the 48 U.S. States ($N = 48$ and $T = 17$, Munnell 1990), the aggregate sectoral data for manufacturing from developed and developing countries ($N = 25$ and $T = 25$, Eberhardt and Teal 2020), and the manufacturing industries across the OECD countries ($N = 82$ and $T = 26$, Eberhardt, Helmers, and Strauss 2013).

4.1.2. Gravity Model of Bilateral Trade Flows

We estimate a gravity model of the bilateral trade flows, where bilateral trade flow is a function of GDP, countries' similarity, relative factor endowment, the real exchange rate as well as the trade union and common currency dummies. We estimate a model for the 91 pairs of EU14 countries from 1960 to 2008 ($N = 91$ and $T = 49$, Serlenga and Shin 2007).

4.1.3. Gasoline Demand Function

To evaluate the price and income elasticities of gasoline demand, we estimate the gasoline demand function using the quarterly data for the 50 U.S. States over the period 1994–2008 ($N = 50$ and $T = 60$, Liu 2014).

4.1.4. Housing Prices

We estimate the income elasticity of real housing prices from 1975 to 2010 using two datasets for the 49 U.S. States ($N = 49$ and $T = 36$, Holly, Pesaran, and Yamagata 2010) and the 384 Metropolitan Statistical Areas ($N = 384$ and $T = 36$, Baltagi and Li 2014).

4.1.5. Technological Spillovers on Productivity

We consider two applications. First, we estimate the effects of domestic and foreign R&D on total factor productivity (TFP) controlling for the human capital. We use a balanced panel of 24 OECD countries over the period 1971–2004 ($N = 24$ and $T = 34$), see Coe, Helpman, and Hoffmaister (2009) and Ertur and Musolesi (2017). Next, we explore the channels through which technological investments affect the productivity performance of industrialized economies by estimating the productivity effects of R&D and Information and Communication Technologies (ICT), controlling for the inputs accumulation as labor and (non-ICT) capital. We use a balanced panel of 49 high-tech OECD industries over the period 1977–2006 ($N = 53$ and $T = 30$, Pieri, Vecchi, and Venturini 2018).

4.1.6. Health Care

We estimate the relationship between healthcare expenditure and income after controlling for public expenditure over total health expenditure. We consider a panel of 167 countries covering the period 1995–2012 ($N = 167$ and $T = 18$, Baltagi et al. 2017).

4.1.7. Demographic and Business Cycle Volatility

We estimate the impact of the age composition of the labor force on business cycle volatility. We employ a balanced panel dataset for 51 countries over the period 1957–2000 ($N = 51$ and $T = 44$, Everaert and Vierke 2016).

4.1.8. Carbon Emissions and Trade

We explore the nexus between carbon emissions and trade using a balanced panel of 32 OECD countries over the period 1990–2013 ($N = 32$ and $T = 24$, Liddle 2018).

4.1.9. Health Care with Small T

We use a balanced panel data for 140 countries from 1993 to 1997 ($N = 140$, $T = 5$, Greene 2004) and regress a composite measure of health care delivery on per capita public and private health care expenditure. Further, we consider a balanced panel data for 2084 individuals from 1995 to 1999 ($N = 2084$, $T = 5$, Winkelmann 2004). We aim to evaluate the impact of the 1997 health care reform in Germany by estimating the regression of the number of doctor visits on the logged gross income.

4.1.10. Wage Equation with Small T

We estimate a reduced form of the wage equation by employing the balanced panel data ($N = 595$ and $T = 7$, Cornwell and Rupert 1988).

4.2. Dynamic Panel Data Framework

4.2.1. Economic Growth

We explore dynamic economic growth. First, we estimate the growth equation proposed by Islam (1995). We employ a balanced dataset of 87 countries from 1960 to 2007 ($N = 87$ and $T = 48$, Ditzen 2018). Next, we use a balanced panel of 47 countries over 1961–2003 ($N = 47$ and $T = 43$) and estimate the dynamic effects of temperature shocks on aggregate output growth, see Dell, Jones, and Olken (2012) and Vos and Everaert (2021). Lastly, we estimate a dynamic growth model to identify the effects of public debt on economic growth. We use a balanced panel data of 33 countries from 1972 to 2010 ($N = 33$ and $T = 42$, Chudik et al. 2017).

4.2.2. Energy Intensities and Urbanization

We estimate the dynamic impact of urbanisation on total energy usage by using a balanced panel of 24 Chinese provinces from 1987 to 2011 ($N = 24$ and $T = 25$, Ma 2015).

4.2.3. Income Inequality and Natural Resources

We analyse the dynamic relationship between natural resource endowments and income inequality using a balanced panel data of 17 OECD countries from 1981 to 2014 ($N = 17$ and $T = 33$, Kim, Chen, and Lin 2020).

Table 6 presents the main estimation and test results for the heterogeneous static panel datasets. The LM_X test can reject the null hypothesis (4) at the 5% significance only two times for the gravity model of bilateral trade flows and the income elasticity of real housing prices for MSA, while rejecting the null at the 10% significance two more times for the U.S. production function and the income elasticity of real housing prices for U.S. States, out of 14 datasets. These results provide convincing evidence that the null hypothesis of conditional independence between factor loadings and the regressors is not rejected in the majority of cases, suggesting that it is still valid to apply the consistent and robust FE estimator to the panel data with IE.

Table 6. Empirical applications for the static heterogeneous panel data.

		β_{MG}	β_P	β_{PC}	LM_X	H_B	H_W	CD	CD*
Production function									
OECD	$\beta_{\frac{k}{l}}$	0.534 ***	0.589 ***	0.633 ***	0.81	0.003	0.71	−2.743	−3.725
		(0.015)	(0.015)	(0.012)	(0.368)	(0.954)	(0.399)	(0.006)	(0.000)
ITA	$\beta_{\frac{k}{l}}$	0.312 ***	0.537 ***	0.343 ***	2.134	0.004	0.707	−2.930	−2.928
		(0.042)	(0.035)	(0.016)	(0.144)	(0.946)	(0.401)	(0.003)	(0.003)
US	$\beta_{\frac{k}{l}}$	0.134 ***	0.165 ***	0.106 ***	3.607	0.001	0.039	−1.111	−2.316
		(0.011)	(0.011)	(0.004)	(0.057)	(0.978)	(0.841)	(0.266)	(0.021)
UNIDO	$\beta_{\frac{k}{l}}$	0.441 ***	0.637 ***	0.563 ***	0.095	0.008	4.091	−2.139	−1.976
		(0.028)	(0.022)	(0.036)	(0.757)	(0.929)	(0.043)	(0.032)	(0.048)
Production and RD	β_l	0.418 ***	0.603 ***	0.487 ***					
$k = 3$		(0.018)	(0.013)	(0.010)					
	β_k	0.068 ***	0.490 ***	0.462 ***					
		(0.024)	(0.020)	(0.012)					
	β_{rd}	0.074 ***	0.068 ***	0.097 ***	1.257	2.175	2.476	−2.053	−3.071
		(0.008)	(0.006)	(0.007)	(0.739)	(0.536)	(0.479)	(0.040)	(0.002)
Gravity model	β_{gdp}	2.643 ***	3.299 ***	1.808 ***					
$k = 6$		(0.050)	(0.039)	(0.010)					
	β_{rer}	−0.094 ***	0.037 ***	−0.041 ***					
		(0.006)	(0.004)	(0.006)					
	β_{sim}	1.350 ***	1.362 ***	1.231 ***					
		(0.036)	(0.022)	(0.057)					
	β_{rlf}	−0.195 ***	0.033 ***	0.014 ***					
		(0.006)	(0.002)	(0.005)					
	β_{cee}	0.088 ***	0.291 ***	0.369 ***					
		(0.006)	(0.004)	(0.006)					
	β_{emu}	0.074 ***	0.254 ***	0.179 ***	15.12	11.41	24.48	−2.386	−2.193
		(0.008)	(0.006)	(0.006)	(0.019)	(0.076)	(0.001)	(0.017)	(0.028)
Gasoline demand	β_p	−0.138 ***	−0.161 ***	−0.101 ***					
$k = 2$		(0.007)	(0.007)	(0.002)					
	β_{inc}	0.536 ***	0.436 ***	0.392 ***	0.838	0.001	0.296	−2.051	−1.066
		(0.022)	(0.020)	(0.012)	(0.657)	(0.999)	(0.862)	(0.040)	(0.286)
Income elasticity									
US States	β_{inc}	0.605 ***	0.832 ***	0.634 ***	3.122	0.361	0.738	−1.712	−1.847
		(0.033)	(0.030)	(0.011)	(0.077)	(0.547)	(0.391)	(0.087)	(0.065)
US MSAs	β_{inc}	0.610 ***	0.595 ***	0.558 ***	6.304	0.127	14.21	30.442	2.017
		(0.004)	(0.004)	(0.002)	(0.012)	(0.721)	(0.001)	(0.000)	(0.044)
Technological spillovers									
OECD countries	β_{sd}	0.020 **	0.057 ***	0.087 ***					
$k = 3$		(0.008)	(0.006)	(0.003)					
	β_{sf}	0.306 ***	0.139 **	0.133 ***					
		(0.067)	(0.061)	(0.044)					
	β_{hc}	0.036 *	0.112 ***	0.013 ***	0.974	0.002	1.657	−3.066	−3.731
		(0.021)	(0.018)	(0.002)	(0.808)	(0.999)	(0.646)	(0.002)	(0.000)
OECD industries	β_l	0.527 ***	0.154 ***	0.185 ***					
$k = 4$		(0.032)	(0.019)	(0.023)					
	β_{nitc}	−0.148 ***	0.491 ***	0.512 ***					
		(0.041)	(0.026)	(0.040)					
	β_{itc}	0.027 ***	0.024 ***	0.084 ***					
		(0.008)	(0.008)	(0.006)					
	β_{rd}	−0.010 ***	0.027 ***	0.031 ***	2.855	0.974	4.121	−1.029	−2.449
		(0.004)	(0.003)	(0.008)	(0.582)	(0.913)	(0.389)	(0.303)	(0.014)
Health	β_{gdp}	0.871 ***	0.721 ***	0.732 ***					
$k = 2$		(0.008)	(0.006)	(0.007)					
	β_{pe}	0.239 ***	0.019 ***	0.021 **	2.610	0.003	3.676	−1.099	0.904
		(0.009)	(0.002)	(0.009)	(0.271)	(0.998)	(0.159)	(0.272)	(0.366)
Volatility	β_s	−0.141	−0.617 *	−0.683 *	0.192	0.003	0.161	−0.233	−2.644
		(0.374)	(0.370)	(0.391)	(0.661)	(0.955)	(0.688)	(0.816)	(0.008)
Carbon emission	β_{gdp}	0.393 ***	0.429	0.427 ***					
$k = 4$		(0.023)	(0.023)	(0.020)					
	β_{tsh}	−0.169 ***	−0.106	−0.096 ***					
		(0.016)	(0.015)	(0.013)					
	β_{ish}	0.004	−0.123	−0.125 *					
		(0.045)	(0.040)	(0.066)					
	β_{fsh}	0.880 ***	0.866 ***	0.987 ***	5.954	0.168	2.323	17.314	−2.482
		(0.038)	(0.037)	(0.053)	(0.202)	(0.996)	(0.676)	(0.000)	(0.013)

NOTE: β_{MG} and β_P denote the mean group and the pooled FE estimator; β_{PC} is the (bias-corrected) iterative principal component estimator (see Cui et al. 2023) applying the IC_{p1} criterion by Bai and Ng (2002), we extract 4 factors in the gravity dataset and 2 factors in the other datasets. The standard errors inside (.) are computed by the nonparametric variance-covariance estimators. LM_X denotes the LM test defined in (16) using the residuals in (15). H_B is the Hausman test statistic proposed by Bai (2009); H_W the Hausman test statistic proposed by Westerlund (2019b); CD is the Pesaran (2015) test; CD* is the bias corrected CD tests proposed by Pesaran and Xie (2021) with 1 factor. The CD tests are applied to the two way FE residuals.

Next, we turn to the results for the CD test proposed by Pesaran (2015) and the bias-corrected CD* test recently advanced by Pesaran and Xie (2021), both of which test the null hypothesis of weak residual CSD against the alternative of strong CSD. Both CD tests are applied to the two way FE residuals. We report the CD* test results computed using one factor only though the results are qualitatively similar when employing the different number of factors. The CD (CD*) test rejects the null hypothesis 10 (12) times out of 14 datasets. We also consider the Hausman tests proposed by Bai (2009) and Westerlund (2019b), denoted H_B and H_W, which test the null hypothesis of additive two-way effects against the alternative hypothesis of the multiplicative IE. The H_B test marginally rejects the null hypothesis only once for the gravity model of bilateral trade flows whilst the H_W test rejects two more times for UNIDO production function and the income elasticity of real housing prices for MSA. Apparently, these results are in conflict, since the CD test indicates the presence of CSD while the Hausman tests reveal the absence of IE in most panels. We have already shown that the H_B and H_W tests become inconsistent if factor loadings and the regressors are conditionally independent (see Monte Carlo simulations in Section S2 in the Online Supplement). In this regard, these conflicting results indicate that non-rejection of the null hypothesis by the Hausman tests is mainly due to the (observed) conditional independence between the regressors and loadings in the panel data with IE, rather than the absence of IE, under which the FE estimator is consistent. Furthermore, notice that for the three datasets where the H_B and H_W tests can reject the null of additive effects, we can also reject our null hypothesis. Combined together, this provides a strong support for the utility and importance of applying our proposed LM specification test in practice.

Next, Table 7 presents the test results for the dynamic heterogeneous panel data. Out of the five datasets, the LM_X test can reject the null hypothesis marginally only once for the effects of public debt on economic growth. Moreover, the CD (CD*) test rejects the null hypothesis of weak residual CSD at the 1% significance in 5 (4) cases. Notice, however, that the theoretical properties of both CD tests have not been established yet for the dynamic panel data model.

Table 8 reports the LM test results for static panel data with small T. We also find that the LM_X test does not reject the null hypothesis for all three datasets we consider.

Table 8. Empirical applications for heterogeneous panel data with small T.

	LM_X
Health expenditure	0.291
	(0.589)
Health care	1.933
	(0.164)
Wage	2.417
	(0.120)

NOTE: See notes to Tables 6.

Finally, following the suggestion by an anonymous referee, we examine whether the LM_X test results are invariant to the case where the correct model specification is dynamic, but the practitioner applies a static model specification, even though we have applied the same model specification as in the original studies. We now consider the dynamic specifications for the gravity model of the bilateral trade flows, the income elasticity of real housing prices, the R&D production function, and the demographic and business cycle volatility, where we add the lagged dependent variable to the original static specifications. We then apply the LM test statistic constructed using the ARDL residuals from the ARDL(p_T, p_T) approximation, (25). We find that the null hypothesis is rejected for the gravity and the housing income elasticity models whereas the null is not rejected for R&D production function and volatility (these results are available upon request), confirming that the LM test results are the same under both the static and dynamic panel data frameworks.

Combining all these results, we may conclude that our proposed LM test will make an essential specification test, given the pervasive evidence in favor of strong CSD. Therefore, we suggest applying the LM test to determine the form of IE that can validate the use of the FE estimator. Following the suggestion by an anonymous referee, we add practical guidelines for step-by-step model specifications in Section S7 in the Online Appendix. In the case where the null hypothesis of conditional independence is not rejected, the FE estimation can still produce consistent estimation and robust inference in a variety of cross-sectionally correlated panels, albeit less efficient than the PC estimator, though we emphasize that the FE estimator is invariant to any complex issues related to selecting the true number of unobserved factors (Moon and Weidner 2015), and to employing inconsistent initial estimates which may not guarantee the convergence of the iterative PC estimator (Hsiao 2018).

5. Conclusions

A large strand of the literature on panel data has focused on analyzing CSD, based on the error components model with IE, which is implicitly understood to bias the two-way FE estimator, due to the potential endogeneity arising from the correlation between regressors and factors/loadings (e.g., Bai 2009).

In this article we have built upon the notion that if the regressors and factor loadings are conditionally independent, then the panel data model with IE can still be consistently estimated by the two-way FE estimator. This suggests that the null hypothesis of conditional independence between the regressors and loadings emerges as an influential feature of the panel data modeling with IE. We have proposed an easy way of verifying the validity of such an important misspecification hypothesis

Table 7. Empirical applications for the dynamic heterogeneous panel data.

	LM_X	CD	CD*
Growth	1.607	2.942	−0.22
	(0.447)	(0.003)	(0.826)
Temperature	2.468	−3.564	−4.434
	(0.116)	(0.000)	(0.001)
Debt	3.461	−1.965	−2.843
	(0.063)	(0.049)	(0.001)
Energy intensity	3.839	−2.908	−3.217
	(0.279)	(0.004)	(0.001)
Oil	1.552	−3.474	−3.831
	(0.671)	(0.001)	(0.001)

NOTE: LM_X denotes the LM test defined in (16), where we construct the ARDL residuals from the ARDL(p_T, p_T) approximation, (25) with $p_T = 3$. See also notes to Table 6.

through the v test, that only requires the use of the FE estimator. Crucially, our proposed test can be easily applicable to either static panels with a small number of time periods or dynamic panels with serially correlated factors. In this regard the LM test can make a valuable addition to the toolkit of applied researchers as it provides a support for the use of the consistent and robust FE estimator under the null hypothesis.

Finally, we apply the LM test to a number of existing panel datasets, and find substantial evidence that the regressors and factor loadings are likely to be conditionally independent in practice. This suggests that the FE estimator can still provide a simple but robust strategy in a variety of cross-sectionally correlated panels including a growing literature in microeconometrics that focuses on estimating causal/counterfactual effects or policy evaluations.

We note a couple avenues for further researches. First, given the pervasive empirical evidence supporting the conditional independence between the regressors and factor loadings, it will be worthwhile to develop the simple consistent estimator for heterogeneous dynamic panel data model with IE under weaker conditions. Next, we can extend our testing approach to high-dimensional and/or multi-dimensional panels, for example, Cameron, Gelbach, and Miller (2011), Kapetanios, Serlenga, and Shin (2023) and Choi, Lin, and Shin (2021). For example, consider the three-dimensional heterogeneous panel data model of Kapetanios, Serlenga, and Shin (2023), given by $y_{ijt} = \boldsymbol{\beta}'_{ij}\boldsymbol{x}_{ijt} + u_{ijt}$, $i = 1,\ldots,N$, $j = 1,\ldots,N$, $t = 1,\ldots,T$, where y_{ijt} is the dependent variable observed across three indices, i being the origin unit, j the destination unit at period t and $\boldsymbol{x}_{ijt}$ is the $m_x \times 1$ vector of covariates with $\boldsymbol{\beta}_{ij}$ being an $m_x \times 1$ vector of parameters. Further, u_{ijt} is given by $u_{ijt} = \boldsymbol{\gamma}'_{ij}\boldsymbol{f}_t + \boldsymbol{\gamma}'_{\circ j}\boldsymbol{f}_{iot} + \boldsymbol{\gamma}'_{io}\boldsymbol{f}_{ojt} + \varepsilon_{ijt}$, where $\boldsymbol{f}_t, \boldsymbol{f}_{iot}$ and $\boldsymbol{f}_{ojt}$ are vectors of unobserved global, origin-specific and destination-specific factors with $\boldsymbol{\gamma}_{ij}$, $\boldsymbol{\gamma}_{\circ j}$ and $\boldsymbol{\gamma}_{io}$ being the corresponding vectors of heterogeneous loadings, and ε_{ijt} are idiosyncratic errors. If loadings and the regressors are conditionally independent, then we conjecture that the appropriate within estimator retains consistency. In this case an extended LM test can be obtained by

$$LM = \left(\frac{1}{\sqrt{TN}}\sum_{i,j=1}^{N}\widehat{\widetilde{\boldsymbol{X}}}'_{ij}\widehat{\boldsymbol{u}}_{ij}\right)'\widehat{\boldsymbol{V}}^{-1}\left(\frac{1}{\sqrt{TN}}\sum_{i,j=1}^{N}\widehat{\widetilde{\boldsymbol{X}}}'_{ij}\widehat{\boldsymbol{u}}_{ij}\right)$$

where $\widehat{\boldsymbol{V}} = \frac{1}{TN^2}\sum_{i,j=1}^{N}\widehat{\widetilde{\boldsymbol{X}}}'_{ij}\widehat{\boldsymbol{u}}_{ij}\widehat{\boldsymbol{u}}'_{ij}\widehat{\widetilde{\boldsymbol{X}}}_{ij}$, $\widehat{\boldsymbol{u}}_{ij}$ is the corresponding residual and $\widehat{\widetilde{\boldsymbol{X}}}_{ij}$ is the fitted value of the regressors when regressed on the first principal components of $\{\boldsymbol{x}_{ijt}\}_{i=1}^{N}$, $\{\boldsymbol{x}_{ijt}\}_{j=1}^{N}$ and $\{\boldsymbol{x}_{ijt}\}_{i,j=1}^{N}$. Clearly, a number of alternative extensions can be envisaged. Notice, however, that there are both conceptually and technically challenging issues to be addressed carefully.

Supplementary Materials

The online supplement contains all technical proofs, additional simulation results the full description of the datasets and the empirical specifications, and practical guidelines.

Acknowledgments

Disclosure Statement

Funding

References

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., and Wager, S. (2021), "Synthetic Difference in Differences," *American Economic Review*, 111, 4088–4118. [5]

Athey, S., and Imbens, G. W. (2022), "Design-based Analysis in Difference-In-Differences Settings with Staggered Adoption," *Journal of Econometrics*, 226, 62–79. [5]

Bai, J. (2009), "Panel Data Models with Interactive Fixed Effects," *Econometrica*, 77, 1229–1279. [1,2,3,4,6,16,17]

Bai, J., and Ng, S. (2002), "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70, 191–221. [16]

Baltagi, B. H., Lagravinese, R., Moscone, F., and Tosetti, E. (2017), "Health Care Expenditure and Income: A Global Perspective," *Health Economics*, 26, 863–874. [15]

Baltagi, B. H., and Li, J. (2014), "Further Evidence on the Spatio-Temporal Model of House Prices in the United States," *Journal of Applied Econometrics*, 29, 515–522. [15]

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011), "Robust Inference with Multiway Clustering," *Journal of Business & Economic Statistics*, 29, 238–249. [18]

Castagnetti, C., Rossi, E., and Trapani, L. (2015), "Testing for no Factor Structures: On the Use of Hausman-Type Statistics," *Economics Letters*, 130, 66–68. [1]

Charbonneau, K. B. (2017), "Multiple Fixed Effects in Binary Response Panel Data Models," *Econometrics Journal*, 20, S1–S13. [1]

Choi, I., Lin, R., and Shin, Y. (2021), "Canonical Correlation-based Model Selection for the Multilevel Factors," *Journal of Econometrics*, 233, 22–44. DOI:10.1016/j.jeconom.2021.09.008. [18]

Chudik, A., Mohaddes, K., Pesaran, M. H., and Raissi, M. (2017), "Is There a Debt-Threshold Effect on Output Growth?" *The Review of Economics and Statistics*, 99, 135–150. [15]

Chudik, A., and Pesaran, M. H. (2015), "Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Data Models with Weakly Exogenous Regressors," *Journal of Econometrics*, 188, 393–420. [1,6,7]

Coakley, J., Fuertes, A.-M., and Smith, R. (2006), "Unobserved Heterogeneity in Panel Time Series Models," *Computational Statistics and Data Analysis*, 50, 2361–2380. [1,3]

Coe, D. T., Helpman, E., and Hoffmaister, A. W. (2009), "International R&D Spillovers and Institutions," *European Economic Review*, 53, 723–741. [15]

Cornwell, C., and Rupert, P. (1988), "Efficient Estimation with Panel Data: An Empirical Comparison of Instrumental Variables Estimators," *Journal of Applied Econometrics*, 3, 149–155. [15]

Cui, G., Hayakawa, K., Nagata, S., and Yamagata, T. (2023), "A Robust Approach to Heteroskedasticity, Error Serial Correlation and Slope Heterogeneity in Linear Models with Interactive Effects for Large Panel Data," *Journal of Business & Economic Statistics*, 41, 862–875. [4,16]

Cui, G., Sarafidis, V., and Yamagata, T. (2022), "IV Estimation of Spatial Dynamic Panels with Interactive Effects: Large Sample Theory and an

Application on Bank Attitude towards Risk," *The Econometrics Journal*, 226, 62–79. [4]

Dell, M., Jones, B. F., and Olken, B. A. (2012), "Temperature Shocks and Economic Growth: Evidence from the Last Half Century," *American Economic Journal: Macroeconomics*, 4, 66–95. [15]

Ditzen, J. (2018), "Estimating Dynamic Common-Correlated Effects in Stata," *The Stata Journal*, 18, 585–617. [15]

Eberhardt, M., Helmers, C., and Strauss, H. (2013), "Do Spillovers Matter When Estimating Private Returns to R&D?" *The Review of Economics and Statistics*, 95, 436–448. [15]

Eberhardt, M., and Teal, F. (2020), "The Magnitude of the Task Ahead: Macro Implications of Heterogeneous Technology," *Review of Income and Wealth*, 66, 334–360. [15]

Ertur, C., and Musolesi, A. (2017), "Weak and Strong Cross-Sectional Dependence: A Panel Data Analysis of International Technology Diffusion," *Journal of Applied Econometrics*, 32, 477–503. [15]

Everaert, G., and Vierke, H. (2016), "Demographics and Business Cycle Volatility: A Spurious Relationship?" *Journal of Applied Econometrics*, 31, 1467–1477. [15]

Fernandez-Val, I., and Weidner, M. (2016), "Individual and Time Effects in Nonlinear Panel Models with Large N, T," *Journal of Econometrics*, 192, 291–312. [1]

Godfrey, L. G. (1989), "*Misspecification Tests in Econometrics: The Lagrange Multiplier Principle and Other Approaches*. Econometric Society Monographs. Cambridge: Cambridge University Press. [5]

Greene, W. (2004), "Distinguishing between Heterogeneity and Inefficiency: Stochastic Frontier Analysis of the World Health Organization's Panel Data on National Health Care Systems," *Health Economics*, 13, 959–980. [15]

Hannan, E. J., and Deistler, M. (1989), "The Statistical Theory of Linear Systems," *Statistical Papers*, 30, 239–242. [7]

Holly, S., Pesaran, M. H., and Yamagata, T. (2010), "A Spatio-Temporal Model of House Prices in the USA," *Journal of Econometrics*, 158, 160–173. [15]

Hsiao, C. (2018), "Panel Models with Interactive Effects," *Journal of Econometrics*, 206, 645–673. [3,17]

Islam, N. (1995), "Growth Empirics: A Panel Data Approach," *The Quarterly Journal of Economics*, 110, 1127–1170. [15]

Juodis, A. (2022), "A Regularization Approach to Common Correlated Effects Estimation," *Journal of Applied Econometrics*, 37, 788–810. [1]

Juodis, A., and Reese, S. (2021), "The Incidental Parameters Problem in Testing for Remaining Cross-section Correlation," *Journal of Business & Economic Statistics*, 40, 1191–1203. [2]

Kapetanios, G., Pesaran, M. H., and Yamagata, T. (2011), "Panels with Nonstationary Multifactor Error Structures," *Journal of Econometrics*, 160, 326–348. [1]

Kapetanios, G., Serlenga, L., and Shin, Y. (2023), "Estimation and Inference for Multi-Dimensional Heterogeneous Panel Datasets with Hierarchical Multi-Factor Error Structure," *Journal of Econometrics*, 220, 504–531. [18]

——— (2023), "Testing for Correlation between the Regressors and Factor Loadings in Heterogeneous Panels with Interactive Effects," *Empirical Economics*, 64, 2611–2659. [1,2,3,4,11,13]

Karabiyik, H., Reese, S., and Westerlund, J. (2017), "On the Role of the Rank Condition in CCE Estimation of Factor-Augmented Panel Regressions," *Journal of Econometrics*, 197, 60–64. [4]

Kim, D.-H., Chen, T.-C., and Lin, S.-C. (2020), "Does Oil Drive Income Inequality? New Panel Evidence," *Structural Change and Economic Dynamics*, 55, 137–152. [15]

Kripfganz, S., and Sarafidis, V. (2021), "Instrumental-Variable Estimation of Large-t Panel-Data Models with Common Factors," *Stata Journal*, 21, 659–686. [2]

Lewis, R. A., and Reinsel, G. C. (1988), "Prediction Error of Multivariate Time Series with Mis-specified Models," *Journal of Time Series Analysis*, 9, 43–57. [7]

Liddle, B. (2018), "Consumption-based Accounting and the Trade-Carbon Emissions Nexus," *Energy Economics*, 69, 71–78. [15]

Liu, W. (2014), "Modeling Gasoline Demand in the United States: A Flexible Semiparametric Approach," *Energy Economics*, 45, 244–253. [15]

Ma, B. (2015), "Does Urbanization Affect Energy Intensities across Provinces in China? Long-Run Elasticities Estimation Using Dynamic Panels with Heterogeneous Slopes," *Energy Economics*, 49, 390–401. [15]

Mastromarco, C., Serlenga, L., and Shin, Y. (2016), "Modelling Technical Efficiency in Cross Sectionally Dependent Stochastic Frontier Panels," *Journal of Applied Econometrics*, 31, 281–297. [2,15]

Moon, H. R., and Weidner, M. (2015), "Linear Regression for Panel with Unknown Number of Factors as Interactive Fixed Effects," *Econometrica*, 83, 1543–1579. [1,6,7,17]

——— (2017), "Dynamic Linear Panel Regression Models with Interactive Fixed Effects," *Econometric Theory*, 33, 158–195. [6]

Munnell, A. (1990), "How Does Public Infrastructure Affect Regional Economic Performance," *New England Economic Review*, 34, 11–33. [15]

Norkute, 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. [7,8]

Pesaran, M., and Smith, R. (1995), "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels," *Journal of Econometrics*, 68, 79–113. [7]

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

——— (2015), "Testing Weak Cross-Sectional Dependence in Large Panels," *Econometric Reviews*, 34, 1089–1117. [1,2,3,16,17]

——— (2021), "General Diagnostic Tests for Cross-Sectional Dependence in Panels," *Empirical Economics*, 60, 13–50. [3]

Pesaran, M. H., and Xie, Y. (2021), "A Bias-Corrected CD Test for Error Cross-Sectional Dependence in Panel Data Models with Latent Factors," arXiv Working Paper 2109.00408. [16,17]

Pieri, F., Vecchi, M., and Venturini, F. (2018), "Modelling the Joint Impact of R&D and ICT on Productivity: A Frontier Analysis Approach," *Research Policy*, 47, 1842–1852. [15]

Sarafidis, V., and Wansbeek, T. (2012), "Cross-Sectional Dependence in Panel Data Analysis," *Econometric Reviews*, 31, 483–531. [1,3]

Sarafidis, V., Yamagata, T., and Robertson, D. (2009), "A Test of Cross Section Dependence for a Linear Dynamic Panel Model with Regressors," *Journal of Econometrics*, 148, 149–161. [1,3]

Serlenga, L., and Shin, Y. (2007), "Gravity Models of Intra-EU Trade: Application of the CCEP-HT Estimation in Heterogeneous Panels with Unobserved Common Time-Specific Factors," *Journal of Applied Econometrics*, 22, 361–381. [15]

Song, M. (2013), "Asymptotic Theory for Dynamic Heterogeneous Panels with Cross-sectional Dependence and Its Applications," Mimeo, Columbia University. [6]

Vos, I. D., and Everaert, G. (2021), "Bias-Corrected Common Correlated Effects Pooled Estimation in Dynamic Panels," *Journal of Business & Economic Statistics*, 39, 294–306. [15]

Westerlund, J. (2019a), "On Estimation and Inference in Heterogeneous Panel Regressions with Interactive Effects," *Journal of Time Series Analysis*, 40, 852–857. [1,3]

——— (2019b), "Testing Additive versus Interactive Effects in Fixed-T Panels," *Economics Letters*, 174, 5–8. [1,3,16,17]

Westerlund, J., and Urbain, J. P. (2013), "On the Estimation and Inference in Factor-Augmented Panel Regressions with Correlated Loadings," *Economics Letters*, 119, 247–250. [3]

——— (2015), "Cross-Sectional Averages versus Principal Components," *Journal of Econometrics*, 185(2):372–377. [1]

Winkelmann, R. (2004), "Health Care Reform and the Number of Doctor Visits: An Econometric Analysis," *Journal of Applied Econometrics*, 19, 455–472. [15]