A Bias-Corrected CD Test for Error Cross-Sectional Dependence in Panel Data Models with Latent Factors†

M. Hashem Pesaran
USC, and Trinity College, Cambridge, UK
and
Yimeng Xie
Xiamen University, China

August 23, 2022

Abstract

In a recent paper Juodis and Reese (2022) (JR) show that the application of the CD test proposed by Pesaran (2004) to residuals from panels with latent factors results in over-rejection. They propose a randomized test statistic to correct for over-rejection, and add a screening component to achieve power. This paper considers the same problem but from a different perspective. It shows that the standard CD test remains valid if the latent factors are weak, and proposes a simple bias-corrected CD test, labelled CD^*, which is shown to be asymptotically normal, irrespective of whether the latent factors are weak or strong. This result is shown to hold for pure latent factor models as well as for panel regressions with latent factors. The case where the errors are serially correlated is also considered. Small sample properties of the CD^* test are investigated by Monte Carlo experiments and are shown to have the correct size and satisfactory power for both Gaussian and non-Gaussian errors. In contrast, it is found that JR's test tends to over-reject in the case of panels with non-Gaussian errors, and has low power against spatial network alternatives. The use of the CD^* test is illustrated with two empirical applications from the literature.

JEL Classifications: C18, C23, C55

Key Words: Latent factor models, strong and weak factors, error cross-sectional dependence, spatial and network alternatives, size and power.

†We would like to thank Natalia Bailey, Alex Chudik, Zhan Gao, Arturas Juodis, Ron Smith, Simon Reese and Qiankun Zhou for helpful comments, and two anonymous referees for constructive comments and suggestions. The research for this paper was completed when Yimeng Xie was a Ph.D. student at the University of Southern California.

1 Introduction

It is now quite standard to use latent multi-factor models to characterize and explain cross-sectional dependence in panels when the cross section dimension (n) and the time series dimension (T) are both large. However, due to uncertainty regarding the nature of error cross-sectional dependence, it is arguable whether the cross-sectional dependence is fully accounted for by latent factors. Some of the factors could be semi-strong, and the errors might have spatial or network features that are not necessarily captured by common factors alone. Chudik et al. (2011) provide an early discussion of the different sources of cross-sectional dependence, where it is shown that for a factor model to capture spatial dependence one needs a weak factor model where the number of weak factors tends to infinity with n. It is clearly desirable to test for error cross-sectional dependence once the common factor effects are filtered out.

In a recent paper Juodis and Reese (2022) (JR) show that the application of the CD test proposed by Pesaran (2004, 2015a) to the residuals from panels with latent factors is invalid and can result in over-rejection of the null of error cross-sectional independence. They propose a randomized CD test statistic as a solution. Their proposed test is constructed in two steps. First, they multiply the residuals from panel regressions with independent randomized weights to obtain their CD_W statistic, which will have a zero mean by construction. In this way they avoid the over-rejection problem of the CD test, but by the very nature of the randomization process they recognize that the CD_W test will lack power. To overcome the problem of lack of power, JR modify the CD_W test statistic by adding to it a screening component proposed by Fan et al. (2015) which is expected to tend to zero with probability approaching one under the null hypothesis, but to diverge at a reasonably fast rate under the alternative. This further modification of CD_W test is denoted by CD_{W+}. Accordingly, it is presumed that the CD_{W+} test can overcome both over-rejection and the low power problems. However, JR do not provide a formal proof establishing conditions under which the screening component tends to zero under the null and diverges sufficiently fast under alternatives, including spatial or network dependence type alternatives. Using theoretical results established by Bailey et al. (2019) for correlation coefficients we show that the screening component in JR need not converge to zero. Also, our Monte Carlo simulations show that the CD_{W+} test tends to over-reject when the errors are non-Gaussian and $n >> T$, and seems to lack power under spatial alternatives, which is likely to be particularly important in empirical applications.

In this paper we show that the standard CD test is in fact valid for testing error cross-sectional dependence in panel data models with weak latent factors. However, when the latent factors are semi-strong or strong the use of CD test will result in over-rejection and will no longer be valid, extending JR's results to panels with semi-strong latent factors. In short, whilst the CD_{W+} is a useful and welcome addition to testing for error cross-sectional dependence, it would be interesting to develop a modified version of the test that simultaneously deals with the over-rejection problem and does not compromise power for a general class of alternatives. To that end, firstly we study testing for error cross-sectional dependence in a pure latent factor model, and derive an explicit expression for the bias of the CD test statistic in terms of factor loadings and error variances. We then propose a bias-corrected version of the CD test statistic, denoted by CD^*, which is shown to have $N(0,1)$ asymptotic distribution under the null hypothesis irrespective of whether the latent factors are weak or strong. When the latent factors are weak the correction tends to zero, CD and CD^* will be asymptotically equivalent. However, $CD - CD^*$ diverges if at least one of the underlying latent factors is strong. We show that CD^* converges to a standard normal distribution when n and T tend to infinity so long as $n/T \to \kappa$, where $0 < \kappa < \infty$, and a test based on CD^* will have the correct size asymptotically. We then consider the application of the CD^* to test error cross-sectional dependence in the case of panel regressions with latent factors, discussed in Pesaran (2006). It is shown that the asymptotic properties of CD^* in the case of pure latent factor models also carry over to panel data models with latent factors. We also investigate the application of the CD^* to panel data models with serially correlated errors, and consider the method proposed by Baltagi et al. (2016) as well as using an autoregressive distributed lag (ARDL) representation which transforms the model with the serially correlated errors to one without error serial correlation.

The finite sample performance of the CD^* test is investigated by Monte Carlo simulations in the case of pure factor models, panel regressions with latent factors with and without error serial correlation. It is found that the CD^* test avoids the over-rejection problem under the null and diverges fast under spatial alternatives, and has desirable small sample properties regardless of whether the errors are Gaussian or not, under different combinations of n and T. Although computation of CD^* requires estimation of factors and their loadings, the simulation results suggest that prior information on the number of latent factors is unnecessary so long as the number of estimated (selected) factors is no less than the true number. We also find that both adjustments for dealing with error serial correlation considered in the paper give desirable small sample properties. Finally, as compared to JR's CD_{W+} test, the proposed bias-corrected CD test is better in controlling the size of the test and has much better power properties against spatial (or network) alternatives.

The use of CD^* is illustrated by two empirical applications studied in literature. In the first application, we examine modeling real house price changes in the U.S. Because it is evident that real house price changes are driven by macroeconomic trends which can be modeled by latent factors, it is necessary to filter out these factors before testing for spillover effect. By applying CD^* to real house price changes in the U.S. we are able to show significant existence of weak cross-sectional dependence in addition to latent factors. In the second application, we consider modeling R&D investment in industries. Because there is knowledge spillover between industries as well as other cross-sectional dependencies, modeling R&D investment needs to include latent factors and researchers usually apply the CCE approaches proposed by Pesaran (2006) to estimate coefficients. With CD^*, we find that the evidence of cross-sectional dependence in the CCE residuals of modeling R&D investment is weak when the number of selected principal components (PCs) is sufficiently large. The robustness of the empirical applications to possible error serial correlation is also investigated using the approach of Baltagi et al. (2016) as well as the proposed ARDL.

The paper is set out as follows. Section 2 considers a pure latent factor model, establishes the limiting properties of the CD test in the presence of latent factors, derives the bias-corrected test statistic, CD^*, and establishes its asymptotic distribution. The extension to more general panel data models with observed covariates as well as latent factors are discussed in Section 3. Then two adjustments of CD^* test for panels with serially correlated errors are discussed in Section 4. Section 5 sets up the Monte Carlo experiments and reports the small sample properties of CD, CD^* and, CD_{W+} tests. Section 6 provides the empirical illustrations. The proofs of the main theorems and the related lemmas are given in the online supplement.

2 CD* test for a pure latent factor model

We consider the following general linear panel data model which explains y_{it}, in terms of observed and latent covariates,

$$y_{it} = \boldsymbol{\alpha}_i'\mathbf{d}_t + \boldsymbol{\beta}_i'\mathbf{x}_{it} + \boldsymbol{\gamma}_i'\mathbf{f}_t + u_{it}, \tag{1}$$

for $i = 1, 2, ..., n$ and $t = 1, 2, ..., T$, where $\mathbf{d}_t$ is a $k_d \times 1$ vector of observed common factors, $\mathbf{x}_{it}$ is a $k_x \times 1$ vector of unit-specific observed covariates, and $\mathbf{f}_t = (f_{1t}, f_{2t}, ..., f_{m_0 t})'$ is an $m_0 \times 1$ vector of *unobserved* factors. $\boldsymbol{\alpha}_i = (\alpha_{i1}, \alpha_{i2}, ..., \alpha_{ik_d})'$, $\boldsymbol{\beta}_i = (\beta_{i1}, \beta_{i2}, ..., \beta_{ik_x})'$ and $\boldsymbol{\gamma}_i = (\gamma_{i1}, \gamma_{i2}, ..., \gamma_{im_0})'$ are the associated vectors of unknown coefficients. u_{it} is the idiosyncratic error for unit i at time t, and its cross-sectional dependence property is the primary object of interest. To simplify the exposition and to highlight the main issue of concern, namely the presence of unobserved factors, initially we focus on the pure factor model,

$$y_{it} = \boldsymbol{\gamma}_i'\mathbf{f}_t + u_{it}, \tag{2}$$

and assume that u_{it} are serially independent. Both of these restrictions will be relaxed. We also assume that m_0, the true number of factors, is known, and make the following standard assumptions:

Assumption 1 *(a) $\mathbf{f}_t$ is a covariance-stationary process with zero means and the covariance matrix, $E(\mathbf{f}_t\mathbf{f}_t') = \boldsymbol{\Sigma}_{ff} > 0$. (b) $T^{-1}\sum_{t=1}^{T}[\|\mathbf{f}_t\|^j - E(\|\mathbf{f}_t\|^j)] \to_p 0$, for $j = 3, 4$, as $T \to \infty$. (c) There exists T_0 such that for all $T > T_0$, $T^{-1}\sum_{t=1}^{T}\mathbf{f}_t\mathbf{f}_t' = T^{-1}\mathbf{F}'\mathbf{F} = \boldsymbol{\Sigma}_{T,ff} > \mathbf{0}$, and $\boldsymbol{\Sigma}_{T,ff} \to_p \boldsymbol{\Sigma}_{ff} > \mathbf{0}$, where $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_T)'$.*

Assumption 2 *$u_{it} \sim IID(0, \sigma_i^2)$, where σ_i^2 are treated as given such that $\sup_i \sigma_i^2 < K$, $inf_i\sigma_i^2 > c$, and $E\left(|u_{it}|^{8+c}\right) < K$. (a)There exists a finite integer T_0 such that for all $T > T_0$, $inf_i\omega_{i,T}^2 = T^{-1}inf_i\left(\mathbf{u}_i'\mathbf{M}_F\mathbf{u}_i\right) > c$,*

$$\sup_i E\left(\frac{\mathbf{u}_i'\mathbf{M}_F\mathbf{u}_i}{T}\right)^{-8-\epsilon} < K, \tag{3}$$

for some small c and $\epsilon > 0$, where $\mathbf{u}_i = (u_{i1}, u_{i2}, ..., u_{iT})'$ and $\mathbf{M}_F = \mathbf{I}_T - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$. (b) u_{it} and $u_{jt'}$ are distributed independently for all $i \neq j$ and $t \neq t'$, such that $E[\lambda_{max}(\mathbf{V}_T)] = O(1)$, where $\mathbf{V}_T = T^{-1}\sum_{t=1}^{T}\mathbf{u}_t\mathbf{u}_t'$, and $\mathbf{u}_t = (u_{1t}, u_{2t}, ..., u_{nt})'$. (c) for all i and t, u_{it} is distributed independently of $\mathbf{f}_{t'}$, for all i, t and t'.

Assumption 3 *The $m_0 \times 1$ vector of factor loadings $\boldsymbol{\gamma}_i$ is bounded such that $\sup_i \|\boldsymbol{\gamma}_i\| < K$, $n^{-1}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i' = \boldsymbol{\Sigma}_{n,\gamma\gamma} \to \boldsymbol{\Sigma}_{\gamma\gamma} > \mathbf{0}$, and*

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T} b_{in} f_{jt} u_{it} \gamma_{ij} = O_p(1), \text{ for } j = 1, 2, ..., m_0, \tag{4}$$

where $\{b_{in}\}$ a sequence of bounded constants such that $n^{-1}\sum_{i=1}^{n} b_{in}^2 = O(1)$. Also

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} a_{i,n}^2 > 0, \tag{5}$$

where

$$a_{i,n} = 1 - \sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i, \tag{6}$$

and $\boldsymbol{\varphi}_n = n^{-1}\sum_{i=1}^{n}\boldsymbol{\gamma}_i/\sigma_i$.

Remark 1 *Let* $\varepsilon_{it} = u_{it}/\sigma_i$ *and* $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{iT})$. *The fact that there exists a finite* T_0 *such that (3) holds can be established readily if it is further assumed that* $\boldsymbol{\varepsilon}_i \sim IIDN(\mathbf{0}, \mathbf{I}_T)$. *In this case* $\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i$ *is distributed as* $\chi^2_{T-m_0}$ *and* $E\left(\frac{T}{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i}\right)^8 < K$ *so long as* $T > m_0 + 16$. *See Section 4 of Smith (1988). Under non-Gaussian errors a larger value of* T *will be typically needed for the moment condition (3) to hold.*

Remark 2 *The sequence of bounded constants,* b_{in}, *is introduced in (4) for convenience and can be readily absorbed as scalars of* f_{jt} *and* γ_{ij}, *since factors and their loadings are only identified up to rotations.*

Remark 3 *Condition (5) is required for the validity of the bias corrected CD test proposed in this paper and, for example, rules out the special case, noted by one of the reviewers, where* $m_0 = 1$, $\sigma_i = \sigma$, $\gamma_i = \gamma = \pm 1$.

To allow one or more of the latent factors to be weak, following Bailey et al. (2021) we denote the strength of factor j by α_j as defined by the rate at which the sum of absolute values of factor loadings rises with n, namely

$$\sum_{i=1}^{n} |\gamma_{ij}| = \ominus\left(n^{\alpha_j}\right), \text{ for } j = 1, 2, \ldots, m_0. \tag{7}$$

The case of strong factors assumed in the principal component analysis (PCA) literature corresponds to $\alpha_j = 1$, for $j = 1, 2, ..., m_0$. The CD test does not require any modifications if the factors are weak, namely when $\alpha_j < 1/2$ for all j. The intermediate case of semi-strong factors where $1/2 < \alpha_j < 1$ leads to additional technical challenges and will not be considered in this paper.

Most of the above assumptions relate closely to those made in the literature on CD tests and large dimensional factor models. See, for example, the assumptions in Pesaran (2004, 2015a), and assumptions L and LFE in Bai and Ng (2008). The zero means in Assumption 1 are not restrictive and will be relaxed when we consider panel data models with observed regressors. Under Assumption 3 all factors are required to be strong. Since $\boldsymbol{\gamma}_i$ and $\mathbf{f}_t$ are identified only up to an $m_0 \times m_0$ non-singular rotation matrix, we set $\boldsymbol{\Sigma}_{\gamma\gamma} = \mathbf{I}_{m_0}$, where $\mathbf{I}_{m_0}$ is an identity matrix of order m_0. However, later we show that our main Theorem 1 continues to hold so long as the maximum factor strength $\alpha = max_j(\alpha_j) = 1$, namely there is at least one strong factor. It is not required that all m_0 latent factors should be strong, as required when Assumption 3 holds. Assumption 2 is a technical assumption, also made for the proof of the asymptotic normality of the standard CD test.

Under the above assumptions the asymptotic results of Bai (2003) apply, and the latent factors and their loadings can be estimated using PCs, given as the solution to the following optimization problem

$$\min_{\mathbf{F},\boldsymbol{\Gamma}} \sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it} - \boldsymbol{\gamma}_i'\mathbf{f}_t\right)^2,$$

where $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_T)'$ and $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_n)'$, with $\hat{\mathbf{F}}$ and $\hat{\boldsymbol{\Gamma}}$ satisfying the normalization restrictions: $n^{-1}\hat{\boldsymbol{\Gamma}}'\hat{\boldsymbol{\Gamma}} = \mathbf{I}_{m_0}$, and $T^{-1}\hat{\mathbf{F}}'\hat{\mathbf{F}}$ being a diagonal matrix. The estimators of factors and their loadings are then given by

$$\hat{\boldsymbol{\Gamma}} = \sqrt{n}\hat{\mathbf{Q}}, \text{ and } \hat{\mathbf{F}} = \frac{1}{\sqrt{n}}\mathbf{Y}\hat{\mathbf{Q}}, \tag{8}$$

where we define $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{iT})'$ for $i = 1, 2, \ldots, n$ so that $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ is the $T \times n$ matrix of observations on y_{it} and $\hat{\mathbf{Q}}$ is $n \times m_0$ matrix of the associated orthonormal eigenvectors of $\mathbf{Y}'\mathbf{Y}$. Then the residuals to be used in the construction of the CD test statistics are given by

$$e_{it} = y_{it} - \hat{\boldsymbol{\gamma}}_i'\hat{\mathbf{f}}_t. \tag{9}$$

2.1 The CD test and its JR modification

The CD test statistic based on the residuals, (9), is given by

$$CD = \sqrt{\frac{2T}{n(n-1)}} \left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \hat{\rho}_{ij,T} \right), \tag{10}$$

where $\hat{\rho}_{ij,T} = T^{-1} \sum_{t=1}^{T} \tilde{e}_{it,T} \tilde{e}_{jt,T}$, $\tilde{e}_{it,T}$ is the scaled residual defined by,

$$\tilde{e}_{it,T} = \frac{e_{it}}{\hat{\sigma}_{i,T}}, \tag{11}$$

and $\hat{\sigma}_{i,T}^2 = T^{-1} \sum_{t=1}^{T} e_{it}^2 > c > 0$. JR consider a panel regression model with latent factors, assuming that all the factors are strong and show that in that case $CD = O_p\left(\sqrt{T}\right)$, and its use will lead to gross over-rejection of the null of error cross-sectional independence. To deal with the over-rejection problem they propose the following randomized CD test based on the random weights, w_i, drawn independently of the residuals, e_{jt}, namely

$$CD_W = \sqrt{\frac{2}{Tn(n-1)}} \sum_{t=1}^{T} \sum_{i=2}^{n} \sum_{j=1}^{i-1} (w_i e_{it}) (w_j e_{jt}), \tag{12}$$

where w_i, for $i = 1, 2, ..., n$ are independently drawn from a Rademacher distribution. The CD_W statistic can also be computed using the scaled residuals, $\tilde{e}_{it,T}$. The test outcomes do not seem to be much affected by whether scaled or unscaled residuals are used. Here we follow JR and define CD_W in terms of unscaled residuals. Because of the random properties of the weights, JR show that CD_W converges to a standard normal distribution regardless of the values of e_{it}, and as a result the over-rejection problem of the standard CD test is avoided if CD_W statistic is used instead. But as recognized by JR, this is achieved at the expense of power. To overcome this limitation, JR construct another power enhanced test statistic by following Fan et al. (2015), and add the screening component, Δ_{nT}, to CD_W, to obtain CD_{W+} defined by

$$CD_{W+} = CD_W + \Delta_{nT}, \tag{13}$$

where

$$\Delta_{nT} = \sum_{i=2}^{n} \sum_{j=1}^{i-1} |\hat{\rho}_{ij,T}| \, \mathbf{1} \left(|\hat{\rho}_{ij,T}| > 2\sqrt{\frac{\ln(n)}{T}} \right). \tag{14}$$

For the CD_{W+} test to have the correct size under $H_0 : \rho_{ij} = 0$, for all $i \neq j$, the screening component Δ_{nT} must converge to zero as n and $T \to \infty$, jointly. To our knowledge, the conditions under which this holds are not investigated by JR. Whilst it is beyond the scope of the present paper to investigate the limiting properties of Δ_{nT} in the case of a general factor model, using results presented in Bailey et al. (2019) (BPS), we will provide sufficient conditions for $\Delta_{nT} \rightarrow_p 0$ in the case of the simple model $y_{it} = \mu_i + \sigma_i \varepsilon_{it}$. By the Cauchy-Schwarz inequality we first note that for all $i \neq j$,

$$
\begin{aligned}
& E\left[|\hat{\rho}_{ij,T}|\, I\left(|\hat{\rho}_{ij,T}| > 2\sqrt{\frac{\ln(n)}{T}}\right)|\rho_{ij} = 0\right] \\
& \leq \left[E\left(|\hat{\rho}_{ij,T}|^2 |\rho_{ij} = 0\right)\right]^{1/2} \left[P\left(|\hat{\rho}_{ij,T}| > 2\sqrt{\frac{\ln(n)}{T}}|\rho_{ij} = 0\right)\right]^{1/2},
\end{aligned} \tag{15}
$$

where $\rho_{ij} = E(\varepsilon_{it}\varepsilon_{jt})$. Hence

$$
\begin{aligned}
& E\left(\Delta_{nT}|\rho_{ij} = 0, \text{ for all } i \neq j\right) \\
& \leq \frac{n(n-1)}{2} \sup_{i \neq j} \left[E\left[|\hat{\rho}_{ij,T}|^2 |\rho_{ij} = 0\right]\right]^{1/2} \sup_{i \neq j} \left[P\left[|\hat{\rho}_{ij,T}| > 2\sqrt{\frac{\ln(n)}{T}}|\rho_{ij} = 0\right]\right]^{1/2}.
\end{aligned} \tag{16}
$$

Now using results (9) and (10) of BPS, we have

$$
E\left[|\hat{\rho}_{ij,T}|^2 |\rho_{ij} = 0\right] = O\left(\frac{1}{T}\right), \tag{17}
$$

and using result (A.4) in the online supplement of BPS, we also have

$$
\sup_{i \neq j} P\left[|\hat{\rho}_{ij,T}| > \frac{C_p(n,\delta)}{\sqrt{T}}|\rho_{ij} = 0\right] = O\left(e^{-\frac{1}{2}\frac{C_p^2(n,\delta)}{\varphi_{\max}}}\right) + O\left(T^{-\frac{(s-1)}{2}}\right),
$$

where $C_p(n,\delta) = \Phi^{-1}\left(1 - \frac{p}{2n\delta}\right)$, $0 < p < 1$, $\Phi^{-1}(\cdot)$ is the inverse of the cumulative distribution of a standard normal variable, $\delta > 0$, $\varphi_{\max} = \sup_{i \neq j} E\left(\varepsilon_{it}^2 \varepsilon_{jt}^2\right)$, and s is such that $\sup_{i \neq j} E|\varepsilon_{it}|^{2s} < K$, for some integer $s \geq 3$ (see Assumption 2 of BPS). Also using results in Lemma 2 in the online supplement of BPS, we have

$$
\lim_{n \to \infty} \frac{C_p^2(n,\delta)}{\ln(n)} = 2\delta, \text{ and } e^{\frac{-\frac{1}{2}C_p^2(n,\delta)}{\varphi_{\max}}} = O\left(n^{-\delta/\varphi_{\max}}\right).
$$

Therefore, $T^{-1/2}C_p(n,\delta)$, and $2\sqrt{T^{-1}\ln(n)}$ have the same limiting properties if we set $\delta = 2$. Overall, it then follows that

$$
\sup_{i \neq j} P\left(|\hat{\rho}_{ij,T}| > 2\sqrt{\frac{\ln(n)}{T}}|\rho_{ij} = 0\right) = O\left(n^{-\frac{2}{\varphi_{\max}}}\right) + O\left(T^{-\frac{(s-1)}{2}}\right). \tag{18}
$$

Using (17) and (18) in (16), we now have

$$
E\left(\Delta_{nT}|\rho_{ij} = 0, \text{ for all } i \neq j\right) = O\left(\frac{n^{2-\frac{1}{\varphi_{\max}}}}{\sqrt{T}}\right) + O\left(n^2 T^{-\frac{(s+1)}{4}}\right). \tag{19}
$$

Therefore, $\Delta_{nT} \to_p 0$, if $n^2 T^{-\frac{s+1}{4}} \to 0$ and $T^{-1/2} n^{2-\frac{1}{\varphi_{\max}}} \to 0$. It is easily seen that both of these conditions will be met as n and $T \to \infty$ and $n = o\left(\sqrt{T}\right)$ if ε_{it} is Gaussian, since under Gaussian errors, $\varphi_{\max} = 1$ and s can be taken to be sufficiently large. But, in general the expansion rate of T relative to n required to ensure $\Delta_{nT} \to_p 0$ will also depend on the degree to which $E\left(\varepsilon_{it}^2 \varepsilon_{jt}^2\right)$ exceeds unity. For example, if ε_{it} has a multivariate t-distribution with degrees of freedom $v > 4$, then letting $T = n^d$, $d > 0$, and using results in Lemma 5 of BPS's online supplement, we have

$$\varphi_{\max} = \sup_{i \neq j} E\left(\varepsilon_{it}^2 \varepsilon_{jt}^2 | \rho_{ij} = 0\right) = \frac{v-2}{v-4}.$$

Hence, $E\left(\Delta_{nT} | \rho_{ij} = 0, \text{ for all } i \neq j\right)$ defined by (19) tends to 0 if $n^{2-\frac{v-4}{v-2}-d/2} \to 0$, or if $d > \frac{2v}{v-2}$. Assumption 1 of JR requires $E\left|\varepsilon_{it}\right|^{8+\epsilon} < K$, for some small positive ϵ, and for this to be satisfied in the case of t-distributed errors we need $v > 9$, which yields $d > 2$ when $v = 10$, requiring T to rise faster than n.

Finally, for the CD_{W+} test to have power it is also necessary to show that Δ_{nT} diverges in n and T sufficiently fast under alternative hypotheses of interest, namely spatial or network dependence. Later in the paper, we provide some Monte Carlo evidence on this issue, which indicates Δ_{nT} need not diverge sufficiently fast and can cause the CD_{W+} test to suffer from low power against spatial or network alternatives. Our Monte Carlo experiments also show that the issue of over-rejection of CD_{W+} when $n >> T$ prevails when the errors are chi-squared distributed and the moment condition in Assumption 1 of JR is met.

2.2 The bias-corrected CD test

As shown by JR, the main reason for the failure of the standard CD test in the case of the latent factor models lies in the fact that both the factors and their loadings are unobserved and need to be estimated, for example by PCA as in (8). Essentially the differences between $\hat{\boldsymbol{\gamma}}_i' \hat{\mathbf{f}}_t$ and $\boldsymbol{\gamma}_i' \mathbf{f}_t$ do not tend to zero at a sufficiently fast rate for the CD test to be valid, unless the latent factors are weak, namely unless $\alpha = max_j(\alpha_j) < 1/2$. Since the errors from estimation of $\boldsymbol{\gamma}_i' \mathbf{f}_t$ are included in the residuals e_{it}, the resultant CD statistic tends to over-state the degree of underlying error cross-sectional dependence. This problem also arises when the latent factors are proxied by cross section averages, as is the case when panel data models are estimated using correlated common effect (CCE) estimators proposed by Pesaran (2006), which we shall address below in Section 3.

We propose a bias-corrected CD test statistic, which we denote by CD^*, that *directly* corrects the asymptotic bias of the CD test using the estimates of the factor loadings and error variances. To obtain the expression for the bias we first write the CD statistic, defined by (10), equivalently as (established in Lemma 8 in the online supplement)

$$CD = \left(\sqrt{\frac{n}{n-1}}\right) \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left[\frac{\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{e_{it}}{\hat{\sigma}_{i,T}}\right)^2 - 1}{\sqrt{2}} \right]. \tag{20}$$

We also introduce the following analogue of CD

$$\widetilde{CD} = \left(\sqrt{\frac{n}{n-1}}\right) \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left[\frac{\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{e_{it}}{\omega_{i,T}}\right)^2 - 1}{\sqrt{2}} \right], \tag{21}$$

where $\omega_{i,T}^2 = T^{-1}\mathbf{u}_i'\mathbf{M}_F\mathbf{u}_i$. Also, using results established in Lemmas 2 and 9 in the online supplement, we have

$$\hat{\sigma}_{i,T}^2 = \omega_{i,T}^2 + O_p\left(\frac{1}{n}\right) + O_p\left(\frac{1}{T}\right),$$

and

$$CD - \widetilde{CD} = o_p(1). \tag{22}$$

Scaling the residuals by $\omega_{i,T}$ instead of $\hat{\sigma}_{i,T}$, we are able to establish a faster rate of convergence which in turn allows us to derive an expression for the asymptotic bias of CD statistic, considering that $\widetilde{CD}$ and CD are asymptotically equivalent.

Now to analyze the asymptotic properties of $\widetilde{CD}$, let $\boldsymbol{\delta}_{i,T} = \boldsymbol{\gamma}_i/\omega_{i,T}$ and $\hat{\boldsymbol{\delta}}_{i,T} = \hat{\boldsymbol{\gamma}}_i/\omega_{i,T}$, and note that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{e_{it}}{\omega_{i,T}} = \psi_{t,nT} - s_{t,nT} \tag{23}$$

where

$$\psi_{t,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{a_{i,nT}u_{it}}{\omega_{i,T}}, \ a_{i,nT} = 1 - \omega_{i,T}\boldsymbol{\varphi}_{nT}'\boldsymbol{\gamma}_i, \tag{24}$$

$$s_{t,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\boldsymbol{\varphi}_{nT}' \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right) u_{it} + \left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right)' \hat{\mathbf{f}}_t + \boldsymbol{\varphi}_{nT}' \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right) \boldsymbol{\gamma}_i'\mathbf{f}_t \right], \tag{25}$$

with $\boldsymbol{\varphi}_{nT} = n^{-1}\sum_{i=1}^{n} \boldsymbol{\delta}_{i,T}$. Using the above results, the $\widetilde{CD}$ statistic defined in (21) can then be decomposed as

$$\begin{aligned}
\widetilde{CD} &= \left(\sqrt{\frac{n}{n-1}}\right) \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{\left(\psi_{t,nT} - s_{t,nT}\right)^2 - 1}{\sqrt{2}} \\
&= \left(\sqrt{\frac{n}{n-1}}\right) \left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\frac{\psi_{t,nT}^2 - 1}{\sqrt{2}}\right) + \frac{1}{\sqrt{2}} \left(T^{-1/2} \sum_{t=1}^{T} s_{t,nT}^2\right) - \sqrt{2}\left(T^{-1/2} \sum_{t=1}^{T} \psi_{t,nT}s_{t,nT}\right) \right].
\end{aligned}$$

Under Assumptions 1-3 the last two terms of $\widetilde{CD}$ are shown in Lemma 4 in the online supplement to be asymptotically negligible, in the sense that they tend to zero in probability as $(n,T) \to \infty$, so long as $n/T \to \kappa$, and $0 < \kappa < \infty$. Hence, $\widetilde{CD} = z_{nT} + o_p(1)$, where $z_{nT} = T^{-1/2}\sum_{t=1}^{T}\left(\frac{\psi_{t,nT}^2 - 1}{\sqrt{2}}\right)$. Also, using (22) it can be shown $CD = z_{nT} + o_p(1)$ and furthermore we have

$$z_{nT} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\frac{\xi_{t,n}^2 - 1}{\sqrt{2}}\right) + o_p\left(1\right), \tag{26}$$

where

$$\xi_{t,n} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} a_{i,n}\varepsilon_{it}, \ a_{i,n} = 1 - \sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i, \tag{27}$$

$\boldsymbol{\varphi}_n = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\delta}_i$, and $\boldsymbol{\delta}_i = \boldsymbol{\gamma}_i/\sigma_i$, which is established in proof of Theorem 1 in the online supplement. Since $a_{i,n}$ are given constants, then $E\left(\xi_{t,n}\right) = 0$,

$$E\left(\xi_{t,n}^2\right) \equiv \omega_n^2 = \frac{1}{n}\sum_{i=1}^{n} a_{i,n}^2 = n^{-1}\sum_{i=1}^{n}\left(1-\sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i\right)^2, \tag{28}$$

and

$$Var\left(\xi_{t,n}^2\right) = 2\left(\frac{1}{n}\sum_{i=1}^{n} a_{i,n}^2\right)^2 - \kappa_2\left(\frac{1}{n^2}\sum_{i=1}^{n} a_{i,n}^4\right), \tag{29}$$

where $\kappa_2 = E\left(\varepsilon_{it}^4\right) - 3$. Clearly, when the errors are Gaussian then $E\left(\varepsilon_{it}^4\right) = 3$, the second term of $Var\left(\xi_{t,n}^2\right)$ defined by (29) is exactly zero. But even for non-Gaussain errors the second term of $Var\left(\xi_{t,n}^2\right)$ is negligible when n is sufficiently large. To see this note that

$$\frac{1}{n^2}\sum_{i=1}^{n} a_{i,n}^4 = \frac{1}{n^2}\sum_{i=1}^{n}(1-\sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i)^4 \leq \frac{K}{n},$$

where K is a positive constant irrespective of whether the underlying factor(s) are strong or weak. Then we can also compute the mean and the variance of z_{nT} as

$$E\left(z_{nT}\right) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\frac{\omega_n^2-1}{\sqrt{2}}\right) = \sqrt{\frac{T}{2}}\left(\omega_n^2-1\right),$$

$$Var\left(z_{nT}\right) = \frac{1}{T}\sum_{t=1}^{T} Var\left(\frac{\xi_{t,n}^2}{\sqrt{2}}\right) = \frac{Var\left(\xi_{t,n}^2\right)}{2}.$$

The above expressions for $E\left(z_{nT}\right)$ give the source of the asymptotic bias of CD as $E\left(z_{nT}\right)$ rises with $\sqrt{T}$, unless

$$\lim_{n\to\infty}\omega_n^2 = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n}\left(1-\sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i\right)^2 = 1.$$

A bias-corrected version of CD can be defined by

$$CD^*(\theta_n) = \frac{CD + \sqrt{\frac{T}{2}}\theta_n}{1-\theta_n}, \tag{30}$$

where

$$\theta_n = 1 - \frac{1}{n}\sum_{i=1}^{n} a_{i,n}^2, \tag{31}$$

with $a_{i,n}$ defined by (27). Because of condition (5) in Assumption 3, $\theta_n = 1$ is ruled out and therefore the existence of $CD^*(\theta_n)$ is guaranteed. The above results are summarized in the following theorem.

Theorem 1 *Consider the model in (2) and assume the number of strong factors included in the panel data model, m_0, is known. Also suppose Assumptions 1-3 hold. Under the null hypothesis of cross-sectional independence as $(n, T) \to \infty$, such that $n/T \to \kappa$, and $0 < \kappa < \infty$, $CD^*(\theta_n)$ defined by (30) has the limiting $N(0, 1)$ distribution.*

Remark 4 *It is clear from the definition of θ_n, given by (31), that it does not go to zero when there is at least one strong factor in the panel data model. And as shown in Section 3 of the online supplement in general $\theta_n = \ominus\left(n^{\alpha-1}\right)$, where $\alpha = max_{j=1,2,\ldots,m_0}(\alpha_j)$, with α_j representing the strength of the latent factor, f_{jt}, defined by (7). Thus, the relationship between CD and $CD^*(\theta_n)$ is essentially controlled by the maximum factor strength α. Also the main difference between CD and $CD^*(\theta_n)$ relates to the correction in the numerator of (30), the order of which is given by*

$$\sqrt{T}\theta_n = O\left(T^{1/2}n^{\alpha-1}\right).$$

Suppose now $T = \ominus\left(n^d\right)$ for some $d > 0$, then $\sqrt{T}\theta_n = \ominus\left(n^{d/2}n^{\alpha-1}\right) = \ominus\left(T^{\alpha+d/2-1}\right)$, and the bias correction becomes negligible if $\alpha < 1 - d/2$. Under the required relative expansion rates of n and T entertained in this paper, we need to set $d = 1$, and for this choice the bias correction term, $\sqrt{T}\theta_n$, becomes negligible if $\alpha < 1/2$, namely if all latent factors are weak. This result also establishes that the standard CD test is still valid if all the latent factors are weak, namely $\alpha < 1/2$, which confirms an earlier finding of Pesaran (2015a) regarding the implicit null of the standard CD test when $d = 1$.

Remark 5 *Although our mathematical derivations are based on results for standard factor models where the factors are assumed to be strong, as we shall see from the Monte Carlo results reported below, the proposed test will be applicable even if some of the latent factors happen to be weak or semi-weak with $\alpha < 2/3$. This is because when a factor is weak it does not matter if its estimation by PCA is not consistent at the standard rate of $\delta_{nT} = \min(n^{1/2}, T^{1/2})$, since a weak factor only affects a few of the units and its inclusion or exclusion from the analysis has no material impact on the CD^* test as n and $T \to \infty$. In effect, in the mathematical derivations it is sufficient to consider strong factors, and absorb the weak factors in the error term. Additional theoretical derivations are required when some of the latent factors are semi-strong, namely when α is close to unity. Such an extension is beyond the scope of the present paper.*

The bias-corrected test statistic, $CD^*(\theta_n)$, depends on the unknown parameter, θ_n, which can be estimated by

$$\hat{\theta}_{nT} = 1 - \frac{1}{n}\sum_{i=1}^{n}\hat{a}_{i,nT}^2 \tag{32}$$

where

$$\hat{a}_{i,nT} = 1 - \hat{\sigma}_{i,T}\left(\hat{\boldsymbol{\varphi}}_{nT}'\hat{\boldsymbol{\gamma}}_i\right), \quad \hat{\boldsymbol{\varphi}}_{nT} = \frac{1}{n}\sum_{i=1}^{n}\hat{\boldsymbol{\delta}}_{i,nT}, \tag{33}$$

and $\hat{\boldsymbol{\delta}}_{i,nT} = \hat{\boldsymbol{\gamma}}_i/\hat{\sigma}_{i,T}$. The following corollary establishes the probability order of the difference between $\hat{\theta}_{nT}$ and θ_n.

Corollary 1 *Consider the bias correction term θ_n in the CD^* statistic given by (31) and its estimator $\hat{\theta}_{nT}$ given by (32). Suppose Assumptions 1-2 hold. Then for $(n, T) \to \infty$, such that $n/T \to \kappa$, where $0 < \kappa < \infty$, we have*

$$\sqrt{T}\left(\hat{\theta}_{nT} - \theta_n\right) = o_p(1). \tag{34}$$

It then readily follows that $CD^*\left(\hat{\theta}_{nT}\right) = CD^*(\theta_n) + o_p(1)$, where

$$CD^*\left(\hat{\theta}_{nT}\right) = CD^* = \frac{CD + \sqrt{\frac{T}{2}}\hat{\theta}_{nT}}{1 - \hat{\theta}_{nT}}. \tag{35}$$

We refer to $CD^*\left(\hat{\theta}_{nT}\right)$, or CD^* for short, as the bias-corrected CD statistic, and the test based on it as the CD^* test. The main result of the paper for pure latent factor models is summarized in the following theorem.

Theorem 2 *Under Assumptions 1-3, CD^* defined by (35) has the limiting $N(0,1)$ distribution, as $(n,T) \to \infty$, such that $n/T \to \kappa$, and $0 < \kappa < \infty$.*

Remark 6 *Estimation of $\hat{\theta}_{nT}$ requires the investigator to decide on the number of latent factors, say $\hat{m}$, when computing the CD^* statistic. Suppose that m_0 denotes the number of strong factors. Then if $\hat{m} > m_0$, the additional assumed number of factors, $\hat{m} - m_0$, must be weak by construction and $CD^* \to_d N(0,1)$ under the null hypothesis. The idea of setting $\hat{m}$ above m_0 was suggested in Pesaran et al. (2013) when testing for unit roots in the context of panels with multi-factor error structure, and was later considered formally by Moon and Weidner (2015) who established that the panel estimators based on m_0 and $\hat{m}$ factors are asymptotically equivalent so long as $\hat{m} > m_0$. We conjecture that the same applies to CD^*, and recommend setting $\hat{m} \geq m_0$. There is no need to have a precise estimate of m_0 which is often unattainable especially when some of the latent factors are semi-strong. In practice the assumed number of factors can be increased to ensure that CD^* test does not result in spurious rejection.*

Remark 7 *Despite the robustness of CD^* test to the choice of $\hat{m}$, so long as $\hat{m} \geq m_0$, it cannot be used to test if the number of latent factor selected is correct. This is because it cannot distinguish whether the cross-sectional dependence is caused by the missing latent factors or other forms of cross-sectional dependence such as spatial error correlations. Our analysis does not contribute to the problem of estimating m_0 addressed in the literature either based on information criterion of Bai and Ng (2002) or eigenvalue ratio test of Ahn and Horenstein (2013).*

3 CD* tests for error cross-sectional dependence in panel data models with interactive effects

Consider now the general panel regression model (1) which can be rewritten as

$$y_{it} = \boldsymbol{\alpha}_i'\mathbf{d}_t + \boldsymbol{\beta}_i'\mathbf{x}_{it} + v_{it},\ v_{it} = \boldsymbol{\gamma}_i'\mathbf{f}_t + u_{it}. \tag{36}$$

To test the cross-sectional independence of error term in a mixed factor model as (36), we need to estimate coefficients $(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)$. When the regressor $\mathbf{x}_{it}$ is independent from both factor structure and error term, a simple least squares regression of y_{it} on $(1, \mathbf{x}_{it})$ for each i would be sufficient. However, in a more general scenario, $\mathbf{x}_{it}$ can be correlated with factor structure. To study this scenario, we adopt the large heterogeneous panel data models discussed in Pesaran (2006), so that the time varying regressor $\mathbf{x}_{it}$ is assumed to be generated as

$$\mathbf{x}_{it} = \mathbf{A}_i'\mathbf{d}_t + \boldsymbol{\Gamma}_i'\mathbf{f}_t + \boldsymbol{\varepsilon}_{xit},$$

where $\mathbf{A}_i$ and $\mathbf{\Gamma}_i$ are $k_d \times k_x$ and $m_0 \times k_x$ factor loading matrices and $\boldsymbol{\varepsilon}_{xit}$ are the specific components of $\mathbf{x}_{it}$, distributed independently of the common effects and across i, but assumed to follow general covariance stationary process. Then in addition to Assumptions 1-3, we make the following assumptions:

Assumption 4 *(a) The $k_d \times 1$ vector $\mathbf{d}_t$ is a covariance stationary process, with absolute summable autocovariance and $\mathbf{d}_t$ is distributed independently of $\mathbf{f}_{t'}$, for all t and t', such that $T^{-1}\mathbf{D}'\mathbf{F} = O_p\left(T^{-1/2}\right)$, where $\mathbf{D} = (\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_T)'$ and $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_T)'$ are matrices of observations on $\mathbf{d}_t$ and $\mathbf{f}_t$. (b) $(\mathbf{d}_t, \mathbf{f}_t)$ is distributed independently of $u_{s,i}$ and $\boldsymbol{\varepsilon}_{xis}$ for all i, t, s.*

Assumption 5 *The unobserved factor loadings $\mathbf{\Gamma}_i$ are bounded, i.e. $\|\mathbf{\Gamma}_i\|_2 < K$ for all i.*

Assumption 6 *The individual-specific errors u_{it} and $\boldsymbol{\varepsilon}_{xi,t'}$ are distributed independently for all i, j, t and t', and $\boldsymbol{\varepsilon}_{xit}$ follows the linear stationary process $\boldsymbol{\varepsilon}_{xit} = \sum_{l=0}^{\infty} \mathbf{S}_{il}\eta_{xi,t-l}$, where for each i, η_{xit} is a $k_x \times 1$ vector of serially uncorrelated random variables with mean zero, the variance matrix $\mathbf{I}_{k_x}$, and finite fourth-order cumulants. For each i, the coefficient matrices $\mathbf{S}_{il}$ satisfy the condition*

$$Var\left(\boldsymbol{\varepsilon}_{xit}\right) = \sum_{l=0}^{\infty} \mathbf{S}_{il}\mathbf{S}_{il}' = \mathbf{\Sigma}_{xi},$$

where $\mathbf{\Sigma}_{xi}$ is a positive definite matrix, such that $\sup_i \|\mathbf{\Sigma}_{xi}\|_2 < K$.

Assumption 7 *Let $\tilde{\mathbf{\Gamma}} = E\left(\boldsymbol{\gamma}_i, \mathbf{\Gamma}_i\right)$. We assume that Rank $\left(\tilde{\mathbf{\Gamma}}\right) = m_0$.*

Assumption 8 *Consider the cross-sectional averages of the individual-specific variables, $\mathbf{z}_{it} = \left(y_{it}, \mathbf{x}_{it}'\right)'$ defined by $\bar{\mathbf{z}}_t = n^{-1}\sum_{i=1}^{n} \mathbf{z}_{it}$, and let $\bar{\mathbf{M}} = \mathbf{I}_T - \bar{\mathbf{H}}\left(\bar{\mathbf{H}}'\bar{\mathbf{H}}\right)^{-1}\bar{\mathbf{H}}'$, and $\mathbf{M}_g = \mathbf{I}_T - \mathbf{G}\left(\mathbf{G}'\mathbf{G}\right)^{-1}\mathbf{G}'$, where $\bar{\mathbf{H}} = \left(\mathbf{D}, \bar{\mathbf{Z}}\right)$, $\mathbf{G} = (\mathbf{D}, \mathbf{F})$, and $\bar{\mathbf{Z}} = (\bar{\mathbf{z}}_1, \bar{\mathbf{z}}_2, \ldots, \bar{\mathbf{z}}_T)$ is the $T \times (k_x + 1)$ matrix of observations on the cross-sectional averages. Let $\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, ..., \mathbf{x}_{iT})'$, then the $k \times k$ matrices $\hat{\mathbf{\Psi}}_{i,T} = T^{-1}\mathbf{X}_i'\bar{\mathbf{M}}\mathbf{X}_i$ and $\mathbf{\Psi}_{ig} = T^{-1}\mathbf{X}_i'\mathbf{M}_g\mathbf{X}_i$ are non-singular, and $\hat{\mathbf{\Psi}}_{i,T}^{-1}$ and $\mathbf{\Psi}_{ig}^{-1}$ have finite second-order moments for all i.*

Remark 8 *The above assumptions are standard in the panel data models with multi-factor error structure. See, for example, Pesaran (2006). But in our setup under Assumption 2 we require the error term, u_{it}, to be serially uncorrelated, since our focus is on testing u_{it} for cross-sectional dependence, and this assumption is needed for asymptotic normality of the bias-corrected CD test. Later in Section 4, we will consider models with serially correlated errors and show that the bias-corrected CD test remains valid. Nevertheless, we allow ε_{xit}, the errors in the $\mathbf{x}_{it}$ equations to be serially correlated. Assumption 4 separates the observed and the latent factors, as in Assumption 11 of Pesaran and Tosetti (2011). This assumption is required to obtain the probability order of estimated residuals needed for computation of CD^* statistic.*

To estimate v_{it} we first filter out the effects of observed covariates using the CCE estimators proposed in Pesaran (2006), namely for each i we estimate β_i by

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

and following Pesaran and Tosetti (2011), estimate $\boldsymbol{\alpha}_i$ by

$$\hat{\boldsymbol{\alpha}}_{CCE,i} = \left(\mathbf{D}'\mathbf{D}\right)^{-1}\mathbf{D}'\left(\mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_{CCE,i}\right).$$

Then we have the following estimator of v_{it}

$$\hat{v}_{it} = y_{it} - \hat{\boldsymbol{\alpha}}'_{CCE,i}\mathbf{d}_t - \hat{\boldsymbol{\beta}}'_{CCE,i}\mathbf{x}_{it}.$$

Using results in Pesaran and Tosetti (2011) (p. 189) it follows that under Assumptions 1-8

$$\hat{v}_{it} = v_{it} + O_p\left(\frac{1}{n}\right) + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{nT}}\right). \tag{37}$$

Note when $\boldsymbol{\alpha}_i = \mathbf{0}$ and $\boldsymbol{\beta}_i = \mathbf{0}$, (36) reduces to the the pure latent factor model, (2), where PCA can be applied to v_{it} directly. In the case of panel regressions $\hat{v}_{it}$ can be used instead of v_{it} to compute the bias-corrected CD statistic given by (35). The errors involved will become asymptotically negligible in view of the fast rate of convergence of $\hat{v}_{it}$ to v_{it}, uniformly for each i and t. Specifically, as in the case of the pure latent factor model, we first compute m_0 PCs of $\{\hat{v}_{it};\ i = 1, \ldots, n;\ \text{and } t = 1, \ldots, T\}$ and the associated factor loadings, $(\hat{\boldsymbol{\gamma}}_i, \hat{\mathbf{f}}_t)$, subject to the normalization $n^{-1}\sum_{i=1}^{n}\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}'_i = \mathbf{I}_{m_0}$. The residuals

$$e_{it} = \hat{v}_{it} - \hat{\boldsymbol{\gamma}}'_i\hat{\mathbf{f}}_t, \tag{38}$$

for $i = 1, \ldots, n$ and $t = 1, \ldots, T$ are then used to compute the standard CD statistic, which is then bias-corrected as before using (35).

Theorem 3 *Consider the panel data model (36) and suppose the true factor number m_0 is known. Also suppose Assumptions 1-8 hold. Then as $(n, T) \to \infty$, such that $n/T \to \kappa$, where $0 < \kappa < \infty$, CD^* has the limiting $N(0, 1)$ distribution.*

Remark 9 *As in the case of the pure latent factor model the CD^* test will be valid so long as the number of estimated factors is at least as large of m_0.*

4 CD^* test for models with serially correlated errors

As shown by Baltagi et al. (2016), when the errors u_{it} in (1) are serially correlated the variance of standard CD test statistic is not unity (even asymptotically) and the test is no longer valid. The same also applies to the CD^* test. To deal with this problem, we propose two solutions which involve different ways of adjusting the CD^* test so that it will become applicable to panels with serially correlated errors. The first method closely follows the variance adjustment proposed by Baltagi et al. (2016), in which CD^* is scaled by ϖ where

$$\varpi^2 = \frac{2T}{n(n-1)}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\tilde{\mathbf{e}}'_{i,T}\left(\tilde{\mathbf{e}}_{j,T} - \tilde{\mathbf{e}}_{(ij),T}\right)\tilde{\mathbf{e}}'_{j,T}\left(\tilde{\mathbf{e}}_{i,T} - \tilde{\mathbf{e}}_{(ij),T}\right), \tag{39}$$

with $\tilde{\mathbf{e}}_{i,T} = (\tilde{e}_{i1,T}, \tilde{e}_{i2,T}, \ldots, \tilde{e}_{iT,T})'$, $\tilde{e}_{it,T}$ defined in (11) and

$$\tilde{\mathbf{e}}_{(ij),T} = \frac{1}{n-2}\sum_{1 \le \tau \ne i,j \le n}\tilde{\mathbf{e}}_{\tau,T}.$$

The expression in (39) is the equivalent to that provided in Theorem 3 of Baltagi et al. (2016) but the factor of 2 in (39) is missing in their paper. The same adjustment is also applied to the CD_{W+} test to allow for serially correlated errors.

Alternatively, following Pesaran (2004), we first transform the panel regressions to eliminate the error serial correlation and then apply the CD^* test to the residuals of the transformed model. This is possible so long as the error serial correlation can be approximated by a finite order stationary autoregressive process. As a simple illustration consider the pure factor model

$$y_{it} = \gamma_i f_t + u_{it},$$

in which factor f_t and loading γ_i are both latent, and the errors u_{it} are generated as $AR(1)$ processes:

$$u_{it} = \rho_i u_{it-1} + \epsilon_{it},$$

where ρ_i is the autoregression coefficient and ϵ_{it} is serially independent, as well as being distributed independently of $f_{t'}$ for all i and $t, t' = 1, 2, \ldots, T$. Testing the cross-sectional independence of u_{it} is equivalent to testing the cross-sectional independence of ϵ_{it} in the following autoregressive distributed lag (ARDL) representation of y_{it}

$$y_{it} = \rho_i y_{i,t-1} + \gamma_i f_t - \rho_i \gamma_i f_{t-1} + \epsilon_{it},$$

which can be written equivalently as a multi-factor AR panel regression

$$y_{it} = \rho_i y_{i,t-1} + \mathring{\boldsymbol{\gamma}}_i' \mathring{\mathbf{f}}_t + \epsilon_{it}, \tag{40}$$

where $\mathring{\mathbf{f}}_t = (f_t, f_{t-1})'$, and $\mathring{\boldsymbol{\gamma}}_i = (\gamma_i, -\rho_i \gamma_i)'$. Since $y_{i,t-1}$ is weakly exogenous, the transformed model satisfies the setup of panel data model (36) with $\mathring{\mathbf{f}}_t$ viewed as a vector of latent variables with the associated factor loadings, $\mathring{\boldsymbol{\gamma}}_i$. It therefore follows that the CD^* test can now be applied to test the cross-sectional independence of ϵ_{it} in (40). We refer to this test as $ARDL$ adjusted CD^* test.

The same approach can also be used for panels with observed covariates. In general, testing cross-sectional independence of u_{it} in model (1) is equivalent to testing the cross-sectional independence of ϵ_{it} in

$$y_{it} = \sum_{s=0}^{S} \boldsymbol{\alpha}_{i,s}' \mathbf{d}_{t-s} + \sum_{s=1}^{S} \rho_{i,s} y_{it-s} + \sum_{s=0}^{S} \boldsymbol{\beta}_{i,s}' \mathbf{x}_{it-s} + \mathbf{g}_i' \mathbf{h}_t + \epsilon_{it}, \tag{41}$$

where $\mathbf{h}_t$ is an extended set of latent factors (that encompass $\mathbf{f}_t$), and $\mathbf{g}_i$ are the associated factor loadings. The number of lags S is determined by the order of the AR specification assumed for u_{it} in (2).

The variance adjustment is simpler to implement but it requires theoretical justification in the context of panel data models with latent factors. The ARDL adjustment is theoretically justified so long as the underlying errors follow finite order AR processes. As we shall see both approaches work well in dealing with serially correlated errors, at least in the context of the limited MC designs that we are considering. Clearly, further theoretical and Monte Carlo investigations are needed for a better understanding of the relative merits of the two approaches.

5 Small sample properties of CD^* and CD_{W+} tests

5.1 Data generating process

We consider the following data generating process

$$y_{it} = \text{a}_i + \sigma_i \left(\beta_{i1} d_t + \beta_{i2} x_{it} + m_0^{-1/2} \boldsymbol{\gamma}_i' \mathbf{f}_t + \varepsilon_{it} \right), \; i = 1, 2, ..., n; t = 1, 2, ..., T, \tag{42}$$

where a_i is a unit-specific effect, d_t is the observed common factor, x_{it} is the observed regressor that varies across i and t, $\mathbf{f}_t$ is the $m_0 \times 1$ vector of unobserved factors, $\boldsymbol{\gamma}_i$ is the vector of associated factor loadings, and ε_{it} are the idiosyncratic errors. The scalar constants, $\sigma_i > 0$, are generated as $\sigma_i^2 = 0.5 + \frac{1}{2}(s_i^2 - 1)$, with $s_i^2 \sim IID\chi^2(2)$, which ensures that $E(\sigma_i^2) = 1$.

5.1.1 DGP under the null hypothesis

Under the null hypothesis, we consider both serially independent errors and serially correlated errors, which are generated by both Gaussian and non-Gaussian distributions:

- Serially independent errors: Gaussian errors, $\varepsilon_{it} \sim IIDN(0, 1)$; chi-squared distributed errors, $\varepsilon_{it} \sim IID\left(\frac{\chi^2(2)-2}{2}\right)$.
- Serially correlated errors: $\varepsilon_{it} = \rho_\varepsilon \varepsilon_{it-1} + \sqrt{1-\rho_\varepsilon^2} e_{\varepsilon it}$, for $i = 1, 2, \ldots, n$ and $t = 1, 2, \ldots, T$, where $\rho_\varepsilon = 0.5$ and $e_{\varepsilon it}$ are generated as Gaussian errors, $e_{\varepsilon it} \sim IIDN(0, 1)$, or chi-squared distributed errors, $e_{\varepsilon it} \sim IID\left(\frac{\chi^2(2)-2}{2}\right)$.

The focus of the experiments is on testing the null hypothesis that ε_{it} are cross-sectional independent, whilst allowing for the presence of m_0 unobserved factors, $\mathbf{f}_t = (f_{1t}, f_{2t}, ..., f_{m_0t})'$. We consider $m_0 = 1$ and $m_0 = 2$, and generate the factor loadings $\boldsymbol{\gamma}_i = (\gamma_{i1}, \gamma_{i2})'$ as:

$$\begin{aligned}
\gamma_{i1} &\sim IIDN\,(0.5, 0.5) \;\text{ for } i = 1, 2, \ldots, [n^{\alpha_1}], \\
\gamma_{i2} &\sim IIDN\,(1, 1) \;\text{ for } i = 1, 2, \ldots, [n^{\alpha_2}], \\
\gamma_{ij} &= 0 \text{ for } i = [n^{\alpha_j}] + 1, [n^{\alpha_j}] + 2, ..., n, \text{ and } j = 1, 2.
\end{aligned}$$

In the one-factor case ($m_0 = 1$), we only include f_{1t} as the latent factor and denote its factor strength by α. Three values of α are considered, namely $\alpha = 1, 2/3, 1/2$, respectively representing strong, semi-strong and weak factor. Similarly, in the two-factor case ($m_0 = 2$), we include both f_{1t} and f_{2t} as the latent factors and consider the following combinations of factor strengths, i.e. $(\alpha_1, \alpha_2) = [(1, 1), (1, 2/3), (2/3, 1/2)]$. The intercepts a_i are generated as $IIDN(1, 2)$ and fixed thereafter. The observed common factor is generated as $d_t = \rho_d d_{t-1} + \sqrt{1-\rho_d^2} v_{dt}$, with $\rho_d = 0.8$, and $v_{dt} \sim IIDN(0, 1)$, thus ensuring that $E(d_t) = 0$ and $Var(d_t) = 1$. The observed unit-specific regressors, x_{it}, for $i = 1, 2, .., n$ are generated to have non-zero correlations with the unobserved factors:

$$x_{it} = \gamma_{xi1} f_{1t} + \gamma_{xi2} f_{2t} + e_{xit}, \tag{43}$$

where $f_{jt} = r_j f_{j,t-1} + \sqrt{1-r_j^2} v_{jt}$, with $r_j = 0.9$ and $v_{jt} \sim IID\left(\frac{\chi^2(2)-2}{2}\right)$, for $j = 1, 2$. The factor loadings in (43) are generated as $\gamma_{xi1} \sim IIDU\,(0.25, 0.75)$ and $\gamma_{xi2} \sim IIDU\,(0.1, 0.5)$.

The error term of (43) is generated as $e_{xit} = \rho_i e_{xi,t-1} + \sqrt{1-\rho_i^2} v_{xit}$, where $\rho_i \sim IIDU(0, 0.95)$ and $v_{xit} \sim IIDN(0, 1)$.

We will examine the small sample properties of the CD and the bias-corrected CD tests for both the pure latent factor model and for the panel regression model which also includes observed covariates.

- In the case of the pure latent factor model we set $\beta_{i1} = \beta_{i2} = 0$.
- In the case of the panel regression model with latent factors, we allow for heterogeneous slopes and generate the slopes of observed covariates, d_t and x_{it}, as $\beta_{i1} \sim IIDN(\mu_{\beta 1}, \sigma^2_{\beta 1})$, and $\beta_{i2} \sim IIDN(\mu_{\beta 2}, \sigma^2_{\beta 2})$ where $\mu_{\beta 1} = \mu_{\beta 2} = 0.5$ and $\sigma^2_{\beta 1} = \sigma^2_{\beta 2} = 0.25$, respectively.

As our theoretical results show the null distribution of the CD and bias-corrected CD tests do not depend on a_i, β_{i1} and β_{i2}, it is therefore innocuous what values are chosen for these parameters. Moreover, the average fit of the panel is controlled in terms of the limiting value of the pooled R-squared defined by

$$PR^2_{nT} = 1 - \frac{(nT)^{-1}\sum_{i=1}^n \sum_{t=1}^T \sigma_i^2 E\left(\varepsilon_{it}^2\right)}{(nT)^{-1}\sum_{i=1}^n \sum_{t=1}^T Var\left(y_{it}\right)}. \tag{44}$$

Since the underlying processes, (42) and (43), are stationary and $E\left(\varepsilon_{it}^2\right) = 1$, we have

$$\lim_{T\to\infty} PR^2_{nT} = PR^2_n = \frac{n^{-1}\sum_{i=1}^n \sigma_i^2 \left[\beta_{i1}^2 + \beta_{i2}^2 Var\left(x_{it}\right) + m_0^{-1}\boldsymbol{\gamma}_i'\boldsymbol{\gamma}_i + 2Cov\left(x_{it}, \boldsymbol{\gamma}_i'\mathbf{f}_t\right)\right]}{n^{-1}\sum_{i=1}^n Var\left(y_{it}\right)}.$$

where $\boldsymbol{\gamma}_i = (\gamma_{i1}, \gamma_{i2})'$, $Var\left(x_{it}\right) = \boldsymbol{\gamma}_{xi}'\boldsymbol{\gamma}_{xi} + 1$, $Cov\left(x_{it}, \boldsymbol{\gamma}_i'\mathbf{f}_t\right) = \boldsymbol{\gamma}_{xi}'\boldsymbol{\gamma}_i$, $\boldsymbol{\gamma}_{xi} = (\gamma_{xi1}, \gamma_{xi2})'$, and

$$Var\left(y_{it}\right) = \sigma_i^2 \left[\beta_{i1}^2 + \beta_{i2}^2 Var\left(x_{it}\right) + m_0^{-1}\boldsymbol{\gamma}_i'\boldsymbol{\gamma}_i + 2m_0^{-1/2}Cov\left(x_{it}, \boldsymbol{\gamma}_i'\mathbf{f}_t\right) + 1\right].$$

Also since σ_i^2 and β_{ij} are independently distributed and $E(\sigma_i^2) = 1$, it then readily follows that $\lim_{n\to\infty} PR^2_n = \eta^2/(1+\eta^2)$, where

$$\eta^2 = \mu^2_{\beta 1} + \sigma^2_{\beta 1} + \left(\mu^2_{\beta 2} + \sigma^2_{\beta 2}\right)\left[1 + E\left(\boldsymbol{\gamma}_{xi}'\boldsymbol{\gamma}_{xi}\right)\right] + \frac{2\mu_{\beta 2}E\left(\boldsymbol{\gamma}_{xi}'\boldsymbol{\gamma}_i\right)}{\sqrt{m_0}} + \frac{E\left(\boldsymbol{\gamma}_i'\boldsymbol{\gamma}_i\right)}{m_0}.$$

By controlling the value of η^2 across the experiments we ensure that the pooled R^2 in large samples will be fixed, regardless of value of σ_i. In particular, in the case of the pure latent model, we have $\eta^2 = m_0^{-1}E\left(\boldsymbol{\gamma}_i'\boldsymbol{\gamma}_i\right) = O\left(n^{\alpha-1}\right)$ where $\alpha = max(\alpha_1, \alpha_2)$.

5.1.2 DGP under the alternative hypothesis

We consider a spatial alternative representation for errors, and generate $\boldsymbol{\varepsilon}_{\circ t} = (\varepsilon_{1t}, \varepsilon_{2t}, ..., \varepsilon_{nt})'$ according to the first order spatial autoregressive process, $\boldsymbol{\varepsilon}_{\circ t} = c\left(\mathbf{I}_n - \rho \boldsymbol{W}\right)^{-1}\boldsymbol{\zeta}_{\circ t}$, where $\boldsymbol{W} = (w_{ij})$, and $\boldsymbol{\zeta}_{\circ t} = (\zeta_{1t}, \zeta_{2t}, \ldots, \zeta_{nt})'$. Similarly to the DGP under the null hypothesis, for the serially uncorrelated errors, we consider both serially independent errors and serially correlated errors, which are generated by both Gaussian and non-Gaussian distributions. For the case where the errors are serially independent we generate them as $\zeta_{it} \sim IIDN\left(0, 1\right)$ or $\zeta_{it} \sim IID[\frac{\chi^2(2)-2}{2}]$. While for the case of serially correlated errors, ζ_{it} are generated as $\zeta_{it} = \rho_\zeta \zeta_{it-1} +$

$\sqrt{1-\rho_{\zeta}^{2}}e_{\zeta it}$, for $i=1,2,\ldots,n$ and $t=1,2,\ldots,T$, where $\rho_{\zeta}=0.5$ and $e_{\zeta it}\sim IIDN\left(0,1\right)$ or $e_{\zeta it}\sim IID[\frac{\chi^{2}(2)-2}{2}]$, covering both Gaussian and non-Gaussian error distributions. For the spatial weights w_{ij}, we first set $w_{ij}^{0}=1$ if $j=i-2,i-1,i+1,i+2$, and zero otherwise. We then row normalize the weights such that $w_{ij}=\left(\sum_{j=1}^{n}w_{ij}^{0}\right)^{-1}w_{ij}^{0}$. We also set $c^{2}=n/\text{tr}\left[\left(\mathbf{I}_{n}-\rho\boldsymbol{W}\right)^{-1}\left(\mathbf{I}_{n}-\rho\boldsymbol{W}\right)^{\prime-1}\right]$, which ensures that $n^{-1}\sum_{i=1}^{n}Var(\varepsilon_{it})=1$, for all values of ρ.

5.2 CD, CD^{*} and CD_{W+} tests

All experiments are carried out for $n=100,200,500,1000$ and $T=100,200,500$ and the number of replications is set to $2,000$. Firstly we consider the DGPs with serially independent errors. For the pure latent factor models, we compute the filtered residuals as $\hat{v}_{it}=y_{it}-\hat{\text{a}}_{i}$, where $\hat{\text{a}}_{i}=T^{-1}\sum_{t=1}^{T}y_{it}$. For the panel regressions with latent factors, the filtered residuals are computed as

$$\hat{v}_{it}=y_{it}-\hat{\text{a}}_{CCE,i}-\hat{\beta}_{CCE,i1}d_{t}-\hat{\beta}_{CCE,i2}x_{it}, \tag{45}$$

where $\left(\hat{\text{a}}_{CCE,i},\hat{\beta}_{CCE,i1},\hat{\beta}_{CCE,i2}\right)$ is the CCE estimator of a_{i}, β_{i1} and β_{i2}, as set out in Pesaran (2006). The CCE estimators are consistent as the relevant rank condition is met, which requires that $m_{0}\leq 1+k=2$ and is satisfied in the case of our Monte Carlo experiments. Then we will compute the first $\hat{m}$ PCs $\{\hat{v}_{it};\ i=1,2,\ldots,n;\text{ and }t=1,2,\ldots,T\}$ and the associated factor loadings, namely $(\hat{\boldsymbol{\gamma}}_{i},\hat{\mathbf{f}}_{t})$, subject to the normalization, $n^{-1}\sum_{i=1}^{n}\hat{\boldsymbol{\gamma}}_{i}\hat{\boldsymbol{\gamma}}_{i}^{\prime}=\mathbf{I}_{\hat{m}}$. Finally the residuals, to be used in the computation of three CD test statistics, are computed as $e_{it}=\hat{v}_{it}-\hat{\boldsymbol{\gamma}}_{i}\hat{\mathbf{f}}_{t}$, for $i=1,2,\ldots,n$ and $t=1,2,\ldots,T$.

As discussed in Section 4, CD^{*} is not valid when the errors are serially correlated. In the simulations, we apply the variance and ARDL adjustments to CD, CD_{W+}, and CD^{*}. The variance adjusted versions are computed by scaling the original statistics by the standard deviation of the CD statistics using the expression in (39) with $e_{it}=\hat{v}_{it}-\hat{\boldsymbol{\gamma}}_{i}\hat{\mathbf{f}}_{t}$. The ARDL adjusted versions of CD, CD^{*}, and CD_{W+}, are computed using the residuals from the following dynamic panel data model with latent factors,

$$y_{it}=a_{i}+\sum_{s=1}^{S}\rho_{is}y_{i,t-s}+\sum_{s=0}^{S}\beta_{i1,s}d_{t-s}+\sum_{s=0}^{S}\beta_{i2,s}x_{i,t-s}+\mathbf{g}_{i}^{\prime}\mathbf{h}_{t}+\epsilon_{it}. \tag{46}$$

In the simulations we set $S=1$, but higher order values can also be considered. The number of latent factors in $\mathbf{h}_{t}$ depends on S and is given by $m_{h}=(S+1)m_{0}$. Accordingly, the number of PCs, $\hat{m}$, should satisfy $\hat{m}\geq(S+1)m_{0}$. In the simulations if $S=0$, we consider $\hat{m}=1$ and 2 if $m_{0}=1$, and $\hat{m}=2$ and 4 if $m_{0}=2$. But if $S=1$ we consider $\hat{m}=2$ and 4 if $m_{0}=1$, and $\hat{m}=4$ and 6 if $m_{0}=2$. Seen from this perspective, the variance adjustment is preferable since it does not require specifying the serial dependence of the error process.

5.3 Simulation results

We first report the simulation results for the DGPs with normally distributed errors, under which the correction term of JR test, namely Δ_{nT} in (13), tends to zero sufficiently fast. Next, we report simulation results for the DGPs with chi-squared distributed errors, allowing us to investigate the robustness of the JR test and our proposed bias-corrected CD test to departures from Gaussianity. Then we report simulation results for the DGPs with serially correlated errors, using the variance and ARDL adjusted CD tests discussed in Section 4. Finally, we consider spatial alternatives for the errors, by setting $\rho = 0.25$. As to be expected power rises very quickly as ρ is raised above 0 ($\rho = 0$ representing the null), and additional simulation results for values of $\rho > 0.25$ do not seem to add much to our investigation.

5.3.1 Serially independent errors: normally distributed errors

The simulation results for the DGPs with the errors following Gaussian distribution are shown in Tables 1-4 in the online supplement. Table 1 reports the test results for the pure single factor models. The top panel gives the results for the case where the number of selected PCs, denoted by $\hat{m}$, is the same as the true number of factors, m_0, while the bottom panel reports the results when $\hat{m} = 2$. As to be expected the standard CD test over-rejects when the factor is strong, namely when $a = 1$. By comparison, the rejection frequencies of both CD^* and CD_{W+} tests under null ($\rho = 0$) are generally around the nominal size of 5 per cent. Under the alternative (when $\rho = 0.25$), the CD^* has satisfactory power properties with significantly high rejection frequencies even when the sample size is small. But CD_{W+} test performs quite poorly under spatial alternatives, especially when T is small.

Table 2 in the online supplement summarizes the size and power results for the pure factor model with $m_0 = 2$, and reports the results when $\hat{m}$(the selected number PCs) is set to 2 (the top panel) and 4 (the bottom panel). The results are qualitatively similar to the ones reported for the single factor model. The CD test over-rejects if at least one of the factors is strong, and the empirical sizes of CD^* and CD_{W+} tests are close to their nominal value of 5 per cent, although we now observe some mild over-rejection when $n = 100$ and the selected number of PCs is 4. In terms of power, the CD^* performs well, although there is some loss of power as the number of factors and selected PCs rise. Similarly, the power of the CD_{W^+} test is now even lower and quite close to 5 per cent when $T < 500$ even if the number of PCs is set to $m_0 = 2$.

Turning to panel regressions with latent factors estimated by CCE, the associated simulation results are summarized in Tables 3 and 4 in the online supplement. As can be seen, the results are very close to the ones reported in Tables 1 and 2 for the pure factor model, and are in line with the asymptotic result in (37) that underlie the use of CCE approach to filter out the effects of observed covariates, as well as latent factors.

5.3.2 Serially independent errors: chi-squared distributed errors

The simulation results for the DGPs with chi-squared errors are provided in Tables 5 to 8 in the online supplement. For standard CD test and its biased-corrected version, CD^*, as shown in Tables 5 and 6, the results are very similar to the ones with Gaussian errors, suggesting that CD^* test is likely to be robust to departures from Gaussianity. As with the experiments with Gaussian errors, the standard CD test continues to over-reject unless $\alpha < 2/3$, and CD^* has the correct size for all n and T combinations, except when the number of selected PCs is large relative to m_0, and $T = 100$. The main difference between the results with and without Gaussian errors is the tendency for the CD_{W^+} test to over-reject when $n > T$, which seems to be a universal feature of this test and holds for *all* choices of m_0 and the number of selected PCs, irrespective of whether the factors are strong or weak. As we discussed in Section 2.1, this could be due to the screening component of CD_{W+} not tending to zero sufficiently fast with

n and T. Furthermore, the CD^* test continues to have satisfactory power, but CD_{W+} clearly lacks power against spatial or network alternatives that are of primary interest.

Similar results are obtained for panel regressions with latent factors, summarized in Tables 7 and 8 in the online supplement.

5.3.3 Serially correlated errors

The results for the DGPs with serially correlated errors are summarized in Tables 9-24 in the online supplement. Tables 9 to 16 give the simulation results for the variance adjusted CD tests, whilst Tables 17 to 24 provide the results for the ARDL adjusted tests. Overall, the results corroborate our earlier findings obtained for DGPs with serially independent errors. Both adjustments for serial error correlation work well, with size and power of the adjusted CD^* tests being quite close to the results already reported for DGPs with serially independent errors. It is also clear that without adjustments for latent factors and error serial correlation, the standard CD test will lead to large size distortions when the latent factors are strong. But in line with our theoretical results, the standard CD test, when adjusted for error serial correlation if needed, tends to have the correct size when the latent factors are weak or even semi-weak, namely when $a < 2/3$.

Comparing the two types of adjustments for error serial correlations (for pure latent factor models as well as for panel regressions with latent factors), the variance adjusted CD^* test works particularly well, and only shows mild over-rejection in the case where $T = 100$ and $n > T$. In contrast, the CD_{W+} with variance adjustment over-rejects for all combinations of n and T.

The ARDL adjusted version of the CD^* test also works well when the number PCs are not too large, and tends to have the correct size for all (n, T) combinations and only shows slight over-rejection when $n = 100$. The CD_{W+} test using ARDL adjustment does better in controlling for the size when the errors are Gaussian, but tends to over-reject when the errors are chi-squared distributed and $n > T$. Both adjusted versions of the CD_{W+} test continue to lack power against spatial or network alternatives.

6 Empirical illustrations

6.1 Are there spill-over effects in house price changes?

In our first illustration of the use of CD tests we consider the problem of spillover effects in regional house price changes. It is well known that house price changes are spatially correlated, but it is unclear if such correlations are mainly due to common factors (national or regional) or arise from spatial spillover effects not related to the common factors, a phenomenon also referred to as the ripple effect. See, for example, Tsai (2015), Chiang and Tsai (2016), Holly et al. (2011), and Bailey et al. (2016). To test for the presence of ripple effects the influence of common factors must first be filtered out and this is often a challenging exercise due to the latent nature of regional and national factors. Therefore, to find if there exist local spillover effects, one needs to test for significant residual cross-sectional dependence once the effects of common factors are filtered out.

We consider quarterly data on real house prices in the U.S. at the metropolitan statistical areas (MSAs). There are 381 MSAs, under the February 2013 definition provided by the U.S.

Office of Management and Budget (OMB). We use quarterly data on real house price changes compiled by Yang (2021) which covers $n = 377$ MSAs from the contiguous United States over the period 1975Q2-2014Q4 ($T = 160$ quarters). To allow for possible regional factors, we also follow Bailey et al. (2016) and start with the Bureau of Economic Analysis eight regional classification, namely New England, Mideast, Great Lakes, Plains, Southeast, Southwest, Rocky Mountain and Far West. But due to the low number of MSAs in New England and Rocky Mountain regions, we combine New England and Mideast, and Southwest and Rocky Mountain as two regions. We end up with a six region classification ($R = 6$), each covering a reasonable number of MSAs.

Initially, we model house price changes without regional groupings and consider the pure latent factor model with deterministic seasonal dummies to allow for seasonal movements in house prices. Specifically, we suppose

$$\pi_{it} = a_i + \sum_{j=1}^{3} \beta_{ij} 1\{q_t = j\} + \boldsymbol{\gamma}_i' \mathbf{f}_t + u_{it}, \tag{47}$$

where π_{it} is real house price in MSA i at time t, and $1\{q_t = j\}$ is the index for quarter j, and $\mathbf{f}_t$ is the $m_0 \times 1$ vector of latent factors. To filter out the seasonal effects we first estimate a_i and β_{ij} by running OLS regression of π_{it} on an intercept and the three seasonal dummies. This is justified since seasonal dummies are independently distributed of the latent factors. We then apply the PCA to $\{\hat{v}_{it} : i = 1, 2, \dots, n, t = 1, 2, \dots, T\}$, where $\hat{v}_{it} = \pi_{it} - \hat{a}_i - \sum_{j=1}^{3} \hat{\beta}_{ij} 1\{q_t = j\}$, to obtain the estimates $\hat{\boldsymbol{\gamma}}_i$ and $\hat{\mathbf{f}}_t$ for different choices of $\hat{m}$. For the case of serially independent errors we compute the standard CD, its bias-corrected version, CD^*, and the CD_{W+} test of JR using the residuals

$$e_{it} = \pi_{it} - \hat{a}_i - \sum_{j=1}^{3} \hat{\beta}_{ij} 1\{q_t = j\} - \hat{\boldsymbol{\gamma}}_i' \hat{\mathbf{f}}_t. \tag{48}$$

For the case with serially correlated errors, we consider the variance and ARDL adjusted versions of the CD, CD^*, and CD_{W+}. The variance adjusted versions are computed by scaling the original three CD statistics using expression in (39) and e_{it} generated from (48), and the ARDL adjusted versions are computed using the residuals from the following dynamic panel data model with latent factors,

$$\pi_{it} = a_i + \rho_i \pi_{it-1} + \sum_{j=1}^{3} \beta_{ij} 1\{q_t = j\} + \sum_{j=1}^{3} \lambda_{ij} 1\{q_{t-1} = j\} + \mathbf{g}_i' \mathbf{h}_t + \epsilon_{it}, \tag{49}$$

where the number of lags is set as one, and $\mathbf{h}_t$ is the vector of latent factors that includes $\mathbf{f}_t$ and $\mathbf{f}_{t-1}$, with associated factor loadings, $\mathbf{g}_i$.

Bailey et al. (2016) also find evidence of regional factors in U.S. house price changes which might not be picked up when using PCA. As a robustness check, we also consider an extended factor model containing observed regional and national factors, as well as latent factors:

$$\pi_{irt} = a_{ir} + \sum_{j=1}^{3} \beta_{ir,j} 1\{q_t = j\} + \delta_{ir,1} \bar{\pi}_{rt} + \delta_{ir,2} \bar{\pi}_t + \boldsymbol{\gamma}_{ir}' \mathbf{f}_t + u_{irt}, \tag{50}$$

where π_{irt} is the real house price changes in MSA i located in region $r = 1, 2, ..., 6$. $\bar{\pi}_{rt} = n_r^{-1} \sum_{i=1}^{n_r} \pi_{irt}$ and $\bar{\pi}_t = n^{-1} \sum_{r=1}^{R} \sum_{i=1}^{n_r} \pi_{irt}$ are proxies for the regional and national factors.

To filter out the effects of seasonal dummies as well as observed factors, we first run the least squares regression of π_{irt} on an intercept and $(1\{q_t = j\}, \bar{\pi}_{rt}, \bar{\pi}_t)$ for each i to generate the residuals

$$\hat{v}_{irt} = \pi_{irt} - \hat{a}_{ir} - \sum_{j=1}^{3} \hat{\beta}_{ir,j} 1\{q_t = j\} - \hat{\delta}_{ir,1}\bar{\pi}_{rt} - \hat{\delta}_{ir,2}\bar{\pi}_t, \tag{51}$$

and then apply PCA to $\{\hat{v}_{irt} : i = 1, \ldots, n, r = 1, \ldots, R, t = 1, \ldots, T\}$ to obtain $\hat{\boldsymbol{\gamma}}_{ir}$ and $\hat{\mathbf{f}}_t$, for different choices of $\hat{m}$, yielding the residuals

$$e_{irt} = \pi_{irt} - \hat{a}_{ir} - \sum_{j=1}^{3} \hat{\beta}_{ir,j} 1\{q_t = j\} - \hat{\delta}_{ir,1}\bar{\pi}_{rt} - \hat{\delta}_{ir,2}\bar{\pi}_t - \hat{\boldsymbol{\gamma}}_{ir}'\hat{\mathbf{f}}_t. \tag{52}$$

We compute CD, CD^* and CD_{W+} statistics using the above residuals, without adjusting for error serial correlation. For the case with serially correlated errors, the variance adjusted versions of the three CD statistics are generated by scaling original test statistics using (39) with e_{it} replaced by e_{irt} generated from (52), while the ARDL adjusted versions are computed using the residuals from the following dynamic panel data model with latent factors,

$$\begin{aligned}\pi_{irt} &= a_{ir} + \rho_{ir}\pi_{irt-1} + \sum_{j=1}^{3} \beta_{ir,j} 1\{q_t = j\} + \delta_{ir,1}\bar{\pi}_{rt} + \delta_{ir,2}\bar{\pi}_t \\ &+ \sum_{j=1}^{3} \lambda_{ir,j} 1\{q_{t-1} = j\} + \omega_{ir,1}\bar{\pi}_{rt-1} + \omega_{ir,2}\bar{\pi}_{t-1} + \mathbf{g}_{ir}'\mathbf{h}_t + \epsilon_{irt}.\end{aligned} \tag{53}$$

The test results are summarized in Table 1 for values of $\hat{m} = 1, 2, ..., 6$. The panel (a) of the table reports the results for models without regional or national factors, and panel (b) for models with regional and national factors. In both panels, the first three columns report CD, CD^* and CD_{W+} statistics that are not adjusted for error serial correlation, while the remaining six columns report the variance adjusted and ARDL adjusted versions of these statistics.

Both CD^* and CD_{W+} tests reject the null hypothesis of cross-sectional independence for all choices of $\hat{m}$, and irrespective of whether we allow for error serial correlation. Adding national and regional factors does not alter the main conclusion that use of latent factors does not seem to be sufficient for dealing with residual error cross-sectional dependence. In contrast, the results based on standard CD test are quite sensitive on the choice of $\hat{m}$ and whether we add national and regional factors to the model. It is also interesting that the values of CD^* statistics fall somewhat as we consider a larger number of latent factors, but even with $\hat{m} = 6$ we still obtain CD^* values in excess of 32 (compared to the critical value of 1.96). The CD_{W+} values are very large indeed, giving values in excess of 545 across all applications.

Overall, the above test results provide strong evidence that in addition to latent factors, spatial modeling of the type carried out in Bailey et al. (2016) is likely to be necessary to account for the remaining cross-sectional dependence.

6.2 Testing error cross-sectional dependence in CCE model of R&D investment

A number of recent empirical studies of R&D investment using panel data have resorted to latent factors to take account of knowledge spillover as well as dependencies across industries,

Table 1: Tests of error cross-sectional dependence for the house price application

Panel (a): Without regional or national factors

	Not adjusted for error serial correlation			Adjusted for error serial correlation: variance adjusted			Adjusted for error serial correlation: ARDL adjusted		
$\hat{m}$\Test	CD	CD^*	CD_{W+}	CD	CD^*	CD_{W+}	CD	CD^*	CD_{W+}
1	138.7	400.7	3679.0	48.8	140.9	1295.7	38.3	170.5	2585.5
2	8.8	108.5	2643.9	3.2	39.7	967.2	9.2	119.7	2197.7
3	15.5	150.4	2449.2	5.8	56.8	925.3	-1.4	93.1	2166.5
4	4.7	99.1	1754.7	1.9	39.9	707.8	-2.1	83.1	2015.4
5	1.3	92.9	1668.0	0.5	38.4	689.6	-3.4	75.4	1930.9
6	0.1	86.8	1534.6	0.0	36.5	646.1	-2.2	82.3	1859.7

Panel (b): With regional and national factors

	Not adjusted for error serial correlation			Adjusted for error serial correlation: variance adjusted			Adjusted for error serial correlation: ARDL adjusted		
$\hat{m}$\Test	CD	CD^*	CD_{W+}	CD	CD^*	CD_{W+}	CD	CD^*	CD_{W+}
1	137.5	139.0	1793.1	58.4	59.0	762.7	61.9	63.4	1742.1
2	107.5	117.6	1582.5	47.2	51.6	693.2	66.2	67.8	1592.0
3	106.5	119.2	1537.7	47.2	52.8	681.5	78.0	79.7	1636.5
4	112.0	122.0	1437.4	50.6	55.1	648.6	80.3	83.2	1559.6
5	124.7	136.6	1315.5	58.3	63.9	613.7	67.9	71.3	1483.1
6	45.8	69.8	1205.8	20.8	31.7	545.5	48.9	54.8	1375.9

Note: In the first panel the tests are applied to residuals from equation (47) for the first six columns and applied to residuals in equation (49) for the last three columns. Both equations de-seasonalize and de-factor real house price change. In the second panel the tests are applied to residuals in equation (50) for the first six columns and applied to residuals in equation (53) for the last three columns. Both equations not only de-seasonalize and de-factor real house price change but also filter out the regional and the national effects. CD denotes the standard CD test statistic, CD^* denotes the bias-corrected CD test statistic, and CD_{W+} denotes JR's power-enhanced randomized CD statistic. The number of selected PCs is denoted by $\hat{m}$.

and have applied the CCE approach of Pesaran (2006) to filter out these effects. For instance, Eberhardt et al. (2013) estimate panel data regressions of 12 manufacturing industries across 10 countries including Denmark, Finland, Germany, Italy, Japan, Netherlands, Portugal, Sweden, United Kingdom, and United States, over the period 1981- 2005. They apply the standard CD test to the residuals of their regressions to check if the CCE approach has been effective in fully capturing the error cross-sectional dependence. Their findings show that the CD test rejects the null hypothesis of error cross-sectional independence. JR revisit test results of Eberhardt et al. (2013) using their randomized CD test CD_{W+}, but again reject the null of error cross-sectional independence.

Here we focus on one of the panel regressions considered by Eberhardt et al. (2013) namely (see their Table 5)

$$\ln(Y_{it}) = a + \beta_1 \ln(L_{it}) + \beta_2 \ln(K_{it}) + \beta_3 \ln(R_{it}) + \boldsymbol{\gamma}_i' \mathbf{f}_t + u_{it}, \tag{54}$$

where Y_{it}, L_{it}, and K_{it} denote production, labor and physical capital inputs, respectively, and R_{it} is R&D capital. We estimate the panel regression over a balanced panel and compute the residuals after the CCE estimation:

$$\hat{v}_{it} = \ln(y_{it}) - \hat{a}_{CCE} - \hat{\beta}_{CCE,1}\ln(L_{it}) - \hat{\beta}_{CCE,2}\ln(K_{it}) - \hat{\beta}_{CCE,3}\ln(R_{it}). \tag{55}$$

In both Eberhardt et al. (2013) and JR the residuals in (55) are furthermore filtered out by cross-sectional average of $(\ln(y_{it}), \ln(L_{it}), \ln(K_{it}), \ln(R_{it}))$, and then the tests of error cross-sectional dependence are applied. Here, we apply PCA to residuals $\{\hat{v}_{it} := i = 1, \ldots, n, t = 1, \ldots, T\}$ to get estimates $\hat{\boldsymbol{\gamma}}_i$ and $\hat{\mathbf{f}}_t$, because PCA is not only required for construction of CD^* but also can present the change of cross-sectional dependence associated with the number of selected PCs to estimate factors, which is denoted by $\hat{m}$. Also, it is implicitly assumed that the number of latent factors in (54) is not larger than the number of time varying regressors (which is 3 in the present application) plus one, to satisfy the rank condition for CCE estimation. Yet it is worth noting that the CCE estimator continues to be consistent even if the rank condition is not met, but requires that factor loadings $\boldsymbol{\gamma}_i$ in (54) are independently distributed across i. See Pesaran (2006), Pesaran (2015b), and the more recent contribution by Juodis et al. (2021).

As before, for the case with serially independent errors the three CD statistics are computed using the residuals

$$e_{it} = \ln(y_{it}) - \hat{a}_{CCE} - \hat{\beta}_{CCE,1}\ln(L_{it}) - \hat{\beta}_{CCE,2}\ln(K_{it}) - \hat{\beta}_{CCE,3}\ln(R_{it}) - \hat{\boldsymbol{\gamma}}_i'\hat{\mathbf{f}}_t, \tag{56}$$

where $\hat{\boldsymbol{\gamma}}_i$ and $\hat{\mathbf{f}}_t$ are the first $\hat{m}$ PC of $\hat{v}_{it}$, for $\hat{m} = 1, 2, 3$, and 4. For the case with serially correlated errors, the variance adjusted versions of the three CD tests are generated by scaling original test statistics using expression in (39) and e_{it} in (56), while the ARDL adjusted versions are computed using the residuals from the following dynamic panel data model

$$\begin{aligned}\ln(Y_{it}) = a + \rho\ln(Y_{it-1}) + \beta_1\ln(L_{it}) + \beta_2\ln(K_{it}) + \beta_3\ln(R_{it})\\ + \lambda_1\ln(L_{it-1}) + \lambda_2\ln(K_{it-1}) + \lambda_3\ln(R_{it-1}) + \mathbf{g}_i'\mathbf{h}_t + \epsilon_{it},\end{aligned} \tag{57}$$

where the number of lags is set as one, and $\mathbf{h}_t$ is the extended vector of latent factors.

The test results are summarized in Table 2 for values of $\hat{m} = 1, 2, 3, 4$. As can be seen, the test outcomes are quite sensitive to the number of PCs selected regardless of whether the CD tests are adjusted for error serial correlation. For instance, the CD and CD^* tests reject the null of cross-sectional independence when $\hat{m} = 3$, but not if $\hat{m} = 4$. In comparison, the CD_{W^+} test rejects the null for all values of $\hat{m}$ when there is no adjustment for serially correlated errors, but the variance adjusted and ARDL adjusted versions of CD_{W+} only reject the null when $\hat{m} = 1$. The CD^* test does not reject the null hypothesis if $\hat{m}$ is set to 4; and given the favorable MC results for the CD^* test, we are inclined to conclude that a latent factor model seems to be adequate for modelling error cross-sectional dependence for this data set.

7 Concluding remarks

In this paper we have revisited the problem of testing error cross-sectional dependence in panel data models with latent factors. Starting with a pure multi-factor model we show that the standard CD test proposed by Pesaran (2004) remains valid if the latent factors are weak, but over-reject when one or more of the latent factors are strong. The over-rejection of the CD test in the case of strong factors is also established by Juodis and Reese (2022), who propose a randomized test statistic to correct for over-rejection and add a screening component to

Table 2: Tests of error cross-sectional dependence for panel regressions of R&D investment

	Not adjusted for error serial correlation			Adjusted for error serial correlation: variance adjusted			ARDL adjusted		
$\hat{m}\backslash$Test	CD	CD^*	CD_{W+}	CD	CD^*	CD_{W+}	CD	CD^*	CD_{W+}
1	0.5	2.1	38.4	0.2	1.0	19.1	0.7	1.4	5.4
2	2.1	3.3	3.9	1.3	1.9	0.6	1.8	2.5	0.1
3	4.1	6.3	3.7	2.6	4.0	0.6	3.1	4.6	0.8
4	-0.8	1.7	8.6	-0.5	1.1	0.4	-1.3	0.3	-0.7

Note: The tests are applied to residuals from equation (54) for the first six columns, and applied to residuals from equation (57) for the last three columns. Both equations model R&D investment. See also the notes to Table 1.

achieve power. However, as we show, JR's CD_{W+} test is not guaranteed to have the correct size and need not be powerful against spatial or network alternatives. Such alternatives are of particular interest in the analyses of ripple effects in housing markets, and clustering of firms within industries in capital or arbitrage asset pricing models. In fact, using Monte Carlo experiments we show that under non-Gaussian errors the JR test continues to over-reject when the cross section dimension (n) is larger than the time dimension (T), and often has power close to size against spatial alternatives. To overcome some of these shortcomings, we propose a simple bias-corrected CD test, labelled CD^*, which is shown to be asymptotically $N(0,1)$ when n and $T \to \infty$ such that $n/T \to \kappa$, for a fixed constant κ. This result holds for pure latent factor models as well as for panel regressions with latent factors. To deal with possible error serial dependence, following Baltagi et al. (2016), we also consider a variance adjusted version of CD^*, as well as an alternative ARDL adjusted version that eliminates the error serial dependence before the application of the CD^* test procedure. Both of these approaches are shown to perform well within the Monte Carlo set up of the paper.

References

Ahn, S. C. and Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, 81(3):1203–1227.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1):135–171.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Bai, J. and Ng, S. (2008). *Large dimensional factor analysis*. Now Publishers Inc.

Bailey, N., Holly, S., and Pesaran, M. H. (2016). A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence. *Journal of Applied Econometrics*, 31(1):249–280.

Bailey, N., Kapetanios, G., and Pesaran, M. H. (2021). Measurement of factor strength: Theory and practice. *Journal of Applied Econometrics*, 36(2):431–453.

Bailey, N., Pesaran, M. H., and Smith, L. V. (2019). A multiple testing approach to the regularisation of large sample correlation matrices. *Journal of Econometrics*, 208(2):507–534.

Baltagi, B. H., Kao, C., and Peng, B. (2016). Testing cross-sectional correlation in large panel data models with serial correlation. *Econometrics*, 4(4):44.

Chiang, M.-C. and Tsai, I.-C. (2016). Ripple effect and contagious effect in the us regional housing markets. *The Annals of Regional Science*, 56(1):55–82.

Chudik, A., Pesaran, M., and Tosetti, E. (2011). Weak and strong cross-section dependence and estimation of large panels. *Econometrics Journal*, 14:C45–C90.

Eberhardt, M., Helmers, C., and Strauss, H. (2013). Do spillovers matter when estimating private returns to r&d? *Review of Economics and Statistics*, 95(2):436–448.

Fan, J., Liao, Y., and Yao, J. (2015). Power enhancement in high-dimensional cross-sectional tests. *Econometrica*, 83(4):1497–1541.

Holly, S., Pesaran, M. H., and Yamagata, T. (2011). The spatial and temporal diffusion of house prices in the uk. *Journal of Urban Economics*, 69(1):2–23.

Juodis, A., Karabiyik, H., and Westerlund, J. (2021). On the robustness of the pooled cce estimator. *Journal of Econometrics*, 220(2):325–348.

Juodis, A. and Reese, S. (2022). The incidental parameters problem in testing for remaining cross-section correlation. *Journal of Business & Economic Statistics*, 40(3):1191–1203.

Moon, H. R. and Weidner, M. (2015). Linear regression for panel with unknown number of factors as interactive fixed effects. *Econometrica*, 83(4):1543–1579.

Pesaran, M. H. (2004). General diagnostic tests for cross-sectional dependence in panels. *University of Cambridge, Cambridge Working Papers in Economics*, 435.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4):967–1012.

Pesaran, M. H. (2015a). Testing weak cross-sectional dependence in large panels. *Econometric reviews*, 34(6-10):1089–1117.

Pesaran, M. H. (2015b). *Time series and panel data econometrics*. Oxford University Press.

Pesaran, M. H., Smith, L. V., and Yamagata, T. (2013). Panel unit root tests in the presence of a multifactor error structure. *Journal of Econometrics*, 175(2):94–115.

Pesaran, M. H. and Tosetti, E. (2011). Large panels with common factors and spatial correlation. *Journal of Econometrics*, 161(2):182–202.

Smith, M. D. (1988). Convergent series expressions for inverse moments of quadratic forms in normal variables. *Australian Journal of Statistics*, 30(2):235–246.

Tsai, I.-C. (2015). Spillover effect between the regional and the national housing markets in the uk. *Regional Studies*, 49(12):1957–1976.

Yang, C. F. (2021). Common factors and spatial dependence: An application to us house prices. *Econometric Reviews*, 40(1):14–50.

Online supplement to

A Bias-Corrected CD Test for Error Cross-Sectional Dependence in Panel Data Model with Latent Factors

by

M. Hashem Pesaran and Yimeng Xie

August 23, 2022

This online supplement is in four parts. Section 1 provides the proofs of Theorems 1 to 3, and Corollary 1 of the paper. Section 2 states and establishes the auxiliary lemmas used in these proofs. Section 3 derives the order of θ_n in terms of the factor strengths. Section 4 gives the summary result tables for the Monte Carlo simulations discussed in Section 5 of the paper.

Notations: For the $n \times n$ matrix $\mathbf{A} = (a_{ij})$, we denote its smallest and largest eigenvalues by $\lambda_{min}(\mathbf{A})$ and $\lambda_{max}(\mathbf{A})$, respectively, its trace by $tr(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$, its spectral radius by $\rho(\mathbf{A}) = |\lambda_{\max}(\mathbf{A})|$, its Frobenius norm by $\|\mathbf{A}\|_F = (tr(\mathbf{A}'\mathbf{A}))^{1/2}$, its spectral norm by $\|\mathbf{A}\| = \lambda_{\max}^{1/2}(\mathbf{A}'\mathbf{A}) \leq \|\mathbf{A}\|_F$, its maximum absolute column sum norm by $\|\mathbf{A}\|_1 = \max_{1\leq j\leq n}\left(\sum_{i=1}^{n}|a_{ij}|\right)$, and its maximum absolute row sum norm by $\|\mathbf{A}\|_\infty = \max_{1\leq i\leq n}\left(\sum_{j=1}^{n}|a_{ij}|\right)$. We write $\mathbf{A} > \mathbf{0}$ when $\mathbf{A}$ is positive definite. We denote the ℓ_p-norm of the random variable x by $\|x\|_p = E(|x|^p)^{1/p}$ for $p \geq 1$, assuming that $E(|x|^p) < K$. $\rightarrow_p$ denotes convergence in probability, $\overset{a.s.}{\rightarrow}$ almost sure convergence, $\rightarrow_d$ convergence in distribution, and $\overset{a}{\sim}$ asymptotic equivalence in distribution. $O_p(\cdot)$ and $o_p(\cdot)$ denote the stochastic order relations. In particular, $o_p(1)$ indicates terms that tend to zero in probability as $(n, T) \rightarrow \infty$, such that $n/T \rightarrow \kappa$, where $0 < \kappa < \infty$. K and c will be used to denote finite large and non-zero small positive numbers, respectively, that do not depend on n and T. They can take different values at different instances. $\delta_{nT} = \min\left(\sqrt{n}, \sqrt{T}\right)$. If $\{f_n\}_{n=1}^{\infty}$ is any real sequence and $\{g_n\}_{n=1}^{\infty}$ is a sequence of positive real numbers, then $f_n = O(g_n)$, if there exists K such that $|f_n|/g_n \leq K$ for all n. $f_n = o(g_n)$ if $f_n/g_n \rightarrow 0$ as $n \rightarrow \infty$. If $\{f_n\}_{n=1}^{\infty}$ and $\{g_n\}_{n=1}^{\infty}$ are both positive sequences of real numbers, then $f_n = \ominus(g_n)$ if there exists $n_0 \geq 1$ and positive finite constants K_0 and K_1, such that $\inf_{n\geq n_0}(f_n/g_n) \geq K_0$, and $\sup_{n\geq n_0}(f_n/g_n) \leq K_1$.

1 Proofs of the Theorems

Proof of Theorem 1. We first note that the residuals of the factor model (2) estimated using PCs, given by (8), can be written as:

$$e_{it} = u_{it} - \boldsymbol{\gamma}_i'\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right) - \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)'\mathbf{f}_t - \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)'\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right). \tag{1}$$

Let $\boldsymbol{\delta}_{i,T} = \boldsymbol{\gamma}_i/\omega_{i,T}$, and $\hat{\boldsymbol{\delta}}_{i,T} = \hat{\boldsymbol{\gamma}}_i/\omega_{i,T}$, where $\omega_{i,T}^2 = T^{-1}\mathbf{u}_i'\mathbf{M}_F\mathbf{u}_i$, then

$$e_{it}/\omega_{i,T} = u_{it}/\omega_{i,T} - \boldsymbol{\delta}_{i,T}'\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right) - \left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right)'\mathbf{f}_t - \left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right)'\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right). \tag{2}$$

As shown in Lemma 1 in Section 2, and recalling that $y_{jt} = \gamma_j' f_t + u_{jt}$, we have

$$\hat{\mathbf{f}}_t - \mathbf{f}_t = \left[n^{-1} \sum_{j=1}^{n} \left(\hat{\boldsymbol{\gamma}}_j - \boldsymbol{\gamma}_j \right) \boldsymbol{\gamma}_j' \right] \mathbf{f}_t + n^{-1} \sum_{j=1}^{n} \left(\hat{\boldsymbol{\gamma}}_j - \boldsymbol{\gamma}_j \right) u_{jt} + n^{-1} \sum_{j=1}^{n} \boldsymbol{\gamma}_j u_{jt}.$$

Using this result in (2) we obtain

$$\begin{aligned} e_{it}/\omega_{i,T} = u_{it}/\omega_{i,T} - \boldsymbol{\delta}_{i,T}' \left(n^{-1} \sum_{j=1}^{n} \boldsymbol{\gamma}_j u_{jt} \right) - \boldsymbol{\delta}_{i,T}' \left[n^{-1} \sum_{j=1}^{n} \left(\hat{\boldsymbol{\gamma}}_j - \boldsymbol{\gamma}_j \right) \boldsymbol{\gamma}_j' \right] \mathbf{f}_t - \boldsymbol{\delta}_{i,T}' \left[n^{-1} \sum_{j=1}^{n} \left(\hat{\boldsymbol{\gamma}}_j - \boldsymbol{\gamma}_j \right) u_{jt} \right] \\ - \left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T} \right)' \mathbf{f}_t - \left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T} \right)' \left(\hat{\mathbf{f}}_t - \mathbf{f}_t \right), \end{aligned} \tag{3}$$

and summing over i we have

$$\begin{aligned} n^{-1/2} \sum_{i=1}^{n} e_{it}/\omega_{i,T} = n^{-1/2} \sum_{i=1}^{n} u_{it}/\omega_{i,T} - \boldsymbol{\varphi}_{nT}' \left(n^{-1/2} \sum_{i=1}^{n} \boldsymbol{\gamma}_i u_{it} \right) - \boldsymbol{\varphi}_{nT}' \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i \right) \boldsymbol{\gamma}_i' \right] \mathbf{f}_t \\ - \boldsymbol{\varphi}_{nT}' \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i \right) u_{it} \right] - \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T} \right)' \right] \mathbf{f}_t \\ - \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T} \right) \right]' \left(\hat{\mathbf{f}}_t - \mathbf{f}_t \right), \end{aligned}$$

where

$$\boldsymbol{\varphi}_{nT} = n^{-1} \sum_{i=1}^{n} \boldsymbol{\delta}_{i,T}. \tag{4}$$

Written more compactly

$$h_{t,nT} = n^{-1/2} \sum_{i=1}^{n} e_{it}/\omega_{i,T} = \psi_{t,nT} - s_{t,nT}, \tag{5}$$

where

$$\psi_{t,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{a_{i,nT} u_{it}}{\omega_{i,T}}, \qquad a_{i,nT} = 1 - \omega_{i,T} \boldsymbol{\varphi}_{nT}' \boldsymbol{\gamma}_i, \tag{6}$$

$$s_{t,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\boldsymbol{\varphi}_{nT}' \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i \right) u_{it} + \left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T} \right)' \hat{\mathbf{f}}_t + \boldsymbol{\varphi}_{nT}' \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i \right) \boldsymbol{\gamma}_i' \mathbf{f}_t \right]. \tag{7}$$

Further, let

$$\xi_{t,n} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} a_{i,n} \varepsilon_{it}, \quad a_{i,n} = 1 - \sigma_i \boldsymbol{\varphi}_n' \boldsymbol{\gamma}_i, \tag{8}$$

where $\varphi_n = n^{-1} \sum_{i=1}^{n} \delta_i$, and $\delta_i = \gamma_i/\sigma_i$. Then $\psi_{t,nT}$, given by (6), can be written as

$$\begin{aligned} \psi_{t,nT} &= \xi_{t,n} + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(1 - \omega_{i,T} \boldsymbol{\varphi}_{nT}' \boldsymbol{\gamma}_i \right) \frac{u_{it}}{\omega_{i,T}} - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(1 - \sigma_i \boldsymbol{\varphi}_n' \boldsymbol{\gamma}_i \right) \varepsilon_{it} \\ &= \xi_{t,n} - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{\varphi}_{nT}' \boldsymbol{\gamma}_i u_{it} + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sigma_i \boldsymbol{\varphi}_n' \boldsymbol{\gamma}_i \varepsilon_{it} + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{u_{it}}{\omega_{i,T}} - \varepsilon_{it} \right). \end{aligned}$$

Define

$$\zeta_{it} = \left(\frac{1}{\left(T^{-1}\boldsymbol{\varepsilon}_i' \mathbf{M}_F \boldsymbol{\varepsilon}_i\right)^{1/2}} - 1 \right) \varepsilon_{it}$$

and since $u_{it} = \sigma_i \varepsilon_{it}$, then we have

$$\psi_{t,nT} = \xi_{t,n} + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \zeta_{it} - \left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)' \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{\gamma}_i \sigma_i \varepsilon_{it}.$$

So $\psi_{t,nT}$ can also be written equivalently as

$$\psi_{t,nT} = \xi_{t,n} + \upsilon_{t,nT} - \left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)' \kappa_{t,n}, \tag{9}$$

where

$$\xi_{t,n} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} a_{i,n} \varepsilon_{it}, \quad a_{i,n} = 1 - \sigma_i \boldsymbol{\varphi}_n' \boldsymbol{\gamma}_i, \tag{10}$$

$$\kappa_{t,n} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{\gamma}_i \sigma_i \varepsilon_{it}, \tag{11}$$

$$\upsilon_{t,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \zeta_{it}. \tag{12}$$

Using (5) in (21) and after some algebra we have

$$\widetilde{CD} = \left(\sqrt{\frac{n}{n-1}}\right) \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\frac{\psi_{t,nT}^2 - 1}{\sqrt{2}} \right) + \left(\sqrt{\frac{n}{n-1}}\right) \left(p_{nT} - q_{nT}\right),$$

where

$$p_{nT} = T^{-1/2} \sum_{t=1}^{T} s_{t,nT}^2, \tag{13}$$

$$q_{nT} = T^{-1/2} \sum_{t=1}^{T} \psi_{t,nT} s_{t,nT}, \tag{14}$$

and by Lemma 4 in Section 2 $p_{nT} = o_p(1)$, and $q_{nT} = o_p(1)$. Hence

$$\widetilde{CD} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\frac{\psi_{t,nT}^2 - 1}{\sqrt{2}} \right) + o_p(1). \tag{15}$$

Now consider $T^{-1/2} \sum_{t=1}^{T} \psi_{t,nT}^2$ and using (9) note that

$$\begin{aligned}
\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \psi_{t,nT}^2 &= \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \xi_{t,n}^2 \right) \left(1 + \frac{2 \sum_{t=1}^{T} \xi_{t,n} \upsilon_{t,nT}}{\sum_{t=1}^{T} \xi_{t,n}^2} \right) + \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \upsilon_{t,nT}^2 \\
&\quad + \sqrt{T} \left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)' \left(\frac{\sum_{t=1}^{T} \kappa_{t,n} \kappa_{t,n}'}{T} \right) \left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right) - 2\sqrt{T} \left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)' \left(\frac{1}{T} \sum_{t=1}^{T} \kappa_{t,n} \upsilon_{t,nT} \right) \\
&\quad - 2\sqrt{T} \left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)' \left(\frac{1}{T} \sum_{t=1}^{T} \kappa_{t,n} \xi_{t,n} \right) \\
&= g_{1,nT} + g_{2,nT} + g_{3,nT} - 2g_{4,nT} - 2g_{5,nT}.
\end{aligned} \tag{16}$$

Starting with the second term, $g_{2,nT}$, note that

$$E\left(g_{2,nT}\right) = E\left(\left|g_{2,nT}\right|\right) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} E\left(\upsilon_{t,nT}^2\right), \tag{17}$$

where $\upsilon_{t,nT}$ is defined by (12). Based on the cross-sectional independence of ε_{it} in Assumption 2 we have

$$E\left(\upsilon_{t,nT}^2\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(\zeta_{it}^2\right) = \frac{1}{n}\sum_{i=1}^{n} E\left[\left(\frac{1}{\left(\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i/T\right)^{1/2}} - 1\right)^2 \varepsilon_{it}^2\right],$$

where

$$E\left(\zeta_{it}^2\right) = E\left(\frac{\varepsilon_{it}^2}{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i/T}\right) - 1 - 2\left[E\left(\frac{\varepsilon_{it}^2}{\left(\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i/T\right)^{1/2}}\right) - 1\right] = O\left(\frac{1}{T}\right), \tag{18}$$

as (96) and (97) in Lemma 7 in Section 2 show

$$E\left(\frac{\varepsilon_{it}^2}{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i/T} - 1\right) = O\left(\frac{1}{T}\right), \text{ and } E\left(\frac{\varepsilon_{it}^2}{\left(\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i/T\right)^{1/2}} - 1\right) = O\left(\frac{1}{T}\right),$$

and therefore

$$E\left(\upsilon_{t,nT}^2\right) = O\left(T^{-1}\right). \tag{19}$$

Using this result in (17) we obtain $E\left(\left|g_{2,nT}\right|\right) = O(T^{-1/2})$, which by Markov inequality establishes that $g_{2,nT} = o_p(1)$. Consider the three remaining terms, $g_{3,nT}$, $g_{4,nT}$ and $g_{5,nT}$ of (16), starting with

$$g_{3,nT} = \frac{1}{\sqrt{T}}\left[\sqrt{T}\left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)'\left(\frac{\sum_{t=1}^{T}\kappa_{t,n}\kappa_{t,n}'}{T}\right)\sqrt{T}\left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)\right],$$

and note that by Lemma 5 we have

$$\sqrt{T}\left(\boldsymbol{\varphi}_n - \boldsymbol{\varphi}_{nT}\right) = O_p\left(n^{-1/2}\right) + O_p\left(T^{-1/2}\right). \tag{20}$$

Furthermore

$$E\left\|T^{-1}\sum_{t=1}^{T}\kappa_{t,n}\kappa_{t,n}'\right\| \leq \frac{1}{T}\sum_{t=1}^{T} E\left\|\kappa_{t,n}\kappa_{t,n}'\right\| = \frac{1}{T}\sum_{t=1}^{T} E\left(\kappa_{t,n}'\kappa_{t,n}\right).$$

Then using (11)

$$E\left(\kappa_{t,n}'\kappa_{t,n}\right) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\boldsymbol{\gamma}_i'\boldsymbol{\gamma}_j\sigma_i\sigma_j E\left(\varepsilon_{it}\varepsilon_{jt}\right) = \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\boldsymbol{\gamma}_i'\boldsymbol{\gamma}_i\right) < K, \tag{21}$$

it then follows that

$$E\left\|T^{-1}\sum_{t=1}^{T}\kappa_{t,n}\kappa_{t,n}'\right\| \leq \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\boldsymbol{\gamma}_i'\boldsymbol{\gamma}_i\right) < K,$$

and $T^{-1}\sum_{t=1}^{T}\kappa_{t,n}\kappa_{t,n}' = O_p(1)$. Using this result together with (20) we then have

$$g_{3,nT} = o_p(1). \tag{22}$$

Similarly, by Cauchy-Schwarz inequality and using (19) and (21) we have

$$\begin{aligned} E\left\| \frac{1}{T}\sum_{t=1}^{T}\kappa_{t,n}\upsilon_{t,nT} \right\| &\le \frac{1}{T}\sum_{t=1}^{T} E\left\|\kappa_{t,n}\upsilon_{t,nT}\right\| \le \frac{1}{T}\sum_{t=1}^{T}\left[E\left(\kappa_{t,n}'\kappa_{t,n}\right)\right]^{1/2}\left[E\left(\upsilon_{t,nT}^2\right)\right]^{1/2}, \\ &\le \left[\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\boldsymbol{\gamma}_i'\boldsymbol{\gamma}_i\right)\right]^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left[E\left(\upsilon_{t,nT}^2\right)\right]^{1/2}\right) = O\left(T^{-1/2}\right). \end{aligned}$$

Then using the above results it also follows that

$$g_{4,nT} = \sqrt{T}\left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)'\left(\frac{1}{T}\sum_{t=1}^{T}\kappa_{t,n}\upsilon_{t,nT}\right) = o_p(1). \tag{23}$$

Similarly,

$$E\left\| \frac{1}{T}\sum_{t=1}^{T}\kappa_{nt}\xi_{t,n} \right\| \le \frac{1}{T}\sum_{t=1}^{T}\left[E\left(\kappa_{t,n}'\kappa_{t,n}\right)\right]^{1/2}\left[E\left(\xi_{t,n}^2\right)\right]^{1/2},$$

where using (8) $E\left(\xi_{t,n}^2\right) = \frac{1}{n}\sum_{i=1}^{n} a_{i,n}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(1-\sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i\right)^2 < K$. Hence, $E\left\|\frac{1}{T}\sum_{t=1}^{T}\kappa_{t,n}\xi_{t,n}\right\| < K$, and again using (20) it follows that

$$g_{5,nT} = \sqrt{T}\left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)'\left(\frac{1}{T}\sum_{t=1}^{T}\kappa_{t,n}\xi_{t,n}\right) = o_p(1). \tag{24}$$

Using results (17) to (24) in (16) now yields

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\psi_{t,nT}^2 = \left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{t,n}^2\right)\left(1 + \frac{2\sum_{t=1}^{T}\xi_{t,n}\upsilon_{t,nT}}{\sum_{t=1}^{T}\xi_{t,n}^2}\right) + o_p(1). \tag{25}$$

as desired. Consider now the (population) bias-corrected version of $\widetilde{CD}$ defined by

$$\widetilde{CD}^* = \frac{\widetilde{CD} + \sqrt{\frac{T}{2}}\theta_n}{1-\theta_n} \tag{26}$$

where $\theta_n = 1 - \frac{1}{n}\sum_{i=1}^{n} a_{i,n}^2$, and $a_{i,n} = 1 - \sigma_i\varphi_n'\gamma_i$. Then using (15) in (26), we have

$$\begin{aligned} \widetilde{CD}^* &= \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\frac{\psi_{t,nT}^2 - 1}{\sqrt{2}}\right] + \sqrt{\frac{T}{2}}\theta_n}{1-\theta_n} + o_p(1) \\ &= \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{\psi_{t,nT}^2}{\sqrt{2}} - \sqrt{\frac{T}{2}}\left(1-\theta_n\right)}{1-\theta_n} + o_p(1). \end{aligned}$$

Now using (25) in the above and after some re-arrangement of the terms we obtain

$$\widetilde{CD}^* = \frac{\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\frac{\xi_{t,n}^2-(1-\theta_n)}{\sqrt{2}}\right)\right]\left(1+\frac{2}{\sqrt{T}}w_{nT}\right)}{1-\theta_n} + \sqrt{2}w_{nT} + o_p\left(1\right), \tag{27}$$

where

$$w_{nT} = \frac{T^{-1/2}\sum_{t=1}^{T}\xi_{t,n}v_{t,nT}}{T^{-1}\sum_{t=1}^{T}\xi_{t,n}^2}. \tag{28}$$

Consider first the denominator of w_{nT} and using (8) note that $\xi_{t,n} = n^{-1/2}\sum_{i=1}^{n} a_{i,n}\varepsilon_{it}$. Also by Assumption 2 and 3, $n^{-1}\sum_{i=1}^{n} a_{i,n}^2$ and $E\left(\varepsilon_{it}^4\right)$ are bounded and strictly positive, then by Lyapunov's central limit theorem we have

$$\xi_{t,n} \rightarrow_d N(0, \omega_\xi^2),$$

where in view of condition (5) of Assumption 3 $\omega_\xi^2 = \lim_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n} a_{i,n}^2\right) > 0$. Hence, the denominator of w_{nT} is $O_p\left(1\right)$. Consider now the numerator of (28), and using (12) we denote $v_{t,nT} = n^{-1/2}\sum_{i=1}^{n}\zeta_{it}$ and $\zeta_{it} = \left[\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{-1/2} - 1\right]\varepsilon_{it}$. Then

$$\xi_{t,n}v_{t,nT} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{in}\varepsilon_{it}\zeta_{jt},$$

where by Assumption 2, ε_{it} and ζ_{jt} are distributed independently for $i \neq j$. Therefore,

$$E\left(\xi_{t,n}v_{t,nT}\right) = \frac{1}{n}\sum_{i=1}^{n} a_{in}E\left[\varepsilon_{it}^2\left(\frac{1}{\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{1/2}} - 1\right)\right].$$

Furthermore, since a_{in} is bounded, and

$$E\left[\varepsilon_{it}^2\left(\frac{1}{\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{1/2}} - 1\right)\right] = O\left(\frac{1}{T}\right)$$

based on result (97), then $E\left(\xi_{t,n}v_{t,nT}\right) = O\left(T^{-1}\right)$. Using these results and denoting the numerator of (28) by r_{nT},

$$E\left(r_{nT}\right) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} E\left(\xi_{t,n}v_{t,nT}\right) = O\left(\frac{1}{\sqrt{T}}\right). \tag{29}$$

Consider now $Var\left(r_{nT}\right)$ and note

$$Var\left(r_{nT}\right) = \frac{1}{T}\sum_{t=1}^{T}\sum_{t'=1}^{T} E\left(\xi_{t,n}v_{t,nT}\xi_{t',n}v_{t',nT}\right) - \left(E\left(r_{nT}\right)\right)^2$$

where $\left(E\left(r_{nT}\right)\right)^2 = O\left(T^{-1}\right)$ based on (29). Also note

$$\begin{aligned}\xi_{t,n}v_{t,nT}\xi_{t',n}v_{t',nT} &\equiv \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{in}a_{jn}\varepsilon_{it}\varepsilon_{jt'}\right)\left(\frac{1}{n}\sum_{r=1}^{n}\sum_{s=1}^{n}\zeta_{rt}\zeta_{st'}\right)\\ &= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{r=1}^{n}\sum_{s=1}^{n} a_{in}a_{jn}\varepsilon_{it}\varepsilon_{jt'}\zeta_{rt}\zeta_{st'},\end{aligned}$$

where ε_{it}, ζ_{rt} are both cross-sectionally independent based on Assumption 2, then we have

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\sum_{t'=1}^{T}E\left(\xi_{t,n}\upsilon_{t,nT}\xi_{t',n}\upsilon_{t',nT}\right) \\
&=\frac{1}{Tn^2}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{r=1}^{n}\sum_{s=1}^{n}E\left(a_{in}a_{jn}\varepsilon_{it}\varepsilon_{jt'}\zeta_{rt}\zeta_{st'}\right) \\
&=\frac{1}{Tn^2}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{i=1}^{n}a_{in}^2E\left(\varepsilon_{it}\varepsilon_{it'}\zeta_{it}\zeta_{it'}\right)+\frac{1}{Tn^2}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{i=1}^{n}\sum_{r\neq i}^{n}a_{in}^2E\left(\varepsilon_{it}\varepsilon_{it'}\right)E\left(\zeta_{rt}\zeta_{rt'}\right)+ \\
&\frac{1}{Tn^2}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n}a_{in}E\left(\varepsilon_{it}\zeta_{it}\right)a_{jn}E\left(\varepsilon_{jt'}\zeta_{jt'}\right)+\frac{1}{Tn^2}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n}a_{in}E\left(\varepsilon_{it}\zeta_{it'}\right)a_{jn}E\left(\varepsilon_{jt'}\zeta_{jt}\right).
\end{aligned} \tag{30}$$

Consider the second term of (30) and using the serial independence of ε_{it} yields

$$\frac{1}{Tn^2}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{i=1}^{n}\sum_{r\neq i}^{n}a_{in}^2E\left(\varepsilon_{it}\varepsilon_{it'}\right)E\left(\zeta_{rt}\zeta_{rt'}\right)=\frac{1}{Tn^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{r\neq i}^{n}a_{in}^2E\left(\varepsilon_{it}^2\right)E\left(\zeta_{rt}^2\right)=O\left(\frac{1}{T}\right), \tag{31}$$

in which the second equation holds because $E\left(\varepsilon_{it}^2\right)=1$ by Assumption 2 and $E\left(\zeta_{rt}^2\right)=O\left(T^{-1}\right)$ based on (18). Next for the third term of (30), since result (97) shows

$$E\left(\varepsilon_{it}\zeta_{it}\right)=E\left[\frac{\varepsilon_{it}^2}{\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{1/2}}-\varepsilon_{it}^2\right]=O\left(\frac{1}{T}\right), \tag{32}$$

then

$$\frac{1}{Tn^2}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n}a_{in}E\left(\varepsilon_{it}\zeta_{it}\right)a_{jn}E\left(\varepsilon_{jt'}\zeta_{jt'}\right)=O\left(\frac{1}{T}\right). \tag{33}$$

For the fourth term of (30), using the serial independence of ε_{it} and result (32) yields

$$\frac{1}{Tn^2}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n}a_{in}E\left(\varepsilon_{it}\zeta_{it'}\right)a_{jn}E\left(\varepsilon_{jt'}\zeta_{jt}\right)=\frac{1}{Tn^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n}a_{in}a_{jn}E\left(\varepsilon_{it}\zeta_{it}\right)E\left(\varepsilon_{jt}\zeta_{jt}\right)=O\left(\frac{1}{T^2}\right). \tag{34}$$

Finally consider the first term of (30) and note

$$\begin{aligned}
&\frac{1}{Tn^2}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{i=1}^{n}a_{in}^2E\left(\varepsilon_{it}\varepsilon_{it'}\zeta_{it}\zeta_{it'}\right) \\
&=\frac{1}{n^2}\sum_{i=1}^{n}a_{in}^2\left(\frac{1}{T}\sum_{t=1}^{T}\sum_{t'=1}^{T}E\left(\varepsilon_{it}\varepsilon_{it'}\zeta_{it}\zeta_{it'}\right)\right) \\
&=\left[\frac{1}{n^2}\sum_{i=1}^{n}a_{in}^2\frac{1}{T}\sum_{t=1}^{T}E\left(\varepsilon_{it}^2\zeta_{it}^2\right)\right]+\frac{1}{n^2}\sum_{i=1}^{n}a_{in}^2\frac{1}{T}\sum_{t=1}^{T}\sum_{t'\neq t}^{T}E\left[\left(\varepsilon_{it}\zeta_{it}\right)\left(\varepsilon_{it'}\zeta_{it'}\right)\right],
\end{aligned}$$

in which the first term is obviously $O\left(n^{-1}\right)$. Consider the second term, and note by Cauchy Schwarz inequality we have

$$E\left[\left(\varepsilon_{it}\zeta_{it}\right)\left(\varepsilon_{it'}\zeta_{it'}\right)\right] = E\left[\varepsilon_{it}^2\varepsilon_{it'}^2\left(\frac{1}{\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{1/2}} - 1\right)^2\right]$$

$$\leq \left[E\left(\varepsilon_{it}^4\varepsilon_{it'}^4\right)\right]^{1/2}\left[E\left(\frac{1}{\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{1/2}} - 1\right)^4\right]^{1/2},$$

where by Assumption 2 we have $E\left(\varepsilon_{it}^4\varepsilon_{it'}^4\right) = E\left(\varepsilon_{it}^4\right)E\left(\varepsilon_{it'}^4\right) = O(1)$. Meanwhile note

$$\begin{aligned} E\left(\frac{1}{\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{1/2}} - 1\right)^4 &= E\left[\frac{1}{\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^2}\right] - 4E\left[\frac{1}{\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{3/2}}\right] + 6E\left[\frac{1}{T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i}\right] \\ &\quad - 4E\left[\frac{1}{\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{1/2}}\right] + 1 \\ &= O\left(\frac{1}{T}\right), \end{aligned}$$

given result (94). Overall we have shown $E\left[\left(\varepsilon_{it}\zeta_{it}\right)\left(\varepsilon_{it'}\zeta_{it'}\right)\right] = O\left(T^{-1/2}\right)$ such that

$$\frac{1}{n^2}\sum_{i=1}^{n} a_{in}^2 \frac{1}{T}\sum_{t=1}^{T}\sum_{t'\neq t}^{T} E\left[\left(\varepsilon_{it}\zeta_{it}\right)\left(\varepsilon_{it'}\zeta_{it'}\right)\right] = O\left(\frac{\sqrt{T}}{n}\right),$$

and so is the order of the first term of (30). Hence combining it with (31), (33), (34) establishes the probability order of (30) is $O\left(T^{1/2}n^{-1}\right)$, then with (29) it follows

$$Var\left(r_{nT}\right) = O\left(\frac{\sqrt{T}}{n}\right),$$

and by apply Chebyshv's inequality we have

$$r_{nT} = O_p\left(\frac{\sqrt{T}}{n}\right).$$

Using this result in (28) and noting the denominator of w_{nT} is $O_p(1)$ then it follows that $w_{nT} = o_p(1)$, and as a result (using (27)) we finally have

$$\widetilde{CD}^* = \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\frac{\xi_{t,n}^2 - (1-\theta_n)}{\sqrt{2}}\right)}{1-\theta_n} + o_p(1),$$

with $E\left(\xi_{t,n}^2\right) = 1 - \theta_n = \frac{1}{n}\sum_{i=1}^{n} a_{i,n}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(1 - \sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i\right)^2 > 0$, and

$$Var\left(\xi_{t,n}^2\right) = 2\left(\frac{1}{n}\sum_{i=1}^{n} a_{i,n}^2\right)^2 - \kappa_2\left(\frac{1}{n^2}\sum_{i=1}^{n} a_{i,n}^4\right),$$

where $\kappa_2 = E\left(\varepsilon_{it}^4\right) - 3$. But since $\frac{1}{n^2}\sum_{i=1}^n a_{i,n}^4 = O(n^{-1})$, then $Var\left(\xi_{t,n}^2\right) = 2\left(1-\theta_n\right)^2 + o(1)$, and

$$\widetilde{CD}^* = \frac{\widetilde{CD} + \sqrt{\frac{T}{2}}\theta_n}{1-\theta_n} + o_p\left(1\right)$$
$$= \frac{1}{\sqrt{T}}\sum_{t=1}^T \left(\frac{\xi_{t,n}^2 - E\left(\xi_{t,n}^2\right)}{\sqrt{Var\left(\xi_{t,n}^2\right)}}\right) + o_p\left(1\right).$$

Recalling that $\xi_{t,n}^2$ for $t = 1, 2, ..., T$ are distributed independently over t, then

$$\frac{1}{\sqrt{T}}\sum_{t=1}^T \left(\frac{\xi_{t,n}^2 - E\left(\xi_{t,n}^2\right)}{\sqrt{Var\left(\xi_{t,n}^2\right)}}\right) \rightarrow_d N(0,1).$$

Also by Lemma 9 in Section 2, $CD = \widetilde{CD} + o_p(1)$.Then it follows that

$$CD^*(\theta_n) = \frac{CD + \sqrt{\frac{T}{2}}\theta_n}{1-\theta_n} = \frac{\widetilde{CD} + \sqrt{\frac{T}{2}}\theta_n}{1-\theta_n} + o_p(1)$$
$$= \frac{1}{\sqrt{T}}\sum_{t=1}^T \left(\frac{\xi_{t,n}^2 - E\left(\xi_{t,n}^2\right)}{\sqrt{Var\left(\xi_{t,n}^2\right)}}\right) + o_p\left(1\right) \rightarrow_d N(0,1),$$

which establishes Theorem 1. ■

Proof of Corollary 1. Note that θ_n define by (31) can be written as $\theta_n = 2g_n - \boldsymbol{\varphi}_n'\mathbf{H}_n\boldsymbol{\varphi}_n$, where $g_n = n^{-1}\sum_{i=1}^n \sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i$, $\mathbf{H}_n = n^{-1}\sum_{i=1}^n \sigma_i\left(\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\right)$, $\boldsymbol{\varphi}_n = n^{-1}\sum_{i=1}^n \boldsymbol{\delta}_i$, and $\boldsymbol{\delta}_i = \boldsymbol{\gamma}_i/\sigma_i$. Similarly using (32) we have $\hat{\theta}_{nT} = 2\hat{g}_{nT} - \hat{\boldsymbol{\varphi}}_{nT}'\hat{\mathbf{H}}_{nT}\hat{\boldsymbol{\varphi}}_{nT}$, where $\hat{g}_{nT} = n^{-1}\sum_{i=1}^n \hat{\sigma}_{i,T}\hat{\boldsymbol{\varphi}}_{nT}'\hat{\boldsymbol{\gamma}}_i$, $\hat{\mathbf{H}}_{nT} = n^{-1}\sum_{i=1}^n \hat{\sigma}_{i,T}^2\left(\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}_i'\right)$, $\hat{\boldsymbol{\varphi}}_{nT} = n^{-1}\sum_{i=1}^n \hat{\boldsymbol{\delta}}_{i,nT}$, and $\hat{\boldsymbol{\delta}}_{i,nT} = \hat{\boldsymbol{\gamma}}_i/\hat{\sigma}_{i,T}$. Then

$$\sqrt{T}\left(\hat{\theta}_{nT} - \theta_n\right) = 2\sqrt{T}\left(\hat{g}_{nT} - g_n\right) - \sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT}'\hat{\mathbf{H}}_{nT}\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}_n'\mathbf{H}_n\boldsymbol{\varphi}_n\right). \tag{35}$$

Consider the first term of the above

$$2\sqrt{T}\left(\hat{g}_{nT} - g_n\right) = 2\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}_n\right)\left(\frac{1}{n}\sum_{i=1}^n \sigma_i\boldsymbol{\gamma}_i\right)$$
$$+ 2\sqrt{T}\left[\hat{\boldsymbol{\varphi}}_{nT}'\left(\frac{1}{n}\sum_{i=1}^n \hat{\sigma}_{i,T}\hat{\boldsymbol{\gamma}}_i - \frac{1}{n}\sum_{i=1}^n \sigma_i\boldsymbol{\gamma}_i\right)\right], \tag{36}$$

and since σ_i and $\boldsymbol{\gamma}_i$ are bounded then $n^{-1}\sum_{i=1}^n \sigma_i\boldsymbol{\gamma}_i = O_p(1)$. Also by (88) of Lemma 5 we have $\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}_n\right) = o_p(1)$, and hence the first term of the above is $o_p(1)$. To establish the probability order of the second term of (36), we first note that

$$2\sqrt{T}\left[\hat{\boldsymbol{\varphi}}_{nT}'\left(\frac{1}{n}\sum_{i=1}^n \hat{\sigma}_{i,T}\hat{\boldsymbol{\gamma}}_i - \frac{1}{n}\sum_{i=1}^n \sigma_i\boldsymbol{\gamma}_i\right)\right] = 2\sqrt{T}\left[\boldsymbol{\varphi}_n'\left(\frac{1}{n}\sum_{i=1}^n \hat{\sigma}_{i,T}\hat{\boldsymbol{\gamma}}_i - \frac{1}{n}\sum_{i=1}^n \sigma_i\boldsymbol{\gamma}_i\right)\right]$$
$$+ 2\sqrt{T}\left[\left(\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}_n\right)'\left(\frac{1}{n}\sum_{i=1}^n \hat{\sigma}_{i,T}\hat{\boldsymbol{\gamma}}_i - \frac{1}{n}\sum_{i=1}^n \sigma_i\boldsymbol{\gamma}_i\right)\right]. \tag{37}$$

But by (117) $\boldsymbol{\varphi}_n = O_p(1)$, and by (112) in Lemma 10 in Section 2 $n^{-1}\sum_{i=1}^{n}\left(\hat{\sigma}_{i,T}\hat{\boldsymbol{\gamma}}_i - \sigma_i\boldsymbol{\gamma}_i\right) = O_p\left(\delta_{nT}^{-2}\right)$, which also establishes that the second term of (37) is $o_p(1)$. Therefore overall we have

$$\sqrt{T}\left(\hat{g}_{nT} - g_n\right) = o_p(1). \tag{38}$$

Consider now the second term of (35) and note that

$$\begin{aligned}\sqrt{T}\left(\hat{\boldsymbol{\varphi}}'_{nT}\hat{\mathbf{H}}_{nT}\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}'_n\mathbf{H}_n\boldsymbol{\varphi}_n\right) &= \sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}_n\right)'\hat{\mathbf{H}}_{nT}\left(\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}_n\right) \\ &\quad + 2\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}_n\right)'\hat{\mathbf{H}}_{nT}\boldsymbol{\varphi}_n + \sqrt{T}\boldsymbol{\varphi}'_n\left(\hat{\mathbf{H}}_{nT} - \mathbf{H}_n\right)\boldsymbol{\varphi}_n, \end{aligned} \tag{39}$$

where $\hat{\mathbf{H}}_{nT} = \frac{1}{n}\sum_{i=1}^{n}\hat{\sigma}_{i,T}^2\left(\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}'_i\right)$, and

$$\begin{aligned}\sqrt{T}\left(\hat{\mathbf{H}}_{nT} - \mathbf{H}_n\right) &= \frac{\sqrt{T}}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}'_i - \boldsymbol{\gamma}_i\boldsymbol{\gamma}'_i\right) + \frac{\sqrt{T}}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_{i,T}^2 - \sigma_i^2\right)\left(\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}'_i - \boldsymbol{\gamma}_i\boldsymbol{\gamma}'_i\right) \\ &\quad + \frac{\sqrt{T}}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_{i,T}^2 - \omega_{i,T}^2\right)\boldsymbol{\gamma}_i\boldsymbol{\gamma}'_i + \frac{\sqrt{T}}{n}\sum_{i=1}^{n}\left(\omega_{i,T}^2 - \sigma_i^2\right)\boldsymbol{\gamma}_i\boldsymbol{\gamma}'_i \\ &= \sum_{j=1}^{4}\mathbf{D}_{j,nT}. \end{aligned} \tag{40}$$

The first two terms of (39) are $o_p(1)$, since $\|\boldsymbol{\varphi}_n\| < K$, and $\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}_n\right) = o_p(1)$, and $n^{-1}\sum_{i=1}^{n}\hat{\sigma}_{i,T}^2\left(\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}'_i\right) = O_p(1)$. To establish the probability order of the third term of (39), since $\|\boldsymbol{\varphi}_n\| < K$ it is sufficient to consider the four terms of $\sqrt{T}\left(\hat{\mathbf{H}}_{nT} - \mathbf{H}_n\right)$. It is clear that $\mathbf{D}_{2,nT}$ is dominated by $\mathbf{D}_{1,nT}$ and by (113) of Lemma 10, $\mathbf{D}_{1,nT} = \frac{\sqrt{T}}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}'_i - \boldsymbol{\gamma}_i\boldsymbol{\gamma}'_i\right) = O_p\left(\frac{\sqrt{T}}{\delta_{nT}^2}\right) = o_p(1)$. Using (56) of Lemma 2 in Section 2 and replacing b_{ni} with $\gamma_{ij}\gamma_{ij'}$ for $j, j' = 1, 2, \ldots, m_0$, it then follows that

$$\mathbf{D}_{3,nT} = \frac{\sqrt{T}}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_{i,T}^2 - \omega_{i,T}^2\right)\boldsymbol{\gamma}_i\boldsymbol{\gamma}'_i = O_p\left(\frac{\sqrt{T}}{\delta_{nT}^2}\right) = o_p(1).$$

Finally, denote the (j, j') element of $\mathbf{D}_{4,nT}$ by $d_{4,nT}(j, j')$ and note that

$$d_{4,nT}(j, j') = \frac{1}{n}\sum_{i=1}^{n}\left(\sigma_i^2\gamma_{ij}\gamma_{ij'}\right)\sqrt{T}\left(\frac{\boldsymbol{\varepsilon}'_i\mathbf{M}_F\boldsymbol{\varepsilon}_i}{T} - 1\right), \text{ for } j, j' = 1, 2, ..., m_0.$$

But under Assumptions 2 and 3, $|\sigma_i^2\gamma_{ij}\gamma_{ij'}| < K$, and $\sqrt{T}\left(T^{-1}\boldsymbol{\varepsilon}'_i\mathbf{M}_F\boldsymbol{\varepsilon}_i - 1\right)$, for $i = 1, 2, ..., n$ are identically and independently distributed across i, with mean $1/\sqrt{T}$ and a finite variance[1]. Then by standard law of large numbers, for each (j, j'), $d_{4,nT}(j, j') \to_p 0$, as n and $T \to \infty$, and hence we also have $\mathbf{D}_{4,nT} = o_p(1)$. Overall, $\hat{\mathbf{H}}_{nT} - \mathbf{H}_n = o_p(1)$, and we have $\sqrt{T}\left(\hat{\boldsymbol{\varphi}}'_{nT}\hat{\mathbf{H}}_{nT}\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}'_n\mathbf{H}_n\boldsymbol{\varphi}_n\right) = o_p(1)$. Using this result and (38) in (35) now yields $\sqrt{T}\left(\hat{\theta}_{nT} - \theta_n\right) = o_p(1)$, as required. ■

[1] The mean and variance of $\sqrt{T}\left(T^{-1}\boldsymbol{\varepsilon}'_i\mathbf{M}_F\boldsymbol{\varepsilon}_i - 1\right)$ can be obtained using (93) and (94) in Lemma 7 in Section 2.

Proof of Theorem 2. Recall from (35) that $CD^*\left(\hat{\theta}_{nT}\right)$ is given by

$$CD^*\left(\hat{\theta}_{nT}\right) = \frac{CD + \sqrt{\frac{T}{2}}\hat{\theta}_{nT}}{1 - \hat{\theta}_{nT}},$$

where $\hat{\theta}_{nT} = 1 - \frac{1}{n}\sum_{i=1}^{n} \hat{a}_{i,n}^2$, $\hat{a}_{i,n} = 1 - \hat{\sigma}_{i,T}\left(\boldsymbol{\varphi}_{nT}^{\prime}\hat{\boldsymbol{\gamma}}_i\right)$, and $\hat{\boldsymbol{\varphi}}_{nT} = n^{-1}\sum_{i=1}^{n} \hat{\boldsymbol{\gamma}}_i/\hat{\sigma}_{i,T}$, subject to the normalization $n^{-1}\sum_{i=1}^{n} \hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}_i^{\prime} = \mathbf{I}_{m_0}$. By Lemma 9 in Section 2 we have $CD = \widetilde{CD} + o_p(1)$. Then $CD^*(\hat{\theta}_{nT})$ can be written as (noting that $1 - \hat{\theta}_{nT} = \frac{1}{n}\sum_{i=1}^{n} \hat{a}_{i,n}^2 > 0$)

$$CD^*(\hat{\theta}_{nT}) = \frac{CD + \sqrt{\frac{T}{2}}\hat{\theta}_{nT}}{1 - \hat{\theta}_{nT}} = \frac{\widetilde{CD} + \sqrt{\frac{T}{2}}\hat{\theta}_{nT}}{1 - \hat{\theta}_{nT}} + o_p\left(1\right).$$

By result (34) of Corollary 1, $\sqrt{T}\left(\hat{\theta}_{nT} - \theta_n\right) = o_p(1)$, and hence

$$\begin{aligned} CD^*(\hat{\theta}_{nT}) &= \frac{\widetilde{CD} + \sqrt{\frac{T}{2}}\theta_n + \sqrt{\frac{T}{2}}\left(\hat{\theta}_{nT} - \theta_n\right)}{1 - \theta_n - \left(\hat{\theta}_{nT} - \theta_n\right)} + o_p(1) \\ &= \frac{\widetilde{CD} + \sqrt{\frac{T}{2}}\theta_n}{1 - \theta_n} + o_p\left(1\right) = CD^*(\theta_n) + o_p\left(1\right) \end{aligned}$$

However, by Theorem 1, $CD^*(\theta_n) \to_d N(0,1)$, which in turn establishes that $CD^*(\hat{\theta}_{nT}) \to_d N(0,1)$, considering that $CD^*(\hat{\theta}_{nT}) - CD^*(\theta_n) = o_p(1)$. ■

Proof of Theorem 3. Let $v_{it} = y_{it} - \boldsymbol{\beta}_i^{\prime}\mathbf{x}_{it}$, and $u_{it} = y_{it} - \boldsymbol{\beta}_i^{\prime}\mathbf{x}_{it} - \boldsymbol{\gamma}_i^{\prime}\mathbf{f}_t = v_{it} - \boldsymbol{\gamma}_i^{\prime}\mathbf{f}_t$, and consider the following two optimization problems

$$\min_{\boldsymbol{\Gamma},\boldsymbol{F}} \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(v_{it} - \boldsymbol{\gamma}_i^{\prime}\mathbf{f}_t\right)^2, \tag{41}$$

$$\min_{\boldsymbol{\Gamma},\boldsymbol{F}} \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{v}_{it} - \boldsymbol{\gamma}_i^{\prime}\mathbf{f}_t\right)^2, \tag{42}$$

where

$$\hat{v}_{it} = y_{it} - \hat{\boldsymbol{\beta}}_{CCE,i}^{\prime}\mathbf{x}_{it} = y_{it} - \boldsymbol{\beta}_i^{\prime}\mathbf{x}_{it} - \left(\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right)^{\prime}\mathbf{x}_{it} = v_{it} - \left(\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right)^{\prime}\mathbf{x}_{it}. \tag{43}$$

We need to show that solving problem (42) is asymptotically equivalent to solving problem (41). First, using the results in Pesaran (2006) and the fact that $\mathbf{x}_{it}$ is (stochastically) bounded[2],

$$\mathbf{x}_{it}^{\prime}\left(\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right) = O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{n}\right) + O_p\left(\frac{1}{\sqrt{nT}}\right), \tag{44}$$

[2]See equation (45) in Pesaran (2006).

then rewrite the criterion for (42) with (43),

$$\begin{aligned}
&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{v}_{it}-\boldsymbol{\gamma}_i'\mathbf{f}_t\right)^2=\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(v_{it}-\boldsymbol{\gamma}_i'\mathbf{f}_t-\left(\hat{\boldsymbol{\beta}}_{CCE,i}-\boldsymbol{\beta}_i\right)'\mathbf{x}_{it}\right)^2\\
&=\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(v_{it}-\boldsymbol{\gamma}_i'\mathbf{f}_t\right)^2+\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\beta}}_{CCE,i}-\boldsymbol{\beta}_i\right)'\mathbf{x}_{it}\mathbf{x}_{it}'\left(\hat{\boldsymbol{\beta}}_{CCE,i}-\boldsymbol{\beta}_i\right)\\
&-2\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(v_{it}-\boldsymbol{\gamma}_i'\mathbf{f}_t\right)\left(\hat{\boldsymbol{\beta}}_{CCE,i}-\boldsymbol{\beta}_i\right)'\mathbf{x}_{it}\\
&=A_{nT}+B_{nT}-2C_{nT}. \qquad (45)
\end{aligned}$$

Therefore, using (44) it is obvious $B_{nT}=O_p\left(\frac{1}{\sqrt{T}}\right)+O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{\sqrt{nT}}\right)$. Also consider the final term of (45) and note that by Cauchy Schwarz inequality,

$$\begin{aligned}
|C_{nT}|&=\left|\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{u}_{it}\left(\hat{\boldsymbol{\beta}}_{CCE,i}-\boldsymbol{\beta}_i\right)'\mathbf{x}_{it}\right|\\
&\leq\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{u}_{it}^2\right)^{1/2}\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\left\|\left(\hat{\boldsymbol{\beta}}_{CCE,i}-\boldsymbol{\beta}_i\right)'\mathbf{x}_{it}\right\|^2\right)^{1/2},
\end{aligned}$$

where $\tilde{u}_{it}=v_{it}-\boldsymbol{\gamma}_i'\mathbf{f}_t=y_{it}-\boldsymbol{\beta}_i'\mathbf{x}_{it}-\boldsymbol{\gamma}_i'\mathbf{f}_t$. Since in both optimization problems $\boldsymbol{\gamma}_i$ and $\mathbf{f}_t$ are only identified up to $m_0\times m_0$ rotation matrices, $\tilde{u}_{it}$ and u_{it} have similar properties and we also have $C_{nT}=O_p\left(\frac{1}{\sqrt{T}}\right)+O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{\sqrt{nT}}\right)$. Overall we have

$$\begin{aligned}
\min_{\boldsymbol{\Gamma},\boldsymbol{F}}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(v_{it}-\boldsymbol{\gamma}_i'\mathbf{f}_t\right)^2\equiv&\min_{\boldsymbol{\Gamma},\boldsymbol{F}}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{v}_{it}-\boldsymbol{\gamma}_i'\mathbf{f}_t\right)^2\\
&+O_p\left(\frac{1}{\sqrt{T}}\right)+O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{\sqrt{nT}}\right).
\end{aligned}$$

Hence, PCs based on $\hat{v}_{it}$ are asymptotically equivalent to those based on v_{it}. The remaining proof of Theorem 3 can follow from the proof of Theorem 2. ■

2 Statement and Proofs of the Lemmas

This section provides auxiliary lemmas and the associated proofs, which are required to establish the main results of the paper. Throughout $\delta_{nT}=min(\sqrt{n},\sqrt{T})$.

Lemma 1 *Suppose that Assumptions 1-3 hold, and the latent factors, $\mathbf{f}_t$, and their loadings, $\boldsymbol{\gamma}_i$, in model (2) are estimated by principal components, $\hat{\mathbf{f}}_t$ and $\hat{\boldsymbol{\gamma}}_i$, given by (8). Then*

$$\left\|\hat{\mathbf{F}}-\mathbf{F}\right\|_F = O_p\left(\frac{\sqrt{T}}{\delta_{nT}}\right), \tag{46}$$

$$\left\|\hat{\mathbf{\Gamma}}-\mathbf{\Gamma}\right\|_F = O_p\left(\frac{\sqrt{n}}{\delta_{nT}}\right), \tag{47}$$

$$\left\|\mathbf{U}'(\hat{\mathbf{F}}-\mathbf{F})\right\|_F = O_p\left(\frac{\sqrt{nT}}{\delta_{nT}}\right), \tag{48}$$

$$\left\|\mathbf{\Gamma}'(\hat{\mathbf{\Gamma}}-\mathbf{\Gamma})\right\| = O_p\left(\frac{n}{\delta_{nT}}\right), \tag{49}$$

$$\left\|\mathbf{F}'(\hat{\mathbf{F}}-\mathbf{F})\right\| = O_p\left(\frac{T}{\delta_{nT}}\right), \tag{50}$$

$$\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{F} = O_p\left(\frac{T}{\delta_{nT}^2}\right), \tag{51}$$

$$\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\hat{\mathbf{F}} = O_p\left(\frac{T}{\delta_{nT}^2}\right), \tag{52}$$

$$\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{u}_i = O_p\left(\frac{T}{\delta_{nT}^2}\right), \tag{53}$$

$$(\hat{\mathbf{\Gamma}}-\mathbf{\Gamma})'\mathbf{u}_t = O_p\left(\frac{n}{\delta_{nT}^2}\right), \tag{54}$$

where $\mathbf{\Gamma} = (\gamma_1, \gamma_2, ..., \gamma_n)'$,$\hat{\mathbf{\Gamma}} = (\hat{\gamma}_1, \hat{\gamma}_2, ..., \hat{\gamma}_n)'$, $\mathbf{u}_i = (u_{i1}, u_{i2}, ..., u_{iT})'$, $\mathbf{u}_t = (u_{1t}, u_{2t}, ..., u_{nt})'$, *and* $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n)$.

Proof. Since Assumptions 1-3 are a sub-set of assumptions made by Bai (2003), results (46) to (49), (51) and (53) follow directly from Lemmas B1, B2 and B3, and Theorem 2 of Bai (2003). The remaining two results, (50) and (54), can be established analogously. ■

Lemma 2 *Consider* $\hat{\sigma}_{i,T}^2 = T^{-1}\mathbf{e}_i'\mathbf{e}_i$*, the estimator of the* σ_i^2*, the error variance of the* i^{th} *unit of the latent factor model, (2), where* $\mathbf{e}_i = (e_{i1}, e_{i2}, ..., e_{iT})'$ *is the principal component estimator of* $\mathbf{u}_i = (u_{i1}, u_{i2}, ..., u_{iT})'$*, namely* $\mathbf{e}_i = \mathbf{M}_{\hat{F}}\mathbf{y}_i$*, where* $\mathbf{M}_{\hat{F}} = \mathbf{I}_T - \hat{\mathbf{F}}(\hat{\mathbf{F}}'\hat{\mathbf{F}})^{-1}\hat{\mathbf{F}}'$*,* $\mathbf{y}_i =$

$(y_{i1}, y_{i2}, ..., y_{iT}\ )'$, *and* $\hat{\mathbf{F}}$ *is given by (8). Suppose that Assumptions 1-3 hold. Then*

$$\hat{\sigma}^2_{i,T} - \omega^2_{i,T} = O_p\left(\frac{1}{\delta^2_{nT}}\right), \tag{55}$$

$$\frac{1}{n}\sum_{i=1}^{n} b_{in}\left(\hat{\sigma}^2_{i,T} - \omega^2_{i,T}\right) = O_p\left(\frac{1}{\delta^2_{nT}}\right), \tag{56}$$

$$\hat{\sigma}_{i,T} - \omega_{i,T} = O_p\left(\frac{1}{\delta^2_{nT}}\right), \tag{57}$$

$$\frac{1}{n}\sum_{i=1}^{n} b_{in}\left(\hat{\sigma}_{i,T} - \omega_{i,T}\right) = O_p\left(\frac{1}{\delta^2_{nT}}\right), \tag{58}$$

$$\frac{1}{\hat{\sigma}_{i,T}} - \frac{1}{\omega_{i,T}} = O_p\left(\frac{1}{\delta^2_{nT}}\right), \tag{59}$$

$$\frac{1}{n}\sum_{i=1}^{n} b_{in}\left(\frac{1}{\hat{\sigma}_{i,T}} - \frac{1}{\omega_{i,T}}\right) = O_p\left(\frac{1}{\delta^2_{nT}}\right), \tag{60}$$

where $\omega^2_{i,T} = T^{-1}\mathbf{u}_i'\mathbf{M}_F\mathbf{u}_i$, $\mathbf{M}_F = \mathbf{I}_T - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$, $\delta^2_{nT} = min(n,T)$, *and* $\{b_{in}\}_{i=1}^{n}$ *is a sequence of fixed bounded constants such that* $n^{-1}\sum_{i=1}^{n} b^2_{in} = O(1)$.

Proof. We first note that

$$\begin{aligned}\mathbf{e}_i &= \mathbf{M}_{\hat{F}}\mathbf{y}_i = \mathbf{M}_{\hat{\mathbf{F}}}\left(\mathbf{F}\boldsymbol{\gamma}_i + \mathbf{u}_i\right)\\ &= \mathbf{M}_F\mathbf{u}_i + \left(\mathbf{M}_{\hat{F}} - \mathbf{M}_F\right)\mathbf{u}_i + \mathbf{M}_{\hat{F}}\mathbf{F}\boldsymbol{\gamma}_i,\end{aligned}$$

which yields the following error variance decomposition

$$\begin{aligned}\hat{\sigma}^2_{i,T} &= \frac{\mathbf{u}_i'\mathbf{M}_F\mathbf{u}_i}{T} + \frac{\boldsymbol{\gamma}_i'\mathbf{F}'\mathbf{M}_{\hat{F}}\mathbf{F}\boldsymbol{\gamma}_i}{T} + \frac{\mathbf{u}_i'\left(\mathbf{M}_{\hat{F}} - \mathbf{M}_F\right)\left(\mathbf{M}_{\hat{F}} - \mathbf{M}_F\right)\mathbf{u}_i}{T}\\ &+ \frac{2\mathbf{u}_i'\mathbf{M}_F\left(\mathbf{M}_{\hat{F}} - \mathbf{M}_F\right)\mathbf{u}_i}{T} + \frac{2\mathbf{u}_i'\mathbf{M}_F\mathbf{M}_{\hat{F}}\mathbf{F}\boldsymbol{\gamma}_i}{T} + \frac{2\mathbf{u}_i'\left(\mathbf{M}_{\hat{F}} - \mathbf{M}_F\right)\mathbf{M}_{\hat{F}}\mathbf{F}\boldsymbol{\gamma}_i}{T}\\ &= \sum_{j=1}^{6} B_{j,iT}.\end{aligned} \tag{61}$$

Starting with the second term, we note that

$$\left\|B_{2,iT}\right\| = \left\|\frac{\boldsymbol{\gamma}_i'\left(\mathbf{F} - \hat{\mathbf{F}}\right)'\mathbf{M}_{\hat{\mathbf{F}}}\left(\mathbf{F} - \hat{\mathbf{F}}\right)\boldsymbol{\gamma}_i}{T}\right\| \leq \left\|\boldsymbol{\gamma}_i\right\|^2 \frac{\left\|\mathbf{F} - \hat{\mathbf{F}}\right\|^2}{T}\left\|\mathbf{M}_{\hat{\mathbf{F}}}\right\|,$$

where $\|\boldsymbol{\gamma}_i\|$ is bounded by Assumption 3 and $\|\mathbf{M}_{\hat{F}}\| = 1$. Then using (46) it follows that $\|B_{2,iT}\| = O_p\left(\delta^{-2}_{nT}\right)$. Before establishing the probability order of the remaining term $B_{3,iT}$ we first observe that

$$\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T} - \frac{\mathbf{F}'\mathbf{F}}{T} = \frac{\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\left(\hat{\mathbf{F}} - \mathbf{F}\right)}{T} + \frac{\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\mathbf{F}}{T} + \frac{\mathbf{F}'\left(\hat{\mathbf{F}} - \mathbf{F}\right)}{T}.$$

Then using results (46) and (50) it follows that

$$\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T} = \frac{\mathbf{F}'\mathbf{F}}{T} + O_p\left(\frac{1}{\delta_{nT}^2}\right), \text{ and } \left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1} = \left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1} + O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{62}$$

and in consequence (given that by assumption $T^{-1}\mathbf{F}'\mathbf{F}$ is a positive definite matrix)

$$\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T} = O_p(1), \text{ and } \left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1} = O_p(1). \tag{63}$$

Consider now $B_{3,iT}$, and note that

$$\begin{aligned}
&(\mathbf{M}_{\hat{F}} - \mathbf{M}_F)(\mathbf{M}_{\hat{F}} - \mathbf{M}_F) \\
&= \left[\hat{\mathbf{F}}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}' - \mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}'\right] + \left(\mathbf{I}_m - \hat{\mathbf{F}}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}'\right)\mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}' \\
&+ \mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}'\left(\mathbf{I}_m - \hat{\mathbf{F}}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}'\right),
\end{aligned}$$

where the term in the first brackets can be furthermore decomposed as

$$\begin{aligned}
\hat{\mathbf{F}}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}' - \mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}' &= \left(\hat{\mathbf{F}} - \mathbf{F}\right)\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\left(\hat{\mathbf{F}} - \mathbf{F}\right)' + \left[\mathbf{F}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\mathbf{F}' - \mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}'\right] \\
&+ \mathbf{F}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\left(\hat{\mathbf{F}} - \mathbf{F}\right)' + \left(\hat{\mathbf{F}} - \mathbf{F}\right)\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\mathbf{F}'.
\end{aligned}$$

Hence we have

$$B_{3,iT} = T^{-1}\mathbf{u}_i'(\mathbf{M}_{\hat{F}} - \mathbf{M}_F)(\mathbf{M}_{\hat{F}} - \mathbf{M}_F)\mathbf{u}_i = \sum_{s=1}^{6} C_{s,iT}, \tag{64}$$

where

$$\begin{aligned}
C_{1,iT} &= \frac{1}{T}\mathbf{u}_i'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\mathbf{u}_i, \\
C_{2,iT} &= \frac{1}{T}\mathbf{u}_i'\left[\mathbf{F}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\mathbf{F}' - \mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}'\right]\mathbf{u}_i, \\
C_{3,iT} &= \frac{1}{T}\mathbf{u}_i'\mathbf{F}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\mathbf{u}_i, \\
C_{4,iT} &= C_{3,iT} = \frac{1}{T}\mathbf{u}_i'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\mathbf{F}'\mathbf{u}_i, \\
C_{5,iT} &= \frac{1}{T}\mathbf{u}_i'\left(\mathbf{I}_m - \hat{\mathbf{F}}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}'\right)\mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}'\mathbf{u}_i, \\
C_{6,iT} &= \frac{1}{T}\mathbf{u}_i'\mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}'\left(\mathbf{I}_m - \hat{\mathbf{F}}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}'\right)\mathbf{u}_i.
\end{aligned}$$

Consider $C_{1,iT}$ and note that

$$\|C_{1,iT}\| \le \frac{\|\mathbf{u}_i\|^2}{T}\left\|\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\right\|\left(\frac{\left\|\hat{\mathbf{F}} - \mathbf{F}\right\|^2}{T}\right).$$

The first two terms are bounded in probability, since $T^{-1}\left\|\mathbf{u}_i\right\|^2 = T^{-1}\sum_{t=1}^{T}u_{it}^2 = \sigma_i^2 + O_p\left(T^{-1/2}\right)$, and given the results in (63). Using (46) it now follows that $\left\|C_{1,iT}\right\| = O_p\left(\delta_{nT}^{-2}\right)$. To establish the order of $C_{2,iT}$, we note that

$$\begin{aligned}
&\mathbf{F}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\mathbf{F}' - \mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}' \\
&= \mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\left[\mathbf{F}'\mathbf{F} - \hat{\mathbf{F}}'\hat{\mathbf{F}}\right]\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\mathbf{F}' \\
&= \mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\left[\mathbf{F}'\mathbf{F} - \left(\hat{\mathbf{F}} - \mathbf{F} + \mathbf{F}\right)'\left(\hat{\mathbf{F}} - \mathbf{F} + \mathbf{F}\right)\right]\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\mathbf{F}' \\
&= \mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\left[-\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\left(\hat{\mathbf{F}} - \mathbf{F}\right) - \left(\hat{\mathbf{F}} - \mathbf{F}\right)'\mathbf{F} - \mathbf{F}'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\right]\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\mathbf{F}'.
\end{aligned}$$

Then we have

$$C_{2,iT} = \frac{\mathbf{u}_i'\mathbf{F}}{\sqrt{T}}\left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1}\left[-\frac{\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\left(\hat{\mathbf{F}} - \mathbf{F}\right)}{T} - \frac{\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\mathbf{F}}{T} - \frac{\mathbf{F}'\left(\hat{\mathbf{F}} - \mathbf{F}\right)}{T}\right]\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\frac{\mathbf{F}'\mathbf{u}_i}{\sqrt{T}},$$

and hence

$$\left\|C_{2,iT}\right\| \le \left\|\frac{\mathbf{u}_i'\mathbf{F}}{\sqrt{T}}\right\|^2\left\|\left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1}\right\|\left\|\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\right\|\left[\frac{\left\|\hat{\mathbf{F}} - \mathbf{F}\right\|^2}{T} + \frac{2\left\|\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\mathbf{F}\right\|}{T}\right].$$

Under Assumption 2, $T^{-1/2}\mathbf{u}_i'\mathbf{F} = O_p(1)$, and using results (46) and (50) it follows that $\left\|C_{2,iT}\right\| = O_p\left(\frac{1}{\delta_{nT}^2}\right)$. Also, by result (53), $\frac{\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{u}_i}{\sqrt{T}} = O_p\left(\frac{\sqrt{T}}{\delta_{nT}^2}\right)$, and therefore

$$\left\|C_{3,iT}\right\| = \left\|C_{4,iT}\right\| \le \frac{1}{T}\left\|\frac{\mathbf{u}_i\mathbf{F}}{\sqrt{T}}\right\|\left\|\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\right\|\left\|\frac{\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\mathbf{u}_i}{\sqrt{T}}\right\| = O_p\left(\frac{1}{\sqrt{T}\delta_{nT}}\right).$$

To establish the order of $C_{5,iT}$, note that

$$\begin{aligned}
&\mathbf{u}_i'\hat{\mathbf{F}}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}'\mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}'\mathbf{u}_i \\
&= \left[\mathbf{u}_i'\left(\hat{\mathbf{F}} - \mathbf{F}\right) + \mathbf{u}_i'\mathbf{F}\right]\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\left[\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\mathbf{F} + \mathbf{F}'\mathbf{F}\right]\left(\mathbf{F}'\mathbf{F}\right)^{-1}\left(\mathbf{F}'\mathbf{u}_i\right),
\end{aligned}$$

and therefore

$$\begin{aligned}
C_{5,iT} = {} & \frac{\mathbf{u}_i'\mathbf{F}}{T}\left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1}\frac{\mathbf{F}'\mathbf{u}_i}{T} \\
& - \left[\frac{\mathbf{u}_i'\left(\hat{\mathbf{F}} - \mathbf{F}\right)}{T} + \frac{\mathbf{u}_i'\mathbf{F}}{T}\right]\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\left[\frac{\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\mathbf{F}}{T} + \frac{\mathbf{F}'\mathbf{F}}{T}\right]\left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1}\left(\frac{\mathbf{F}'\mathbf{u}_i}{T}\right),
\end{aligned}$$

hence

$$
\begin{aligned}
\left\|C_{5,iT}\right\| \leq \; & \frac{1}{T}\left\|\frac{\mathbf{u}_i'\mathbf{F}}{\sqrt{T}}\right\|^2\left\|\left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1}\right\| \\
& + \left[\left\|\frac{\mathbf{u}_i'\left(\hat{\mathbf{F}}-\mathbf{F}\right)}{T}\right\| + \left\|\frac{\mathbf{u}_i'\mathbf{F}}{T}\right\|\right]\left\|\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\right\| \\
& \left[\left\|\frac{\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{F}}{T}\right\| + \left\|\frac{\mathbf{F}'\mathbf{F}}{T}\right\|\right]\left\|\left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{u}_i'\mathbf{F}}{T}\right\|,
\end{aligned}
$$

and it follows that $\left\|C_{5,iT}\right\| = O_p\left(T^{-1}\right)$. Similarly, the probability order of $C_{6,iT}$ is also established to be $O_p\left(T^{-1}\right)$. Overall, $B_{3,iT} = O_p\left(\frac{1}{\delta_{nT}^2}\right)$. Consider now the fourth term of (61),

$$
\begin{aligned}
B_{4,iT} &= \frac{\mathbf{u}_i'\left(\mathbf{I}_m-\mathbf{P}_F\right)\left(\mathbf{P}_F-\mathbf{P}_{\hat{F}}\right)\mathbf{u}_i}{T} = \frac{\mathbf{u}_i'\left(\mathbf{P}_F-\mathbf{P}_{\hat{F}}+\mathbf{P}_F\mathbf{P}_{\hat{F}}\right)\mathbf{u}_i}{T} \\
&= \frac{\mathbf{u}_i'\left(\mathbf{P}_F-\mathbf{P}_{\hat{F}}\right)\mathbf{u}_i}{T} - \frac{\mathbf{u}_i'\mathbf{P}_F\mathbf{u}_i}{T} + \frac{\mathbf{u}_i'\mathbf{P}_F\mathbf{P}_{\hat{F}}\mathbf{u}_i}{T},
\end{aligned} \tag{65}
$$

where $\mathbf{P}_F = \mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}'$, and $\mathbf{P}_{\hat{F}} = \hat{\mathbf{F}}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)\hat{\mathbf{F}}'$. The order of the first term of (65) is the same as that of (64), namely $O_p\left(\delta_{nT}^{-2}\right)$. Since $\mathbf{F}$ is distributed independently of $\mathbf{u}_i$, using (63), then it readily follows that the second term is $O_p\left(T^{-1}\right)$. The third term of (65) can be written as

$$
\begin{aligned}
& \left\|\frac{\mathbf{u}_i'\mathbf{F}}{T}\left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1}\left(\frac{\mathbf{F}'\hat{\mathbf{F}}}{T}\right)\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\frac{\hat{\mathbf{F}}'\mathbf{u}_i}{T}\right\| \\
&= \left\|\frac{1}{\sqrt{T}}\left(\frac{\mathbf{u}_i'\mathbf{F}}{\sqrt{T}}\right)\left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1}\left(\frac{\mathbf{F}'\mathbf{F}}{T}+\frac{\mathbf{F}'\left(\hat{\mathbf{F}}-\mathbf{F}\right)}{T}\right)\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\left(\frac{\mathbf{F}'\mathbf{u}_i}{T}+\frac{\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{u}_i}{T}\right)\right\| \\
&\leq \frac{1}{\sqrt{T}}\left\|\frac{\mathbf{u}_i'\mathbf{F}}{\sqrt{T}}\right\|\left\|\left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1}\right\|\left(\left\|\frac{\mathbf{F}'\mathbf{F}}{T}\right\|+\left\|\frac{\mathbf{F}'\left(\hat{\mathbf{F}}-\mathbf{F}\right)}{T}\right\|\right)\left\|\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\right\| \\
& \left(\left\|\frac{\mathbf{F}'\mathbf{u}_i}{T}\right\|+\left\|\frac{\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{u}_i}{T}\right\|\right) = \frac{1}{\sqrt{T}}O_p(1)\times O_p\left(\frac{1}{\sqrt{T}}\right) = O_p\left(\frac{1}{T}\right).
\end{aligned}
$$

Consider now $B_{5,iT}$ and note that

$$\frac{\boldsymbol{\gamma}_i'\mathbf{F}'\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{M}_F\mathbf{u}_i}{T} = -\frac{\boldsymbol{\gamma}_i'\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{M}_{\hat{F}}\mathbf{u}_i}{T} - \frac{\boldsymbol{\gamma}_i'\mathbf{F}'\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}'\mathbf{u}_i}{T}$$
$$= -\frac{\boldsymbol{\gamma}_i'\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{u}_i}{T} + \frac{\boldsymbol{\gamma}_i'\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\hat{\mathbf{F}}}{T}\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\frac{\hat{\mathbf{F}}'\mathbf{u}_i}{T}$$
$$+\boldsymbol{\gamma}_i'\frac{\mathbf{F}'\mathbf{M}_{\hat{\mathbf{F}}}\left(\hat{\mathbf{F}}-\mathbf{F}\right)}{T}\left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1}\frac{\mathbf{F}'\mathbf{u}_i}{T}. \tag{66}$$

But, using results in Lemma 1, we have

$$\left\|\frac{\boldsymbol{\gamma}_i'\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{u}_i}{T}\right\| \leq \|\boldsymbol{\gamma}_i\|\left\|\frac{\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{u}_i}{T}\right\| = O_p\left(\frac{1}{\delta_{nT}^2}\right),$$

$$\left\|\frac{\boldsymbol{\gamma}_i'\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\hat{\mathbf{F}}}{T}\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\frac{\hat{\mathbf{F}}'\mathbf{u}_i}{T}\right\| \leq \|\boldsymbol{\gamma}_i\|\left\|\frac{\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\hat{\mathbf{F}}}{T}\right\|\left\|\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\right\|\left\|\frac{\hat{\mathbf{F}}'\mathbf{u}_i}{T}\right\|$$
$$= O_p\left(\frac{1}{\delta_{nT}^2\sqrt{T}}\right),$$

and

$$\left\|\boldsymbol{\gamma}_i'\frac{\mathbf{F}'\mathbf{M}_{\hat{F}}\left(\hat{\mathbf{F}}-\mathbf{F}\right)}{T}\left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1}\frac{\mathbf{F}'\mathbf{u}_i}{T}\right\| \leq \|\boldsymbol{\gamma}_i\|\frac{\left\|\hat{\mathbf{F}}-\mathbf{F}\right\|^2}{T}\left\|\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}'\mathbf{u}_i}{T}\right\|$$
$$= O_p\left(\frac{1}{\delta_{nT}^2\sqrt{T}}\right).$$

Thus, $B_{5,iT} = T^{-1}\boldsymbol{\gamma}_i'\mathbf{F}'\mathbf{M}_{\hat{F}}\mathbf{M}_F\mathbf{u}_i = O_p\left(\delta_{nT}^{-2}\right)$. Similarly $B_{6,iT} = O_p\left(\delta_{nT}^{-2}\right)$. The above results now establish (55). Result (56) can be obtained similarly, either by directly considering the weighted average of (55) with weights b_{in}, or by noting that $\hat{\sigma}_{i,T}^2$ and $\omega_{i,T}^2$ are both integrable processes and the probability order of the average will be the same as the probability order of the underlying units. Results (57) and (59) follow from (55), noting that under Assumption 2 there exists T_0 such that for all $T > T_0$, $\hat{\sigma}_{i,T}^2 > c > 0$ and $\omega_{iT}^2 > c > 0$. Furthermore, $\hat{\sigma}_{i,T} + \omega_{i,T} = O_p(1)$ and $\omega_{i,T}\hat{\sigma}_{i,T} = O_p(1)$. More specifically, to establish (57) note that $|\hat{\sigma}_{i,T} - \omega_{i,T}| \leq \left|\hat{\sigma}_{i,T}^2 - \omega_{i,T}^2\right| / \left(\hat{\sigma}_{i,T} + \omega_{i,T}\right) \leq c^{-1}\left|\hat{\sigma}_{i,T}^2 - \omega_{i,T}^2\right|$, and hence by (55) we have $|\hat{\sigma}_{i,T} - \omega_{i,T}| = O_p\left(\delta_{nT}^{-2}\right)$. Similarly, result (59) is established noting that

$$\left|\frac{1}{\hat{\sigma}_{i,T}} - \frac{1}{\omega_{i,T}}\right| \leq \frac{|\hat{\sigma}_{i,T} - \omega_{i,T}|}{\omega_{i,T}\hat{\sigma}_{i,T}} \leq c^{-1}|\hat{\sigma}_{i,T} - \omega_{i,T}| = O_p\left(\frac{1}{\delta_{nT}^2}\right). \tag{67}$$

Finally, results (58) and (60) can be obtained, respectively, in a similar way to the proof of (56), since under Assumption 2, $\hat{\sigma}_{i,T}$, $\omega_{i,T}$, $\hat{\sigma}_{i,T}^{-1}$ and $\omega_{i,T}^{-1}$ are also integrable processes. ■

Lemma 3 *Suppose that the latent factors, $\mathbf{f}_t$, and their loadings, $\boldsymbol{\gamma}_i$, in model (2) are estimated by principal components, $\hat{\mathbf{F}}$ and $\hat{\boldsymbol{\gamma}}_i$, given by (8). Then under Assumptions 1-3 with n and $T \to \infty$, such that $n/T \to \kappa$, for $0 < \kappa < \infty$, we have*

$$\mathbf{d}_{1,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} b_{in}(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) = O_p\left(\frac{1}{\delta_{nT}}\right), \tag{68}$$

$$\mathbf{d}_{2,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\omega_{i,T} - \sigma_i)(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) = O_p\left(\frac{1}{\delta_{nT}}\right), \tag{69}$$

$$\mathbf{d}_{3,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{1}{\omega_{iT}} - \frac{1}{\sigma_i}\right)(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) = O_p\left(\frac{1}{\delta_{nT}}\right), \tag{70}$$

$$\mathbf{d}_{4,nT} = \frac{1}{n} \sum_{i=1}^{n} (\hat{\sigma}_{i,T} - \omega_{i,T})(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{71}$$

$$\mathbf{d}_{5,nT} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{\hat{\sigma}_{i,T}} - \frac{1}{\omega_{i,T}}\right)(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) = O_p\left(\frac{1}{\delta_{nT}^2}\right) \tag{72}$$

$$\mathbf{d}_{6,nT} = \frac{1}{n} \sum_{i=1}^{n} (\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}) = O_p\left(\frac{1}{\delta_{nT}}\right), \tag{73}$$

$$\mathbf{d}_{7,nT} = \frac{1}{n} \sum_{i=1}^{n} b_{in}(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i)\boldsymbol{\gamma}_i' = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{74}$$

where $\{b_{in}\}_{i=1}^{n}$ is a sequence of fixed values bounded in n, such that $n^{-1}\sum_{i=1}^{n} b_{in}^2 = O(1)$, $\boldsymbol{\delta}_{i,T} = \boldsymbol{\gamma}_i/\omega_{iT}$, $\hat{\boldsymbol{\delta}}_{i,T} = \hat{\boldsymbol{\gamma}}_i/\omega_{iT}$, and $\omega_{iT} = T^{-1}\mathbf{u}_i'\mathbf{M}_F\mathbf{u}_i$.

Proof. Note that in general

$$\begin{aligned}
\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i &= \left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\left(\frac{\hat{\mathbf{F}}'\mathbf{F}\boldsymbol{\gamma}_i}{T} + \frac{\hat{\mathbf{F}}'\mathbf{u}_i}{T}\right) - \left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)\boldsymbol{\gamma}_i \\
&= -\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\left[\frac{\hat{\mathbf{F}}'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\boldsymbol{\gamma}_i}{T}\right] + \left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\left(\frac{\hat{\mathbf{F}}'\mathbf{u}_i}{T}\right),
\end{aligned} \tag{75}$$

and we have

$$\begin{aligned}
\mathbf{d}_{1,nT} &= n^{-1/2}\sum_{i=1}^{n} b_{in}(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) \\
&= -\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\left[\frac{\sqrt{n}}{T}\hat{\mathbf{F}}'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\right]\left(\frac{1}{n}\sum_{i=1}^{n} b_{in}\boldsymbol{\gamma}_i\right) + \left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1} T^{-1}\left(n^{-1/2}\sum_{i=1}^{n} b_{in}\hat{\mathbf{F}}'\mathbf{u}_i\right).
\end{aligned} \tag{76}$$

By result (63) $\left(T^{-1}\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1} = O_p(1)$, and by result (50) we have

$$\frac{\sqrt{n}}{T}\hat{\mathbf{F}}'\left(\hat{\mathbf{F}} - \mathbf{F}\right) = O_p\left(\frac{\sqrt{n}T}{T\delta_{nT}^2}\right) = O_p\left(\frac{\sqrt{n}}{\min(n, T)}\right) = O_p\left(\frac{1}{\delta_{nT}}\right). \tag{77}$$

Also since by assumption $\|\boldsymbol{\gamma}_i\| < K$, and

$$\left\| \frac{1}{n} \sum_{i=1}^{n} b_{in} \boldsymbol{\gamma}_i \right\| \leq \left(\frac{1}{n} \sum_{i=1}^{n} b_{in}^2 \right)^{1/2} \left\| \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\gamma}_i \boldsymbol{\gamma}_i^{'} \right\|^{1/2}$$
$$\leq \left(\frac{1}{n} \sum_{i=1}^{n} b_{in}^2 \right)^{1/2} \left(\frac{1}{n} \sum_{i=1}^{n} \|\boldsymbol{\gamma}_i\|^2 \right)^{1/2} < K.$$

Hence, the first term of (76) is $O_p\left(\delta_{nT}^{-1}\right)$. For the second term of (76), since $\left(T^{-1}\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1} = O_p(1)$, we note that

$$T^{-1}\left(\hat{\mathbf{F}} - \mathbf{F} + \mathbf{F}\right)' \left(n^{-1/2} \sum_{i=1}^{n} b_{in} \mathbf{u}_i\right) = n^{-1/2} \sum_{i=1}^{n} b_{in} \left(\frac{\hat{\mathbf{F}} - \mathbf{F}}{T}\right)' \mathbf{u}_i + \frac{1}{T\sqrt{n}} \sum_{i=1}^{n} b_{in} \mathbf{F}^{'} \mathbf{u}_i.$$

It is clear that the first term is dominated by the second term, and under Assumptions 2, we have

$$\frac{1}{T\sqrt{n}} \sum_{i=1}^{n} b_{in} \mathbf{F}^{'} \mathbf{u}_i = \left(\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} b_{in} \mathbf{f}_t u_{it} \right) \frac{1}{\sqrt{T}} = O_p\left(\frac{1}{\sqrt{T}}\right). \tag{78}$$

Result (68) now follows using (77) and (78) in (76), and noting that by assumption n and T are of the same order. To prove (69) we first write it as

$$\mathbf{d}_{2,nT} = \left(\sqrt{\frac{n}{T}}\right) \frac{1}{n} \sum_{i=1}^{n} q_{iT} \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right),$$

where

$$q_{iT} = \sqrt{T}\left(\omega_{i,T} - \sigma_i\right) = \sigma_i \sqrt{T} \left[\left(\frac{\boldsymbol{\varepsilon}_i' \mathbf{M}_F \boldsymbol{\varepsilon}_i}{T} \right)^{1/2} - 1 \right],$$

and conditional on $\mathbf{F}$ and σ_i, q_{iT} are independently distributed across i. Using results in Lemma 7 it is easily seen that $E\left(q_{iT}\right) = O(T^{-1/2})$ and $Var\left(q_{iT}\right) = O(1)$, and hence $n^{-1}\sum_{i=1}^{n} q_{iT}^2 = O_p(1)$. Also by Cauchy-Schwarz inequality we have

$$|\mathbf{d}_{2,nT}| \leq \left(\sqrt{\frac{n}{T}}\right) \left(n^{-1} \sum_{i=1}^{n} q_{iT}^2 \right)^{1/2} \left(n^{-1/2} \left\| \hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma} \right\| \right),$$

where $T^{-1}n = \ominus(1)$, and by (47) $n^{-1/2}\left\|\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma}\right\| = O_p(\delta_{nT}^{-1})$, and (69) is established. Result (70) follows similarly, with q_{iT} re-defined as $q_{iT} = \sigma_i^{-1}\sqrt{T}\left[\left(\frac{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i}{T}\right)^{-1/2} - 1\right]$, and noting that $sup_i(1/\sigma_i^2) < K$, and using results in Lemma 7. To establish (71), using Cauchy-Schwarz inequality, we have[3]

$$|\mathbf{d}_{4,nT}| \leq \left[n^{-1} \sum_{i=1}^{n} \left(\hat{\sigma}_{i,T} - \omega_{i,T}\right)^2 \right]^{1/2} \left(n^{-1/2} \left\| \hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma} \right\| \right),$$

[3] The proofs of (69) and (70) are different from that of (71) and (72) due to the fact that $\omega_{iT}^2 - \sigma_i^2 = O_p\left(T^{-1/2}\right)$, but $\hat{\sigma}_{iT}^2 - \omega_{iT}^2 = O_p\left(n^{-1}\right) + O_p\left(T^{-1}\right)$.

where by (47) $n^{-1/2}\left\|\hat{\mathbf{\Gamma}}-\mathbf{\Gamma}\right\| = O_p\left(\delta_{nT}^{-1}\right)$, and by (57) $n^{-1}\sum_{i=1}^{n}\left(\hat{\sigma}_{i,T}-\omega_{i,T}\right)^2 = O_p(\delta_{nT}^{-2})$. Similarly by (47) and (59) we have

$$
\begin{aligned}
|\mathbf{d}_{5,nT}| &\leq \left[n^{-1}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{i,T}}-\frac{1}{\omega_{i,T}}\right)^2\right]^{1/2}\left(n^{-1/2}\left\|\hat{\mathbf{\Gamma}}-\mathbf{\Gamma}\right\|\right), \\
&= O_p\left(\frac{1}{\delta_{nT}}\right)O_p\left(\frac{1}{\delta_{nT}}\right) = O_p\left(\frac{1}{\delta_{nT}^2}\right).
\end{aligned}
$$

Consider now (73), and note that it can be written as

$$
\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\hat{\boldsymbol{\delta}}_{i,T}-\boldsymbol{\delta}_{i,T}) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\hat{\boldsymbol{\gamma}}_i}{\omega_{i,T}}-\frac{\boldsymbol{\gamma}_i}{\omega_{i,T}}\right) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i}{\sigma_i}\right)\left(1-\frac{\omega_{i,T}-\sigma_i}{\omega_{i,T}}\right) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i}{\sigma_i}\right)-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i}{\sigma_i}\right)\left(1-\frac{T}{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i}\right).
\end{aligned}
$$

The first term of the above has the same form as (68), and becomes identical to it if we replace a_i in (68) with $1/\sigma_i$, since by assumption $inf_i(\sigma_i) > c$. Hence, the order of the first term is $O_p(\delta_{nT}^{-1})$. Also the second term is dominated by the first term, since $1-\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{-1} = O_p(T^{-1})$ based on result (94). Therefore (73) is established as required. Finally, consider (74) and note that

$$
\begin{aligned}
n^{-1/2}\sum_{i=1}^{n}b_{in}\left(\hat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i' &= n^{-1/2}\sum_{i=1}^{n}b_{in}\left[-\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}'\left(\hat{\mathbf{F}}-\mathbf{F}\right)\boldsymbol{\gamma}_i+\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}'\mathbf{u}_i\right]\boldsymbol{\gamma}_i' \\
&= -\sqrt{n}\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\frac{\hat{\mathbf{F}}'\left(\hat{\mathbf{F}}-\mathbf{F}\right)}{T}\left(n^{-1}\sum_{i=1}^{n}b_{in}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\right)+\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}T^{-1}\hat{\mathbf{F}}'\left(n^{-1/2}\sum_{i=1}^{n}b_{in}\mathbf{u}_i\boldsymbol{\gamma}_i'\right) \\
&= -\sqrt{n}\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\frac{\hat{\mathbf{F}}'\left(\hat{\mathbf{F}}-\mathbf{F}\right)}{T}\left(n^{-1}\sum_{i=1}^{n}b_{in}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\right) \\
&+\sqrt{n}\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\left(n^{-1}\sum_{i=1}^{n}T^{-1}b_{in}\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{u}_i\boldsymbol{\gamma}_i'\right)+\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\left(n^{-1/2}\sum_{i=1}^{n}T^{-1}b_{in}\mathbf{F}'\mathbf{u}_i\boldsymbol{\gamma}_i'\right).
\end{aligned} \tag{79}
$$

Recall that $\left(T^{-1}\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1} = O_p(1)$, and $n^{-1}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i' = O_p(1)$. Also note that b_{in} is bounded in n. Then using (52) it follows that (n and T being of the same order)

$$
\sqrt{n}\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\frac{\hat{\mathbf{F}}'\left(\hat{\mathbf{F}}-\mathbf{F}\right)}{T}\left(n^{-1}\sum_{i=1}^{n}b_{in}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\right) = O_p\left(\frac{\sqrt{n}}{\min(n,T)}\right) = O_p(\delta_{nT}^{-1}).
$$

Similarly, using (53)

$$
\sqrt{n}\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\left(n^{-1}\sum_{i=1}^{n}T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}\right)'b_{in}\mathbf{u}_i\boldsymbol{\gamma}_i'\right) = O_p\left(\frac{\sqrt{n}}{\delta_{nT}^2}\right) = O_p(\delta_{nT}^{-1}).
$$

Also, the last term of (79) can be written as $T^{-1/2}\left(T^{-1}\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\left(n^{-1/2}T^{-1/2}\sum_{i=1}^{n}b_{in}\mathbf{F}'\mathbf{u}_i\boldsymbol{\gamma}_i'\right)$, where by assumption the $m_0\times m_0$ matrix, $n^{-1/2}T^{-1/2}\sum_{i=1}^{n}b_{in}\mathbf{F}'\mathbf{u}_i\boldsymbol{\gamma}_i'=n^{-1/2}T^{-1/2}\sum_{i=1}^{n}\sum_{t=1}^{T}$ $b_{in}f_{jt}u_{it}\gamma_{ij'}=O_p(1)$, and hence this last term is also $O_p(\delta_{nT}^{-1})$. Thus result (74) is established, as required. ■

Lemma 4 *Suppose that Assumptions 1-3 hold, and as* $(n,T)\rightarrow\infty$, $n/T\rightarrow\kappa$, *with* $0<\kappa<\infty$. *Then we have*

$$p_{nT}=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}s_{t,nT}^{2}=o_p(1), \tag{80}$$

$$q_{nT}=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\psi_{t,nT}s_{t,nT}=o_p(1), \tag{81}$$

where $\psi_{t,nT}$ *and* $s_{t,nT}$ *are defined by (6) and (7), respectively.*

Proof. Using (7), recall that

$$s_{t,nT}=\boldsymbol{\varphi}_{nT}'\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)u_{it}\right]+\boldsymbol{\varphi}_{nT}'\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i'\right]\mathbf{f}_t$$
$$+\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T}-\boldsymbol{\delta}_{i,T}\right)'\right]\mathbf{f}_t+\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T}-\boldsymbol{\delta}_{i,T}\right)'\right]\left(\hat{\mathbf{f}}_t-\mathbf{f}_t\right). \tag{82}$$

We also note that using (9), $\psi_{t,nT}$ can be written as

$$\psi_{t,nT}=\xi_{t,n}-\left(\boldsymbol{\varphi}_{nT}-\boldsymbol{\varphi}_n\right)'\kappa_{t,n}+\upsilon_{t,nT} \tag{83}$$

where

$$\xi_{t,n}=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}a_{i,n}\varepsilon_{it},\ \ a_{i,n}=1-\sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i, \tag{84}$$

$$\kappa_{t,n}=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\sigma_i\varepsilon_{it}, \tag{85}$$

$$\upsilon_{t,nT}=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{1}{\left(\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i/T\right)^{1/2}}-1\right]\varepsilon_{it}. \tag{86}$$

After squaring $s_{t,nT}$, we end up with $p_{nT}=\sum_{j=1}^{10}A_{j,nT}$, composed of four squared terms and six cross product terms. For the first square term we have

$$A_{1,nT}=\sqrt{T}\boldsymbol{\varphi}_{nT}'\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{b}_{t,n}\mathbf{b}_{t,n}'\right)\boldsymbol{\varphi}_{nT},$$

where $\mathbf{b}_{t,n}=n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)u_{it}=n^{-1/2}\left(\hat{\boldsymbol{\Gamma}}-\boldsymbol{\Gamma}\right)'\mathbf{u}_t$. Then

$$\left|A_{1,nT}\right|\leq\frac{\sqrt{T}}{n}\left\|\boldsymbol{\varphi}_{nT}\right\|^2\left\|\left(\hat{\boldsymbol{\Gamma}}-\boldsymbol{\Gamma}\right)\right\|^2\left\|\mathbf{V}_T\right\|,$$

where $\mathbf{V}_T = T^{-1}\sum_{t=1}^{T}\mathbf{u}_t\mathbf{u}_t^{\prime}$. But by Assumption 2, $\|\mathbf{V}_T\| = \lambda_{max}(\mathbf{V}_T) = O_p(1)$, and using (47) $\left\|\left(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma}\right)\right\| = O_p\left(\frac{\sqrt{n}}{\delta_{nT}}\right)$. Note that $\left\|\left(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma}\right)\right\| \leq \left\|\left(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma}\right)\right\|_F$. Also $\|\boldsymbol{\varphi}_{nT}\| = O_p(1)$. Then $|A_{1,nT}| = \frac{\sqrt{T}}{n}O_p\left(\frac{\sqrt{n}}{\delta_{nT}}\right) = \sqrt{\frac{T}{n}}O_p\left(\frac{1}{\delta_{nT}}\right) = O_p\left(\frac{1}{\delta_{nT}}\right)$, since n and T are of the same order. For the second squared term we have

$$A_{2,nT} = \sqrt{T}\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right)^{\prime}\right]\left(T^{-1}\mathbf{F}^{\prime}\mathbf{F}\right)\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right)\right].$$

By assumption $T^{-1}\mathbf{F}^{\prime}\mathbf{F} = O_p(1)$, and using (73) $n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right) = O_p\left(\delta_{nT}^{-1}\right)$. Hence, $A_{2,nT} = O_p\left(\sqrt{T}\delta_{nT}^{-2}\right) = o_p(1)$. Similarly,

$$\begin{aligned}
A_{3,nT} &= \sqrt{T}\boldsymbol{\varphi}_{nT}^{\prime}\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i^{\prime}\right]\left(T^{-1}\sum_{t=1}^{T}\mathbf{f}_t\mathbf{f}_t^{\prime}\right)\left[n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\left(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_i\right)^{\prime}\right]\boldsymbol{\varphi}_{nT} \\
&= \sqrt{T}\boldsymbol{\varphi}_{nT}^{\prime}\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i^{\prime}\right]\left(T^{-1}\mathbf{F}^{\prime}\mathbf{F}\right)\left[n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\left(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_i\right)^{\prime}\right]\boldsymbol{\varphi}_{nT},
\end{aligned}$$

where $\|\boldsymbol{\varphi}_{nT}\|$ is bounded, and by (74) $n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i^{\prime} = O_p\left(\delta_{nT}^{-1}\right)$. Hence

$$\begin{aligned}
A_{3,nT} &= \sqrt{T}\boldsymbol{\varphi}_{nT}^{\prime}O_p\left(\frac{1}{\min(\sqrt{n},\sqrt{T})}\right)\left(T^{-1}\mathbf{F}^{\prime}\mathbf{F}\right)O_p\left(\frac{1}{\min(\sqrt{n},\sqrt{T})}\right)\boldsymbol{\varphi}_{nT} \\
&= O_p\left(\frac{\sqrt{T}}{\min(n,T)}\right) = o_p(1).
\end{aligned}$$

Next

$$A_{4,nT} = \sqrt{T}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right)\right]^{\prime}\left(\frac{\left\|\hat{\mathbf{F}} - \mathbf{F}\right\|^2}{T}\right)\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right)\right],$$

where by (46) $T^{-1}\left\|\hat{\mathbf{F}} - \mathbf{F}\right\|_F^2 = O_p\left(\delta_{nT}^{-2}\right)$, and by (73) $n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right) = O_p(1)$. Hence, $A_{4,nT} = \sqrt{T}O_p\left(\delta_{nT}^{-2}\right) = o_p(1)$. The probability orders of the cross product terms of p_{nT}, namely $A_{5,NT}, \ldots, A_{10,NT}$, are also easily seen to be $o_p(1)$, by application of the Cauchy Schwarz inequality to the product pairs of the terms $A_{1,NT}, A_{2,NT}, A_{3,NT}$, and $A_{4,NT}$. Thus, overall $p_{nT} = o_p(1)$, as required. Consider now q_{nT} and note that it can be written as (using (83) in (81))

$$q_{nT} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}s_{t,nT}\xi_{t,n} - \left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)^{\prime}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}s_{t,nT}\kappa_{t,n} + \frac{1}{\sqrt{T}}\sum_{t=1}^{T}s_{t,nT}\upsilon_{t,nT},$$

where $\xi_{t,n}$, $\upsilon_{t,nT}$ and $\kappa_{t,n}$ are given by (84), (85) and (86), respectively. Consider the first term of the above and using (84) write it as

$$\begin{aligned}
\frac{1}{\sqrt{T}}\sum_{t=1}^{T}s_{t,nT}\xi_{t,n} &= \boldsymbol{\varphi}_{nT}'\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{t,n}u_{it}\right] \\
&+\boldsymbol{\varphi}_{nT}'\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i'\right]\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{t,n}\mathbf{f}_t\right) \\
&+\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T}-\boldsymbol{\delta}_{i,T}\right)'\right]\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{t,n}\mathbf{f}_t\right) \\
&+\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T}-\boldsymbol{\delta}_{i,T}\right)'\right]\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{t,n}\left(\hat{\mathbf{f}}_t-\mathbf{f}_t\right)\right] \\
&=\sum_{j=1}^{4}B_{j,nT}.
\end{aligned}$$

Using (84), $B_{1,nT}$ can be written as

$$B_{1,nT}=\boldsymbol{\varphi}_{nT}'\left[\frac{\sqrt{T}}{\sqrt{n}}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\frac{1}{\sqrt{n}}\sum_{j=1}^{n}a_{j,n}\left(\frac{1}{T}\sum_{t=1}^{T}\sigma_i\varepsilon_{jt}\varepsilon_{it}\right)\right]$$

where $a_{i,n}=1-\sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i$. Since ε_{it} are independently distributed over i and t; and n and T are of the same order, and $\boldsymbol{\varphi}_{nT}=O_p(1)$, then

$$B_{1,nT}=O_p\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}a_{i,n}\sigma_i\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\right).$$

Further, letting $b_{in}=a_{i,n}\sigma_i$ and noting that $n^{-1}\sum_{i=1}^{n}\sigma_i^2\left(1-\sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i\right)^2<K$, it follows from (68) that $n^{-1/2}\sum_{i=1}^{n}a_{i,n}\sigma_i\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)=O_p(\delta_{nT}^{-1})$, which in turn establishes that $B_{1,nT}=o_p(1)$. Similarly, using (84), $B_{2,nT}$ can be written as

$$B_{2,nT}=\boldsymbol{\varphi}_{nT}'\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i'\right]\left(\frac{1}{\sqrt{nT}}\sum_{j=1}^{n}\sum_{t=1}^{T}a_{j,n}\mathbf{f}_t\varepsilon_{jt}\right).$$

Recall that $\boldsymbol{\varphi}_{nT}=O_p(1)$, and by (74) $n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i'=O_p(\delta_{nT}^{-1})$. Also, under Assumption 2 $\frac{1}{\sqrt{nT}}\sum_{j=1}^{n}\sum_{t=1}^{T}a_{j,n}\mathbf{f}_t\varepsilon_{jt}=O_p(1)$. Then it follows that $B_{2,nT}=o_p(1)$. Similarly, it is established that $B_{3,nT}=o_p(1)$, noting that by (73) we have $n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T}-\boldsymbol{\delta}_{i,T}\right)=O_p(\delta_{nT}^{-1})$. The fourth term, $B_{4,nT}$, is dominated by the third term and is also $o_p(1)$. Thus overall, $T^{-1/2}\sum_{t=1}^{T}s_{t,nT}\xi_{t,n}=o_p(1)$. Using the same line of reasoning, it is also readily established that $T^{-1/2}\sum_{t=1}^{T}s_{t,nT}\kappa_{t,n}=o_p(1)$, considering that, $\kappa_{t,n}=n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\sigma_i\varepsilon_{it}$ has the same format as $\xi_{t,n}$, and in addition by (87) $\boldsymbol{\varphi}_{nT}-\boldsymbol{\varphi}_n=O_p(n^{-1/2}T^{-1/2})+O_p(T^{-1})$. Finally, the last term of q_{nT} is given by

$$\begin{aligned}
\frac{1}{\sqrt{T}}\sum_{t=1}^{T} s_{t,nT}\upsilon_{t,nT} &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\upsilon_{t,nT}\boldsymbol{\varphi}_{nT}'\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)u_{it}\right] \\
&+\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\upsilon_{t,nT}\boldsymbol{\varphi}_{nT}'\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i'\right]\mathbf{f}_t \\
&+\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\upsilon_{t,nT}\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T}-\boldsymbol{\delta}_{i,T}\right)'\right]\mathbf{f}_t \\
&+\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\upsilon_{t,nT}\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T}-\boldsymbol{\delta}_{i,T}\right)'\right]\left(\hat{\mathbf{f}}_t-\mathbf{f}_t\right) \\
&=\sum_{j=1}^{4} C_{j,nT}.
\end{aligned}$$

Using (86) we have

$$\begin{aligned}
C_{1,nT} &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\frac{1}{\left(\boldsymbol{\varepsilon}_j'\mathbf{M}_F\boldsymbol{\varepsilon}_j/T\right)^{1/2}}-1\right)\varepsilon_{jt}\boldsymbol{\varphi}_{nT}'\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)u_{it}\right] \\
&=\sqrt{\frac{T}{n}}\boldsymbol{\varphi}_{nT}'\sum_{j=1}^{n}\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\frac{1}{T}\sum_{t=1}^{T}\sigma_i\left(\frac{1}{\left(\boldsymbol{\varepsilon}_j'\mathbf{M}_F\boldsymbol{\varepsilon}_j/T\right)^{1/2}}-1\right)\varepsilon_{it}\varepsilon_{jt}\right].
\end{aligned}$$

Again, since ε_{it} is distributed independently over i and t, then

$$\frac{1}{T}\sum_{t=1}^{T}\sigma_i\left(\frac{1}{\left(\boldsymbol{\varepsilon}_j'\mathbf{M}_F\boldsymbol{\varepsilon}_j/T\right)^{1/2}}-1\right)\varepsilon_{it}\varepsilon_{jt}\rightarrow_p 0,\text{ if } i\neq j,$$

and

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\sigma_i\left(\frac{1}{\left(\boldsymbol{\varepsilon}_j'\mathbf{M}_F\boldsymbol{\varepsilon}_j/T\right)^{1/2}}-1\right)\varepsilon_{it}\varepsilon_{jt} \\
&\rightarrow_p \lim_{T}\frac{1}{T}\sum_{t=1}^{T}\sigma_i\left\{E\left[\varepsilon_{it}^2\left(\frac{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i}{T}\right)^{-1/2}\right]-1\right\},\text{ if } i=j
\end{aligned}$$

Also by (97), $E\left[\varepsilon_{it}^2\left(\frac{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i}{T}\right)^{-1/2}\right]=1+O\left(\frac{1}{T}\right)$, n and T being of the same order, and by (68) $n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)=O_p(\delta_{nT}^{-1})$. It then follows that $C_{1,nT}=o_p(1)$. Similarly to $B_{2,nT}$, we have

$$\begin{aligned}
C_{2,nT} &= \boldsymbol{\varphi}_{nT}'\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i'\right]\left[\frac{1}{\sqrt{nT}}\sum_{j=1}^{n}\sum_{t=1}^{T}\left(\frac{1}{\left(\boldsymbol{\varepsilon}_j'\mathbf{M}_F\boldsymbol{\varepsilon}_j/T\right)^{1/2}}-1\right)\varepsilon_{jt}\mathbf{f}_t\right] \\
&=O_p(\delta_{nT}^{-1})O_p(1)=o_p(1).
\end{aligned}$$

The same line of reasoning as used for $B_{3,nT}$ and $B_{4,nT}$ can be used to establish $C_{j,nT}=o_p(1)$ for $j=3$ and 4. Hence, $T^{-1/2}\sum_{t=1}^{T}s_{t,nT}\upsilon_{t,nT}=o_p(1)$, and overall we have $q_{nT}=o_p(1)$, as required. ■

Lemma 5 *Under Assumptions 1-3, and as $(n,T)\to\infty$, such that $n/T\to\kappa$, with $0<\kappa<\infty$, we have*

$$\sqrt{T}\left(\boldsymbol{\varphi}_n-\boldsymbol{\varphi}_{nT}\right)=O_p\left(n^{-1/2}\right)+O_p\left(T^{-1/2}\right)=o_p(1), \tag{87}$$

$$\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT}-\boldsymbol{\varphi}_n\right)=o_p(1) \tag{88}$$

where $\boldsymbol{\varphi}_n=n^{-1}\sum_{i=1}^n\boldsymbol{\gamma}_i/\sigma_i$, $\boldsymbol{\varphi}_{nT}=n^{-1}\sum_{i=1}^n\boldsymbol{\gamma}_i/\omega_{i,T}$, $\hat{\boldsymbol{\varphi}}_{nT}=n^{-1}\sum_{i=1}^n\hat{\boldsymbol{\gamma}}_i/\hat{\sigma}_{iT}$, $\omega_{iT}=T^{-1}\mathbf{u}_i'\mathbf{M}_F\mathbf{u}_i$, $\hat{\sigma}_{iT}^2=T^{-1}\boldsymbol{y}_i'\mathbf{M}_{\hat{F}}\boldsymbol{y}_i$, *and* $\hat{\boldsymbol{\gamma}}_i$ *and* $\hat{\mathbf{F}}$ *are the principal component estimators of* $\boldsymbol{\gamma}_i$ *and* $\mathbf{F}$.

Proof. First note that

$$\begin{aligned}\sqrt{T}\left(\boldsymbol{\varphi}_n-\boldsymbol{\varphi}_{nT}\right)&=\frac{\sqrt{T}}{n}\sum_{i=1}^n\frac{\boldsymbol{\gamma}_i}{\sigma_i}\left\{\left(1-\frac{\sigma_i}{\omega_{i,T}}\right)-\left[1-E\left(\frac{\sigma_i}{\omega_{i,T}}\right)\right]\right\}+\frac{\sqrt{T}}{n}\sum_{i=1}^n\frac{\boldsymbol{\gamma}_i}{\sigma_i}\left[1-E\left(\frac{\sigma_i}{\omega_{i,T}}\right)\right],\\&=\mathbf{d}_{1,nT}+\mathbf{d}_{2,nT}.\end{aligned}$$

where

$$\begin{aligned}\mathbf{d}_{1,nT}&=-\frac{1}{n}\sum_{i=1}^n\sqrt{T}\left[\frac{\sigma_i}{\omega_{i,T}}-E\left(\frac{\sigma_i}{\omega_{i,T}}\right)\right]\frac{\boldsymbol{\gamma}_i}{\sigma_i},\\\mathbf{d}_{2,nT}&=\frac{\sqrt{T}}{n}\sum_{i=1}^n\left[1-E\left(\frac{\sigma_i}{\omega_{i,T}}\right)\right]\frac{\boldsymbol{\gamma}_i}{\sigma_i}.\end{aligned}$$

Since $\sigma_i/\omega_{i,T}=\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{-1/2}$, $\|\boldsymbol{\gamma}_i\|<K$, $\sigma_i<K$, then using result (94) in Lemma 7 we have $E\left(\frac{\sigma_i}{\omega_{i,T}}\right)=1+O\left(T^{-1}\right)$, and $\mathbf{d}_{2,nT}=O\left(T^{-1/2}\right)$. The first term can be written as $\mathbf{d}_{1,nT}=n^{-1}\sum_{i=1}^n\boldsymbol{\gamma}_i\sigma_i^{-1}\chi_{i,T}$, where $\chi_{i,T}=-\sqrt{T}\left[\sigma_i/\omega_{i,T}-E\left(\sigma_i/\omega_{i,T}\right)\right]$. Conditional on $\mathbf{F}$ and σ_i, $\chi_{i,T}$ are distributed independently over i with mean zero and bounded variances:[4]

$$Var\left(\chi_{i,T}\right)=T\left[E\left(\frac{T}{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i}\right)-\left[E\left(\frac{\sigma_i}{\omega_{i,T}}\right)\right]^2\right]=T\left[1+O\left(\frac{1}{T}\right)-\left[1+O\left(\frac{1}{T}\right)\right]^{1/2}\right]=O(1).$$

Hence, $\mathbf{d}_{1,nT}=O_p\left(n^{-1/2}\right)$, and the desired result (87) follows. Consider now (88) and note that it can be decomposed as

$$\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT}-\boldsymbol{\varphi}_n\right)=\sqrt{T}\left(\boldsymbol{\varphi}_{nT}-\boldsymbol{\varphi}_n\right)+\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT}-\boldsymbol{\varphi}_{nT}\right), \tag{89}$$

where it is already established that the first term is $o_p(1)$. Consider now the probability order of the second term and note that it can be written as

$$\begin{aligned}\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT}-\boldsymbol{\varphi}_{nT}\right)&=\frac{\sqrt{T}}{n}\sum_{i=1}^n\left(\frac{\hat{\boldsymbol{\gamma}}_i}{\hat{\sigma}_{i,T}}-\frac{\boldsymbol{\gamma}_i}{\omega_{iT}}\right)\\&=\frac{\sqrt{T}}{n}\sum_{i=1}^n\boldsymbol{\gamma}_i\left(\frac{1}{\hat{\sigma}_{i,T}}-\frac{1}{\omega_{i,T}}\right)+\frac{\sqrt{T}}{n}\sum_{i=1}^n\left(\frac{1}{\hat{\sigma}_{i,T}}-\frac{1}{\omega_{i,T}}\right)\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right).\end{aligned}$$

[4] When ε_{it} are normally distributed we have the exact result $E\left(\frac{T}{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i}\right)=T/(T-m_0-2)$.

Now using (60) of Lemma 2 we have

$$\frac{\sqrt{T}}{n}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{i,T}}-\frac{1}{\omega_{i,T}}\right)\boldsymbol{\gamma}_i = O_p\left(\frac{\sqrt{T}}{\delta_{nT}^2}\right) = o_p(1).$$

Also by Cauchy–Schwarz inequality using (47) and (57) we have

$$\left\|\frac{\sqrt{T}}{n}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{i,T}}-\frac{1}{\omega_{i,T}}\right)(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i)\right\| \leq \sqrt{T}\left[\frac{1}{n}\sum_{i=1}^{n}(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i)'(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i)\right]^{1/2}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{i,T}}-\frac{1}{\omega_{i,T}}\right)^2\right]^{1/2},$$
$$= \sqrt{T}O_p\left(\frac{1}{\delta_{nT}}\right)\times O_p\left(\frac{1}{\delta_{nT}^2}\right) = o_p(1).$$

Using the above results, we have $\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT}-\boldsymbol{\varphi}_{nT}\right) = o_p(1)$, which in turn establishes (88), as required. ■

Lemma 6 *Suppose that* $\boldsymbol{\varepsilon} \sim IID(\mathbf{0},\mathbf{I}_T)$, *where* $\boldsymbol{\varepsilon} = (\varepsilon_1,\varepsilon_2,...,\varepsilon_T)'$, $\kappa_2 = E\left(\varepsilon_t^4\right)-3$, *and* $\mathbf{A}=(a_{ij})$ *and* $\mathbf{B}=(b_{ij})$ *are* $T\times T$ *real symmetric matrices and* τ_T *is a* $T\times 1$ *vector of ones. Then*

$$E\left(\boldsymbol{\varepsilon}'\mathbf{A}\boldsymbol{\varepsilon}\right) = tr(\mathbf{A}) \tag{90}$$
$$E\left[\left(\boldsymbol{\varepsilon}'\mathbf{A}\boldsymbol{\varepsilon}\right)\left(\boldsymbol{\varepsilon}'\mathbf{B}\boldsymbol{\varepsilon}\right)\right] = \kappa_2 Tr\left[(\mathbf{A}\odot\mathbf{B})\right]+Tr\left(\mathbf{A}\right)Tr\left(\mathbf{B}\right)+2Tr\left(\mathbf{AB}\right) \tag{91}$$

where $\mathbf{A}\odot\mathbf{B}=\mathbf{B}\odot\mathbf{A}$ *denotes Hadamard product with elements* $a_{ij}b_{ij}$.

Proof. See Appendix A.5 of Ullah (2004). ■

Lemma 7 *Suppose that* $\boldsymbol{\varepsilon}\,|\mathbf{F} \sim IID(\mathbf{0},\mathbf{I}_T)$, $\boldsymbol{\varepsilon}=(\varepsilon_1,\varepsilon_2,...,\varepsilon_T)'$, *where* ε_t *is independent over* t, $\sup_t E\left(|\varepsilon_t|^{8+\epsilon}\right)$ *exists for small* $\epsilon>0$ *and let* $\mathbf{M}_F=\mathbf{I}_T-\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$, *where* $\mathbf{F}$ *is* $T\times m_0$ *matrix such that* $\mathbf{F}'\mathbf{F}$ *is non-singular. Also let* $v=T-m_0$ *and suppose*

$$E\left[\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v}\right)^{-8-\epsilon}\right] < K. \tag{92}$$

for some small $\epsilon>0$. *Then there exists* v_0 *such that for all* $v>v_0$ *we have*

$$E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v}\right) = 1, \tag{93}$$
$$E\left[\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v}\right)^{s/2}\right] = 1+O\left(\frac{1}{v}\right),\ \textit{for } s=4,1,-1,-2,-3,-4, \tag{94}$$
$$E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v}\right)\right] = 1+O\left(\frac{1}{v}\right), \tag{95}$$
$$E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v}\right)^{-1}\right] = 1+O\left(\frac{1}{v}\right), \tag{96}$$
$$E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v}\right)^{-1/2}\right] = 1+O\left(\frac{1}{v}\right). \tag{97}$$

Proof. Firstly consider result (93) and note it follows immediately from (90) in Lemma 6 as $tr(\mathbf{M}_F) = T - m$. For result (94), when $s = 4$, it follows using (91) in Lemma 6, by setting $\mathbf{A} = \mathbf{B} = \mathbf{M}_F$, and noting that $tr\left[(\mathbf{M}_F \odot \mathbf{M}_F)\right] = \sum_{t=1}^{T} m^2 = O(T)$, $tr\left(\mathbf{M}_F^2\right) = tr\left(\mathbf{M}_F\right) = T - m_0$. Next to establish the result for $s = 1, -1, -2, -3, -4$, we first note that

$$\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v} = 1 + \frac{1}{\sqrt{v}} q_v \tag{98}$$

where q_v is defined as

$$q_v = \sqrt{v}\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v} - 1\right) = \frac{\sum_{t=1}^{v}\left(\eta_t^2 - 1\right)}{\sqrt{v}} \tag{99}$$

with $\eta_t \sim IID(0,1)$. Noting that $\eta_t^2 - 1$ is independently distributed over t, with zero means and finite variances, it also follows that there exists a v_0 such that for all $v > v_0$, q_v is approximately Gaussian. Then based on (98) and (99), and using the Taylor Theorem $\left(v^{-1}\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}\right)^{s/2}$ can be expanded in terms of q_v as

$$\begin{aligned}\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v}\right)^{s/2} &= \left(1 + \frac{1}{\sqrt{v}} q_v\right)^{s/2} \\ &= 1 + \frac{s}{2}\frac{q_v}{\sqrt{v}} + \left[\frac{1}{2}\frac{s}{2}\left(\frac{s-2}{2}\right)\frac{1}{v}\right] R_{s,v},\end{aligned} \tag{100}$$

with the term $R_{s,v}$ given by

$$R_{s,v} = \left(1 + \frac{\bar{q}_v}{\sqrt{v}}\right)^{(s-4)/2} q_v^2,$$

where $\bar{q}_v$ lies in the interval between 0 and q_v. If $q_v > 0$, then $\bar{q}_v < q_v$, and if $q_v < 0$, then $\bar{q}_v > q_v$. Note also that q_v tends to a mean-zero normal variate as $v \to \infty$, and $q_v = O_p(1)$. Taking expectations of both sides of (100) we have (noting that $E(q_v) = 0$)

$$E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v}\right)^{s/2} = 1 + \frac{1}{v}\frac{s\,(s-2)}{8} E\left(R_{s,v}\right).$$

Also by Cauchy-Schwarz inequality we have

$$E\left(R_{s,v}\right) \leq \left\{E\left[\left(1 + \frac{\bar{q}_v}{\sqrt{v}}\right)^{(s-4)}\right]\right\}^{1/2}\left[E\left(q_v^4\right)\right]^{1/2}. \tag{101}$$

Since q_v is approximately Gaussian, then there exists v_0 such that for all $v > v_0$, $E\left(q_v^4\right) < K$. Consider now the first term of (101), and note that for values of s $(= 1, -1, -2, -3, -4)$, under consideration, $s - 4 < 0$, and $\left(1 + \frac{\bar{q}}{\sqrt{v}}\right)^{-(4-s)} < K$, for values of $\bar{q}_v > 0$, and $E\left(R_v\right) < K$. When $\bar{q}_v < 0$, for $s - 4 < 0$ we have (recalling that in this case $q_v < \bar{q}_v$)

$$E\left[\left(1 + \frac{\bar{q}_v}{\sqrt{v}}\right)^{(s-4)}\right] \leq E\left[\left(1 + \frac{q_v}{\sqrt{v}}\right)^{(s-4)}\right] = E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v}\right)^{s-4}.$$

Therefore, it follows that $E\left(R_v\right) \leq K$, irrespective of the sign of $\bar{q}_v$, so long as $E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v}\right)^{s-4} < K$, and results in (94) follow under the lemma's assumptions. Result (95) follows using (91) in Lemma 6, and setting $\varepsilon_t^2 = \boldsymbol{\varepsilon}'\mathbf{A}\boldsymbol{\varepsilon}$, where $\mathbf{A}$ has only one none-zero element on its diagonal. Result (96) can be established using a result due to Lieberman (1994) (see Lemmas 5 and 21 of Pesaran and Yamagata (2017)). Finally, to establish (97), using (100) note that

$$E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v}\right)^{-1/2}\right] = E\left(\varepsilon_t^2\right) - \frac{1}{2}\frac{E\left(q_v\varepsilon_t^2\right)}{\sqrt{v}} + \frac{3}{8v}E\left(R_{e,v}\right), \tag{102}$$

where

$$R_{e,v} = \left(1+\frac{\bar{q}_v}{\sqrt{v}}\right)^{-5/2} q_v^2\varepsilon_t^2.$$

$E(\varepsilon_t^2) = 1$, and using (95) we have

$$E\left[\varepsilon_t^2\left(\frac{q_v}{\sqrt{v}}\right)\right] = E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{v} - 1\right)\right] = O\left(\frac{1}{v}\right). \tag{103}$$

Finally, by Cauchy-Schwarz inequality

$$\begin{aligned} E\left(R_{e,v}\right) &\leq \left[E\left(1+\frac{\bar{q}_v}{\sqrt{v}}\right)^{-5}\right]^{1/2}\left[E\left(\varepsilon_t^4 q_v^4\right)\right]^{1/2} \\ &\leq \left[E\left(1+\frac{\bar{q}_v}{\sqrt{v}}\right)^{-5}\right]^{1/2}\left[E\left(\varepsilon_t^8\right)\right]^{1/4}\left[E\left(q_v^8\right)\right]^{1/4}. \end{aligned}$$

Using a similar line of reasoning it is easily seen that $E\left(1+\frac{\bar{q}_v}{\sqrt{v}}\right)^{-5} < K$, under condition (92). Also, since q_v is asymptotically normally distributed, then there exists v_0 such that for all $v > v_0$, $E\left(q_v^8\right) < K$, which completes the proof of (97). ■

Lemma 8 *The CD statistic defined by (10) can be written equivalently as,*

$$CD = \left(\sqrt{\frac{n}{n-1}}\right)\frac{1}{\sqrt{2T}}\sum_{t=1}^{T}\left[\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{e_{it}}{\hat{\sigma}_{i,T}}\right)^2 - 1\right]. \tag{104}$$

Proof. Using $\hat{\rho}_{ij,T} = \left(\frac{1}{T}\sum_{t=1}^{T} e_{it}e_{jt}\right)/\hat{\sigma}_{i,T}\hat{\sigma}_{j,T}$ in (10) we have:

$$CD = \sqrt{\frac{2T}{n(n-1)}}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{\frac{1}{T}\sum_{t=1}^{T}e_{it}e_{jt}}{\hat{\sigma}_{i,T}\hat{\sigma}_{j,T}} = \sqrt{\frac{2T}{n(n-1)}}\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(\frac{e_{it}}{\hat{\sigma}_{i,T}}\right)\left(\frac{e_{jt}}{\hat{\sigma}_{j,T}}\right)\right). \tag{105}$$

Further, we note that

$$\frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(\frac{e_{it}}{\hat{\sigma}_{i,T}}\right)\left(\frac{e_{jt}}{\hat{\sigma}_{j,T}}\right) = \frac{1}{2}\left[\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{e_{it}}{\hat{\sigma}_{i,T}}\right)^2 - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{e_{it}}{\hat{\sigma}_{i,T}}\right)^2\right].$$

Then using this result in (105), and after some algebra, we have

$$CD = \sqrt{\frac{2Tn^2}{n(n-1)}} \frac{1}{2T} \sum_{t=1}^{T} \left[\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{e_{it}}{\hat{\sigma}_{i,T}} \right)^2 - \frac{1}{n} \sum_{i=1}^{n} \left(\frac{e_{it}}{\hat{\sigma}_{i,T}} \right)^2 \right]$$

$$= \sqrt{\frac{2Tn^2}{n(n-1)}} \frac{1}{2} \left[\frac{1}{T} \sum_{t=1}^{T} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{e_{it}}{\hat{\sigma}_{i,T}} \right)^2 - \frac{1}{n} \sum_{i=1}^{n} \frac{1}{T} \sum_{t=1}^{T} \left(\frac{e_{it}}{\hat{\sigma}_{i,T}} \right)^2 \right]$$

$$= \left(\sqrt{\frac{n}{n-1}} \right) \frac{1}{\sqrt{2T}} \sum_{t=1}^{T} \left[\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{e_{it}}{\hat{\sigma}_{i,T}} \right)^2 - 1 \right],$$

as required. ■

Lemma 9 *Consider the CD and $\widetilde{CD}$ statistics defined by (20) and (21), respectively and suppose that Assumptions 1-3 hold. Then, as $(n,T) \to \infty$, such that $n/T \to \kappa$, where $0 < \kappa < \infty$, we have*

$$CD = \widetilde{CD} + o_p(1). \tag{106}$$

Proof. Using (20) and (21) we first note that

$$\left(\sqrt{\frac{2(n-1)}{n}} \right) \left(CD - \widetilde{CD} \right) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left[\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{e_{it}}{\hat{\sigma}_{i,T}} \right)^2 - \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{e_{it}}{\omega_{i,T}} \right)^2 \right]. \tag{107}$$

Also note that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{e_{it}}{\hat{\sigma}_{i,T}} = h_{t,nT} + g_{t,nT} \tag{108}$$

where (see also (5))

$$h_{t,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{e_{it}}{\omega_{i,T}} = \frac{\mathbf{c}'_{nT}\mathbf{e}_t}{\sqrt{n}}, \text{ and } g_{t,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} e_{it} \left(\frac{1}{\hat{\sigma}_{i,T}} - \frac{1}{\omega_{i,T}} \right) = \frac{\mathbf{d}'_{nT}\mathbf{e}_t}{\sqrt{n}}$$

$\mathbf{e}_t = (e_{1t}, e_{2t}, ..., e_{nt})'$, $\mathbf{c}_{nT} = (\omega_{1,T}^{-1}, \omega_{2,T}^{-1}, ..., \omega_{n,T}^{-1})'$, $\mathbf{d}_{nT} = (d_{1T}, d_{2T}, ..., d_{nT})'$, and $d_{iT} = \hat{\sigma}_{i,T}^{-1} - \omega_{i,T}^{-1}$. Then squaring both sides of (108) and using the result in (107) we have

$$\left(\sqrt{\frac{2(n-1)}{n}} \right) \left(CD - \widetilde{CD} \right) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} g_{t,nT}^2 + \frac{2}{\sqrt{T}} \sum_{t=1}^{T} h_{t,nT}\, g_{t,nT}$$

$$= \sqrt{\frac{T}{n}} \left(\frac{1}{\sqrt{n}} \mathbf{d}'_{nT} \mathbf{V}_{eT} \mathbf{d}_{nT} + \frac{2}{\sqrt{n}} \mathbf{c}'_{nT} \mathbf{V}_{eT} \mathbf{d}_{nT} \right), \tag{109}$$

where $\mathbf{V}_{eT} = T^{-1} \sum_{t=1}^{T} \mathbf{e}_t \mathbf{e}'_t$. Now using (1), the error vector $\mathbf{e}_t$ can be written as

$$\mathbf{e}_t = \mathbf{u}_t - \boldsymbol{\Gamma} \left(\hat{\mathbf{f}}_t - \mathbf{f}_t \right) - \left(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma} \right) \mathbf{f}_t - \left(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma} \right) \left(\hat{\mathbf{f}}_t - \mathbf{f}_t \right).$$

Using this expression we now have

$$\mathbf{V}_{eT} = T^{-1}\sum_{t=1}^{T}\mathbf{u}_t\mathbf{u}_t' + \mathbf{\Gamma}\left[T^{-1}\sum_{t=1}^{T}\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right)\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right)'\right]\mathbf{\Gamma}' + \left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)\left(T^{-1}\sum_{t=1}^{T}\mathbf{f}_t\mathbf{f}_t'\right)\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)'$$
$$+\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)\left[T^{-1}\sum_{t=1}^{T}\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right)\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right)'\right]\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)' - \left[T^{-1}\sum_{t=1}^{T}\mathbf{u}_t\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right)'\right]\mathbf{\Gamma}'$$
$$-\left[T^{-1}\sum_{t=1}^{T}\mathbf{u}_t\mathbf{f}_t'\right]\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)' - \left[T^{-1}\sum_{t=1}^{T}\mathbf{u}_t\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right)'\right]\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)'$$
$$+\mathbf{\Gamma}\left[T^{-1}\sum_{t=1}^{T}\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right)\mathbf{f}_t'\right]\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)' + \mathbf{\Gamma}\left[T^{-1}\sum_{t=1}^{T}\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right)\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right)'\right]\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)'$$
$$+\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)\left[T^{-1}\sum_{t=1}^{T}\mathbf{f}_t\left(\hat{\mathbf{f}}_t - \mathbf{f}_t\right)'\right]\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)'.$$

or in matrix forms

$$\mathbf{V}_{eT} = \mathbf{V}_T + \mathbf{\Gamma}\left[T^{-1}\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\right]\mathbf{\Gamma}' + \left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)\mathbf{\Sigma}_{T,ff}\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)'$$
$$+\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)\left[T^{-1}\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\right]\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)' - T^{-1}\mathbf{U}'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\mathbf{\Gamma}'$$
$$-T^{-1}\mathbf{U}'\mathbf{F}\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)' - T^{-1}\mathbf{U}'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)' + \mathbf{\Gamma}\left[T^{-1}\mathbf{F}'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\right]\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)'$$
$$+\mathbf{\Gamma}\left[T^{-1}\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\right]\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)' + \left[T^{-1}\mathbf{F}'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\right]\left(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right)'$$

where $\mathbf{V}_T = T^{-1}\sum_{t=1}^{T}\mathbf{u}_t\mathbf{u}_t'$. Since $\|\mathbf{V}_T\| = \lambda_{\max}^{1/2}\left(\mathbf{V}_T\mathbf{V}_T'\right) = \lambda_{\max}(\mathbf{V}_T)$, then under Assumption 2 we have

$$E\|\mathbf{V}_T\| = E\left[\lambda_{\max}(\mathbf{V}_T)\right] = O(1),$$

and therefore $\|\mathbf{V}_T\| = O_p(1)$ by Markov inequality. Also by results in Lemma 1 all other terms of the $\mathbf{V}_{eT}$ are either $O_p(1)$ or of lower order, and we also have $\mathbf{V}_{eT} = O_p(1)$. Consider now the terms in (109) and note that

$$\left(\sqrt{\frac{2(n-1)}{n}}\right)\left|CD - \widetilde{CD}\right| < K\|\mathbf{V}_{eT}\|\left[\left(\frac{1}{\sqrt{n}}\|\mathbf{d}_{nT}\|^2\right) + \left(\frac{2}{\sqrt{n}}\|\mathbf{c}_{nT}\|\right)\|\mathbf{d}_{nT}\|\right].$$

But

$$\frac{1}{\sqrt{n}}\|\mathbf{c}_{nT}\| = \left(n^{-1}\sum_{i=1}^{n}\omega_{iT}^{-2}\right)^{1/2},$$
$$\|\mathbf{d}_{nT}\| = \sqrt{n}\left(n^{-1}\sum_{i=1}^{n}\left(\hat{\sigma}_{i,T}^{-1} - \omega_{i,T}^{-1}\right)^2\right)^{1/2}.$$

By assumption $\omega_{iT} > c > 0$, and $\omega_{iT}^{-2} < c^{-1} < \infty$, and hence $n^{-1/2}\|\mathbf{c}_{nT}\| = O_p(1)$. Also, using (67) we have $\left(\hat{\sigma}_{i,T}^{-1} - \omega_{i,T}^{-1}\right)^2 = O_p\left(\frac{1}{\delta_{nT}^4}\right)$, and it follows that $\|\mathbf{d}_{nT}\| = O_p\left(\frac{\sqrt{n}}{\delta_{nT}^2}\right) = o_p(1)$, recalling that n and T are of the same order. Hence, $\left|CD - \widetilde{CD}\right| = o_p(1)$, as required. ■

Lemma 10 *Consider the latent factor loadings, $\boldsymbol{\gamma}_i$, in model (2) and their estimates $\hat{\boldsymbol{\gamma}}_i$ given by (8). Then under Assumptions 1-3 with n and $T \to \infty$, such that $n/T \to \kappa$, for $0 < \kappa < \infty$, we have*

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i}{\sigma_i} = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{110}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\sigma_i = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{111}$$

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\sigma}_{i,T}\hat{\boldsymbol{\gamma}}_i - \frac{1}{n}\sum_{i=1}^{n}\sigma_i\boldsymbol{\gamma}_i = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{112}$$

$$\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}_i' - \boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\right) = O_p\left(\frac{1}{\delta_{nT}^2}\right). \tag{113}$$

Proof. Results (110) and (111) follow directly from (68) by setting $b_{in} = \sigma_i^{-1}$ and $b_{in} = \sigma_i$, respectively. To prove (112) note that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\hat{\sigma}_{i,T}\hat{\boldsymbol{\gamma}}_i - \frac{1}{n}\sum_{i=1}^{n}\sigma_i\boldsymbol{\gamma}_i \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\left(\hat{\sigma}_{i,T} - \omega_{i,T}\right) + \omega_{i,T}\right]\left[\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i + \boldsymbol{\gamma}_i\right] - \frac{1}{n}\sum_{i=1}^{n}\sigma_i\boldsymbol{\gamma}_i \\
&= \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\left(\omega_{i,T} - \sigma_i\right) + \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\left(\hat{\sigma}_{i,T} - \omega_{i,T}\right) + \frac{1}{n}\sum_{i=1}^{n}\sigma_i\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right) \\
&+ \frac{1}{n}\sum_{i=1}^{n}\left(\omega_{i,T} - \sigma_i\right)\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right) + \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_{i,T} - \omega_{i,T}\right)\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right) \\
&= \mathbf{A}_{1,nT} + \mathbf{A}_{2,nT} + \mathbf{A}_{3,nT} + \mathbf{A}_{4,nT} + \mathbf{A}_{5,nT}.
\end{aligned} \tag{114}$$

Recall also that under Assumptions 2 and 3 σ_i and $\boldsymbol{\gamma}_i$ are bounded and $\omega_{i,T} = T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i$, for $i = 1, 2, ..., n$ are distributed independently across i, and from σ_i and $\boldsymbol{\gamma}_i$. Starting with $\mathbf{A}_{1,nT}$, and using (94) we have

$$E\left(\sqrt{nT}\mathbf{A}_{1,nT}\right) = \frac{\sqrt{nT}}{n}\sum_{i=1}^{n}(\boldsymbol{\gamma}_i\sigma_i)E\left[\left(\frac{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i}{T}\right)^{1/2} - 1\right] = O\left(\frac{\sqrt{nT}}{T}\right),$$

Since n and T are assumed to be of the same order then $E\left(\sqrt{nT}\mathbf{A}_{1,nT}\right) = O(1)$. Also, using results (93) and (94)

$$Var\left(\sqrt{nT}\mathbf{A}_{1,nT}\right) = \frac{nT}{n^2}\sum_{i=1}^{n}\left(\sigma_i^2\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\right)Var\left[\left(\frac{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i}{T}\right)^{1/2}\right] = O(1).$$

Therefore, $\sqrt{nT}\mathbf{A}_{1,nT} = O_p(1)$ and it follows that $\mathbf{A}_{1,nT} = O_p\left[(nT)^{-1/2}\right]$. Further, using (58) setting $b_{in} = \gamma_{ij}$, for $j = 1, 2, ..., m_0$, then it follows that

$$\mathbf{A}_{2,nT} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\left(\hat{\sigma}_{i,T} - \omega_{i,T}\right) = O_p\left(\frac{1}{\sqrt{n}\delta_{nT}}\right).$$

Since $\mathbf{A}_{3,nT}$ is the same as the result in (111), which is already established, then $\mathbf{A}_{3,nT} = O_p\left(\delta_{nT}^{-2}\right)$. Using result (69) it follows that

$$\mathbf{A}_{4,nT} = \frac{1}{n}\sum_{i=1}^{n}\left(\omega_{i,T} - \sigma_i\right)\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right) = O_p\left(\frac{1}{\delta_{nT}^2}\right).$$

Using result (71) we have

$$\mathbf{A}_{5,nT} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_{i,T} - \omega_{i,T}\right)\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right) = O_p\left(\frac{1}{\delta_{nT}^2}\right).$$

Result (112) now follows since $\mathbf{A}_{j,nT} = O_p(\delta_{nT}^{-2})$, for $j = 1, 2, ..., 5$. Finally, consider (113) and note that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}_i' - \boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\right) &= \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)' \\ &+ \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i' + \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\boldsymbol{\gamma}_i\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)'. \end{aligned} \tag{115}$$

Since σ_i^2 is bounded, then

$$\begin{aligned}\left\|\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)'\right\| &\leq \left(\sup_{1\leq i\leq n}\sigma_i^2\right)\left\|\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)'\right\| \\ &\leq \left(\sup_{1\leq i\leq n}\sigma_i^2\right)\left(\frac{1}{n}\sum_{i=1}^{n}\left\|\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right\|^2\right),\end{aligned}$$

and using (47) it follows that

$$\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)' = O_p\left(\frac{1}{\delta_{nT}^2}\right).$$

Now using (74), setting $b_{in} = \sigma_i$, we have

$$\left\|\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i'\right\| = \left\|\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\boldsymbol{\gamma}_i\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)'\right\| = O_p\left(\frac{1}{\sqrt{n}\delta_{nT}}\right),$$

and (113) follows. ■

3 Derivation of θ_n in terms of factor strengths

Consider θ_n defined by (31), and note that it can be written as

$$\begin{aligned}\theta_n &= 1 - \frac{1}{n}\sum_{i=1}^{n}a_{i,n}^2 = 1 - \frac{1}{n}\sum_{i=1}^{n}\left(1 - \sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i\right)^2 \\ &= 2\boldsymbol{\varphi}_n'\left(\frac{1}{n}\sum_{i=1}^{n}\sigma_i\boldsymbol{\gamma}_i\right) - \boldsymbol{\varphi}_n'\left(\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\right)\boldsymbol{\varphi}_n,\end{aligned} \tag{116}$$

where $\boldsymbol{\varphi}_n = n^{-1}\sum_{i=1}^{n}\boldsymbol{\gamma}_i/\sigma_i$. Then

$$|\theta_n| \leq sup_i(\sigma_i^2)\left\|\boldsymbol{\varphi}_n\right\|_1^2\left(\frac{1}{n}\sum_{i=1}^{n}\left\|\boldsymbol{\gamma}_i\right\|_1^2\right) + 2sup_i(\sigma_i)\left\|\boldsymbol{\varphi}_n\right\|_1\left(\frac{1}{n}\sum_{i=1}^{n}\left\|\boldsymbol{\gamma}_i\right\|_1\right).$$

$\left\|\boldsymbol{\gamma}_i\right\|_1 = \sum_{j=1}^{m_0}|\gamma_{ij}|$, and

$$\left\|\boldsymbol{\varphi}_n\right\|_1 \leq inf_i(\sigma_i)\left(\frac{1}{n}\sum_{i=1}^{n}\left\|\boldsymbol{\gamma}_i\right\|_1\right) = inf_i(\sigma_i)\left[\sum_{j=1}^{m_0}\left(\frac{1}{n}\sum_{i=1}^{n}|\gamma_{ij}|\right)\right]. \tag{117}$$

Since by assumption $inf_i(\sigma_i) > c > 0$, and $sup_i(\sigma_i^2) < K < \infty$, then the order of $|\theta_n|$ is determined by $\sum_{j=1}^{m_0}\left(\frac{1}{n}\sum_{i=1}^{n}|\gamma_{ij}|\right)$, where m_0 is a fixed integer. Hence, $|\theta_n| = \ominus\left(n^{\alpha-1}\right)$ as required, where $\alpha = max_j(\alpha_j)$, and α_j is defined by $\sum_{i=1}^{n}|\gamma_{ij}| = \ominus\left(n^{\alpha_j}\right)$. See (7).

4 Simulation results

This section provides Monte Carlo simulation results for the experiments discussed in Section 5 of the main paper. Tables 1-8 report the results for the DGPs with serially independent errors. Tables 9-16 report the results for the DGPs with serially correlated errors using variance adjustment. Tables 17-24 report the results for the DGPs with serially correlated errors using ARDL adjustment.

Table 1: Size and power of tests of error cross-sectional dependence for pure single factor model ($m_0 = 1$) with serially independent Gaussian errors

$\hat{m} = 1$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	64.7	88.1	97.5	5.8	9.7	22.5	5.3	5.9	9.4	23.8	37.5	55.7	68.9	86.7	97.7	81.0	93.8	99.4
	200	67.7	92.3	99.4	5.3	7.1	14.2	5.9	5.4	7.0	16.0	31.0	50.6	75.4	93.1	99.8	84.9	97.6	100.0
	500	71.0	95.2	100.0	5.1	4.2	8.6	6.2	4.9	4.3	10.9	22.9	46.3	82.3	97.5	100.0	89.6	98.9	100.0
	1000	69.1	95.2	100.0	5.1	4.5	5.7	6.1	5.4	4.5	9.4	20.0	44.6	84.0	98.1	100.0	89.7	99.0	100.0
CD^*	100	5.7	3.9	4.4	4.8	5.2	5.8	5.9	5.9	5.5	58.0	82.0	98.4	86.1	98.8	100.0	88.6	98.7	100.0
	200	5.5	4.9	5.3	5.5	5.1	5.2	5.9	5.2	5.1	59.3	81.1	98.9	84.8	98.3	100.0	88.8	98.8	100.0
	500	5.3	5.2	4.4	5.7	5.0	4.8	6.3	5.1	4.6	57.9	83.4	99.4	87.2	98.7	100.0	90.4	99.1	100.0
	1000	4.4	5.3	5.1	5.5	4.8	4.6	6.1	5.5	4.9	60.1	84.0	99.4	86.7	99.2	100.0	90.1	99.2	100.0
CD_{W+}	100	5.8	5.1	5.0	5.5	5.5	5.9	5.8	5.4	7.6	6.9	7.8	49.5	6.4	9.7	59.8	6.9	8.5	64.5
	200	6.1	5.7	5.5	4.7	5.4	4.5	5.9	7.1	5.9	6.5	7.3	51.9	5.7	7.8	61.0	6.9	8.0	60.3
	500	5.4	5.4	5.2	5.7	5.6	5.4	5.1	5.5	5.3	5.5	5.9	52.9	5.9	6.7	56.9	5.8	6.6	57.4
	1000	5.1	4.6	4.8	4.6	5.9	5.3	4.7	5.9	6.0	4.9	5.6	51.7	4.6	6.2	54.2	5.6	5.6	52.8

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	65.2	87.9	97.6	5.9	10.2	21.8	5.6	6.2	8.8	26.6	44.2	62.7	59.4	73.8	85.8	70.5	83.8	92.2
	200	68.3	91.7	99.5	5.1	7.6	14.6	5.8	5.1	7.1	17.3	34.5	56.9	70.5	89.3	98.3	81.1	95.2	99.4
	500	70.7	94.9	100.0	5.0	4.5	8.2	6.3	5.0	5.0	12.0	25.2	49.7	80.2	96.8	100.0	88.2	98.3	100.0
	1000	68.9	95.1	100.0	4.8	4.8	5.3	5.9	5.7	4.9	9.3	20.7	45.7	83.0	97.4	100.0	88.7	98.9	100.0
CD^*	100	5.7	4.7	5.8	5.5	6.3	6.2	5.8	6.5	6.3	57.8	81.8	98.8	83.9	98.4	100.0	85.8	98.3	100.0
	200	5.1	5.0	5.9	6.1	5.3	5.7	5.8	5.3	5.1	57.9	81.0	98.8	84.2	98.1	100.0	87.8	98.6	100.0
	500	5.3	5.3	4.3	5.6	5.1	4.7	6.5	5.0	4.3	57.4	82.5	99.4	86.8	98.6	100.0	89.3	98.9	100.0
	1000	4.3	5.1	5.0	5.8	5.1	5.0	6.3	5.8	4.6	59.7	83.6	99.4	86.2	99.1	100.0	89.4	99.2	100.0
CD_{W+}	100	5.1	5.4	6.5	3.8	5.4	6.1	4.8	5.4	8.4	5.4	8.1	33.8	5.1	8.2	36.7	5.9	8.2	44.1
	200	5.5	5.3	6.1	4.9	5.8	5.3	6.2	5.3	5.5	5.5	6.6	38.8	5.7	7.1	41.7	6.2	7.0	43.6
	500	4.7	4.5	4.3	5.3	5.1	4.7	5.3	5.7	4.1	4.9	6.1	45.2	5.5	5.9	48.2	6.2	6.2	47.0
	1000	3.9	4.6	4.9	4.9	5.7	4.6	5.9	4.6	4.8	3.7	5.2	43.9	5.3	6.6	48.3	6.0	6.1	47.3

Note: The DGP is given by (42) with $\beta_{i1} = \beta_{i2} = 0$ and contains a single latent factor with factor strength $\alpha = 1, 2/3, 1/2$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 2: Size and power of tests of error cross-sectional dependence for pure two factors model ($m_0 = 2$) with serially independent Gaussian errors

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	99.9	100.0	100.0	98.3	99.9	100.0	8.3	15.4	40.7	99.4	100.0	100.0	90.3	98.0	99.4	56.7	65.4	81.3
	200	100.0	100.0	100.0	99.4	100.0	100.0	6.6	8.6	24.1	99.8	100.0	100.0	93.1	99.2	100.0	71.5	84.3	97.8
	500	100.0	100.0	100.0	99.7	100.0	100.0	5.7	5.8	13.0	99.9	100.0	100.0	93.4	99.8	100.0	83.1	95.4	100.0
	1000	100.0	100.0	100.0	99.9	100.0	100.0	7.6	4.9	7.8	100.0	100.0	100.0	94.4	100.0	100.0	85.5	97.7	100.0
CD^*	100	5.5	4.7	3.8	5.6	5.3	4.9	8.8	7.2	6.8	23.8	35.2	60.0	31.3	49.3	78.9	83.7	97.8	100.0
	200	5.4	5.0	4.9	6.3	4.6	4.6	8.3	5.5	5.3	22.4	33.1	60.6	32.1	49.4	81.8	86.4	98.2	100.0
	500	4.8	4.7	6.0	5.8	4.7	4.1	7.4	5.9	5.1	21.1	35.1	64.6	33.5	51.3	83.4	88.4	98.6	100.0
	1000	5.1	4.5	4.8	5.4	5.4	4.7	8.3	6.2	5.2	23.5	34.2	63.1	36.2	52.8	85.2	88.4	98.7	100.0
CD_{W+}	100	4.4	6.8	5.9	5.5	5.9	7.8	5.8	5.3	9.4	5.4	9.4	38.1	6.4	9.3	48.1	7.1	10.1	63.8
	200	5.5	5.4	5.3	6.4	5.6	6.3	5.9	5.9	5.6	6.0	6.1	44.7	6.4	6.9	46.9	6.9	7.3	56.7
	500	5.3	5.6	5.1	5.5	5.1	4.8	5.4	4.7	6.4	5.9	6.6	49.5	5.7	5.8	50.1	6.3	5.0	57.3
	1000	4.9	4.6	4.6	5.1	4.3	5.4	6.6	5.4	5.1	5.0	5.5	50.1	4.9	5.2	50.1	6.7	6.3	53.4

$\hat{m} = 4$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	99.8	100.0	100.0	98.8	100.0	100.0	7.4	15.9	40.8	99.4	100.0	100.0	93.6	98.8	99.8	39.9	43.0	55.4
	200	100.0	100.0	100.0	99.6	100.0	100.0	5.7	8.9	24.3	99.8	100.0	100.0	95.4	99.5	100.0	61.7	73.3	87.8
	500	100.0	100.0	100.0	99.6	100.0	100.0	5.9	5.7	12.8	99.9	100.0	100.0	94.6	99.9	100.0	79.2	93.4	99.6
	1000	100.0	100.0	100.0	100.0	100.0	100.0	7.4	4.7	8.0	99.9	100.0	100.0	94.2	100.0	100.0	83.3	96.7	100.0
CD^*	100	7.6	7.9	15.4	7.0	6.8	13.2	9.5	8.4	10.1	26.4	39.2	67.5	33.7	56.4	86.4	79.8	96.6	100.0
	200	5.5	6.0	6.6	6.8	6.3	7.0	7.6	6.3	6.5	22.8	36.8	66.0	33.3	51.7	85.1	84.4	98.0	100.0
	500	5.1	4.7	6.3	6.2	5.1	4.8	7.8	6.8	5.1	21.0	35.2	66.1	34.3	51.9	84.1	87.2	98.1	100.0
	1000	5.8	5.0	4.7	5.4	5.0	4.8	8.6	5.7	5.7	23.4	34.5	63.4	35.9	52.4	84.9	87.3	98.7	100.0
CD_{W+}	100	5.5	6.8	26.2	5.9	6.3	15.7	6.2	6.1	11.8	6.2	8.8	39.5	6.2	8.7	40.5	7.3	7.9	34.9
	200	4.5	5.7	6.4	5.2	4.7	5.3	5.5	5.4	6.7	5.4	7.2	37.4	5.6	6.5	29.6	6.3	6.2	31.2
	500	5.6	5.0	5.3	5.9	5.1	5.7	5.9	5.3	4.8	6.3	6.1	44.2	6.7	6.4	38.8	5.9	5.9	39.0
	1000	4.6	5.8	4.9	5.3	4.9	4.6	5.9	5.5	5.9	5.0	6.6	45.6	4.8	5.0	42.4	6.1	5.7	43.7

Note: The DGP is given by (42) with $\beta_{i1} = \beta_{i2} = 0$, and contains two latent factors with factor strengths $(\alpha_1, \alpha_2) = [(1,1), (1,2/3), (2/3,1/2)]$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 3: Size and power of tests of error cross-sectional dependence for panel regression with one latent factor ($m_0 = 1$) with serially independent Gaussian errors

$\hat{m} = 1$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	67.9	88.6	98.5	6.7	9.5	20.5	6.9	6.3	9.0	25.1	37.6	55.6	71.1	86.6	97.7	81.4	93.9	99.3
	200	68.9	92.5	99.7	5.5	6.7	13.4	6.8	5.3	5.9	17.2	29.7	49.7	77.2	94.0	99.7	86.5	98.1	100.0
	500	67.4	94.7	100.0	4.9	6.1	8.5	5.3	5.0	5.9	11.8	22.1	45.7	83.2	97.0	100.0	89.5	99.0	100.0
	1000	69.0	95.1	100.0	5.9	4.3	6.6	6.8	5.8	5.6	10.1	19.7	44.8	85.8	98.3	100.0	89.8	99.0	100.0
CD^*	100	5.1	5.6	4.6	7.4	5.5	6.1	7.8	6.1	6.2	57.5	81.3	98.2	85.9	98.0	100.0	88.9	98.8	100.0
	200	5.5	6.0	4.2	6.2	5.8	4.6	6.8	5.7	4.9	58.9	83.0	99.1	85.8	98.3	100.0	89.2	99.2	100.0
	500	5.0	5.0	4.5	5.5	5.7	5.5	5.5	5.3	5.5	59.0	83.7	99.3	87.4	98.6	100.0	90.6	99.4	100.0
	1000	4.5	4.6	5.4	6.3	5.0	5.6	7.0	5.9	5.9	60.0	83.3	99.5	88.3	99.0	100.0	90.1	99.2	100.0
CD_{W+}	100	5.4	5.3	6.4	5.3	5.7	6.2	4.5	5.9	6.4	5.9	8.6	46.4	5.7	8.9	60.1	6.2	8.8	64.6
	200	5.8	5.1	5.2	5.3	4.9	5.3	6.1	6.4	4.8	6.0	6.6	51.8	6.0	6.3	60.2	6.2	7.6	59.3
	500	5.1	5.4	5.9	5.6	4.5	5.1	5.6	6.4	5.7	5.4	6.7	55.0	5.9	6.1	57.4	6.2	7.0	58.0
	1000	4.9	5.5	4.5	6.0	4.6	4.7	5.1	6.0	4.8	4.6	5.6	48.9	5.4	5.3	52.8	4.8	5.9	51.9

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	67.5	88.9	98.6	7.7	9.3	21.8	6.8	6.8	9.4	28.3	43.6	63.0	59.0	75.0	86.3	69.5	84.5	90.9
	200	69.0	92.2	99.5	5.6	6.7	13.3	6.8	5.4	5.8	18.7	32.3	55.7	71.5	89.4	98.3	80.8	96.1	99.5
	500	68.0	94.9	100.0	5.0	5.7	8.0	5.1	5.1	5.5	12.4	23.7	47.6	80.9	96.0	100.0	87.6	98.6	100.0
	1000	69.9	95.0	100.0	5.9	4.2	6.7	6.7	5.7	5.7	11.0	20.4	46.2	84.1	97.9	100.0	88.4	98.9	100.0
CD^*	100	5.6	6.5	6.4	7.9	5.8	7.1	8.0	6.6	6.7	57.0	81.0	98.4	83.3	97.8	100.0	86.6	98.6	100.0
	200	5.8	5.7	4.9	6.2	6.9	5.7	7.0	5.6	4.7	58.2	82.6	98.9	84.4	98.1	100.0	87.0	99.0	100.0
	500	5.5	5.0	4.4	5.6	5.3	5.3	5.6	5.3	5.5	58.0	83.4	99.4	86.5	98.7	100.0	89.1	99.3	100.0
	1000	4.6	4.7	5.3	6.3	5.1	5.6	7.0	6.2	6.1	58.5	83.0	99.4	87.1	98.9	100.0	89.4	99.2	100.0
CD_{W+}	100	6.1	5.8	6.6	6.9	5.9	6.4	5.5	6.5	7.4	7.6	7.9	32.6	8.0	8.0	36.4	7.2	7.8	41.4
	200	5.9	4.8	5.6	5.2	6.1	5.6	5.3	5.4	4.4	5.8	6.1	39.4	6.1	6.2	43.4	5.8	6.8	42.8
	500	5.2	4.9	5.3	5.1	5.1	5.7	5.7	5.1	5.3	5.1	5.6	48.6	5.4	6.5	47.9	5.9	6.3	47.6
	1000	5.1	5.1	4.4	5.1	4.7	5.6	6.2	4.8	4.9	5.1	5.7	47.7	5.9	4.6	49.9	5.6	5.3	48.8

Note: The DGP is given by (42) with β_{i1} and β_{i2} both generated from normal distribution, and contains a single latent factor with factor strength $\alpha = 1, 2/3, 1/2$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 4: Size and power of tests of error cross-sectional dependence for panel regression with two latent factors ($m_0 = 2$) with serially independent Gaussian errors

$\hat{m} = 2$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	100.0	100.0	100.0	97.8	100.0	100.0	7.4	13.9	40.1	99.2	99.9	100.0	89.6	96.9	99.5	56.1	66.1	81.6
	200	100.0	100.0	100.0	99.3	100.0	100.0	5.3	8.2	23.8	99.5	100.0	100.0	91.8	99.2	100.0	71.6	86.3	98.1
	500	100.0	100.0	100.0	99.4	100.0	100.0	6.2	5.4	11.0	100.0	100.0	100.0	92.2	99.8	100.0	83.2	95.3	100.0
	1000	100.0	100.0	100.0	99.9	100.0	100.0	7.8	5.9	7.0	100.0	100.0	100.0	94.9	99.8	100.0	86.5	97.7	100.0
CD^*	100	5.8	5.3	5.3	5.9	5.8	5.4	9.4	6.4	5.9	23.1	34.2	63.5	33.7	50.1	79.6	83.7	97.7	100.0
	200	6.0	5.4	5.4	5.3	4.9	4.8	7.4	6.9	5.7	23.0	36.3	64.6	33.9	52.2	82.0	86.3	97.8	100.0
	500	5.7	5.4	5.5	6.6	4.3	4.9	7.9	6.1	5.7	22.3	35.7	65.1	35.0	52.6	82.8	88.9	98.8	100.0
	1000	5.1	5.4	5.7	5.1	6.3	4.8	9.4	7.0	5.8	23.5	34.6	64.8	35.2	53.6	85.2	89.9	99.1	100.0
CD_{W+}	100	5.6	4.9	6.8	5.5	5.5	8.1	6.4	6.3	8.6	6.8	7.9	38.7	6.5	8.1	44.4	8.9	10.7	62.1
	200	6.2	5.9	5.6	5.6	5.0	5.1	5.8	4.6	5.8	6.1	7.9	44.8	6.1	6.9	47.9	7.0	6.6	57.9
	500	6.4	6.3	5.5	5.5	5.4	4.3	5.5	5.3	4.9	6.3	6.6	49.4	5.7	6.4	50.1	5.9	6.2	55.9
	1000	5.2	5.1	4.7	6.6	5.1	4.6	6.0	6.4	4.3	5.3	5.6	49.9	6.8	6.3	49.8	6.2	7.2	49.9

$\hat{m} = 4$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	100.0	100.0	100.0	97.9	100.0	100.0	7.4	15.4	40.8	99.3	100.0	100.0	92.3	98.8	99.9	37.6	43.8	54.7
	200	100.0	100.0	100.0	99.3	100.0	100.0	5.9	9.4	23.5	99.6	100.0	100.0	93.4	99.5	100.0	61.3	74.7	89.3
	500	100.0	100.0	100.0	99.5	100.0	100.0	6.4	5.6	10.7	100.0	100.0	100.0	93.7	99.8	100.0	77.8	93.1	99.8
	1000	100.0	100.0	100.0	99.8	100.0	100.0	7.0	6.3	7.0	100.0	100.0	100.0	94.8	99.8	100.0	84.3	96.8	100.0
CD^*	100	7.9	8.1	16.0	7.4	8.6	13.7	10.3	9.6	10.4	27.6	43.1	77.2	37.2	56.5	86.6	79.2	95.9	100.0
	200	6.1	6.6	7.9	6.6	6.3	6.6	8.6	7.7	6.7	26.1	41.3	70.8	34.6	54.7	85.9	82.7	97.2	100.0
	500	5.7	5.7	6.2	7.0	4.8	4.7	8.0	6.6	5.8	22.7	36.4	67.4	34.9	51.9	83.4	87.3	98.3	100.0
	1000	5.0	5.1	5.5	5.8	6.8	5.0	8.8	7.1	6.0	24.2	35.0	66.3	34.8	52.8	85.9	87.9	98.8	100.0
CD_{W+}	100	6.3	7.1	25.2	5.6	6.3	17.0	6.3	5.8	10.1	6.7	7.5	47.8	5.1	8.6	42.3	7.0	7.4	33.4
	200	6.1	5.3	6.1	5.7	5.9	5.4	6.6	5.8	6.0	5.9	6.4	31.8	6.3	7.4	28.5	7.7	6.9	28.2
	500	5.9	5.7	4.7	6.1	6.3	5.3	5.9	6.1	5.3	6.2	6.1	38.5	6.3	6.4	38.8	6.5	7.3	39.0
	1000	6.1	5.4	5.1	5.9	4.8	5.3	6.3	4.5	4.9	6.6	6.2	42.5	5.6	5.5	42.9	6.5	5.6	44.1

Note: The DGP is given by (42) with β_{i1} and β_{i2} both generated from normal distribution, and contains two latent factors with factor strengths $(\alpha_1, \alpha_2) = [(1,1), (1, 2/3), (2/3, 1/2)]$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 5: Size and power of tests of error cross-sectional dependence for pure single factor model ($m_0 = 1$) with serially independent non-Gaussian errors

$\hat{m} = 1$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	64.5	86.4	97.7	5.3	8.9	23.1	6.1	6.6	8.0	23.5	37.1	57.5	69.0	87.0	97.8	80.4	94.3	99.4
	200	68.8	93.1	99.6	5.2	7.6	13.5	5.7	5.3	5.5	16.0	29.2	50.0	78.1	93.3	100.0	86.7	97.3	100.0
	500	69.1	94.0	100.0	4.8	5.8	9.8	5.8	6.0	5.4	11.1	23.8	46.5	84.5	97.5	100.0	90.2	98.8	100.0
	1000	69.4	95.4	100.0	4.9	5.0	6.9	6.0	5.0	5.1	10.0	19.1	43.8	86.3	97.9	100.0	91.8	99.4	100.0
CD^*	100	5.1	5.4	5.6	5.2	5.0	5.8	7.1	6.2	5.8	58.8	82.4	98.8	86.2	98.4	100.0	88.4	99.0	100.0
	200	3.8	4.4	4.4	5.8	5.4	5.1	5.5	5.3	5.0	58.4	81.4	99.3	86.7	98.7	100.0	89.6	99.1	100.0
	500	5.4	5.6	5.1	5.3	5.5	5.3	6.0	6.2	5.5	58.6	83.5	99.5	88.4	99.1	100.0	91.1	99.3	100.0
	1000	5.4	5.2	4.6	5.3	5.4	5.3	6.4	5.0	4.9	60.6	84.4	99.4	88.8	99.0	100.0	92.4	99.4	100.0
CD_{W+}	100	8.3	5.9	6.0	7.4	6.0	5.4	7.9	6.2	7.4	9.8	11.7	55.1	10.0	12.4	66.7	11.4	12.2	70.9
	200	8.6	5.8	6.0	9.4	7.5	5.6	9.7	6.2	5.3	10.8	11.7	66.9	11.4	14.6	72.1	12.9	12.5	71.4
	500	17.3	10.1	6.1	16.7	10.9	5.1	17.6	9.8	5.3	18.3	17.7	79.8	18.1	17.6	82.4	19.3	16.7	82.9
	1000	27.8	14.9	6.1	28.5	16.3	6.2	27.0	16.6	6.4	27.8	21.2	85.4	28.1	23.6	87.1	27.1	24.6	85.9

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	65.8	87.0	97.6	5.9	10.4	22.8	6.3	6.9	9.4	26.5	44.2	62.5	58.9	75.0	86.1	70.2	84.5	91.1
	200	68.2	93.2	99.6	5.2	6.9	13.9	5.2	5.1	5.9	18.5	32.6	55.7	73.2	90.1	98.8	83.1	95.2	99.6
	500	68.8	93.7	100.0	4.7	5.1	9.5	5.6	6.1	5.8	11.4	24.9	50.2	82.0	96.9	100.0	88.4	98.6	100.0
	1000	68.4	95.3	100.0	5.1	5.1	6.5	6.1	5.5	5.1	10.3	20.1	45.9	85.4	97.9	100.0	91.2	99.2	100.0
CD^*	100	6.2	5.9	7.2	5.9	5.5	5.9	7.1	6.7	5.8	57.5	81.7	98.3	83.8	97.8	100.0	86.4	98.6	100.0
	200	4.7	4.6	5.1	6.1	5.7	5.8	6.2	5.3	5.3	58.8	81.3	99.1	85.1	97.9	100.0	88.7	99.2	100.0
	500	5.4	5.7	5.6	5.4	5.8	5.2	6.3	6.5	5.4	58.4	82.6	99.5	87.5	99.0	100.0	90.8	99.1	100.0
	1000	5.5	5.3	4.4	5.9	5.9	5.5	6.2	5.5	4.7	60.8	83.8	99.2	88.4	98.9	100.0	91.9	99.5	100.0
CD_{W+}	100	6.4	6.4	6.1	6.5	5.8	5.6	7.0	5.6	8.1	7.8	10.9	37.6	9.1	10.2	42.1	8.6	9.5	50.1
	200	7.3	5.9	6.6	10.1	6.3	5.0	9.8	5.5	5.6	9.3	10.5	52.0	11.3	10.1	54.9	11.1	9.8	54.8
	500	15.8	8.3	5.8	13.8	9.0	5.9	15.1	8.4	5.9	16.8	15.0	73.3	15.5	16.5	73.3	17.2	15.0	74.7
	1000	27.1	15.0	6.1	27.4	15.9	6.6	27.6	16.4	6.7	26.4	21.5	81.3	26.9	22.6	82.1	27.9	23.5	82.4

Note: The DGP is given by (42) with $\beta_{i1} = \beta_{i2} = 0$ and contains a single latent factor with factor strength $\alpha = 1, 2/3, 1/2$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 6: Size and power of tests of error cross-sectional dependence for pure two factors model ($m_0 = 2$) with serially independent non-Gaussian errors

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	99.9	100.0	100.0	98.4	99.9	100.0	9.9	17.3	43.1	99.1	99.9	100.0	88.9	97.5	99.3	55.8	64.5	79.5
	200	100.0	100.0	100.0	99.2	100.0	100.0	7.6	8.9	25.3	99.8	100.0	100.0	92.2	99.1	100.0	69.8	85.8	98.3
	500	100.0	100.0	100.0	99.5	100.0	100.0	7.4	5.3	10.8	100.0	100.0	100.0	92.0	99.8	100.0	82.7	95.9	99.9
	1000	100.0	100.0	100.0	99.7	100.0	100.0	7.6	5.1	7.9	100.0	100.0	100.0	93.5	99.8	100.0	87.4	97.9	100.0
CD^*	100	5.3	5.5	5.0	5.9	4.8	5.1	10.8	6.7	5.4	22.8	35.1	61.3	34.1	51.8	79.6	83.5	97.4	100.0
	200	5.4	4.7	4.1	6.5	5.1	5.8	9.0	5.9	5.5	20.8	33.8	62.6	34.3	49.5	80.7	84.9	98.5	100.0
	500	5.6	6.1	5.9	6.3	5.1	5.2	9.0	5.9	5.0	22.0	35.1	63.6	33.8	53.8	83.8	88.9	98.9	100.0
	1000	5.8	4.7	4.7	4.8	5.1	5.9	8.6	5.7	4.5	23.4	35.1	64.5	35.9	53.3	84.1	90.4	99.3	100.0
CD_{W+}	100	7.1	6.7	7.0	7.3	5.9	7.8	8.2	6.1	10.8	9.4	11.8	43.5	8.9	10.5	52.6	11.4	13.0	70.0
	200	9.1	7.0	5.9	8.8	6.5	5.1	9.6	6.8	5.7	10.6	13.1	58.1	10.5	11.6	60.6	12.9	12.7	70.8
	500	15.4	9.3	5.8	16.0	8.6	5.1	17.5	9.2	6.1	17.3	16.5	76.8	17.5	15.4	75.8	19.1	16.4	80.1
	1000	26.9	18.2	5.6	29.0	16.0	5.9	30.3	17.7	6.7	27.4	24.0	82.6	27.6	22.2	83.6	30.0	26.3	85.6

$\hat{m} = 4$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	99.8	100.0	100.0	98.4	99.9	100.0	9.3	16.4	44.1	99.3	99.9	100.0	91.5	98.4	99.6	38.6	43.6	57.3
	200	100.0	100.0	100.0	99.4	100.0	100.0	6.3	8.5	27.1	99.8	100.0	100.0	93.1	99.2	100.0	60.3	73.9	88.8
	500	100.0	100.0	100.0	99.5	100.0	100.0	7.2	5.1	11.4	100.0	100.0	100.0	92.8	99.8	100.0	79.6	93.3	99.7
	1000	100.0	100.0	100.0	99.8	100.0	100.0	7.7	5.3	8.6	100.0	100.0	100.0	93.9	99.8	100.0	85.7	96.8	100.0
CD^*	100	7.4	8.4	14.0	7.5	6.2	8.0	10.5	8.8	8.0	28.5	43.3	75.0	35.7	53.8	81.7	80.0	95.3	100.0
	200	6.2	5.8	7.2	7.2	5.6	6.7	8.4	6.6	7.4	24.1	37.2	69.4	34.5	50.5	82.1	82.2	98.0	100.0
	500	7.0	6.1	6.9	6.3	5.3	5.5	9.1	5.9	5.1	24.2	37.5	65.5	34.2	52.7	84.1	87.8	98.5	100.0
	1000	5.9	4.7	4.4	5.4	5.1	6.2	9.2	5.4	5.0	24.8	37.6	65.1	36.9	52.9	84.4	89.3	99.2	100.0
CD_{W+}	100	6.5	7.8	26.9	6.6	6.9	10.7	6.3	6.6	9.7	7.8	10.1	51.2	7.1	9.9	44.2	7.9	9.2	37.8
	200	8.6	5.5	6.8	9.0	7.1	5.7	8.3	6.5	5.3	8.8	9.7	39.5	9.6	12.6	50.0	10.1	10.1	38.3
	500	15.5	8.7	5.1	14.9	9.0	5.8	14.8	9.5	6.2	15.9	13.7	63.9	14.9	13.9	69.8	16.8	14.1	64.9
	1000	29.8	14.1	7.0	27.1	16.1	6.1	28.1	14.7	5.5	27.5	20.5	76.4	26.6	23.2	80.5	28.0	21.7	78.5

Note: The DGP is given by (42) with $\beta_{i1} = \beta_{i2} = 0$ and contains two latent factors with factor strengths $(\alpha_1, \alpha_2) = [(1,1), (1, 2/3), (2/3, 1/2)]$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 7: Size and power of tests of error cross-sectional dependence for panel regression with one latent factor ($m_0 = 1$) with serially independent non-Gaussian errors

$\hat{m} = 1$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	65.7	87.9	98.4	6.1	9.3	20.3	6.9	6.6	9.0	26.1	38.0	56.0	70.9	87.5	97.5	81.8	95.2	99.4
	200	67.6	91.4	99.6	5.8	7.2	13.0	5.5	6.0	6.3	16.6	30.8	49.6	80.0	93.9	99.9	86.6	98.5	100.0
	500	67.4	94.9	99.9	5.3	5.0	7.4	6.9	5.8	4.9	12.0	22.3	45.6	85.3	97.4	100.0	91.3	98.9	100.0
	1000	68.7	95.3	100.0	6.3	4.9	7.0	6.7	6.3	5.0	10.0	19.3	44.1	86.1	98.1	100.0	90.9	99.3	100.0
CD^*	100	5.3	4.9	5.1	5.9	5.8	5.9	7.6	6.8	6.1	57.0	82.4	98.4	86.0	98.2	100.0	89.1	98.8	100.0
	200	6.0	5.4	5.1	6.2	6.1	5.3	6.3	6.8	5.5	58.5	83.0	99.2	87.6	98.8	100.0	89.6	99.2	100.0
	500	5.5	5.1	4.5	6.2	5.0	4.5	7.3	6.1	4.9	60.2	83.2	99.5	90.0	98.8	100.0	92.2	99.1	100.0
	1000	5.5	5.4	4.8	7.0	5.6	5.3	6.7	6.5	5.4	58.4	84.3	99.2	88.9	99.2	100.0	91.3	99.2	100.0
CD_{W+}	100	7.5	7.0	5.3	7.0	5.3	5.9	7.2	6.7	7.5	9.0	11.7	53.8	8.5	11.1	64.9	9.6	13.8	69.1
	200	9.0	7.6	4.5	8.6	7.8	5.1	8.8	7.0	5.9	10.3	13.4	64.9	10.7	14.1	70.4	11.6	16.0	71.9
	500	15.9	8.9	6.1	15.8	9.5	5.8	16.1	9.7	5.7	16.6	15.8	77.5	17.8	15.7	80.4	18.2	15.8	81.0
	1000	26.0	14.9	6.0	26.1	16.9	5.8	26.8	14.4	6.1	24.5	22.6	84.5	25.7	23.5	85.2	25.4	21.3	86.0

$\hat{m} = 2$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	67.2	88.4	98.4	5.9	9.2	21.3	7.0	6.0	10.0	29.7	45.1	64.4	60.3	73.3	83.5	71.4	85.2	90.6
	200	67.3	90.6	99.8	5.5	6.9	13.0	6.1	5.3	6.5	18.1	35.1	55.5	72.9	88.9	98.7	80.8	95.8	99.4
	500	67.3	94.7	100.0	6.1	4.8	7.6	6.8	6.3	4.6	12.6	23.4	48.8	83.9	96.5	99.9	89.1	98.3	100.0
	1000	68.4	95.0	100.0	6.2	4.7	7.3	7.0	6.4	5.1	10.0	19.4	45.9	84.8	97.9	100.0	89.6	98.9	100.0
CD^*	100	6.1	5.9	6.5	6.2	6.3	6.4	7.7	7.0	6.9	56.1	81.2	98.0	84.9	97.8	100.0	86.8	98.3	100.0
	200	6.8	5.8	5.8	6.3	6.1	5.1	6.9	6.3	5.8	58.3	82.9	99.1	85.5	98.5	100.0	87.7	99.0	100.0
	500	5.7	5.0	5.1	6.8	5.1	4.1	7.2	6.5	4.9	59.6	83.5	99.6	89.4	98.7	100.0	90.9	99.0	100.0
	1000	5.5	5.4	5.0	7.2	5.7	5.6	7.0	6.8	5.6	57.8	84.5	99.2	87.3	99.0	100.0	90.6	99.1	100.0
CD_{W+}	100	6.5	6.4	7.1	6.3	5.8	6.6	7.3	5.9	8.5	8.9	10.0	38.2	8.1	8.6	40.8	8.8	10.0	48.9
	200	8.6	6.5	5.2	9.0	6.9	5.8	8.8	6.5	6.0	10.2	10.7	50.6	11.4	11.3	54.8	9.5	11.9	56.2
	500	14.1	9.5	5.0	15.5	10.0	5.6	14.9	7.7	4.9	15.8	13.4	70.7	15.8	15.4	72.8	15.9	12.8	73.1
	1000	26.0	15.3	6.9	24.0	14.5	6.4	25.1	14.5	6.0	25.0	20.8	79.3	24.1	21.7	81.0	25.1	20.5	81.1

Note: The DGP is given by (42) with β_{i1} and β_{i2} both generated from normal distribution, and contains a single latent factor with factor strength $\alpha = 1, 2/3, 1/2$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 8: Size and power of tests of error cross-sectional dependence for panel regression with two latent factors ($m_0 = 2$) with serially independent non-Gaussian errors

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	n/T	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	100.0	100.0	100.0	97.9	100.0	100.0	8.6	13.7	39.7	98.9	100.0	100.0	89.5	97.3	99.6	57.7	68.5	80.9
	200	100.0	100.0	100.0	99.0	100.0	100.0	6.6	7.4	22.9	99.7	100.0	100.0	91.4	99.1	100.0	72.8	87.4	98.3
	500	100.0	100.0	100.0	99.5	100.0	100.0	7.0	6.0	11.8	100.0	100.0	100.0	92.6	99.6	100.0	83.0	95.9	99.9
	1000	100.0	100.0	100.0	99.6	100.0	100.0	6.6	5.9	7.6	100.0	100.0	100.0	93.0	99.9	100.0	85.7	97.5	100.0
CD^*	100	5.3	5.4	5.4	6.9	5.7	4.9	9.3	6.3	6.2	23.0	34.5	63.1	32.6	49.6	78.3	83.7	97.7	100.0
	200	5.2	5.1	5.1	6.3	4.7	5.5	8.2	6.7	5.4	21.2	35.9	64.3	34.5	50.2	81.1	87.6	98.5	100.0
	500	5.9	4.7	5.8	5.9	5.3	4.8	9.2	5.9	5.9	21.9	37.6	64.9	36.2	53.4	83.9	88.4	99.0	100.0
	1000	6.8	4.8	5.5	6.5	4.5	5.0	7.6	6.4	5.0	23.4	33.6	63.2	34.2	52.7	86.4	88.9	98.9	100.0
CD_{W+}	100	7.1	6.4	5.6	7.1	5.8	6.8	7.2	5.6	9.9	8.4	10.5	42.2	9.0	9.9	50.6	10.7	13.1	67.6
	200	10.1	7.0	5.1	9.9	7.5	4.9	9.5	5.9	6.1	10.8	12.1	58.1	10.9	12.4	58.8	12.4	12.0	69.0
	500	16.2	10.4	6.1	15.8	9.0	5.4	16.3	9.8	5.8	16.5	15.2	74.4	16.0	14.2	74.8	17.7	17.6	79.8
	1000	26.6	14.7	6.2	28.1	18.0	5.2	28.1	15.0	5.7	25.4	21.5	84.1	27.5	24.2	82.8	27.4	22.2	85.1

$\hat{m} = 4$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	n/T	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
CD	100	99.9	100.0	100.0	98.4	99.9	100.0	8.1	15.3	41.7	99.2	100.0	100.0	91.6	98.1	99.5	38.4	45.2	54.9
	200	100.0	100.0	100.0	99.1	100.0	100.0	6.1	8.6	23.5	99.7	100.0	100.0	92.6	99.2	100.0	61.2	76.4	88.2
	500	100.0	100.0	100.0	99.5	100.0	100.0	7.4	5.6	11.6	100.0	100.0	100.0	92.8	99.8	100.0	77.9	94.0	99.9
	1000	100.0	100.0	100.0	99.7	100.0	100.0	6.8	6.1	7.4	100.0	100.0	100.0	93.0	99.9	100.0	83.2	97.2	100.0
CD^*	100	8.2	8.9	15.0	8.9	8.6	11.0	9.8	7.9	9.0	29.3	41.4	75.1	34.7	52.6	82.0	80.0	95.7	100.0
	200	5.8	6.0	7.1	6.9	5.5	6.9	8.9	8.1	7.0	25.1	39.8	70.7	35.3	51.0	82.4	84.1	97.7	100.0
	500	6.1	5.2	5.8	6.5	5.4	5.4	9.6	6.7	6.2	24.3	39.4	66.7	36.3	53.9	84.4	85.8	98.3	100.0
	1000	7.0	4.8	5.6	7.0	5.0	5.4	7.9	6.6	5.1	25.1	35.2	64.8	35.4	53.2	86.7	86.7	98.7	100.0
CD_{W+}	100	7.3	7.6	26.2	6.8	6.8	15.5	5.5	6.4	12.0	8.9	11.1	48.9	7.6	9.6	43.3	6.2	8.6	38.0
	200	8.2	6.8	5.9	9.0	7.3	5.8	9.5	5.9	5.7	9.6	10.1	38.7	8.7	11.7	47.8	10.4	9.7	38.0
	500	14.5	9.6	4.9	15.1	8.8	6.6	13.3	8.1	5.3	15.8	13.6	61.3	15.6	13.8	68.8	14.6	12.5	62.7
	1000	26.3	13.2	6.9	25.2	14.4	6.3	24.7	16.0	5.6	25.1	19.8	75.1	26.4	21.2	77.8	23.4	21.5	77.5

Note: The DGP is given by (42) with β_{i1} and β_{i2} both generated from normal distribution, and contains two latent factors with factor strengths $(\alpha_1, \alpha_2) = [(1, 1), (1, 2/3), (2/3, 1/2)]$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 9: Size and power of variance adjusted tests of error cross-sectional dependence for pure single factor model ($m_0 = 1$) with serially correlated Gaussian errors

$\hat{m} = 1$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	41.0	71.4	93.8	5.3	6.6	12.9	6.7	6.5	5.9	12.8	26.1	44.1	59.8	74.9	91.8	67.5	85.9	98.3
	200	37.0	75.8	97.2	5.2	5.0	7.5	6.0	5.9	5.5	7.7	18.8	34.8	65.7	83.0	97.7	74.5	90.1	99.9
	500	33.5	76.6	98.6	6.7	4.8	5.7	6.5	5.4	5.6	5.3	14.4	31.9	74.2	90.1	99.3	73.0	94.2	99.7
	1000	28.8	76.6	99.6	10.1	5.9	4.5	6.9	6.6	4.2	4.4	10.4	27.4	77.9	92.3	100.0	78.7	93.7	100.0
Variance adjusted CD^*	100	5.6	5.3	5.6	6.2	6.2	6.3	6.8	6.7	4.9	40.9	61.7	89.3	69.4	88.9	99.7	73.2	92.4	99.8
	200	5.0	4.7	5.9	6.5	4.4	6.1	6.6	6.5	5.4	45.6	66.3	92.9	72.8	90.6	100.0	75.8	93.4	100.0
	500	6.3	5.0	5.4	7.2	5.5	4.5	6.8	5.8	5.6	46.2	63.5	93.5	77.1	93.7	99.9	73.7	94.8	99.8
	1000	6.7	5.2	5.1	11.4	6.3	4.0	6.9	6.4	4.4	49.9	66.7	94.2	79.9	93.4	100.0	78.8	94.2	100.0
Variance adjusted CD_{W+}	100	16.0	15.5	12.8	16.0	14.5	14.9	15.7	16.9	17.0	22.8	39.1	86.8	23.6	38.8	93.9	26.5	43.3	94.4
	200	20.2	25.3	19.8	22.8	26.1	20.6	20.5	26.8	22.4	29.1	55.8	98.6	29.6	57.0	99.8	35.7	59.1	99.5
	500	30.4	54.2	51.5	30.5	54.7	51.1	40.1	58.0	51.7	35.1	79.6	100.0	34.9	78.9	100.0	50.9	82.4	100.0
	1000	36.6	80.8	85.5	39.5	78.3	87.0	53.9	82.6	86.9	42.5	92.6	100.0	42.9	92.2	100.0	58.1	94.6	100.0

$\hat{m} = 2$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	41.7	72.2	93.3	4.2	5.6	12.8	6.1	6.3	6.1	15.5	28.1	50.0	52.4	66.5	81.2	61.4	79.6	89.2
	200	38.2	75.3	96.9	4.8	5.0	7.4	6.4	5.2	5.4	8.1	19.4	37.4	63.2	80.7	96.3	71.4	88.2	99.1
	500	34.1	77.0	98.7	7.3	4.6	5.6	6.8	5.2	5.6	5.1	15.5	33.9	73.5	88.6	99.1	71.3	93.3	99.7
	1000	29.5	77.0	99.6	11.2	5.8	3.9	7.0	6.5	4.4	4.5	10.4	28.5	77.8	90.9	100.0	77.2	93.2	100.0
Variance adjusted CD^*	100	4.9	4.8	5.7	5.9	6.7	5.8	7.0	7.4	6.2	41.1	62.3	88.8	69.1	89.6	99.5	71.6	92.0	99.9
	200	5.6	4.2	6.1	6.8	4.9	7.2	7.1	6.0	5.5	45.5	65.7	93.6	71.5	90.4	99.7	75.3	93.2	100.0
	500	5.6	5.5	5.3	8.2	5.1	5.1	7.0	5.3	5.8	45.2	63.4	93.9	76.7	93.7	99.9	73.1	94.0	99.8
	1000	6.9	5.7	5.2	11.7	7.7	4.1	7.1	6.7	4.3	50.2	67.3	94.4	80.3	93.5	100.0	77.8	94.1	100.0
Variance adjusted CD_{W+}	100	13.4	15.6	14.1	14.9	14.3	13.4	15.0	15.6	16.9	19.3	34.4	80.7	19.8	32.9	84.6	19.6	35.0	86.1
	200	18.3	22.9	20.2	19.9	25.2	18.3	17.4	27.0	21.8	26.3	47.6	97.9	24.5	49.3	98.5	27.5	50.0	98.9
	500	27.3	52.1	49.7	27.0	54.1	47.9	33.7	55.6	51.0	31.3	77.7	100.0	33.4	78.3	100.0	40.7	80.5	100.0
	1000	36.3	80.2	82.9	36.5	78.7	83.5	50.1	84.4	83.8	40.6	93.5	100.0	42.0	92.2	100.0	54.4	95.6	100.0

Note: The DGP is given by (42) with $\beta_{i1} = \beta_{i2} = 0$ and contains a single latent factor with factor strength $\alpha = 1, 2/3, 1/2$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 10: Size and power of variance adjusted tests of error cross-sectional dependence for pure two factors model ($m_0 = 2$) with serially correlated Gaussian errors

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	98.2	100.0	100.0	87.5	98.7	100.0	8.4	6.4	22.4	92.9	99.1	100.0	67.6	90.6	97.8	54.0	61.0	69.8
	200	99.0	100.0	100.0	88.7	99.8	100.0	9.0	5.2	9.8	95.4	99.7	100.0	66.6	94.6	99.4	64.8	77.8	94.0
	500	99.2	100.0	100.0	88.2	100.0	100.0	12.2	8.2	7.2	95.1	100.0	100.0	63.4	96.8	100.0	80.8	90.8	99.2
	1000	98.0	100.0	100.0	84.6	99.9	100.0	13.4	8.4	4.8	92.6	100.0	100.0	52.4	95.4	100.0	81.6	93.0	99.8
Variance adjusted CD^*	100	5.0	5.0	4.5	5.9	5.4	4.4	12.2	5.8	6.0	17.4	22.0	40.3	27.0	34.8	57.8	71.6	91.4	99.6
	200	5.9	4.2	4.8	7.1	5.7	4.2	10.8	6.6	6.8	19.2	24.1	42.1	27.8	35.8	64.6	75.0	91.0	99.8
	500	8.4	4.8	4.3	9.0	6.2	4.5	14.2	9.6	6.6	25.3	27.1	45.3	30.0	41.4	71.2	83.8	93.0	99.8
	1000	13.7	4.7	5.1	12.5	7.7	5.6	14.4	11.6	6.6	33.6	28.4	46.4	42.6	41.2	65.8	83.6	94.8	100.0
Variance adjusted CD_{W+}	100	17.2	16.5	14.1	16.8	18.7	20.9	16.8	17.8	23.6	23.3	37.8	85.9	21.2	33.4	76.6	26.0	43.6	95.4
	200	20.9	29.2	24.2	19.4	27.6	24.6	28.2	30.6	26.2	28.0	53.0	98.3	21.2	39.2	90.8	35.4	58.2	99.6
	500	32.4	54.9	51.7	29.1	55.4	52.7	45.4	63.2	55.4	38.1	79.6	100.0	27.6	51.6	99.4	50.4	80.8	100.0
	1000	39.4	80.9	85.9	38.5	81.3	87.7	59.8	87.6	88.0	43.8	94.0	100.0	36.4	79.0	100.0	60.2	96.4	100.0

$\hat{m} = 4$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	98.7	100.0	100.0	88.7	98.8	100.0	5.8	6.2	20.4	95.4	99.7	99.9	73.2	94.0	99.4	44.4	46.6	46.2
	200	99.5	100.0	100.0	91.2	99.8	100.0	6.6	5.2	12.2	96.3	99.9	100.0	69.4	95.2	99.6	59.8	69.6	87.6
	500	99.2	100.0	100.0	88.4	100.0	100.0	11.4	7.6	7.6	95.5	100.0	100.0	61.4	97.6	100.0	77.4	87.6	98.4
	1000	98.1	100.0	100.0	83.5	99.7	100.0	12.2	8.2	5.8	92.4	100.0	100.0	51.6	96.6	100.0	78.8	91.8	99.8
Variance adjusted CD^*	100	5.6	6.8	9.3	7.8	7.0	6.6	9.2	6.2	8.8	20.7	27.5	51.6	28.4	43.4	65.8	71.0	89.4	99.8
	200	5.5	5.8	5.7	7.1	6.9	5.1	10.2	7.2	8.4	20.3	26.9	48.1	29.4	38.2	67.2	73.8	90.6	100.0
	500	8.9	5.4	4.1	9.3	6.6	5.4	13.0	10.8	6.4	25.9	28.0	46.6	34.0	43.2	71.6	82.6	92.8	99.6
	1000	14.3	4.6	5.1	13.5	7.6	5.5	14.2	10.8	7.2	34.9	29.1	47.9	42.2	43.6	66.0	81.0	93.4	100.0
Variance adjusted CD_{W+}	100	13.4	17.7	34.6	13.2	16.7	32.5	14.0	20.6	25.0	17.4	30.5	82.4	18.2	30.2	80.6	18.2	34.2	80.8
	200	18.8	26.3	23.4	20.6	24.0	25.8	22.8	25.6	22.4	22.7	45.5	93.6	23.6	43.0	95.6	27.6	45.2	95.4
	500	29.6	53.6	48.4	28.9	52.6	49.9	32.6	54.4	53.4	33.9	76.1	100.0	34.0	77.0	99.8	35.6	77.4	100.0
	1000	38.9	79.3	84.4	40.9	79.7	82.9	45.6	83.0	81.2	43.0	92.6	100.0	45.6	91.4	100.0	49.8	93.0	100.0

Note: The DGP is given by (42) with $\beta_{i1} = \beta_{i2} = 0$, and contains two latent factors with factor strengths $(\alpha_1, \alpha_2) = [(1,1), (1, 2/3), (2/3, 1/2)]$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 11: Size and power of variance adjusted tests of error cross-sectional dependence for panel regression with one latent factor ($m_0 = 1$) with serially correlated Gaussian errors

$\hat{m} = 1$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	44.1	73.2	94.5	8.0	5.6	11.5	6.4	5.6	6.6	14.5	24.7	43.0	62.8	78.4	93.3	69.0	85.3	97.5
	200	39.6	74.0	97.0	5.2	5.4	8.3	6.1	7.1	4.6	9.3	18.9	36.3	68.1	87.3	98.1	74.6	91.7	99.9
	500	32.9	77.8	96.3	8.9	5.1	5.0	8.7	6.0	4.9	5.3	13.4	29.4	75.4	90.8	99.6	77.8	94.5	99.8
	1000	27.7	76.2	77.2	8.4	6.4	5.3	7.3	5.3	4.0	5.0	9.3	27.9	80.0	93.1	99.7	79.0	95.4	100.0
Variance adjusted CD^*	100	5.0	4.1	5.4	8.6	5.9	5.8	7.0	6.2	5.4	40.8	62.1	88.3	71.7	90.9	99.7	74.6	91.5	99.7
	200	5.3	6.6	5.0	5.9	6.5	5.8	6.4	7.0	5.1	42.5	64.7	91.8	74.6	93.1	99.9	76.9	93.6	100.0
	500	5.7	5.1	4.9	9.5	6.3	4.8	8.8	6.7	5.0	48.1	65.1	93.8	78.1	93.6	99.9	78.4	95.3	99.9
	1000	8.9	4.0	4.2	9.2	6.8	5.1	7.3	5.2	4.2	51.7	66.8	93.5	81.4	94.9	100.0	79.3	95.7	100.0
Variance adjusted CD_{W+}	100	14.4	16.1	12.5	14.7	15.8	15.0	13.6	16.2	16.1	19.6	35.6	88.5	20.3	39.3	93.8	19.1	40.5	94.1
	200	16.1	25.2	20.6	18.2	25.1	22.8	20.5	26.9	19.8	22.6	50.7	99.4	23.8	51.8	99.3	26.4	54.6	99.6
	500	26.6	50.1	50.7	26.8	48.8	52.6	26.0	54.1	50.8	31.0	75.4	100.0	31.4	75.9	100.0	30.0	78.5	100.0
	1000	29.0	75.5	76.0	34.1	74.3	85.6	34.4	78.2	82.1	35.2	90.0	100.0	36.1	90.1	100.0	37.9	93.4	100.0

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	43.4	72.9	93.8	7.2	5.7	11.3	6.0	5.9	7.3	15.6	28.3	51.4	52.8	67.8	82.6	60.8	77.5	88.4
	200	40.1	73.7	96.9	4.7	5.8	8.0	5.4	6.8	4.7	10.4	20.3	38.6	62.6	84.0	96.3	68.4	87.4	98.9
	500	34.4	77.7	96.3	7.5	5.0	5.1	7.8	5.5	5.2	5.1	14.0	31.5	71.3	89.6	99.4	74.2	93.9	99.7
	1000	27.1	75.2	75.9	7.1	5.3	4.8	6.0	5.8	3.8	5.5	10.1	29.5	76.2	92.6	99.7	74.3	95.3	100.0
Variance adjusted CD^*	100	5.5	4.4	5.7	8.4	6.5	6.6	6.9	7.1	7.0	40.4	63.3	90.4	67.9	90.3	99.9	72.7	92.0	99.9
	200	5.2	6.0	5.8	6.0	6.9	5.3	6.0	6.9	5.3	41.4	64.5	92.4	72.1	92.6	100.0	74.6	91.9	100.0
	500	6.0	5.3	5.1	8.3	6.4	5.7	7.8	5.7	5.3	47.7	65.4	94.6	75.3	93.2	99.9	75.7	95.4	100.0
	1000	9.1	4.3	4.4	7.5	6.0	5.6	6.2	5.8	4.1	51.2	66.5	93.7	78.0	94.8	99.9	75.5	95.7	100.0
Variance adjusted CD_{W+}	100	11.6	15.2	14.8	12.5	15.7	15.5	14.5	14.7	14.7	16.4	30.5	80.4	17.6	31.3	84.5	18.4	31.1	85.0
	200	19.3	22.3	19.7	18.3	21.9	20.8	15.3	23.8	22.6	21.8	44.1	97.3	24.0	46.0	97.6	20.4	48.8	97.0
	500	23.8	46.9	50.5	25.7	47.0	46.7	26.9	50.6	52.3	28.7	71.3	100.0	30.0	70.3	100.0	31.8	75.9	100.0
	1000	30.4	75.9	76.0	31.2	74.4	83.5	32.0	74.4	83.8	32.8	91.6	100.0	33.2	90.0	100.0	36.2	88.4	100.0

Note:The DGP is given by (42) with β_{i1} and β_{i2} both generated from normal distribution, and contains a single latent factor with factor strength $\alpha = 1, 2/3, 1/2$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 12: Size and power of variance adjusted tests of error cross-sectional dependence for panel regression with two latent factors ($m_0 = 2$) with serially correlated Gaussian errors

$\hat{m} = 2$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	98.6	100.0	100.0	89.0	98.6	100.0	6.3	8.1	18.1	93.2	99.5	99.9	68.4	90.3	98.6	54.7	65.3	76.1
	200	99.6	100.0	100.0	91.0	99.5	100.0	8.6	6.2	9.9	94.6	100.0	100.0	68.9	92.8	99.6	67.1	79.0	95.2
	500	99.2	100.0	100.0	88.3	99.9	100.0	7.5	6.1	6.2	95.3	100.0	100.0	61.8	95.7	100.0	75.1	91.6	98.9
	1000	97.4	100.0	100.0	84.5	99.8	100.0	11.3	7.3	3.5	91.2	100.0	100.0	55.5	96.1	100.0	81.1	90.8	100.0
Variance adjusted CD^*	100	4.4	4.4	5.6	6.4	4.9	4.2	9.0	9.0	6.0	18.0	26.0	37.6	26.4	34.6	57.6	71.8	90.2	99.4
	200	6.0	3.2	4.8	6.2	5.6	4.3	10.3	7.2	6.0	19.1	22.0	43.8	27.6	37.3	60.9	75.7	92.6	99.8
	500	8.4	5.2	4.2	7.6	5.4	5.2	9.0	8.6	6.0	25.2	26.3	44.3	34.0	42.3	65.2	78.8	95.5	99.9
	1000	15.0	4.8	5.6	13.6	6.6	5.7	12.5	8.9	5.3	35.6	27.4	45.4	40.3	43.5	68.0	82.9	92.0	100.0
Variance adjusted CD_{W+}	100	14.6	14.6	15.6	15.7	15.8	16.6	17.3	19.2	22.8	19.0	36.1	86.3	20.6	36.1	88.5	23.9	43.0	95.5
	200	19.2	23.6	23.8	18.9	25.1	22.8	23.8	28.6	25.9	24.4	50.3	98.3	26.0	50.5	98.8	30.3	56.6	99.6
	500	28.6	52.8	48.2	26.8	50.8	48.7	38.2	53.3	50.1	33.2	76.8	100.0	32.8	75.0	100.0	44.1	79.6	100.0
	1000	36.4	76.4	84.0	33.0	74.3	84.7	49.0	81.4	83.9	40.0	89.9	100.0	36.8	89.0	100.0	51.2	84.0	100.0

$\hat{m} = 4$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	98.6	100.0	100.0	90.3	98.7	100.0	6.8	8.2	18.0	94.3	99.5	99.9	73.5	93.0	99.0	39.4	44.4	50.2
	200	99.0	100.0	100.0	92.2	99.6	100.0	5.4	4.7	12.2	96.5	100.0	100.0	72.0	94.7	99.6	58.0	73.5	85.7
	500	99.6	100.0	100.0	89.7	99.8	100.0	6.1	5.7	5.7	94.8	100.0	100.0	64.5	96.3	100.0	72.8	88.6	99.2
	1000	96.6	100.0	100.0	84.1	99.8	100.0	10.0	7.3	3.6	91.2	100.0	100.0	55.3	96.7	100.0	78.0	88.9	100.0
Variance adjusted CD^*	100	6.0	5.8	8.0	6.6	7.4	8.0	9.9	9.5	7.8	20.0	30.3	49.9	27.5	40.2	65.4	67.9	88.3	99.3
	200	6.6	3.6	5.2	7.4	6.5	5.6	7.8	7.5	6.3	19.8	24.2	49.0	27.8	39.3	67.6	72.5	90.8	99.8
	500	9.4	5.6	4.4	7.8	5.5	5.1	7.7	7.6	5.8	26.2	27.6	46.6	33.3	42.7	67.5	78.3	93.7	99.9
	1000	16.0	4.6	5.4	12.4	6.5	5.4	10.2	9.3	5.5	35.5	27.9	47.8	42.1	42.8	68.7	80.1	91.0	100.0
Variance adjusted CD_{W+}	100	13.8	17.2	31.8	10.8	16.0	30.2	13.5	17.9	23.8	16.2	26.3	81.2	14.2	26.9	78.1	16.8	31.6	78.4
	200	13.2	22.6	23.8	15.4	18.4	24.6	18.6	24.0	22.9	19.5	40.6	94.6	20.2	37.9	93.7	22.4	44.9	95.4
	500	27.0	46.8	46.8	24.5	46.3	49.6	28.5	50.4	49.5	30.3	70.1	99.9	27.3	69.5	100.0	32.4	72.9	100.0
	1000	32.6	75.2	82.4	33.3	73.5	81.0	38.5	78.9	82.4	36.9	88.4	100.0	35.8	87.2	100.0	44.7	79.9	100.0

Note: The DGP is given by (42) with β_{i1} and β_{i2} both generated from normal distribution, and contains two latent factors with factor strengths $(\alpha_1, \alpha_2) = [(1,1), (1, 2/3), (2/3, 1/2)]$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 13: Size and power of variance adjusted tests of error cross-sectional dependence for pure single factor model ($m_0 = 1$) with serially correlated non-Gaussian errors

$\hat{m} = 1$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	41.6	70.9	93.4	4.6	7.0	10.6	6.4	5.8	6.3	14.1	26.1	41.6	59.7	76.6	94.1	66.4	87.0	97.4
	200	38.4	73.2	96.2	6.0	5.7	7.1	6.3	5.6	5.2	9.3	16.4	34.6	67.9	85.3	98.8	72.2	91.8	99.8
	500	36.3	76.5	98.8	7.2	6.4	6.3	7.3	7.2	5.5	5.4	12.1	31.0	74.7	89.6	100.0	76.0	93.8	100.0
	1000	29.4	75.6	99.5	8.9	6.6	4.5	6.8	6.7	4.0	5.3	9.5	26.8	78.2	92.5	99.9	78.6	91.8	100.0
Variance adjusted CD^*	100	5.4	5.0	3.3	5.8	5.7	4.2	6.7	5.5	5.3	41.0	61.3	91.7	70.4	89.8	99.9	72.1	93.1	99.9
	200	4.5	6.2	6.2	6.9	6.6	5.7	6.6	6.3	6.7	42.6	66.0	92.9	74.7	92.3	99.8	74.2	94.3	99.9
	500	5.1	5.2	4.7	8.4	6.8	4.6	7.5	7.2	5.6	44.1	66.1	93.8	76.8	93.1	100.0	76.9	94.1	100.0
	1000	7.6	5.9	4.0	9.3	6.9	4.1	6.8	6.8	4.2	48.9	67.9	93.8	79.5	94.8	100.0	79.1	92.0	100.0
Variance adjusted CD_{W+}	100	25.3	19.5	14.1	30.1	23.4	15.6	28.2	23.0	20.1	37.7	47.5	89.8	42.2	55.2	94.5	42.6	58.4	96.2
	200	57.4	50.8	26.6	57.1	52.4	29.1	60.2	56.4	30.1	66.3	81.6	99.4	64.6	81.2	99.9	69.7	85.2	99.6
	500	95.6	96.5	75.9	95.0	95.3	77.8	95.7	96.3	79.8	96.4	99.5	100.0	96.4	99.5	100.0	97.4	99.9	100.0
	1000	100.0	100.0	99.3	99.9	100.0	99.5	99.9	100.0	99.6	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	42.9	71.5	93.4	6.0	6.4	10.7	5.3	5.2	7.2	14.1	29.8	46.6	53.4	66.7	82.2	60.4	78.0	88.5
	200	38.2	74.2	96.6	6.3	5.0	7.9	6.1	6.0	6.2	9.0	18.1	40.0	63.9	81.4	97.0	69.5	90.2	99.4
	500	35.9	76.2	98.8	7.6	6.0	6.3	7.9	7.3	6.0	5.4	11.7	33.6	73.0	88.3	99.8	75.8	92.7	100.0
	1000	28.4	75.2	99.4	9.3	6.4	4.2	7.1	6.8	4.1	4.9	9.1	27.4	77.6	91.8	99.9	76.9	91.6	100.0
Variance adjusted CD^*	100	5.5	5.4	4.2	7.2	6.2	5.2	6.5	6.0	6.4	40.5	62.9	92.4	69.9	89.7	99.7	72.1	93.3	99.9
	200	4.7	6.3	6.2	7.0	6.4	5.6	6.9	7.1	7.6	42.8	65.6	92.7	75.5	91.8	99.6	73.7	93.6	99.8
	500	5.6	5.2	5.2	9.2	7.1	4.8	8.3	7.4	5.7	45.0	65.7	94.3	77.9	92.3	100.0	76.7	93.8	100.0
	1000	8.9	6.0	4.5	9.9	7.0	4.3	7.2	6.8	4.3	50.2	67.6	93.5	80.5	93.9	100.0	77.4	91.8	100.0
Variance adjusted CD_{W+}	100	21.2	17.9	17.3	20.9	21.0	16.2	25.2	21.1	19.8	30.0	39.0	81.3	31.8	45.3	84.7	35.0	44.4	88.8
	200	47.0	43.0	26.0	48.4	44.3	27.7	48.3	46.9	29.6	56.7	72.8	98.5	57.1	73.2	99.2	58.2	75.6	98.9
	500	93.1	92.9	75.0	93.0	92.6	72.4	94.1	94.5	75.2	95.0	98.9	100.0	95.1	98.6	100.0	94.9	99.2	100.0
	1000	99.8	99.9	99.2	100.0	100.0	99.1	99.9	100.0	99.6	99.9	100.0	100.0	99.9	100.0	100.0	99.9	100.0	100.0

Note: The DGP is given by (42) with $\beta_{i1} = \beta_{i2} = 0$ and contains a single latent factor with factor strength $\alpha = 1, 2/3, 1/2$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 14: Size and power of variance adjusted tests of error cross-sectional dependence for pure two factors model ($m_0 = 2$) with serially correlated non-Gaussian errors

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	98.6	100.0	100.0	89.0	98.6	100.0	6.3	8.1	18.1	93.2	99.5	99.9	68.4	90.3	98.6	54.7	65.3	76.1
	200	99.6	100.0	100.0	91.0	99.5	100.0	8.6	6.2	9.9	94.6	100.0	100.0	68.9	92.8	99.6	67.1	79.0	95.2
	500	99.2	100.0	100.0	88.3	99.9	100.0	7.5	6.1	6.2	95.3	100.0	100.0	61.8	95.7	100.0	75.1	91.6	98.9
	1000	97.4	100.0	100.0	84.5	99.8	100.0	11.3	7.3	3.5	91.2	100.0	100.0	55.5	96.1	100.0	81.1	90.8	100.0
Variance adjusted CD^*	100	4.4	4.4	5.6	6.4	4.9	4.2	9.0	9.0	6.0	18.0	26.0	37.6	26.4	34.6	57.6	71.8	90.2	99.4
	200	6.0	3.2	4.8	6.2	5.6	4.3	10.3	7.2	6.0	19.1	22.0	43.8	27.6	37.3	60.9	75.7	92.6	99.8
	500	8.4	5.2	4.2	7.6	5.4	5.2	9.0	8.6	6.0	25.2	26.3	44.3	34.0	42.3	65.2	78.8	95.5	99.9
	1000	15.0	4.8	5.6	13.6	6.6	5.7	12.5	8.9	5.3	35.6	27.4	45.4	40.3	43.5	68.0	82.9	92.0	100.0
Variance adjusted CD_{W+}	100	14.6	14.6	15.6	15.7	15.8	16.6	17.3	19.2	22.8	19.0	36.1	86.3	20.6	36.1	88.5	23.9	43.0	95.5
	200	19.2	23.6	23.8	18.9	25.1	22.8	23.8	28.6	25.9	24.4	50.3	98.3	26.0	50.5	98.8	30.3	56.6	99.6
	500	28.6	52.8	48.2	26.8	50.8	48.7	38.2	53.3	50.1	33.2	76.8	100.0	32.8	75.0	100.0	44.1	79.6	100.0
	1000	36.4	76.4	84.0	33.0	74.3	84.7	49.0	81.4	83.9	40.0	89.9	100.0	36.8	89.0	100.0	51.2	84.0	100.0

$\hat{m} = 4$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	98.6	100.0	100.0	90.3	98.7	100.0	6.8	8.2	18.0	94.3	99.5	99.9	73.5	93.0	99.0	39.4	44.4	50.2
	200	99.0	100.0	100.0	92.2	99.6	100.0	5.4	4.7	12.2	96.5	100.0	100.0	72.0	94.7	99.6	58.0	73.5	85.7
	500	99.6	100.0	100.0	89.7	99.8	100.0	6.1	5.7	5.7	94.8	100.0	100.0	64.5	96.3	100.0	72.8	88.6	99.2
	1000	96.6	100.0	100.0	84.1	99.8	100.0	10.0	7.3	3.6	91.2	100.0	100.0	55.3	96.7	100.0	78.0	88.9	100.0
Variance adjusted CD^*	100	6.0	5.8	8.0	6.6	7.4	8.0	9.9	9.5	7.8	20.0	30.3	49.9	27.5	40.2	65.4	67.9	88.3	99.3
	200	6.6	3.6	5.2	7.4	6.5	5.6	7.8	7.5	6.3	19.8	24.2	49.0	27.8	39.3	67.6	72.5	90.8	99.8
	500	9.4	5.6	4.4	7.8	5.5	5.1	7.7	7.6	5.8	26.2	27.6	46.6	33.3	42.7	67.5	78.3	93.7	99.9
	1000	16.0	4.6	5.4	12.4	6.5	5.4	10.2	9.3	5.5	35.5	27.9	46.5	42.1	42.8	68.7	80.1	91.0	100.0
Variance adjusted CD_{W+}	100	13.8	17.2	31.8	10.8	16.0	30.2	13.5	17.9	23.8	16.2	26.3	81.2	14.2	26.9	78.1	16.8	31.6	78.4
	200	13.2	22.6	23.8	15.4	18.4	24.6	18.6	24.0	22.9	19.5	40.6	94.6	20.2	37.9	93.7	22.4	44.9	95.4
	500	27.0	46.8	46.8	24.5	46.3	49.6	28.5	50.4	49.5	30.3	70.1	99.9	27.3	69.5	100.0	32.4	72.9	100.0
	1000	32.6	75.2	82.4	33.3	73.5	81.0	38.5	78.9	82.4	36.9	88.4	100.0	35.8	87.2	100.0	44.7	79.9	100.0

Note: The DGP is given by (42) with $\beta_{i1} = \beta_{i2} = 0$, and contains two latent factors with factor strengths $(\alpha_1, \alpha_2) = [(1,1), (1,2/3), (2/3,1/2)]$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 15: Size and power of variance adjusted tests of error cross-sectional dependence for panel regression with one latent factor ($m_0 = 1$) with serially correlated non-Gaussian errors

$\hat{m} = 1$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	40.0	70.6	93.3	6.1	7.0	13.3	7.2	6.0	7.3	15.9	24.7	44.2	60.5	79.6	92.6	70.4	86.6	98.2
	200	38.4	71.7	96.5	6.4	5.3	7.9	6.2	5.1	5.0	7.4	16.5	36.7	69.3	85.4	99.1	73.8	92.3	99.9
	500	32.6	75.5	99.2	8.0	6.3	5.9	9.4	7.0	3.8	6.3	13.4	29.9	76.7	91.3	99.7	76.9	94.7	100.0
	1000	28.3	69.0	69.0	8.3	4.8	4.5	7.1	5.4	4.8	5.5	9.1	27.1	80.5	93.9	100.0	78.0	96.0	100.0
Variance adjusted CD^*	100	5.6	5.8	4.2	7.6	6.5	5.5	8.0	7.0	6.4	42.8	62.4	89.9	72.2	91.0	99.9	75.9	93.5	100.0
	200	4.6	5.1	6.1	8.1	5.7	5.6	6.7	5.6	5.4	43.1	64.9	92.6	75.8	91.7	99.9	75.6	94.7	100.0
	500	5.8	5.2	5.4	8.5	7.2	4.7	9.7	7.3	3.6	46.2	67.6	93.3	79.2	94.3	100.0	77.5	95.4	100.0
	1000	7.2	5.9	5.9	9.2	5.5	5.1	7.2	5.6	5.9	51.0	65.7	95.4	82.3	96.4	100.0	78.4	96.2	100.0
Variance adjusted CD_{W+}	100	22.5	17.1	13.5	24.5	22.2	19.1	21.7	21.9	17.8	33.0	43.8	90.2	34.4	52.4	95.8	35.7	51.9	96.4
	200	48.2	44.6	27.9	51.9	48.0	26.4	52.7	48.0	28.8	61.0	74.4	99.3	64.1	77.8	100.0	62.9	78.2	99.9
	500	93.1	93.5	73.9	91.7	94.7	76.6	91.7	94.2	77.4	94.1	98.5	100.0	93.9	99.2	100.0	94.3	99.1	100.0
	1000	98.1	100.0	100.0	99.7	100.0	99.5	99.9	100.0	99.5	99.9	100.0	100.0	99.9	99.9	100.0	99.7	100.0	100.0

$\hat{m} = 2$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	41.2	70.6	93.1	4.4	5.5	13.6	6.8	5.6	7.3	15.5	27.2	50.2	51.8	69.9	79.2	57.7	77.0	86.9
	200	39.2	72.4	97.0	5.7	4.6	6.2	5.2	3.9	5.3	8.7	17.9	38.0	63.2	81.6	96.7	67.9	89.8	99.2
	500	32.0	75.6	99.3	6.7	6.1	5.9	7.4	6.4	3.3	6.3	12.8	31.1	71.9	89.6	99.6	72.3	94.1	100.0
	1000	28.9	69.1	69.1	7.2	5.0	4.5	7.0	6.1	5.2	5.2	9.5	28.7	76.6	94.7	99.9	73.2	95.3	100.0
Variance adjusted CD^*	100	5.4	4.9	5.2	6.7	6.7	6.8	8.6	7.2	7.0	40.5	61.4	89.5	69.1	89.6	99.8	70.6	92.3	99.9
	200	4.6	4.2	6.7	6.8	4.2	5.6	6.3	4.8	6.3	41.4	64.8	92.7	72.8	91.5	99.9	73.0	94.4	100.0
	500	6.1	5.2	4.7	7.6	6.9	4.9	7.8	6.8	4.0	48.5	66.5	93.6	75.0	92.6	100.0	73.3	95.3	100.0
	1000	8.1	6.3	6.3	8.0	5.3	4.4	7.1	6.0	5.5	51.0	65.4	95.2	79.4	95.5	100.0	78.4	96.2	100.0
Variance adjusted CD_{W+}	100	21.0	16.9	14.7	17.4	18.9	16.2	19.1	17.8	17.8	25.6	36.7	80.4	26.0	41.3	86.5	28.8	40.3	86.3
	200	42.2	40.4	26.5	45.7	40.8	25.7	44.4	39.2	25.1	50.7	64.5	98.2	55.0	68.7	99.2	53.8	69.4	98.5
	500	88.5	91.7	72.7	89.2	92.3	74.3	86.6	92.0	73.7	89.8	98.2	100.0	91.2	97.8	100.0	89.9	98.3	100.0
	1000	98.0	100.0	100.0	99.7	100.0	99.3	99.4	100.0	99.0	99.3	100.0	100.0	99.4	100.0	100.0	99.7	100.0	100.0

Note: The DGP is given by (42) with β_{i1} and β_{i2} both generated from normal distribution, and contains a single latent factor with factor strength $\alpha = 1, 2/3, 1/2$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 16: Size and power of variance adjusted tests of error cross-sectional dependence for panel regression with two latent factors ($m_0 = 2$) with serially correlated non-Gaussian errors

$\hat{m} = 2$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	97.9	100.0	100.0	87.6	99.0	100.0	8.2	6.7	20.8	91.5	99.2	100.0	69.1	90.7	98.4	53.1	65.7	73.2
	200	98.4	100.0	100.0	88.1	100.0	100.0	7.4	6.0	11.6	93.2	99.9	100.0	67.6	93.1	99.7	66.7	79.6	93.0
	500	98.4	100.0	100.0	88.2	100.0	100.0	10.7	7.5	5.6	93.9	100.0	100.0	62.3	95.8	100.0	75.1	88.9	98.2
	1000	98.1	100.0	100.0	82.3	100.0	100.0	11.0	8.7	5.0	92.7	100.0	100.0	54.3	96.1	100.0	75.8	88.7	100.0
Variance adjusted CD^*	100	6.3	4.5	5.2	6.5	5.3	4.3	9.8	7.7	6.7	16.6	21.9	36.8	24.3	35.3	54.4	71.6	91.0	99.4
	200	5.9	4.1	5.1	7.3	4.9	4.1	10.0	8.2	5.6	20.7	24.9	40.4	27.1	34.8	63.7	75.3	91.9	100.0
	500	8.9	6.1	4.4	9.0	5.4	5.1	11.9	9.6	5.2	25.7	26.8	44.5	32.6	38.9	65.7	79.0	92.7	99.0
	1000	14.2	4.9	5.4	13.5	6.9	5.6	12.0	10.5	6.2	33.5	29.1	44.1	41.2	42.0	66.4	78.3	91.2	100.0
Variance adjusted CD_{W+}	100	21.9	19.4	14.0	22.1	19.2	18.9	25.6	24.6	25.0	28.2	40.6	83.1	27.9	40.3	90.3	37.2	53.9	95.8
	200	45.7	43.7	24.3	46.9	43.3	28.2	54.7	50.6	30.1	54.7	72.3	98.5	52.7	74.1	99.2	64.8	79.8	99.7
	500	90.1	94.7	74.3	89.1	93.4	74.0	93.2	95.8	77.5	92.0	98.1	100.0	91.4	98.7	100.0	94.1	98.8	99.9
	1000	99.9	100.0	100.0	99.7	100.0	100.0	99.7	100.0	100.0	99.5	100.0	100.0	99.8	100.0	100.0	88.5	96.4	100.0

$\hat{m} = 4$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
Variance adjusted CD	100	97.8	100.0	100.0	88.9	99.3	100.0	6.8	6.2	20.9	93.3	99.4	100.0	73.2	93.4	99.3	40.1	45.4	49.5
	200	98.6	100.0	100.0	89.2	100.0	100.0	6.8	5.6	11.3	94.8	99.9	100.0	70.5	94.8	99.5	57.3	71.1	87.2
	500	98.4	100.0	100.0	87.9	100.0	100.0	8.5	6.8	5.1	94.3	100.0	100.0	62.0	96.4	100.0	71.3	85.8	97.9
	1000	97.5	100.0	100.0	81.1	99.8	100.0	10.1	7.9	5.1	91.5	100.0	100.0	51.8	96.2	100.0	73.8	87.3	100.0
Variance adjusted CD^*	100	6.7	6.7	7.1	7.4	7.1	8.0	9.1	9.0	8.8	22.9	29.4	47.0	28.3	40.5	64.6	67.9	89.1	99.2
	200	7.7	5.7	5.7	6.9	6.1	5.2	9.1	7.8	4.9	22.3	28.0	46.0	27.8	37.7	66.7	72.5	91.3	99.9
	500	9.0	7.0	5.2	9.9	6.0	6.5	10.5	8.8	6.4	26.8	28.6	45.4	34.9	41.1	69.3	77.5	90.4	98.5
	1000	16.0	6.5	5.2	15.1	6.6	5.3	11.4	9.2	6.8	36.1	29.0	44.5	43.0	44.4	66.5	75.1	89.3	100.0
Variance adjusted CD_{W+}	100	18.6	17.3	33.9	14.4	16.6	28.9	16.7	18.8	24.6	23.1	31.5	82.3	19.6	31.6	81.7	21.6	36.5	80.3
	200	35.2	36.8	23.9	34.9	35.0	26.5	38.3	37.4	26.3	39.4	57.4	95.8	41.4	58.8	96.5	49.0	63.9	96.5
	500	83.6	88.1	69.0	84.6	89.7	70.5	84.8	92.4	71.6	87.3	96.6	100.0	86.6	97.3	100.0	88.8	96.8	99.5
	1000	99.4	99.9	100.0	99.6	100.0	100.0	99.4	100.0	100.0	99.3	100.0	100.0	99.2	100.0	100.0	87.8	96.4	100.0

Note: The DGP is given by (42) with β_{i1} and β_{i2} both generated from normal distribution, and contains two latent factors with factor strengths $(\alpha_1, \alpha_2) = [(1,1), (1,2/3), (2/3,1/2)]$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 17: Size and power of ARDL adjusted tests of error cross-sectional dependence for pure single factor model ($m_0 = 1$) with serially correlated Gaussian errors

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	68.9	91.3	98.7	6.5	9.0	20.8	6.2	5.7	9.4	28.6	45.2	67.0	54.3	70.5	82.6	68.3	80.6	87.8
	200	70.6	94.4	99.8	5.1	7.0	14.3	5.7	5.5	5.7	19.4	35.4	61.0	69.3	88.8	98.1	81.4	95.7	99.4
	500	71.2	96.5	100.0	4.5	5.1	7.4	7.3	6.0	5.1	12.8	27.3	55.6	79.9	95.9	100.0	88.6	98.0	100.0
	1000	73.0	97.2	100.0	5.9	5.4	5.6	7.2	5.1	5.2	10.1	22.6	52.1	81.7	96.7	100.0	87.4	99.2	100.0
ARDL adjusted CD^*	100	5.0	5.3	5.9	6.8	6.3	5.8	7.4	6.9	7.1	57.1	83.9	99.1	82.3	98.0	100.0	85.0	98.5	100.0
	200	5.5	5.4	5.4	5.8	5.9	4.7	6.3	6.3	5.5	59.1	84.9	99.4	84.6	98.3	100.0	88.7	99.2	100.0
	500	5.0	5.5	5.3	5.5	5.3	5.5	7.4	6.7	5.2	60.7	85.5	99.5	86.0	98.6	100.0	90.4	98.9	100.0
	1000	5.4	4.5	4.6	6.1	6.1	4.9	7.7	5.3	5.1	61.0	85.6	99.8	85.1	98.3	100.0	88.6	99.4	100.0
ARDL adjusted CD_{W+}	100	5.6	5.4	7.2	5.5	5.6	5.2	6.5	5.3	7.5	7.0	6.5	28.9	6.3	7.9	34.0	6.7	8.1	31.8
	200	5.0	6.1	5.4	4.4	5.3	5.7	4.2	5.0	5.5	6.0	6.5	36.2	5.1	7.1	41.7	4.7	7.0	38.3
	500	6.3	5.1	4.8	5.4	4.6	5.1	4.3	5.2	5.7	5.5	6.3	45.7	5.6	5.7	46.8	5.1	6.6	46.1
	1000	4.7	6.1	5.0	5.6	4.6	4.3	4.3	5.5	5.5	4.4	5.7	47.0	5.9	5.6	45.7	4.7	5.6	47.3

$\hat{m} = 4$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	68.9	91.6	98.6	6.2	9.2	22.6	6.9	5.8	10.2	33.7	56.9	78.1	38.4	47.3	62.0	49.3	58.9	69.0
	200	71.5	94.0	99.7	5.4	7.1	13.9	5.5	5.4	6.0	22.3	42.7	69.3	58.9	79.1	91.5	72.2	90.4	95.5
	500	70.8	96.9	100.0	5.0	5.8	7.6	7.3	5.9	4.8	14.7	30.8	60.9	75.9	94.7	99.9	84.6	97.6	100.0
	1000	72.4	97.5	100.0	5.9	5.9	6.0	6.9	5.3	5.7	10.8	24.7	54.6	78.6	95.8	100.0	84.1	98.6	100.0
ARDL adjusted CD^*	100	7.0	7.1	12.0	7.9	8.1	10.3	8.9	9.7	12.0	58.0	84.7	99.5	80.3	97.2	100.0	80.8	98.1	100.0
	200	6.4	5.8	6.6	6.4	6.5	5.7	6.7	6.8	6.8	57.5	84.2	99.3	82.1	97.8	100.0	85.8	98.8	100.0
	500	5.4	5.9	5.2	5.3	5.3	5.5	7.9	6.7	5.2	58.6	85.2	99.6	85.0	98.4	100.0	88.1	98.4	100.0
	1000	4.9	4.9	4.7	6.3	6.1	5.3	7.3	5.3	5.6	59.9	85.0	99.8	82.8	97.9	100.0	85.7	99.0	100.0
ARDL adjusted CD_{W+}	100	5.9	5.9	8.2	5.5	6.1	7.4	5.7	6.2	7.8	6.8	7.0	20.5	5.8	7.4	18.9	5.6	8.0	18.6
	200	5.7	5.1	5.2	4.8	5.1	4.8	5.0	5.7	5.1	5.9	6.1	20.2	6.0	5.1	22.1	5.1	5.5	19.8
	500	4.7	4.5	5.7	5.7	5.4	4.7	5.4	5.9	5.9	5.1	4.9	31.8	5.8	5.1	33.2	5.6	6.1	33.7
	1000	6.7	4.9	4.5	5.7	4.4	4.6	6.5	5.1	5.5	5.8	5.8	40.2	6.4	4.8	40.7	6.2	5.9	38.2

Note: The DGP is given by (42) with $\beta_{i1} = \beta_{i2} = 0$ and contains a single latent factor with factor strength $\alpha = 1, 2/3, 1/2$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 18: Size and power of ARDL adjusted tests of error cross-sectional dependence for pure two factors model ($m_0 = 2$) with serially correlated Gaussian errors

$\hat{m} = 4$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	100.0	100.0	100.0	98.9	99.9	100.0	7.4	11.9	38.0	99.6	100.0	100.0	93.6	99.0	99.9	37.3	39.4	50.1
	200	100.0	100.0	100.0	99.5	100.0	100.0	6.9	8.8	21.8	99.8	100.0	100.0	94.3	99.7	100.0	59.4	72.1	86.3
	500	100.0	100.0	100.0	99.9	100.0	100.0	7.0	5.7	10.7	99.9	100.0	100.0	93.3	99.9	100.0	76.3	93.4	99.8
	1000	100.0	100.0	100.0	99.8	100.0	100.0	7.1	5.5	7.5	100.0	100.0	100.0	93.8	99.9	100.0	81.9	96.8	100.0
ARDL adjusted CD^*	100	8.0	9.3	14.5	7.8	7.2	10.5	10.2	9.9	9.9	31.8	53.7	87.9	39.8	62.2	93.0	81.2	96.7	100.0
	200	5.7	7.2	7.9	6.6	6.2	7.1	8.1	7.6	6.3	30.2	50.0	86.0	38.2	60.0	92.3	81.7	97.7	100.0
	500	6.1	6.1	6.5	7.0	6.2	5.0	8.6	6.1	4.7	29.7	49.5	83.5	40.5	61.4	92.4	85.3	98.3	100.0
	1000	5.5	6.0	5.9	6.3	6.1	5.0	8.2	7.2	6.4	28.3	48.1	81.8	40.4	60.9	92.7	86.4	98.7	100.0
ARDL adjusted CD_{W+}	100	6.3	6.9	8.5	5.4	6.2	10.2	5.8	5.6	9.2	6.1	8.7	21.1	8.4	14.4	75.6	6.8	8.1	25.8
	200	5.3	5.9	6.5	5.5	6.0	5.1	6.0	5.5	6.3	5.5	7.0	25.3	6.7	8.0	25.5	6.4	6.3	26.9
	500	5.4	4.0	4.4	5.3	4.7	4.4	5.9	4.7	6.2	5.6	5.1	36.1	5.8	6.8	23.2	6.0	4.9	35.9
	1000	5.4	5.3	5.0	5.1	5.9	5.1	5.2	5.4	6.3	5.6	5.1	40.7	5.7	6.2	33.6	5.4	6.1	40.6

$\hat{m} = 6$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	100.0	100.0	100.0	98.8	99.9	100.0	7.4	12.4	38.6	99.7	100.0	100.0	95.3	99.3	99.9	27.0	29.7	50.3
	200	100.0	100.0	100.0	99.6	100.0	100.0	6.5	8.9	22.4	100.0	100.0	100.0	95.0	99.8	100.0	47.8	60.6	70.8
	500	100.0	100.0	100.0	99.9	100.0	100.0	6.5	5.1	10.5	100.0	100.0	100.0	94.2	99.9	100.0	72.9	91.1	99.6
	1000	100.0	100.0	100.0	99.8	100.0	100.0	7.1	5.5	7.7	100.0	100.0	100.0	94.0	100.0	100.0	78.0	95.9	100.0
ARDL adjusted CD^*	100	11.4	15.6	30.3	9.8	12.5	24.1	12.3	13.2	17.3	37.0	62.5	93.1	42.7	69.2	96.0	78.5	95.5	100.0
	200	7.5	9.1	12.7	6.9	7.2	10.5	8.8	7.9	8.6	32.1	53.4	89.8	38.8	61.5	94.3	78.6	97.0	100.0
	500	6.6	6.6	7.4	7.1	6.8	5.6	8.3	5.9	5.6	29.8	50.2	84.6	39.5	60.7	92.9	82.9	98.0	100.0
	1000	6.0	6.4	6.1	6.8	5.8	4.9	8.2	7.0	6.4	28.3	47.0	81.6	40.0	60.7	92.5	83.6	98.7	100.0
ARDL adjusted CD_{W+}	100	5.5	7.4	16.2	6.1	6.9	16.5	5.5	6.3	10.4	5.9	8.4	26.2	6.7	8.1	27.8	5.8	8.3	21.5
	200	5.5	5.5	6.5	5.9	5.5	6.7	5.4	5.7	6.3	6.2	6.5	17.8	6.8	6.4	16.5	6.3	6.6	14.3
	500	5.6	4.9	5.1	6.2	4.7	5.1	6.3	6.6	4.7	5.1	6.3	27.9	6.6	4.9	27.8	6.5	6.4	26.6
	1000	5.2	5.4	5.3	5.4	5.0	4.8	6.3	5.5	5.7	5.5	6.0	35.9	5.5	5.3	34.8	6.4	6.8	36.6

Note: The DGP is given by (42) with $\beta_{i1} = \beta_{i2} = 0$, and contains two latent factors with factor strengths $(\alpha_1, \alpha_2) = [(1,1),(1,2/3),(2/3,1/2)]$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 19: Size and power of ARDL adjusted tests of error cross-sectional dependence for panel regression with one latent factor ($m_0 = 1$) with serially correlated Gaussian errors

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	69.1	91.0	98.9	7.0	9.2	22.1	8.1	7.4	9.4	29.5	49.6	67.5	58.3	70.8	82.5	66.6	79.9	86.2
	200	69.6	93.5	99.9	7.0	6.6	11.9	7.4	6.8	5.6	18.5	34.7	60.2	71.6	89.6	98.4	82.2	94.9	99.5
	500	67.2	96.4	100.0	7.1	4.6	7.9	8.0	6.6	5.1	11.4	25.4	53.4	82.8	96.6	100.0	87.2	98.7	100.0
	1000	69.0	96.8	100.0	7.6	6.0	6.2	8.0	6.7	5.1	10.5	22.2	51.2	85.5	98.3	100.0	87.3	99.1	100.0
ARDL adjusted CD^*	100	5.7	6.1	7.8	7.9	7.6	6.9	9.4	8.6	7.4	55.8	81.6	98.9	82.6	98.1	100.0	85.3	99.0	100.0
	200	5.1	5.7	5.1	8.3	6.7	5.3	8.4	8.3	5.4	58.7	84.3	99.6	85.9	98.4	100.0	88.4	98.9	100.0
	500	6.5	6.2	5.2	8.6	6.5	5.8	8.0	7.4	5.7	61.8	85.6	99.6	87.6	99.0	100.0	89.2	99.1	100.0
	1000	6.2	4.7	4.6	8.3	6.0	6.4	8.2	7.0	5.9	60.1	85.3	99.5	88.3	99.0	100.0	88.3	99.3	100.0
ARDL adjusted CD_{W+}	100	5.4	6.4	6.4	5.3	5.9	5.9	5.6	5.5	6.3	5.4	8.0	29.1	6.7	7.8	33.7	6.8	8.5	32.7
	200	6.5	5.9	5.3	6.6	6.4	5.0	5.3	5.7	5.5	7.0	7.4	34.7	7.3	8.2	40.7	5.9	6.4	39.0
	500	5.8	5.6	4.7	5.4	6.2	5.1	5.1	5.4	4.4	6.0	5.8	44.3	5.5	6.6	47.5	5.3	6.0	46.4
	1000	6.7	5.6	4.9	5.9	5.3	5.8	6.7	6.0	5.2	6.5	5.5	42.0	5.5	6.3	48.2	6.7	6.8	47.9

$\hat{m} = 4$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	69.6	90.3	99.2	6.7	10.4	23.7	7.1	7.3	10.8	38.9	60.2	80.9	38.7	47.8	60.3	44.8	58.5	65.8
	200	69.5	93.8	100.0	6.7	6.9	11.6	7.1	5.9	6.8	23.4	42.0	70.1	58.1	77.6	91.6	68.5	87.9	95.8
	500	67.9	96.0	100.0	6.3	4.7	7.8	6.9	6.6	5.1	14.1	28.0	58.8	74.9	95.0	100.0	80.3	97.9	100.0
	1000	68.5	96.6	100.0	7.0	5.7	6.1	8.0	6.5	5.8	12.2	24.8	54.7	77.9	97.3	100.0	81.7	98.8	100.0
ARDL adjusted CD_*	100	6.7	8.3	12.0	9.0	10.3	11.9	9.3	10.5	12.6	55.9	83.0	99.4	77.2	96.7	99.9	80.3	97.9	100.0
	200	5.4	6.6	6.6	8.1	7.6	6.4	8.6	8.8	8.1	57.4	84.1	99.5	81.5	97.5	100.0	85.1	98.3	100.0
	500	6.6	7.0	5.4	7.1	6.2	6.1	7.6	7.6	6.2	58.9	84.8	99.7	83.6	98.7	100.0	85.3	98.8	100.0
	1000	6.6	5.4	5.1	7.6	6.1	6.5	8.2	7.1	6.2	59.1	84.2	99.6	82.6	98.7	100.0	83.7	98.9	100.0
ARDL adjusted CD_{W+}	100	7.0	6.5	9.5	5.8	4.2	7.3	6.6	6.6	7.6	7.1	8.3	20.5	6.4	6.1	17.8	6.5	8.1	17.8
	200	5.3	5.9	4.9	6.5	5.5	5.5	6.3	5.7	5.4	5.7	6.6	21.6	7.0	6.7	20.5	7.0	6.8	20.0
	500	5.5	5.1	5.3	4.8	5.7	4.9	6.4	5.5	5.1	5.6	6.1	30.0	5.8	5.5	34.4	6.5	6.5	32.1
	1000	6.0	5.6	5.5	5.8	5.4	6.3	5.6	5.0	5.5	6.3	6.5	39.7	5.4	5.0	40.3	5.8	5.1	38.5

Note:The DGP is given by (42) with β_{i1} and β_{i2} both generated from normal distribution, and contains a single latent factor with factor strength $\alpha = 1, 2/3, 1/2$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 20: Size and power of ARDL adjusted tests of error cross-sectional dependence for panel regression with two latent factors ($m_0 = 2$) with serially correlated Gaussian errors

$\hat{m} = 4$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	100.0	100.0	100.0	98.4	100.0	100.0	8.8	12.0	37.8	99.4	100.0	100.0	92.7	98.7	100.0	35.9	41.8	49.9
	200	100.0	100.0	100.0	99.3	100.0	100.0	7.7	8.4	19.6	99.8	100.0	100.0	92.3	99.6	100.0	60.6	74.2	87.6
	500	100.0	100.0	100.0	99.6	100.0	100.0	6.8	6.1	8.3	99.9	100.0	100.0	92.2	99.8	100.0	78.0	95.1	99.7
	1000	100.0	100.0	100.0	100.0	100.0	100.0	8.3	6.0	6.9	100.0	100.0	100.0	91.6	99.9	100.0	82.0	97.1	100.0
ARDL adjusted CD^*	100	7.7	7.7	11.9	7.8	8.8	11.6	11.9	11.4	12.4	29.3	51.8	84.9	37.2	64.2	92.2	77.8	96.4	100.0
	200	6.1	7.0	6.0	6.6	6.3	6.1	11.6	9.3	6.7	29.3	48.8	81.9	40.1	61.5	91.4	83.4	97.7	100.0
	500	7.2	5.7	6.0	7.4	5.7	4.7	10.1	8.6	6.6	29.6	41.9	78.6	40.0	60.9	92.4	86.4	98.6	100.0
	1000	5.9	5.8	5.2	7.3	5.6	5.5	9.0	8.4	6.5	27.1	41.6	76.6	42.6	62.1	91.9	86.5	98.4	100.0
ARDL adjusted CD_{W+}	100	6.4	6.8	10.2	5.7	6.9	9.4	5.7	6.3	9.1	8.3	9.1	25.2	6.4	7.2	24.9	7.0	7.7	24.9
	200	5.8	5.9	6.0	5.1	5.9	5.3	5.8	4.7	5.5	6.2	6.8	24.3	5.6	6.9	24.1	5.5	5.5	25.1
	500	5.9	5.8	5.5	6.4	5.5	5.6	5.8	5.7	5.5	5.9	7.0	34.3	6.9	5.9	36.0	5.6	5.7	37.1
	1000	5.9	5.2	4.9	6.5	5.6	6.2	5.5	5.6	5.7	6.2	6.1	38.7	6.6	6.7	38.2	5.9	6.1	40.1

$\hat{m} = 6$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	100.0	100.0	100.0	98.5	100.0	100.0	9.0	13.0	38.7	99.8	100.0	100.0	93.5	99.3	100.0	23.9	32.2	48.7
	200	100.0	100.0	100.0	99.4	100.0	100.0	8.3	8.5	20.2	99.8	100.0	100.0	94.0	99.8	100.0	48.2	62.1	73.2
	500	100.0	100.0	100.0	99.7	100.0	100.0	7.3	5.7	8.4	100.0	100.0	100.0	92.4	99.8	100.0	70.7	91.5	99.5
	1000	100.0	100.0	100.0	100.0	100.0	100.0	6.9	6.1	7.0	100.0	100.0	100.0	92.6	99.9	100.0	75.9	96.2	100.0
ARDL adjusted CD^*	100	10.1	13.8	28.7	10.5	14.2	26.9	13.1	15.3	19.1	32.0	59.4	91.6	41.9	69.9	96.2	73.7	95.8	100.0
	200	7.4	8.0	9.7	8.0	8.0	9.0	11.4	10.4	9.3	30.1	50.9	86.9	40.4	62.9	93.7	78.4	96.2	100.0
	500	7.8	5.7	6.8	8.0	6.0	5.4	9.5	8.7	7.3	29.3	42.4	79.4	39.8	61.2	93.2	82.1	98.3	100.0
	1000	6.3	6.2	5.1	8.3	6.5	5.6	8.4	7.7	6.4	27.7	42.4	77.8	41.2	63.1	91.1	82.1	98.2	100.0
ARDL adjusted CD_{W+}	100	6.3	6.8	22.5	6.9	7.2	16.9	5.5	7.1	10.7	6.9	8.6	32.1	7.9	8.4	28.0	6.7	8.9	20.8
	200	5.9	5.9	5.9	6.5	5.5	6.4	6.6	5.3	5.6	6.2	7.0	16.8	6.7	5.9	15.8	6.9	6.2	16.0
	500	7.4	5.8	5.5	6.7	5.8	5.7	5.7	4.9	4.2	6.9	6.3	28.4	7.1	6.5	26.4	6.5	5.7	25.5
	1000	5.9	5.9	5.2	6.4	6.3	5.3	6.1	5.8	5.3	6.1	6.5	33.1	6.2	6.8	32.9	6.6	5.9	34.4

Note: The DGP is given by (42) with β_{i1} and β_{i2} both generated from normal distribution, and contains two latent factors with factor strengths $(\alpha_1, \alpha_2) = [(1,1),(1,2/3),(2/3,1/2)]$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 21: Size and power of ARDL adjusted tests of error cross-sectional dependence for pure single factor model ($m_0 = 1$) with serially correlated non-Gaussian errors

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	62.2	86.3	97.2	6.7	9.7	23.6	6.6	6.4	10.0	28.5	46.4	67.0	55.4	69.2	81.4	67.0	79.1	86.4
	200	66.2	91.6	96.4	4.9	7.0	14.2	5.1	5.4	6.9	19.2	34.4	58.8	70.1	88.1	98.1	80.1	95.0	99.2
	500	66.0	88.8	97.3	5.6	5.8	9.5	7.3	5.9	5.9	12.2	27.9	52.8	81.3	95.8	100.0	88.9	98.5	100.0
	1000	63.1	96.8	100.0	6.4	5.7	7.4	7.2	5.2	5.0	9.6	21.7	50.4	82.4	97.0	100.0	87.6	99.3	100.0
ARDL adjusted CD^*	100	5.5	5.9	5.7	7.5	6.2	7.6	8.6	7.6	6.9	58.8	84.5	99.1	81.7	97.4	100.0	85.9	98.3	100.0
	200	5.5	5.4	6.3	5.6	5.5	5.7	6.1	6.3	6.2	59.6	86.1	99.3	85.1	98.2	100.0	87.9	99.1	100.0
	500	5.8	5.3	5.8	6.4	5.5	5.6	7.6	6.6	5.8	61.7	85.5	99.4	86.4	98.6	100.0	90.4	99.0	100.0
	1000	5.5	6.2	5.4	7.0	6.0	4.8	7.4	5.6	5.8	62.3	85.8	99.5	86.4	98.8	100.0	88.7	99.6	100.0
ARDL adjusted CD_{W+}	100	6.5	7.0	6.4	6.2	5.1	5.7	5.9	5.7	7.3	7.9	10.9	33.0	7.0	9.2	37.8	8.0	9.2	35.1
	200	8.5	6.5	6.1	8.5	6.5	5.4	8.1	5.5	5.1	10.4	10.7	47.1	9.6	11.5	53.4	9.2	10.4	51.9
	500	13.5	10.9	6.9	15.9	8.1	5.3	14.1	7.8	5.0	16.9	16.0	67.8	16.4	14.9	72.1	15.8	13.6	68.7
	1000	24.3	13.9	5.9	27.4	14.2	7.0	26.0	16.4	6.5	25.2	19.9	81.7	26.3	20.7	81.4	24.7	23.8	82.0

$\hat{m} = 4$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	63.6	86.2	97.5	6.7	9.5	24.6	7.2	6.7	10.7	35.2	60.2	79.8	38.6	46.8	61.5	50.3	59.0	66.0
	200	66.5	90.9	96.3	5.4	7.6	14.5	4.8	5.5	6.7	20.8	41.9	67.9	60.0	78.7	90.7	70.1	87.4	94.8
	500	64.4	88.4	97.3	6.1	5.6	9.2	6.9	6.3	5.8	13.8	29.5	58.7	77.5	94.7	99.9	85.0	97.9	100.0
	1000	62.0	96.8	100.0	6.5	5.4	7.8	7.0	5.7	5.8	10.1	23.5	53.8	80.6	97.2	100.0	86.4	99.0	100.0
ARDL adjusted CD^*	100	7.6	7.0	10.7	9.8	8.5	11.3	9.0	8.3	8.6	58.0	83.8	99.7	78.9	96.9	100.0	83.2	97.3	100.0
	200	7.2	6.6	7.1	6.5	6.3	7.1	6.1	6.4	6.6	59.9	86.1	99.1	81.7	97.7	100.0	85.0	98.7	100.0
	500	6.4	5.3	5.7	7.0	6.4	5.6	7.2	7.0	5.7	61.3	85.6	99.6	85.5	98.3	100.0	87.9	98.8	100.0
	1000	6.5	6.1	5.3	7.0	5.9	5.1	7.4	6.2	6.0	61.6	85.6	99.5	85.6	98.7	100.0	88.3	99.3	100.0
ARDL adjusted CD_{W+}	100	6.7	6.1	7.4	6.7	7.0	7.4	6.3	5.8	8.0	8.2	8.3	22.0	7.6	8.0	18.7	6.5	7.3	18.5
	200	6.8	5.9	6.3	6.3	6.3	6.9	7.0	6.1	5.3	7.6	7.5	28.2	7.7	9.7	29.5	8.6	7.6	26.2
	500	13.5	10.4	6.1	14.1	8.4	6.4	14.0	9.3	5.7	15.4	13.4	52.7	15.5	12.3	54.7	14.3	13.2	57.5
	1000	22.8	13.4	5.3	24.3	14.9	6.3	23.3	14.5	6.2	24.0	18.6	73.0	24.0	20.0	73.2	23.8	19.9	71.9

Note: The DGP is given by (42) with $\beta_{i1} = \beta_{i2} = 0$ and contains a single latent factor with factor strength $\alpha = 1, 2/3, 1/2$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 22: Size and power of ARDL adjusted tests of error cross-sectional dependence for pure two factors model ($m_0 = 2$) with serially correlated non-Gaussian errors

$\hat{m} = 4$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	100.0	100.0	100.0	98.6	100.0	100.0	8.3	14.6	39.4	99.6	100.0	100.0	92.4	98.4	99.9	36.6	41.7	51.0
	200	100.0	100.0	100.0	99.4	100.0	100.0	5.9	8.6	24.1	99.8	100.0	100.0	92.6	99.5	100.0	58.4	70.3	84.7
	500	100.0	100.0	100.0	99.7	100.0	100.0	7.5	5.2	11.5	99.9	100.0	100.0	93.2	99.7	100.0	78.0	93.3	99.8
	1000	100.0	100.0	100.0	99.7	100.0	100.0	7.4	6.1	7.7	100.0	100.0	100.0	91.9	100.0	100.0	82.9	96.4	100.0
ARDL adjusted CD^*	100	9.0	8.9	14.1	8.0	8.9	12.1	10.8	9.0	10.4	32.8	54.9	87.1	41.0	64.1	92.6	79.4	95.8	100.0
	200	6.5	7.1	8.0	6.4	6.9	7.1	8.2	7.4	7.0	29.8	49.2	86.8	39.1	59.7	91.1	82.5	97.5	100.0
	500	6.4	6.1	6.2	6.6	6.0	5.6	9.6	7.9	6.5	31.3	53.1	85.0	42.4	62.4	91.8	86.3	98.6	100.0
	1000	6.2	5.7	5.1	7.1	6.2	4.6	8.9	6.5	4.8	30.5	46.8	82.2	41.9	61.9	92.4	87.5	98.3	100.0
ARDL adjusted CD_{W+}	100	6.9	7.4	8.1	6.0	6.4	12.1	6.1	6.4	10.4	7.9	9.7	25.0	6.4	8.3	29.4	6.7	8.8	26.4
	200	8.1	6.3	5.2	7.8	6.9	6.0	7.8	6.7	6.4	9.7	9.9	31.8	9.4	9.0	32.6	9.0	8.3	34.0
	500	13.0	7.6	5.2	15.0	8.6	5.7	14.7	8.4	5.5	13.6	11.9	59.4	15.4	12.0	56.7	14.6	13.8	59.3
	1000	25.8	13.7	6.0	26.9	13.0	6.2	24.1	14.6	6.3	24.2	19.9	74.9	25.5	18.6	72.6	23.8	21.4	76.4

$\hat{m} = 6$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	99.9	100.0	100.0	98.5	100.0	100.0	8.2	15.0	39.4	99.5	100.0	100.0	94.6	99.2	99.9	26.1	32.5	51.7
	200	100.0	100.0	100.0	99.2	100.0	100.0	5.8	8.5	24.3	99.9	100.0	100.0	94.5	99.7	100.0	49.3	60.5	67.7
	500	100.0	100.0	100.0	99.5	100.0	100.0	7.6	5.7	10.9	100.0	100.0	100.0	92.8	99.7	100.0	75.4	90.9	99.6
	1000	100.0	100.0	100.0	99.6	100.0	100.0	8.0	5.9	7.4	100.0	100.0	100.0	91.9	100.0	100.0	80.4	95.3	100.0
ARDL adjusted CD^*	100	13.1	14.5	30.0	11.9	14.3	25.1	12.7	12.2	17.2	38.6	61.7	93.3	45.5	69.5	96.2	76.1	95.8	100.0
	200	7.6	8.4	11.8	7.9	8.1	10.4	8.8	9.0	9.1	32.9	54.6	90.5	41.3	62.5	93.8	80.7	97.1	100.0
	500	6.8	6.8	6.7	6.9	5.9	6.0	10.0	8.3	6.5	31.9	52.8	84.4	43.6	62.7	92.5	85.4	98.3	100.0
	1000	7.1	6.2	5.7	8.3	7.0	4.9	9.3	6.6	4.9	33.0	47.4	82.4	43.0	62.8	92.8	86.0	98.2	100.0
ARDL adjusted CD_{W+}	100	6.0	6.9	15.6	6.4	6.3	17.1	6.8	6.1	10.6	6.6	7.6	27.7	6.8	8.2	29.0	7.4	8.3	19.2
	200	7.7	6.3	7.1	8.0	5.7	5.5	7.4	5.6	5.7	7.7	8.6	19.1	7.9	7.5	20.5	8.1	7.7	18.3
	500	12.1	8.2	5.8	13.9	8.4	6.3	12.9	8.4	5.8	14.2	11.9	44.1	14.1	11.5	44.6	13.8	12.1	44.9
	1000	24.0	12.8	5.9	26.5	14.4	5.5	23.2	13.5	6.9	24.2	18.2	67.7	24.9	19.9	64.9	24.2	18.5	67.7

Note: The DGP is given by (42) with $\beta_{i1} = \beta_{i2} = 0$, and contains two latent factors with factor strengths $(\alpha_1, \alpha_2) = [(1,1), (1,2/3), (2/3,1/2)]$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 23: Size and power of ARDL adjusted tests of error cross-sectional dependence for panel regression with one latent factor ($m_0 = 1$) with serially correlated non-Gaussian errors

$\hat{m} = 2$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	66.7	90.9	98.9	7.2	10.1	20.4	7.9	7.0	9.9	28.5	49.2	68.2	59.1	70.9	82.7	69.2	78.1	86.2
	200	68.7	93.2	99.9	7.5	5.8	13.6	8.6	5.5	6.4	18.7	36.4	58.7	71.8	90.2	98.1	81.0	95.7	99.3
	500	68.1	96.0	100.0	8.2	7.7	8.9	8.0	5.6	5.8	11.1	25.0	53.0	82.6	96.1	100.0	87.3	98.9	100.0
	1000	66.6	95.9	100.0	7.2	5.6	5.2	7.9	6.9	4.6	10.9	21.6	51.7	84.8	97.9	100.0	89.5	98.8	100.0
ARDL adjusted CD^*	100	5.8	6.3	7.4	8.6	8.2	6.8	8.8	7.2	7.4	57.8	84.2	99.5	83.6	97.6	100.0	87.2	97.9	100.0
	200	6.8	6.7	5.5	9.4	6.9	7.0	9.9	7.2	6.5	60.0	85.2	99.3	84.9	98.5	100.0	87.9	99.2	100.0
	500	6.1	5.5	5.1	9.5	7.5	6.5	8.5	6.7	6.1	60.3	84.7	99.6	87.5	99.1	100.0	89.5	99.4	100.0
	1000	6.2	5.7	5.0	8.5	6.8	5.1	8.2	7.4	4.8	60.6	85.7	99.6	88.0	99.1	100.0	90.6	98.9	100.0
ARDL adjusted CD_{W+}	100	7.6	6.3	6.2	6.9	6.4	7.0	6.7	6.2	6.4	8.4	9.9	32.9	8.8	11.4	40.0	8.3	10.1	37.5
	200	9.0	5.9	5.1	8.9	6.0	6.0	9.2	6.7	5.7	9.8	9.9	46.7	10.3	10.7	52.8	10.4	11.4	52.3
	500	15.0	9.1	5.3	13.2	9.3	5.5	12.8	9.7	5.3	15.0	14.1	67.8	15.3	14.5	71.5	13.8	15.2	71.7
	1000	25.9	14.0	6.2	23.8	15.6	5.9	22.8	14.2	5.8	24.1	19.4	79.9	23.3	22.3	79.3	22.9	20.7	78.0

$\hat{m} = 4$

		Size ($H_o : \rho = 0$)									Power ($H_1 : \rho = 0.25$)								
		$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$			$\alpha = 1$			$\alpha = 2/3$			$\alpha = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	68.1	91.0	98.7	6.9	10.2	22.4	7.8	7.2	10.4	37.8	61.2	80.5	38.1	46.9	60.4	49.4	57.4	66.1
	200	67.9	93.2	99.8	6.7	6.6	13.0	8.4	6.0	7.4	23.0	43.1	69.7	59.0	78.4	89.8	68.2	88.6	94.3
	500	67.8	95.8	100.0	8.0	7.0	9.0	7.4	6.7	5.7	12.8	29.3	57.8	75.1	93.8	99.9	80.8	97.8	100.0
	1000	66.1	95.9	100.0	6.6	5.6	5.0	7.2	6.6	4.7	11.1	23.0	55.2	79.2	96.8	100.0	85.0	98.1	100.0
ARDL adjusted CD^*	100	7.6	8.0	11.8	10.0	9.8	10.3	10.7	10.0	12.9	56.4	84.6	99.7	79.9	96.6	100.0	82.0	97.0	100.0
	200	7.8	7.2	6.6	8.4	7.1	8.5	9.8	8.1	8.6	57.4	84.8	99.4	80.7	97.5	100.0	84.3	98.4	100.0
	500	6.8	5.6	5.8	9.1	7.7	6.9	7.9	7.8	6.5	59.9	84.4	99.5	84.0	98.6	100.0	85.5	99.2	100.0
	1000	6.9	5.9	5.5	7.6	6.5	5.9	7.7	7.6	5.3	60.1	85.8	99.7	84.2	98.7	100.0	87.0	98.8	100.0
ARDL adjusted CD_{W+}	100	7.0	6.7	7.4	6.3	6.1	7.8	6.6	6.6	8.3	7.8	7.6	21.8	8.2	8.7	20.2	8.1	8.0	20.7
	200	7.3	5.9	5.9	8.3	5.8	6.3	7.9	5.8	6.0	8.6	8.5	27.2	9.0	8.6	29.3	8.7	8.2	25.5
	500	13.1	9.8	5.1	14.4	7.8	6.2	11.9	8.7	5.5	14.4	13.6	53.8	14.6	12.0	55.3	13.7	12.2	54.4
	1000	23.9	14.3	6.5	24.1	11.9	5.7	22.6	13.3	6.0	23.0	19.3	71.8	21.8	19.4	71.6	21.7	18.7	71.0

Note: The DGP is given by (42) with β_{i1} and β_{i2} both generated from normal distribution, and contains a single latent factor with factor strength $\alpha = 1, 2/3, 1/2$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

Table 24: Size and power of ARDL adjusted tests of error cross-sectional dependence for panel regression with two latent factors ($m_0 = 2$) with serially correlated non-Gaussian errors

$\hat{m} = 4$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	99.9	100.0	100.0	98.4	100.0	100.0	8.2	12.3	38.5	99.4	100.0	100.0	91.5	98.6	100.0	36.2	43.2	50.4
	200	100.0	100.0	100.0	99.3	100.0	100.0	8.3	7.0	21.5	99.6	100.0	100.0	92.2	99.7	100.0	59.8	75.4	85.4
	500	100.0	100.0	100.0	99.6	100.0	100.0	7.5	7.0	10.7	99.9	100.0	100.0	91.5	99.5	100.0	77.3	94.8	99.8
	1000	100.0	100.0	100.0	99.6	100.0	100.0	9.7	6.1	7.1	99.9	100.0	100.0	91.5	99.9	100.0	82.1	97.9	100.0
ARDL adjusted CD^*	100	8.4	8.4	12.0	8.3	9.4	9.6	11.2	12.1	9.3	33.3	50.8	83.4	40.5	64.4	92.1	79.4	96.5	100.0
	200	6.4	6.8	6.5	7.5	7.8	6.3	11.5	9.2	8.1	29.1	46.2	80.7	39.8	60.7	91.8	82.6	98.4	100.0
	500	7.0	6.7	5.8	6.6	5.6	5.4	10.1	9.2	8.2	30.3	45.0	80.1	43.5	61.9	91.9	85.9	98.3	100.0
	1000	6.4	5.0	5.8	7.5	6.6	5.8	10.5	9.0	7.2	29.3	44.2	78.7	42.4	64.2	90.6	86.6	99.2	100.0
ARDL adjusted CD_{W+}	100	7.2	7.2	10.3	7.3	6.6	10.2	6.1	6.4	9.9	7.9	9.4	27.9	7.6	8.6	29.4	7.7	10.2	28.1
	200	8.5	6.2	6.3	7.6	7.0	5.9	7.6	5.9	5.9	9.8	9.3	32.8	10.0	9.5	32.0	8.0	9.0	34.0
	500	12.1	8.8	5.6	12.2	8.2	5.8	12.8	9.0	4.8	13.9	13.2	59.5	14.7	12.5	57.6	14.5	14.0	60.7
	1000	24.6	15.6	5.9	25.0	12.8	6.6	23.5	12.6	6.5	23.7	22.2	73.1	24.2	19.1	73.0	23.3	18.9	75.6

$\hat{m} = 6$

		Size ($H_o: \rho = 0$)									Power ($H_1: \rho = 0.25$)								
		$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$			$\alpha_1 = 1, \alpha_2 = 1$			$\alpha_1 = 1, \alpha_2 = 2/3$			$\alpha_1 = 2/3, \alpha_2 = 1/2$		
Tests	$n \setminus T$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
ARDL adjusted CD	100	100.0	100.0	100.0	97.9	100.0	100.0	8.1	13.2	40.7	99.6	100.0	100.0	93.9	99.3	100.0	24.7	32.6	50.5
	200	100.0	100.0	100.0	99.1	100.0	100.0	7.6	7.7	22.0	99.9	100.0	100.0	92.9	99.7	100.0	49.2	61.7	71.7
	500	100.0	100.0	100.0	99.4	100.0	100.0	8.6	6.5	11.3	99.9	100.0	100.0	92.5	99.6	100.0	70.8	91.6	99.2
	1000	100.0	100.0	100.0	99.5	100.0	100.0	8.9	6.1	6.8	99.9	100.0	100.0	91.4	99.9	100.0	76.6	96.6	100.0
ARDL adjusted CD^*	100	12.0	14.0	29.5	11.0	14.9	25.8	12.3	16.3	16.7	39.0	59.8	91.6	44.9	70.5	96.1	74.4	95.5	99.9
	200	7.3	8.4	9.6	8.8	8.6	9.2	11.5	10.3	9.5	32.5	49.6	85.1	42.4	64.2	94.0	78.1	97.1	100.0
	500	7.7	7.1	6.5	6.9	6.4	5.5	10.3	9.5	8.3	31.9	46.9	80.5	44.4	62.5	92.7	82.5	98.0	100.0
	1000	7.0	5.3	6.5	7.9	6.7	5.9	9.8	8.3	7.1	30.5	45.1	79.5	43.5	63.5	91.1	83.5	98.4	100.0
ARDL adjusted CD_{W+}	100	6.5	7.4	21.7	7.0	7.1	17.3	7.1	7.7	10.4	6.9	9.4	34.3	7.1	8.0	32.4	8.0	9.1	22.7
	200	6.9	5.8	6.8	7.6	6.4	7.0	7.4	6.3	6.2	8.0	9.2	21.9	8.0	7.8	20.1	8.6	7.7	19.6
	500	12.9	7.5	6.3	13.2	8.3	5.7	12.6	9.2	6.1	14.5	11.8	45.6	12.6	10.8	46.9	13.7	12.4	44.9
	1000	23.5	14.9	5.6	24.8	12.7	6.3	23.4	12.5	6.2	23.5	19.7	66.5	24.6	18.1	65.0	22.5	18.5	64.7

Note: The DGP is given by (42) with β_{i1} and β_{i2} both generated from normal distribution, and contains two latent factors with factor strengths $(\alpha_1, \alpha_2) = [(1,1), (1,2/3), (2/3,1/2)]$. ρ is the spatial autocorrelation coefficient of the error term ε_{it}. m_0 is the true number of factors and $\hat{m}$ is the number of selected PCs used to estimate factors. CD denotes the standard CD test statistic while CD^* denotes the bias-corrected CD test statistic. CD_{W+} denotes the power-enhanced randomized CD test statistic.

References

Bai, J. (2003). Inferential theory for factor models of large dimensions. Econometrica, 71, 135-171.

Baltagi, B. H., Kao, C. & Peng, B. (2016). Testing cross-sectional correlation in large panel data models with serial correlation. Econometrics, 4(4), 44.

Lieberman, O. (1994). A Laplace approximation to the moments of a ratio of quadratic forms. Biometrika, 81, 681-690.

Pesaran, M. Hashem and Yamagata, Takashi, Testing for Alpha in Linear Factor Pricing Models with a Large Number of Securities (March 11, 2017). USC-INET Research Paper No. 17-13, available at SSRN: https://ssrn.com/abstract
=2943640 or http://dx.doi.org/10.2139 /ssrn.2943640.

Ullah, A. et al. (2004). Finite sample econometrics. Oxford University Press, Oxford.