UNIVERSITY OF CAMBRIDGE
Faculty of Economics

CAMBRIDGE WORKING PAPERS IN ECONOMICS

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

M. Hashem Pesaran
University of Southern California and Trinity College, Cambridge

Yimeng Xie
Xiamen University

Abstract

In a recent paper Juodis and Reese (2021) (JR) show that the application of the CD test proposed by Pesaran (2004) to residuals from panels with latent factors results in over-rejection and 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 and 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. 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 have low power against spatial network alternatives. The use of the CD* test is illustrated with two empirical applications from the literature.

Reference Details

CWPE	2158
Published	31 July 2021
Key Words	Latent factor models, strong and weak factors, error cross-sectional dependence, spatial and network alternatives
JEL Codes	C18, C23, C55
Website	www.econ.cam.ac.uk/cwpe

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

M. Hashem Pesaran

University of Southen California, USA, and Trinity College, Cambridge, UK

Yimeng Xie

Department of Statistics and Data Science, School of Economics,

and Wang Yanan Institute for Studies in Economics (WISE), Xiamen University

August 18, 2021

Abstract

In a recent paper Juodis and Reese (2021) (JR) show that the application of the CD test proposed by Pesaran (2004) to residuals from panels with latent factors results in over-rejection and 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 and 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. 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 have low power against spatial network alternatives. The use of the CD^* test is illustrated with two empirical applications from the literature.

JELClassifications: C18, C23, C55

KeyWords: Latent factor models, strong and weak factors, error cross-sectional dependence, spatial and network alternatives

*We are grateful to Natalia Bailey, Alex Chudik, Arturas Juodis, Ron Smith, and Simon Reese for helpful comments and constructive suggestions. The research for this paper was completed when Yimeng Xie was a Ph.D. student at USC.

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.[1] It is, therefore, desirable to test for error cross-sectional dependence once the common factor effects are filtered out.

In a recent paper Juodis and Reese (2021) (JR) show that the application of the CD test proposed by Pesaran (2004, 2015a)[2] to 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-Gaussain and $n >> T$,[3] and seems to lack power under spatial alternatives, which is likely to be particularly important in empirical applications.

In this paper we consider the same problem and 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

[1] See, for example, Chudik et al. (2011) where the different sources of cross-sectional dependence are discussed. 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 the cross section dimension, n.

[2] For a published version of Pesaran (2004) see Pesaran (2021).

[3] The experiments under non-Gaussian errors continue to satisfy JR's moment condition (specified in their Assumption 1) since the errors are generated as chi-squared variates.

over-rejection and will no longer be valid, extending JR's results to panels with semi-strong latent factors.[4] 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.

The finite sample performance of the CD^* test is investigated by Monte Carlo simulations. It is found that 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 of the number of latent factors is unnecessary so long as the number of estimated (selected) factors is no less than the true number. It is also shown that 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

[4] The concepts of weak, semi-strong and strong factors are formalized and discussed by Chudik et al. (2011).

dependence in the CCE residuals of modeling R&D investment is weak when the number of selected principal components (PCs) is sufficiently large.

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. Extension to panel data models with latent factors are discussed in Section 3. Section 4 sets up the Monte Carlo experiments and reports the small sample properties of CD, CD^* and, CD_{W+} tests. Section 5 provides the empirical illustrations. Technical discussions, formal proofs and additional empirical findings are relegated to the Appendix and online supplement.

Notations: For the $n \times n$ matrix $\boldsymbol{A} = (a_{ij})$, we denote its smallest and largest eigenvalues by $\lambda_{min}(\boldsymbol{A})$ and $\lambda_{max}(\boldsymbol{A})$, respectively, its trace by $\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}$, its spectral radius by $\rho(\boldsymbol{A}) = |\lambda_{\max}(\boldsymbol{A})|$, its Frobenius norm by $\|\boldsymbol{A}\|_F$, its spectral norm by $\|\boldsymbol{A}\| = \lambda_{\max}^{1/2}(\boldsymbol{A}'\boldsymbol{A}) \leq \|\boldsymbol{A}\|_F$, its maximum absolute column sum norm by $\|\boldsymbol{A}\|_1 = \max_{1\leq j\leq n}\left(\sum_{i=1}^{n}|a_{ij}|\right)$, and its maximum absolute row sum norm by $\|\boldsymbol{A}\|_\infty = \max_{1\leq i\leq n}\left(\sum_{j=1}^{n}|a_{ij}|\right)$. We write $\boldsymbol{A} > \mathbf{0}$ when $\boldsymbol{A}$ is positive definite. We denote the ℓ_p-norm of the random variable x by $\|x\|_p = E\left(|x|^p\right)^{1/p}$ for $p \geq 1$, assuming that $E\left(|x|^p\right) < K$. $\rightarrow_p$ 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 $(n,T) \to \infty$, such that $n/T \to \kappa$, 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. 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 $f_n/g_n \to 0$ as $n \to \infty$. 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 $\sup_{n\geq n_0}(f_n/g_n) \leq K_1$.

2 Tests of error cross-sectional dependence for a pure latent factor model

2.1 Pure latent factor model

Initially, we consider the following pure multi-factor model,

$$y_{it} = \boldsymbol{\gamma}_i'\mathbf{f}_t + u_{it}, \tag{1}$$

for $i = 1, 2, ..., n$; and $t = 1, 2, ..., T$, where $\mathbf{f}_t = (f_{1t}, f_{2t}, ..., f_{m_0 t})'$ is an $m_0 \times 1$ vector of *unobserved* factors and $\boldsymbol{\gamma}_i = (\gamma_{i1}, \gamma_{i2}, ..., \gamma_{im_0})'$ is the associated vector of unknown coefficients.

Initially we also assume that m_0, the true number of factors, is known, and make the following assumptions:

Assumption 1 *(a) $\mathbf{f}_t$ is a covariance-stationary stochastic process with zero means and covariance matrix, $E(\mathbf{f}_t\mathbf{f}_t') = \mathbf{\Sigma}_{ff} > 0$, where $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_T)'$. (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 $T > T_0$, $T^{-1}\sum_{t=1}^{T}\mathbf{f}_t\mathbf{f}_t' = T^{-1}\mathbf{F}'\mathbf{F} = \mathbf{\Sigma}_{T,ff} > \mathbf{0}$, and $\mathbf{\Sigma}_{T,ff} \to_p \mathbf{\Sigma}_{ff} > \mathbf{0}$.*

Assumption 2 *The $m_0 \times 1$ vector of factor loadings $\boldsymbol{\gamma}_i$ is bounded such that $\sup_i \|\boldsymbol{\gamma}_i\| < K$, $n^{-1}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i' = \mathbf{\Sigma}_{n,\gamma\gamma} \to \mathbf{\Sigma}_{\gamma\gamma} > \mathbf{0}$.*

Assumption 3 *$u_{it} \sim IID(0, \sigma_i^2)$, $sup_i\sigma_i^2 < K$, $inf_i\sigma_i^2 > c$, and $E\left(|u_{it}|^{8+c}\right) < K$. (a) u_{it} is symmetrically distributed around its mean $E(u_{it}) = 0$, and there exists a finite integer T_0 such that for all $T > T_0$, $\omega_{i,T}^2 = T^{-1}\boldsymbol{u}_i'\mathbf{M_F}\boldsymbol{u}_i > c > 0$,*

$$E\left(\frac{\boldsymbol{u}_i'\mathbf{M_F}\boldsymbol{u}_i}{T}\right)^{-4-c} < K, \tag{2}$$

for all i, where $\boldsymbol{u}_i = (u_{i1}, u_{i2}, ..., u_{iT})'$ 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 $\lambda_{max}(\mathbf{V}_T) = O_p(1)$, 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 $\mathbf{f}_{t'}$ and $\boldsymbol{\gamma}_j$, for all i, j, t and t. (d) for a sequence of bounded constants, b_{in}, such that $n^{-1}\sum_{i=1}^{n}b_{in}^2 = O(1)$,

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}b_{in}f_{jt}u_{it}\gamma_{ij} = O_p(1), \text{ for } j = 1, 2, ..., m_0. \tag{3}$$

Remark 1 *Under the above assumptions, the self-normalized error, $\zeta_{it,T}$, defined by*

$$\zeta_{it,T} = \frac{u_{it}}{\omega_{i,T}} = \varepsilon_{it}/(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M_F}\boldsymbol{\varepsilon}_i)^{1/2},$$

where $\varepsilon_{it} = u_{it}/\sigma_i$, and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iT})'$ exists, and is also symmetrically distributed with $E(\zeta_{it,T}) = 0$.

Remark 2 *The fact that there exists a finite T_0 such that (2) 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 $E\left(\frac{1}{\boldsymbol{\varepsilon}_i'\mathbf{M_F}\boldsymbol{\varepsilon}_i}\right)^4 < K$ so long as $T > m_0 + 8$. Under non-Gaussian errors a larger value of T will be needed for the moment condition (2) to hold.*

Remark 3 *The sequence of bounded constants, b_{in}, is introduced in (3) 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.*

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

$$\sum_{i=1}^{n} |\gamma_{ij}| = \ominus\left(n^{\alpha_j}\right), \text{ for } j = 1, 2, \ldots, m_0. \tag{4}$$

The case of strong factors assumed in the principal component analysis (PCA) literature corresponds to $\alpha_j = 1$, for $j = 1, 2, ..., m_0$. Under the weak factor case discussed below, $\alpha_j < 1/2$ for all j. We also denote the maximum value of α_j by $\alpha = max_j(\alpha_j)$.

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 2 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 2 holds. Assumption 3 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 the estimates $\hat{\mathbf{F}}$ and $\hat{\boldsymbol{\Gamma}}$ satisfying the normalization restrictions:

$$\frac{\hat{\boldsymbol{\Gamma}}'\hat{\boldsymbol{\Gamma}}}{n} = \mathbf{I}_{m_0}, \text{ and } \frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T} \text{ is a diagonal matrix.}$$

Then estimators of factors and their loadings are given by

$$\hat{\boldsymbol{\Gamma}} = \sqrt{n}\hat{\mathbf{Q}}, \text{ and } \hat{\mathbf{F}} = \frac{1}{\sqrt{n}}\mathbf{Y}\hat{\mathbf{Q}}, \tag{5}$$

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

$$e_{it} = y_{it} - \hat{\boldsymbol{\gamma}}_i' \hat{\mathbf{f}}_t. \tag{6}$$

2.2 The CD test and its JR modification

The CD test statistic based on the residuals, (6), is given by

$$CD = \sqrt{\frac{2T}{n(n-1)}} \left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \hat{\rho}_{ij,T} \right), \tag{7}$$

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,

$$\tilde{e}_{it,T} = \frac{e_{it}}{\hat{\sigma}_{i,T}}, \tag{8}$$

and $\hat{\sigma}_{i,T}^2 = T^{-1} \sum_{t=1}^{T} e_{it}^2 > c > 0$.[5] 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[6]

$$CD_W = \sqrt{\frac{2}{Tn(n-1)}} \sum_{t=1}^{T} \sum_{i=2}^{n} \sum_{j=1}^{i-1} (w_i e_{it}) (w_j e_{jt}). \tag{9}$$

where w_i, for $i = 1, 2, ..., n$ are independently drawn from a Rademacher distribution of which the probability mass function is

$$f(w_i) = \begin{cases} \frac{1}{2}, & \text{if } w_i = -1 \\ \frac{1}{2}, & \text{if } w_i = 1. \end{cases}$$

Because of the random properties of the weights, JR show that CD_W converges to a standard normal distribution regardless of the values of $\tilde{e}_{it,T}$, and as a result the over-rejection problem of 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},

[5] This condition ensures that $1/\hat{\sigma}_{i,T}^2 < K < \infty$, which is assumed throughout.

[6] 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.

to CD_W, to obtain CD_{W+} defined by

$$CD_{W+} = CD_W + \Delta_{nT}, \tag{10}$$

where

$$\Delta_{nT} = \sum_{i=2}^{n}\sum_{j=1}^{i-1} |\hat{\rho}_{ij,T}| \mathbf{1}\left(|\hat{\rho}_{ij,T}| > 2\sqrt{\ln(n)/T}\right). \tag{11}$$

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 $\Delta_{nT} \to_p 0$ 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 $i \neq j$,

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

where $\rho_{ij} = E\left(\varepsilon_{it}\varepsilon_{jt}\right)$. Then

$$\begin{aligned} & E\left(\Delta_{nT}|\rho_{ij} = 0, \text{ for all } i \neq j\right) \\ & \leq \frac{n^2}{2} \sup_{i\neq j} E\left[|\hat{\rho}_{ij,T}|^2 |\rho_{ij} = 0\right] \times \sup_{i\neq j} P\left[|\hat{\rho}_{ij,T}| > 2\sqrt{\frac{\ln(n)}{T}}|\rho_{ij} = 0\right]. \end{aligned} \tag{12}$$

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

$$E\left[|\hat{\rho}_{ij,T}|^2 |\rho_{ij} = 0\right] = O\left(\frac{1}{T}\right) \tag{13}$$

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 $\sup_{i\neq j} E\left|\varepsilon_{it}\right|^{2s} < K$, for some integer $s \geq 3$ (see Assumption 2 of BPS). Also using results in

Lemma 2 of 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, $\frac{C_p(n,\delta)}{\sqrt{T}}$ and $2\sqrt{\frac{\ln(n)}{T}}$ have the same limiting properties if we set $\delta = 2$. Overall, it then follows that

$$\sup_{i\neq j} P\left(|\hat{\rho}_{ij,T}| > 2\sqrt{\frac{\ln(n)}{T}}|\rho_{ij} = 0\right) = O\left(n^{-\frac{2}{\varphi_{\max}}}\right) + O\left(T^{-\frac{(s-1)}{2}}\right). \tag{14}$$

Using (13) and (14) in (12), we now have

$$E\left(\Delta_{nT}|\rho_{ij} = 0, \text{ for all } i \neq j\right) = O\left(\frac{n^{2-\frac{2}{\varphi_{\max}}}}{\sqrt{T}}\right) + O\left(n^2 T^{-s/2}\right) \tag{15}$$

Therefore, $\Delta_{nT} \to_p 0$, if $n^2T^{-s/2} \to 0$ and $T^{-1/2}n^{2\left(1-\frac{1}{\varphi_{\max}}\right)} \to 0$. It is easily seen that both of these conditions will be met as n and $T \to \infty$ 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 $E\left(\varepsilon_{it}^2\varepsilon_{jt}^2\right)$ exceed unity. For example, if ε_{it} has a multivariate-t distribution with $v > 4$ degrees of freedom, 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 (15) tends to 0 if $n^{2\left(1-\frac{v-4}{v-2}\right)-d/2} \to 0$, or if $d > \frac{8}{v-2}$. 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 $d > 1$ when $v = 10$, requiring T to rise faster than n.[7]

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.

[7] We are grateful to JR who draw our attention to the moment requirement of their Assumption 1.

2.3 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 (5). 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 (7), equivalently as (established in Lemma S.8 of the online supplement)

$$CD = \left(\sqrt{\frac{n}{n-1}}\right)\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\frac{\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{e_{it}}{\hat{\sigma}_{i,T}}\right)^2-1}{\sqrt{2}}\right]. \tag{16}$$

We also introduce the following analogue of CD

$$\widetilde{CD} = \left(\sqrt{\frac{n}{n-1}}\right)\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\frac{\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{e_{it}}{\omega_{i,T}}\right)^2-1}{\sqrt{2}}\right], \tag{17}$$

where $\omega_{i,T}^2 = T^{-1}\mathbf{u}_i'\mathbf{M}_{\mathbf{F}}\mathbf{u}_i$, with the following key results (established in Lemmas S.2 and S.9 in the online supplement):

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

and

$$CD - \widetilde{CD} = o_p(1). \tag{18}$$

By 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

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{e_{it}}{\omega_{i,T}}=\psi_{t,nT}-s_{t,nT} \tag{19}$$

where

$$\psi_{t,nT}=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}a_{i,nT}\zeta_{it,T},\;\; a_{i,nT}=1-\omega_{i,T}\boldsymbol{\varphi}_{nT}'\boldsymbol{\gamma}_i, \tag{20}$$

$$s_{t,nT}=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\boldsymbol{\varphi}_{nT}'\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)u_{it}+\left(\hat{\boldsymbol{\delta}}_{i,T}-\boldsymbol{\delta}_{i,T}\right)'\hat{\mathbf{f}}_t+\boldsymbol{\varphi}_{nT}'\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i'\mathbf{f}_t\right], \tag{21}$$

with $\boldsymbol{\varphi}_{nT}=n^{-1}\sum_{i=1}^{n}\boldsymbol{\delta}_{i,T}$. Using the above results, the $\widetilde{CD}$ statistic defined in (17) can then be decomposed as

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

Under Assumptions 1-3 the last two terms of $\widetilde{CD}$ are shown in Lemma S.4 of 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 (18) we have $CD=z_{nT}+o_p(1)$. Furthermore, as established in proof of Theorem 1, we have

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

where

$$\xi_{t,n}=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}a_{i,n}\varepsilon_{it},\;\; a_{i,n}=1-\sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i, \tag{22}$$

$\boldsymbol{\varphi}_n=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\delta}_i$, and $\boldsymbol{\delta}_i=\boldsymbol{\gamma}_i/\sigma_i$. Since $a_{i,n}$ are given constants, then $E\left(\xi_{t,n}\right)=0$,

$$E\left(\xi_{t,n}^2\right)\equiv\omega_n^2=\frac{1}{n}\sum_{i=1}^{n}a_{i,n}^2=n^{-1}\sum_{i=1}^{n}\left(1-\sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i\right)^2, \tag{23}$$

and

$$Var\left(\xi_{t,n}^2\right) = 2\left(\frac{1}{n}\sum_{i=1}^{n} a_{i,n}^2\right)^2 - \kappa_2\left(\frac{1}{n^2}\sum_{i=1}^{n} a_{i,n}^4\right), \tag{24}$$

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 (24) 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

$$CD^*(\theta_n) = \frac{CD + \sqrt{\frac{T}{2}}\theta_n}{1-\theta_n}, \tag{25}$$

where

$$\theta_n = 1 - \frac{1}{n}\sum_{i=1}^{n} a_{i,n}^2, \tag{26}$$

with $a_{i,n}$ defined by (22). The above results are summarized in the following theorem.

Theorem 1 *Consider the model in (1) and assume the factor number m_0 is known. Also suppose Assumptions 1-3 hold. (a) 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 (25) has the limiting $N(0,1)$ distribution. (b) $\theta_n = \ominus\left(n^{\alpha-1}\right)$, where θ_n is defined by (26), and $\alpha = max_{j=1,2,\ldots,m_0}(\alpha_j)$, with α_j representing the strength of the latent factor, f_{jt}, defined by (4).*

Remark 4 *Part (b) of the above theorem establishes that the relationship between CD and $CD^*\left(\theta_n\right)$ 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 (25), 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 $d = 1$.

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

$$\hat{\theta}_{nT} = 1 - \frac{1}{n}\sum_{i=1}^{n}\hat{a}_{i,nT}^2 \tag{27}$$

where

$$\hat{a}_{i,nT} = 1 - \hat{\sigma}_{i,T}\left(\hat{\boldsymbol{\varphi}}_{nT}'\hat{\boldsymbol{\gamma}}_i\right), \quad \hat{\boldsymbol{\varphi}}_{nT} = \frac{1}{n}\sum_{i=1}^{n}\hat{\boldsymbol{\delta}}_{i,nT}, \tag{28}$$

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 (26) and its estimator $\hat{\theta}_{nT}$ given by (27). Suppose Assumptions 1-3 hold. Then for $(n,T) \to \infty$, such that $n/T \to \kappa$, where $0 < \kappa < \infty$, we have*

$$\sqrt{T}\left(\hat{\theta}_{nT} - \theta_n\right) = o_p(1). \tag{29}$$

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

$$CD^*\left(\hat{\theta}_{nT}\right) =_{.} CD^* = \frac{CD + \sqrt{\frac{T}{2}}\hat{\theta}_{nT}}{1 - \hat{\theta}_{nT}}. \tag{30}$$

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 (30) has the limiting $N(0,1)$ distribution. as $(n,T) \to \infty$, such that $n/T \to \kappa$, and $0 < \kappa < \infty$.*

Remark 5 *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 the $CD^* \to_d N(0,1)$ under the null hypothesis. Therefore, to control the size of the CD^* test the number of factors assumed when estimating $\hat{\theta}_{nT}$ should be set to ensure that $\hat{m} \geq m_0$. 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 6 *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 Tests of error cross-sectional dependence for a panel data model with latent factors

Consider now the following general panel regression model that explains the scalar variables y_{it}, for $i = 1, 2, ..., n$; and $t = 1, 2, ..., T$, in terms of observed and latent covariates:

$$y_{it} = \boldsymbol{\alpha}_i' \mathbf{d}_t + \boldsymbol{\beta}_i' \mathbf{x}_{it} + v_{it},\ v_{it} = \boldsymbol{\gamma}_i' \mathbf{f}_t + u_{it}, \tag{31}$$

where $\mathbf{d}_t$ is a $k_d \times 1$ vector of observed common factors which can be either constant or covariance stationary, $\mathbf{x}_{it}$ is a $k_x \times 1$ vector of unit-specific regressors, 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 vector 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 test the cross-sectional independence of error term in a mixed factor model as (31), 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

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

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_{is} and $\boldsymbol{\varepsilon}_{xis}$ for all i, t, s.*

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

Assumption 6 *The individual-specific errors u_{it} and $\boldsymbol{\varepsilon}_{xjs}$ are distributed independently for all i, j, t and s, and for each i, $\boldsymbol{\varepsilon}_{xjs}$ follows a linear stationary process with absolute summable autocovariances given by*

$$\boldsymbol{\varepsilon}_{xit} = \sum_{l=0}^{\infty} \mathbf{S}_{il}\eta_{xit-l}$$

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 $\sup_i \left\|\mathbf{\Sigma}_{xi}\right\|_2 < K$.

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 section 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 7 *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 1 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. 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 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

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

Note when $\boldsymbol{\alpha}_i = \mathbf{0}$ and $\boldsymbol{\beta}_i = \mathbf{0}$, (31) reduces the the pure latent factor model, (1), 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 (30). 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 $\{\hat{v}_{it};\ i = 1, \ldots, n;$ 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 $e_{it} = \hat{v}_{it} - \hat{\boldsymbol{\gamma}}_i'\hat{\mathbf{f}}_t$, 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 (30).

Theorem 3 *Consider the panel data model (31) 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 8 *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 Small sample properties of CD^* and CD_{W^+} tests

4.1 Data generating process

We consider the following data generating process

$$y_{it} = \mathrm{a}_i + \sigma_i \left(\beta_{i1} d_t + \beta_{i2} x_{it} + m_0^{-1/2} \boldsymbol{\gamma}_i' \mathbf{f}_t + \varepsilon_{it} \right), \ i = 1, 2, ..., n; t = 1, 2, ..., T, \tag{32}$$

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 $\sigma_i^2 = 0.5 + \frac{1}{2} s_i^2$, with $s_i^2 \sim IID\chi^2(2)$, which ensures that $E(\sigma_i^2) = 1$.

4.1.1 DGP under the null hypothesis

Under the null hypothesis the errors ε_{it} are generated as $IID(0, 1)$, we consider both Gaussian and non-Gaussian distributions for ε_{it}:

- Gaussian errors: $\varepsilon_{it} \sim IIDN(0, 1)$,
- Chi-squared distributed errors: $\varepsilon_{it} \sim IID\left(\frac{\chi^2(2)-2}{2}\right)$.

The focus of the experiments is on testing the null hypothesis that ε_{it} are IID, whilst allowing for the presence of m_0 unobserved factors, $\mathbf{f}_t = (f_{1t}, f_{2t}, ..., f_{m_0 t})'$. 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.

$$(\alpha_1, \alpha_2) = [(1, 1), (1, 2/3), (2/3, 1/2)]\,.$$

The intercepts a_i are generated as $IIDN(1, 2)$ and fixed thereafter. The observed common factor is generated as an AR(1) process:

$$d_t = \rho_d d_{t-1} + \sqrt{1 - \rho_d^2} v_{dt},$$

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:

$$x_{it} = \gamma_{xi1} f_{1t} + \gamma_{xi2} f_{2t} + e_{xit}, \tag{33}$$

where

$$f_{jt} = r_j f_{j,t-1} + \sqrt{1 - r_j^2} v_{jt},$$

$r_j = 0.9$ and $v_{jt} \sim IID\left(\frac{\chi^2(2)-2}{2}\right)$, for $j = 1, 2$. The factor loadings in (33) are generated as $\gamma_{xi1} \sim IIDU\,(0.25, 0.75)$ and $\gamma_{xi2} \sim IIDU\,(0.1, 0.5)$. The error term of (33) is generated as a stationary process:

$$e_{xit} = \rho_i e_{xi,t-1} + \sqrt{1 - \rho_i^2} v_{xit},\; i = 1, 2 \ldots, n,$$

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 panel regression model with latent factors, we allow for heterogeneous slopes and generate the slopes of observed covariates, d_t and x_{it}, as $\beta_{i1} \sim IIDN(\mu_{\beta 1}, \sigma_{\beta 1}^2)$, and $\beta_{i2} \sim IIDN(\mu_{\beta 2}, \sigma_{\beta 2}^2)$ where $\mu_{\beta 1} = \mu_{\beta 2} = 0.5$ and $\sigma_{\beta 1}^2 = \sigma_{\beta 2}^2 = 0.25$, 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

$$PR_{nT}^2 = 1 - \frac{(nT)^{-1} \sum_{i=1}^n \sum_{t=1}^T \sigma_i^2 E\left(\varepsilon_{it}^2\right)}{(nT)^{-1} \sum_{i=1}^n \sum_{t=1}^T Var\left(y_{it}\right)}. \tag{34}$$

Since the underlying processes, (32) and (33), are stationary and $E\left(\varepsilon_{it}^2\right) = 1$, we have

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

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

$lim_{n\to\infty} PR_n^2 = \eta^2/(1+\eta^2)$, where

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

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 $\alpha = max(\alpha_1, \alpha_2)$.

4.1.2 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 following 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 errors, ζ_{it}, we consider both Gaussian and non-Gaussian distributions, namely $\zeta_{it} \sim IIDN\left(0,1\right)$ and $\zeta_{it} \sim IID[\frac{\chi^2(2)-2}{2}]$. 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

$$c^2 = \frac{n}{\mathrm{tr}\left[\left(\mathbf{I}_n - \rho \boldsymbol{W}\right)^{-1}\left(\mathbf{I}_n - \rho \boldsymbol{W}\right)'^{-1}\right]},$$

which ensures that $n^{-1}\sum_{i=1}^n Var(\varepsilon_{it}) = 1$, irrespective of the choice of ρ.

4.2 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 $2,000$. For the pure latent factor models, we compute the filtered residuals as $\hat{v}_{it} = y_{it} - \hat{\mathrm{a}}_i$, where $\hat{\mathrm{a}}_i = T^{-1}\sum_{t=1}^T y_{it}$. For the panel regressions with latent factors, the filtered residuals are computed as

$$\hat{v}_{it} = y_{it} - \hat{\mathrm{a}}_{CCE,i} - \hat{\beta}_{CCE,i1} d_t - \hat{\beta}_{CCE,i2} x_{it}, \tag{35}$$

where $\left(\hat{\mathrm{a}}_{CCE,i}, \hat{\beta}_{CCE,i1}, \hat{\beta}_{CCE,i2}\right)$ is the CCE estimator of a_i, β_{i1} and β_{i2}, as set out in Pesaran (2006). The CCE estimators are consistent so long as the relevant rank condition is met, which requires that $m_0 \leq 1+k = 2$, which is satisfied in the case of our Monte Carlo experiments. Then we will compute the first m PCs $\{\hat{v}_{it};\ i = 1, 2, \ldots, n;$ and $t = 1, 2, \ldots, T\}$ and the associated factor loadings, namely $(\hat{\boldsymbol{\gamma}}_i, \hat{\mathbf{f}}_t)$, subject to the normalization, $n^{-1}\sum_{i=1}^n \hat{\boldsymbol{\gamma}}_i \hat{\boldsymbol{\gamma}}_i' = \mathbf{I}_m$. Finally the residuals, to be used in the computation of 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 $t = 1, 2, \ldots, T$.

In practice, the true number of factors m_0 is not known and in carrying the various CD tests we need to set m such that $m \geq m_0$. To that end, in addition to reporting the results with $m = m_0$, we also consider $m = 2$ for one-factor specification and $m = 4$ for the two-factor specification. JR apply a similar procedure to obtain $\hat{v}_{it}$ as shown in (35), but they differ from us in computing the residual e_{it} as the latent factors are estimated by cross-sectional averages.

4.3 Simulation results

4.3.1 Gaussian errors

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 (10), tends to zero sufficiently fast and our Assumption 3 is met. Next, we report simulation results for the DGPs with chi-squared distributed errors, where the errors do not satisfy the symmetry requirement of Assumption 3, and also allows us to investigate the robustness of the JR test and our proposed bias-corrected CD test to departures from Gaussianity. We consider spatial alternatives such that the size is examined by setting the spatial coefficient $\rho = 0$, and report power for $\rho = 0.25$. As to be expected power rises with ρ and additional simulation results for values of $\rho > 0.25$ do not seem to add much to our investigation.

The simulation results for the DGPs with the errors following standard normal distribution are shown in Table 1-4. 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 m, is the same as the true number of factors, m_0, while the bottom panel reports the results when $m = 2$. 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 summarizes the size and power results for the pure factor model with $m_0 = 2$, and reports the results when 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 efficiency 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, 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 our asymptotic results.

4.3.2 Chi-squared distributed errors

The simulation results for the DGPs with chi-squared errors are provided in Tables 5 to 8. For standard CD test and its biased-corrected version, CD^*, the results are very similar to the ones with Gaussian errors, suggesting that CD^* test is robust to the symmetry assumption that underlie our theoretical derivations. 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; and irrespective of whether the factors are strong or weak. As we discussed in Section 2.2, 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 primarily interest.

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

5 Empirical illustrations

5.1 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.

Table 1: Size and power of tests of error cross-sectional dependence for a pure single factor model ($m_0 = 1$) with standard normal errors and spatial correlated alternative

$m = 1$

		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/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

$m = 2$

		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/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 (32) 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 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 a pure two factors model ($m_0 = 2$) with standard normal errors and spatial correlated alternative

$m = 2$

		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	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

$m = 4$

		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	99.8	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 (32) 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 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 data model with a latent factor ($m_0 = 1$) with standard normal errors and spatial correlated alternative

$m = 1$

		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/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

$m = 2$

		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/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 (32) 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 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 data model with two latent factors ($m_0 = 2$) with standard normal errors and spatially correlated alternative

$m = 2$

		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	97.8	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

$m = 4$

		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	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 (32) 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 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 a pure single factor model ($m_0 = 1$) with chi-squared errors with 2 degrees of freedom and spatially correlated alternative

$m = 1$

		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/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

$m = 2$

		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/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 (32) 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 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 a pure two factors model ($m_0 = 2$) with chi-squared errors with 2 degrees of freedom and spatially correlated alternative

$m = 2$

		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	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

$m = 4$

		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	99.8	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 (32) 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 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 data model with one latent factor ($m_0 = 1$) with chi-squared errors with 2 degrees of freedom and spatially correlated alternative

$m = 1$

		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/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

$m = 2$

		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/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 (32) 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 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 data model with two latent factors ($m_0 = 2$) with chi-squared errors with 2 degrees of freedom and spatially correlated alternative

$m = 2$

		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

$m = 4$

		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 (32) 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 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.

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

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

where π_{it} is real house price in MSA i at time t, and $\mathrm{l}\{q_t = j\}$ is the index for quarter j, and $\mathbf{f}_t$ is the $m \times 1$ 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 $\{\hat{v}_{it} : i = 1, 2, \ldots, n, t = 1, 2, \ldots, T\}$, where $\hat{v}_{it} = \pi_{it} - \hat{a}_i - \sum_{j=1}^{3} \hat{\beta}_{ij} \mathrm{l}\{q_t = j\}$, to obtain the estimates $\hat{\boldsymbol{\gamma}}_i$ and $\hat{\mathbf{f}}_t$ for different choices of m (selected number of PCs).[8] Then the standard CD, its bias-corrected version, CD^*, and the CD_{W+} test of JR are computed using the de-seasonalized and de-factored series given by

$$e_{it} = \pi_{it} - \hat{a}_i - \sum_{j=1}^{3} \hat{\beta}_{ij} \mathrm{l}\left\{q_t = j\right\} - \hat{\boldsymbol{\gamma}}_i' \hat{\mathbf{f}}_t. \tag{36}$$

The CD statistics are reported in the panel (a) of Table 9 for values of $m = 1, 2, ..., 6$. Generally all three CD tests reject the null hypothesis of cross-sectional independence irrespective of m, with exception of the standard CD test when $m = 5$. It can also be observed that CD is always less than CD^*, indicating CD is negatively biased.

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

[8] $\hat{a}_i$ and $\hat{\beta}_{ij}$ are estimated by OLS regression of π_{it} on the intercept and seasonal dummies, that are independent of the latent factors.

factor model containing observed regional and national factors, as well as latent factors:

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

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 $\left(1\left\{q_t = j\right\}, \bar{\pi}_{rt}, \bar{\pi}_t\right)$ for each i to generate the residuals

$$\hat{v}_{irt} = \pi_{irt} - \hat{a}_{ir} - \sum_{j=1}^{3} \hat{\beta}_{ir,j} 1\left\{q_t = j\right\} - \hat{\delta}_{ir,1}\bar{\pi}_{rt} - \hat{\delta}_{ir,2}\bar{\pi}_t, \tag{37}$$

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 choice of m, and the residuals

$$e_{irt} = \pi_{irt} - \hat{a}_{ir} - \sum_{j=1}^{3} \hat{\beta}_{ir,j} 1\left\{q_t = j\right\} - \hat{\delta}_{ir,1}\bar{\pi}_{rt} - \hat{\delta}_{ir,2}\bar{\pi}_t - \hat{\boldsymbol{\gamma}}_{ir}^{\prime}\hat{\mathbf{f}}_t. \tag{38}$$

The CD, CD^* and CD_{W+} test statistics based on these residuals are reported in the panel (b) of Table 9, again for $m = 1, 2, ..., 6$. All three CD tests reject the null hypothesis of cross-sectional independence for all choices of m. CD is still less than CD^* for each m, but compared to the model without regional and national effects, now the difference between CD and CD^* is much smaller. Intuitively regional and national effects account for some of the latent factors such that after filtering these effects the cross-sectional dependence in $\hat{v}_{irt}$ of (37) becomes weaker. Since the bias in CD decreases with the strengths of latent factors that are included in $\hat{v}_{irt}$, the standard CD test and the bias-corrected CD test become closer. 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.

5.2 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, 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

Table 9: Tests of error cross-sectional dependence for real house price changes

Panel (a): Without regional and national factors

Test\\m	1	2	3	4	5	6
CD	8.9	15.5	4.7	1.3	0.1	-3.9
CD^*	108.8	150.9	99.4	93.2	87.1	50.8
CD_{w+}	2659.2	2456.8	1765.3	1677.5	1547.4	1348.1

Panel (b): With regional and national factors

Test\\m	1	2	3	4	5	6
CD	107.9	106.8	112.4	125.0	46.0	42.4
CD^*	117.9	119.5	122.4	137.0	70.0	67.0
CD_{w+}	1593.3	1546.8	1441.5	1321.9	1211.2	1150.9

Note: In the first panel the tests are applied to residuals in equation (36) where we de-seasonalize and de-factor real house price change. In the second panel the tests are applied to residuals in equation (38) where we 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 m.

10 countries[9] over the period 1981- 2005, and 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. They find that the CD test rejects the null hypothesis of error cross-sectional independence. JR revisit Eberhardt et al. (2013) test results 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)

$$\ln(Y_{it}) = \beta_0 + \beta_1 \ln(L_{it}) + \beta_2 \ln(K_{it}) + \beta_3 \ln(R_{it}) + \boldsymbol{\gamma}_i' \mathbf{f}_t + u_{it} \tag{39}$$

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:

$$\hat{v}_{it} = \ln(y_{it}) - \hat{\beta}_{CCE,0} - \hat{\beta}_{CCE,1} \ln(L_{it}) - \hat{\beta}_{CCE,2} \ln(K_{it}) - \hat{\beta}_{CCE,3} \ln(R_{it}). \tag{40}$$

In both Eberhardt et al. (2013) and JR the residuals in (40) 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

[9]The countries include Denmark, Finland, Germany, Italy, Japan, Netherlands, Portugal, Sweden, United Kingdom, and United States.

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 as m. Also, because rank condition is required for the consistency of CCE estimators, it is implicitly assumed that the number of latent factors in (39) is not larger than the number of time varying regressors (in the present application 3) plus one.[10] Accordingly, we apply PCA to $\hat{v}_{it}$, with the number of selected PCs set to $m = 1, 2, 3$, and 4. The results are summarized in Table 10. As can be seen, the test outcomes are quite sensitive to the number of PCs selected. The CD and CD^* tests reject the null of cross-sectional independence when $m = 3$, but not if $m = 4$. In comparison, the CD_{W^+} test rejects the null for all values of m.

Table 10: Tests of error cross-sectional dependence for panel regressions of R&D investment

	1	2	3	4
CD	0.5	2.1	4.1	-0.8
CD^*	2.1	3.3	6.3	1.7
CD_{W+}	38.4	3.9	3.7	8.6

Note: The tests are applied to residuals in equation (40) where we model R&D investment. See also the notes to Table 9.

6 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 CD test in the case of strong factors is also established by Juodis and Reese (2021), who propose a randomized test statistic to correct for over-rejection and add a screening component to 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

[10] It is worth noting that the CCE estimator continues to be consistent even with failure of rank condition, but requires additional assumptions such that factor loadings $\boldsymbol{\gamma}_i$ in (39) are independently and identically distributed across i, see Pesaran (2006) and Pesaran (2015b).

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 tend to infinity 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. Our analysis is confined to static panels and further research is required before the CD* can be considered for dynamic panels with latent factors.

Appendix

This appendix provides the proofs of Theorems 1 to 3, and Corollary 1. The auxiliary lemmas used in these proofs are stated and established in Section S1 of the online supplement.

Proof of Theorem 1. We first note that the residuals of the factor model (1) estimated using PCs, given by (5), can be written as:

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

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

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

As shown in Lemma S.1 of the online supplement, and recalling that $y_{jt} = \boldsymbol{\gamma}_j'\mathbf{f}_t = u_{jt}$, we have

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

Using this result in (A.2) we obtain

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

and summing over i we have

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

where

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

Written more compactly

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

where

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

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

Further, let

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

where $\boldsymbol{\varphi}_n = n^{-1}\sum_{i=1}^{n}\boldsymbol{\delta}_i$, and $\boldsymbol{\delta}_i = \boldsymbol{\gamma}_i/\sigma_i$. Then $\psi_{t,nT}$, given by (A.6), can be written as

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

and since $\omega_{i,T}\zeta_{it,T} = u_{it} = \sigma_i\varepsilon_{it}$, (recall that $\zeta_{it,T} = \varepsilon_{it} / \left(\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i/T\right)^{1/2}$) then

$$\begin{aligned}
\psi_{t,nT} &= \xi_{t,n} + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\zeta_{it,T} - \varepsilon_{it}\right) - \left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)'\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\sigma_i\varepsilon_{it}. \\
&= \xi_{t,n} + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{1}{\left(\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i/T\right)^{1/2}} - 1\right)\varepsilon_{it} - \left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)'\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\sigma_i\varepsilon_{it}\right).
\end{aligned}$$

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

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

where

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

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

$$\upsilon_{t,nT} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{1}{\left(\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i/T\right)^{1/2}} - 1\right)\varepsilon_{it}. \tag{A.12}$$

Now recall from (17) that

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

Then using (A.5), 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

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

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

and by Lemma S.4 of the online supplement $p_{nT} = o_p(1)$, and $q_{nT} = o_p(1)$. Hence

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

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

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

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

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

where $\upsilon_{t,nT}$ is defined by (A.12), and we have

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

and hence

$$\begin{aligned}
E\left(\upsilon_{t,nT}^2\right) &= \frac{1}{n} \sum_{i=1}^{n} \left[E\left(\frac{\varepsilon_{it}^2}{\boldsymbol{\varepsilon}_i' \mathbf{M}_F \boldsymbol{\varepsilon}_i / T} \right) - 1 \right] - 2\frac{1}{n} \sum_{i=1}^{n} \left[E\left(\frac{\varepsilon_{it}^2}{\left(\boldsymbol{\varepsilon}_i' \mathbf{M}_F \boldsymbol{\varepsilon}_i / T\right)^{1/2}} \right) - 1 \right] \\
&\leq \frac{1}{n} \sum_{i=1}^{n} \left| E\left(\frac{\varepsilon_{it}^2}{\boldsymbol{\varepsilon}_i' \mathbf{M}_F \boldsymbol{\varepsilon}_i / T} - \right) \right| + 2\frac{1}{n} \sum_{i=1}^{n} \left| E\left(\frac{\varepsilon_{it}^2}{\left(\boldsymbol{\varepsilon}_i' \mathbf{M}_F \boldsymbol{\varepsilon}_i / T\right)^{1/2}} - 1 \right) - 1 \right|.
\end{aligned}$$

Meanwhile using (S.48) and (S.52) in Lemma S.7 of the online supplement, we also have

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

and therefore

$$E\left(\upsilon_{t,nT}^2\right) = O\left(T^{-1}\right). \tag{A.18}$$

Using this result in (A.17) we obtain $E\left(|g_{2,nT}|\right) = O(T^{-1/2})$, 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 (A.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 S.5 we have

$$\sqrt{T}\left(\boldsymbol{\varphi}_{n}-\boldsymbol{\varphi}_{nT}\right) = O_p\left(n^{-1/2}\right) + O_p\left(T^{-1/2}\right). \tag{A.19}$$

Furthermore

$$\begin{aligned} E\left\|T^{-1}\sum_{t=1}^{T}\kappa_{t,n}\kappa_{t,n}'\right\| &\leq \frac{1}{T}\sum_{t=1}^{T}E\left\|\kappa_{t,n}\kappa_{t,n}'\right\| = \frac{1}{T}\sum_{t=1}^{T}E\left\|\kappa_{t,n}\right\|^2 \\ &= \frac{1}{T}\sum_{t=1}^{T}E\left(\kappa_{t,n}'\kappa_{t,n}\right). \end{aligned}$$

Then using (A.11)

$$E\left(\kappa_{t,n}'\kappa_{t,n}\right) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\boldsymbol{\gamma}_i'\boldsymbol{\gamma}_j\sigma_i\sigma_j E\left(\varepsilon_{it}\varepsilon_{jt}\right) = \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\boldsymbol{\gamma}_i'\boldsymbol{\gamma}_i\right) < K, \tag{A.20}$$

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 (A.19) we then have

$$g_{3,nT} = o_p(1). \tag{A.21}$$

Similarly, by Cauchy-Schwarz inequality and using (A.18) and (A.20) we have

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

Then using the above results it also follows that

$$g_{4,nT} = \sqrt{T}\left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)' \left(\frac{1}{T}\sum_{t=1}^{T} \kappa_{t,n} \upsilon_{t,nT}\right) = o_p(1). \tag{A.22}$$

Similarly,

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

where using (A.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 (A.19) it follows that

$$g_{5,nT} = \sqrt{T}\left(\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n\right)' \left(\frac{1}{T}\sum_{t=1}^{T} \kappa_{t,n}\xi_{t,n}\right) = o_p(1). \tag{A.23}$$

Using results (A.17) to (A.23) in (A.16) now yields

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\psi_{t,nT}^2 = \left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{t,n}^2\right)\left(1 + \frac{2\sum_{t=1}^{T}\xi_{t,n}\upsilon_{t,nT}}{\sum_{t=1}^{T}\xi_{t,n}^2}\right) + o_p(1). \tag{A.24}$$

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

$$\widetilde{CD}^* = \frac{\widetilde{CD} + \sqrt{\frac{T}{2}}\theta_n}{1-\theta_n} \tag{A.25}$$

where $\theta_n = 1 - \frac{1}{n}\sum_{i=1}^{n} a_{i,n}^2$, and $a_{i,n} = 1 - \sigma_i\boldsymbol{\varphi}_n'\boldsymbol{\gamma}_i$. Then using (A.15) in (A.25), we have

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

Now using (A.24) in the above and after some re-arrangement of the terms we obtain

$$\widetilde{CD}^* = \frac{\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\frac{\xi_{t,n}^2-(1-\theta_n)}{\sqrt{2}}\right)\right]\left(1+\frac{2}{\sqrt{T}}w_{nT}\right)}{1-\theta_n} + \sqrt{2}w_{nT} + o_p\left(1\right), \tag{A.26}$$

where

$$w_{nT} = \frac{T^{-1}\sum_{t=1}^{T}\xi_{t,n}\left(\sqrt{T}\upsilon_{t,nT}\right)}{T^{-1}\sum_{t=1}^{T}\xi_{t,n}^2}. \tag{A.27}$$

It is clear that the denominator of w_{nT} is $O_p(1)$, and using (A.8) and (A.12) we have

$$\xi_{t,n} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} a_{i,n} \varepsilon_{it} \rightarrow_d N(0, \omega_\xi^2),$$

where $\omega_\xi^2 = lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{i=1}^{n} a_{i,n}^2 \right) < K$. Consider now the numerator of (A.27), and note that there exists T_0 such that for all $T > T_0$, $\xi_{t,n} \left(\sqrt{T} \upsilon_{t,nT} \right)$ are serially independent with zero means and finite variances. Also using (A.12I note that

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

where it is easily seen that

$$E \left(\sqrt{T} \upsilon_{t,nT} \right) = 0, \text{ and } Var \left(\sqrt{T} \upsilon_{t,nT} \right) = \frac{1}{n} \sum_{i=1}^{n} E \left[T \left(\frac{1}{\left(\boldsymbol{\varepsilon}_i' \mathbf{M}_F \boldsymbol{\varepsilon}_i / T \right)^{1/2}} - 1 \right)^2 \varepsilon_{it}^2 \right].$$

Also using (S.54) and (S.55) in Lemma S.7 of the online supplement, we have

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

Hence, for all $T > T_0$ by application of standard central limit theorem to $\sqrt{T} \upsilon_{t,nT}$ we have $\sqrt{T} \upsilon_{t,nT} \rightarrow_d N(0, \omega_\upsilon^2)$, as $n \rightarrow \infty$, where

$$\omega_\upsilon^2 = lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^{n} E \left[T \left(\frac{1}{\left(\boldsymbol{\varepsilon}_i' \mathbf{M}_F \boldsymbol{\varepsilon}_i / T \right)^{1/2}} - 1 \right)^2 \varepsilon_{it}^2 \right] < K.$$

Then it readily follows that

$$Var \left[T^{-1} \sum_{t=1}^{T} \xi_{t,n} \left(\sqrt{T} \upsilon_{t,nT} \right) \right] = T^{-2} \sum_{t=1}^{T} E \left(\xi_{t,n}^2 T \upsilon_{t,nT}^2 \right) = O(T^{-1}),$$

and hence $T^{-1} \sum_{t=1}^{T} \xi_{t,n} \left(\sqrt{T} \upsilon_{t,nT} \right) = o_p(1)$. Using this result in (A.27) and noting that its numerator is bounded then it follows that $w_{nT} = o_p(1)$, and as a result (using (A.26)) 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

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

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

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

Also by Lemma S.9 in the online supplement, $CD = \widetilde{CD} + o_p(1)$.Then it follows that

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

which establishes part (a) of Theorem 1. To prove part (b) of the theorem we first note that

$$\begin{aligned}\theta_n &= 1 - \frac{1}{n}\sum_{i=1}^n a_{i,n}^2 = 1 - \frac{1}{n}\sum_{i=1}^n \left(1-\sigma_i \boldsymbol{\varphi}_n' \boldsymbol{\gamma}_i\right)^2 \\ &= 2\boldsymbol{\varphi}_n'\left(\frac{1}{n}\sum_{i=1}^n \sigma_i \boldsymbol{\gamma}_i\right) - \boldsymbol{\varphi}_n'\left(\frac{1}{n}\sum_{i=1}^n \sigma_i^2 \boldsymbol{\gamma}_i \boldsymbol{\gamma}_i'\right)\boldsymbol{\varphi}_n,\end{aligned} \tag{A.28}$$

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

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

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

$$\|\boldsymbol{\varphi}_n\|_1 \leq inf_i(\sigma_i) \left(\frac{1}{n} \sum_{i=1}^{n} \|\boldsymbol{\gamma}_i\|_1 \right) = inf_i(\sigma_i) \left[\sum_{j=1}^{m_0} \left(\frac{1}{n} \sum_{i=1}^{n} |\gamma_{ij}| \right) \right]. \tag{A.29}$$

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 (4). ■

Proof of Corollary 1. Note that θ_n define by (26) can be written as

$$\theta_n = 2g_n - \boldsymbol{\varphi}_n' \mathbf{H}_n \boldsymbol{\varphi}_n,$$

where

$$g_n = \frac{1}{n} \sum_{i=1}^{n} \sigma_i \boldsymbol{\varphi}_n' \boldsymbol{\gamma}_i, \;\; \mathbf{H}_n = \frac{1}{n} \sum_{i=1}^{n} \sigma_i \left(\boldsymbol{\gamma}_i \boldsymbol{\gamma}_i' \right),$$

$$\boldsymbol{\varphi}_n = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\delta}_i, \;\; \text{and } \boldsymbol{\delta}_i = \frac{\boldsymbol{\gamma}_i}{\sigma_i}.$$

Similarly using (27) we have

$$\hat{\theta}_{nT} = 2\hat{g}_{nT} - \hat{\boldsymbol{\varphi}}_{nT}' \hat{\mathbf{H}}_{nT} \hat{\boldsymbol{\varphi}}_{nT},$$

where

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

$$\hat{\boldsymbol{\varphi}}_{nT} = \frac{1}{n} \sum_{i=1}^{n} \hat{\boldsymbol{\delta}}_{i,nT}, \;\; \text{and } \hat{\boldsymbol{\delta}}_{i,nT} = \frac{\hat{\boldsymbol{\gamma}}_i}{\hat{\sigma}_{i,T}}.$$

Then

$$\sqrt{T}\left(\hat{\theta}_{nT} - \theta_n\right) = 2\sqrt{T}\left(\hat{g}_{nT} - g_n\right) - \sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT}' \hat{\mathbf{H}}_{nT} \hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}_n' \mathbf{H}_n \boldsymbol{\varphi}_n\right). \tag{A.30}$$

Consider the first term of the above

$$\begin{aligned} 2\sqrt{T}\left(\hat{g}_{nT} - g_n\right) &= 2\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}_n\right) \left(\frac{1}{n} \sum_{i=1}^{n} \sigma_i \boldsymbol{\gamma}_i \right) \\ &\quad + 2\sqrt{T} \left[\hat{\boldsymbol{\varphi}}_{nT}' \left(\frac{1}{n} \sum_{i=1}^{n} \hat{\sigma}_{i,T} \hat{\boldsymbol{\gamma}}_i - \frac{1}{n} \sum_{i=1}^{n} \sigma_i \boldsymbol{\gamma}_i \right) \right], \end{aligned} \tag{A.31}$$

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 (S.43) of Lemma S.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 (A.31), we first note that

$$2\sqrt{T}\left[\hat{\boldsymbol{\varphi}}_{nT}'\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\sigma}_{i,T}\hat{\boldsymbol{\gamma}}_i-\frac{1}{n}\sum_{i=1}^{n}\sigma_i\boldsymbol{\gamma}_i\right)\right]=2\sqrt{T}\left[\boldsymbol{\varphi}_n'\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\sigma}_{i,T}\hat{\boldsymbol{\gamma}}_i-\frac{1}{n}\sum_{i=1}^{n}\sigma_i\boldsymbol{\gamma}_i\right)\right]$$
$$+2\sqrt{T}\left[\left(\hat{\boldsymbol{\varphi}}_{nT}-\boldsymbol{\varphi}_n\right)'\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\sigma}_{i,T}\hat{\boldsymbol{\gamma}}_i-\frac{1}{n}\sum_{i=1}^{n}\sigma_i\boldsymbol{\gamma}_i\right)\right]. \tag{A.32}$$

But by (A.29) $\boldsymbol{\varphi}_n=O_p(1)$, and by (S.68) in Lemma S.10 of the online supplement (recall that $\delta_{nT}=min(\sqrt{n},\sqrt{T})$).

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

which also establishes that the second term of (A.32) is $o_p(1)$.Therefore overall we have

$$\sqrt{T}\left(\hat{g}_{nT}-g_n\right)=o_p(1). \tag{A.33}$$

Consider now the second term of (A.30) and note that

$$\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT}'\hat{\mathbf{H}}_{nT}\hat{\boldsymbol{\varphi}}_{nT}-\boldsymbol{\varphi}_n'\mathbf{H}_n\boldsymbol{\varphi}_n\right)=\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT}-\boldsymbol{\varphi}_n\right)'\hat{\mathbf{H}}_{nT}\left(\hat{\boldsymbol{\varphi}}_{nT}-\boldsymbol{\varphi}_n\right)$$
$$+2\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT}-\boldsymbol{\varphi}_n\right)'\hat{\mathbf{H}}_{nT}\boldsymbol{\varphi}_n+\sqrt{T}\boldsymbol{\varphi}_n'\left(\hat{\mathbf{H}}_{nT}-\mathbf{H}_n\right)\boldsymbol{\varphi}_n, \tag{A.34}$$

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

$$\begin{aligned}\sqrt{T}\left(\hat{\mathbf{H}}_{nT}-\mathbf{H}_n\right)&=\frac{\sqrt{T}}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}_i'-\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\right)+\frac{\sqrt{T}}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_{i,T}^2-\sigma_i^2\right)\left(\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}_i'-\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\right)\\&+\frac{\sqrt{T}}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_{i,T}^2-\omega_{i,T}^2\right)\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'+\frac{\sqrt{T}}{n}\sum_{i=1}^{n}\left(\omega_{i,T}^2-\sigma_i^2\right)\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\\&=\sum_{j=1}^{4}\mathbf{D}_{j,nT}.\end{aligned} \tag{A.35}$$

The first two terms of (A.34) 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 (A.34), 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 (S.69) of Lemma S.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 (S.11) of Lemma S.2 in the online supplement 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 $m_0/\sqrt{T}$ and a finite variance.[11] Then by standard law of large numbers, for each (j,j'), $d_{4,nT}(j,j') \rightarrow_p 0$, as n and $T \rightarrow \infty$, 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 (A.33) in (A.30) now yields $\sqrt{T}\left(\hat{\theta}_{nT} - \theta_n\right) = o_p(1)$, as required. ∎

Proof of Theorem 2. Recall from (30) 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 $\hat{\theta}_{nT} = 1 - \frac{1}{n}\sum_{i=1}^{n}\hat{a}_{i,n}^2$, $\hat{a}_{i,n} = 1 - \hat{\sigma}_{i,T}\left(\boldsymbol{\varphi}_{nT}'\hat{\boldsymbol{\gamma}}_i\right)$, and $\hat{\boldsymbol{\varphi}}_{nT} = n^{-1}\sum_{i=1}^{n}\hat{\boldsymbol{\gamma}}_i/\hat{\sigma}_{i,T}$, subject to the normalization $n^{-1}\sum_{i=1}^{n}\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}_i' = \mathbf{I}_{m_0}$. By Lemma S.9 of the online supplement we have $CD = \widetilde{CD} + o_p(1)$. Then $CD^*(\hat{\theta}_{nT})$ can be written as (noting that $1 - \hat{\theta}_{nT} = \frac{1}{n}\sum_{i=1}^{n}\hat{a}_{i,n}^2 > 0$)

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

By result (29) of Corollary 1, $\sqrt{T}\left(\hat{\theta}_{nT} - \theta_n\right) = o_p(1)$, and hence

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

[11]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 (S.47) and (S.48) in Lemma S.7 of the online supplement.

However, by Theorem 1, $CD^{*}(\theta_n) \to_p N(0,1)$, which in turn establishes that $CD^{*}(\hat{\theta}_{nT}) \to_p N(0,1)$, 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

$$\min_{\boldsymbol{\Gamma},\boldsymbol{F}} \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(v_{it} - \boldsymbol{\gamma}_i^{\prime}\mathbf{f}_t\right)^2, \tag{A.36}$$

$$\min_{\boldsymbol{\Gamma},\boldsymbol{F}} \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{v}_{it} - \boldsymbol{\gamma}_i^{\prime}\mathbf{f}_t\right)^2, \tag{A.37}$$

where

$$\hat{v}_{it} = y_{it} - \hat{\boldsymbol{\beta}}_{CCE,i}^{\prime}\mathbf{x}_{it} = y_{it} - \boldsymbol{\beta}_i^{\prime}\mathbf{x}_{it} - \left(\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right)^{\prime}\mathbf{x}_{it} = v_{it} - \left(\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right)^{\prime}\mathbf{x}_{it}. \tag{A.38}$$

We need to show that solving problem (A.37) is asymptotically equivalent to solving problem (A.36). First, using (A.38) the criterion for (A.37) can be written as

$$\begin{aligned}
&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{v}_{it} - \boldsymbol{\gamma}_i^{\prime}\mathbf{f}_t\right)^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(v_{it} - \boldsymbol{\gamma}_i^{\prime}\mathbf{f}_t - \left(\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right)^{\prime}\mathbf{x}_{it}\right)^2 \\
&\equiv \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(v_{it} - \boldsymbol{\gamma}_i^{\prime}\mathbf{f}_t\right)^2 + \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right)^{\prime}\mathbf{x}_{it}\mathbf{x}_{it}^{\prime}\left(\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right) \\
&- 2\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(v_{it} - \boldsymbol{\gamma}_i^{\prime}\mathbf{f}_t\right)\left(\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right)^{\prime}\mathbf{x}_{it} \\
&= A_{nT} + B_{nT} - 2C_{nT}.
\end{aligned} \tag{A.39}$$

Note now that

$$|B_{nT}| = B_{nT} \leq \lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{x}_{it}\mathbf{x}_{it}^{\prime}\right) \times sup_i \left\|\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right\|,$$

where under Assumptions 1-8, we have $\sup_i \lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{x}_{it}\mathbf{x}_{it}^{\prime}\right) < K$, and[12]

$$\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i = O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{n}\right) + O_p\left(\frac{1}{\sqrt{nT}}\right). \tag{A.40}$$

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

Therefore, we also have $B_{nT} = O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{n}\right) + O_p\left(\frac{1}{\sqrt{nT}}\right)$.[13] Consider now the final term of (A.39) and note that

$$C_{nT} = \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{u}_{it}\left(\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right)^{\prime}\mathbf{x}_{it} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right)^{\prime}\frac{1}{T}\sum_{t=1}^{T}\tilde{u}_{it}\mathbf{x}_{it},$$

where $\tilde{u}_{it} = v_{it} - \boldsymbol{\gamma}_i^{\prime}\mathbf{f}_t = y_{it} - \boldsymbol{\beta}_i^{\prime}\mathbf{x}_{it} - \boldsymbol{\gamma}_i^{\prime}\mathbf{f}_t$. 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$\left\| T^{-1}\sum_{t=1}^{T}\tilde{u}_{it}\mathbf{x}_{it}\right\| = O_p\left(T^{-1/2}\right)$, with C_{nT} dominated by B_{nT}. Overall we have

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

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. ■

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, forthcoming.*

[13] Since the result for $\hat{\boldsymbol{\beta}}_{CCE,i}$ in the literature relates to $\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i$, for a given i and not $sup_i\left\|\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right\|$, to to be conservative one can consider using

$$sup_i\left\|\hat{\boldsymbol{\beta}}_{CCE,i} - \boldsymbol{\beta}_i\right\| = O_p\left(\frac{ln(n)}{\sqrt{T}}\right) + O_p\left(\frac{ln(n)}{n}\right) + O_p\left(\frac{ln(n)}{\sqrt{nT}}\right).$$

Our asymptotic results continue to hold even this more conservative rate is used.

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.

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. and Reese, S. (2021). The incidental parameters problem in testing for remaining cross-section correlation. *Journal of Business & Economic Statistics, forthcoming.*

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, Oxford.

Pesaran, M. H. (2021). General diagnostic tests for cross-sectional dependence in panels. *Empirical Economics*, 60:13–50.

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

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 18, 2021

This online supplement provides proofs of the lemmas used in the main paper, and states and establishes a number of auxiliary lemmas used for these proofs.

S1 Proof of 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 S.1 *Suppose that Assumptions 1-3 hold, and the latent factors,* $\mathbf{f}_t$, *and their loadings,* $\boldsymbol{\gamma}_i$, *in model (1) are estimated by principle components,* $\hat{\mathbf{f}}_t$ *and* $\hat{\boldsymbol{\gamma}}_i$, *given by (5). Then*

$$\left\|\hat{\mathbf{F}} - \mathbf{F}\right\|_F = O_p\left(\frac{\sqrt{T}}{\delta_{nT}}\right), \tag{S.1}$$

$$\left\|\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma}\right\|_F = O_p\left(\frac{\sqrt{n}}{\delta_{nT}}\right), \tag{S.2}$$

$$\left\|\mathbf{U}^{'}(\hat{\mathbf{F}} - \mathbf{F})\right\|_F = O_p\left(\frac{\sqrt{nT}}{\delta_{nT}}\right), \tag{S.3}$$

$$\left\|\boldsymbol{\Gamma}'(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma})\right\| = O_p\left(\frac{n}{\delta_{nT}}\right), \tag{S.4}$$

$$\left\|\mathbf{F}'(\hat{\mathbf{F}}-\mathbf{F})\right\| = O_p\left(\frac{T}{\delta_{nT}}\right), \tag{S.5}$$

$$\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{F} = O_p\left(\frac{T}{\delta_{nT}^2}\right), \tag{S.6}$$

$$\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\hat{\mathbf{F}} = O_p\left(\frac{T}{\delta_{nT}^2}\right), \tag{S.7}$$

$$\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{u}_i = O_p\left(\frac{T}{\delta_{nT}^2}\right), \tag{S.8}$$

$$(\hat{\mathbf{\Gamma}}-\mathbf{\Gamma})'\mathbf{u}_t = O_p\left(\frac{n}{\delta_{nT}^2}\right), \tag{S.9}$$

where $\mathbf{u}_i = (u_{i1}, u_{i2}, ..., u_{iT})'$, $\mathbf{u}_t = (u_{1t}, u_{2t}, ..., u_{nt})'$, *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 (S.1) to (S.4), (S.6) and (S.8) follow directly from Lemmas B1, B2 and B3, and Theorem 2 of Bai (2003). The remaining two results, (S.5) and (S.9), can be established analogously. ■

Lemma S.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, (1), where* $\mathbf{e}_i = (e_{i1}, e_{i2}, ..., e_{iT})'$ *is the principle component estimator of* $\mathbf{u}_i = (u_{i1}, u_{i2}, ..., u_{iT})'$, *namely* $\mathbf{e}_i = \mathbf{M}_{\hat{\mathbf{F}}}\mathbf{y}_i$, *where* $\mathbf{M}_{\hat{\mathbf{F}}} = \mathbf{I}_T - \hat{\mathbf{F}}(\hat{\mathbf{F}}'\hat{\mathbf{F}})^{-1}\hat{\mathbf{F}}'$, $\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{iT}\ )'$, *and* $\hat{\mathbf{F}}$ *is given by (5). Suppose that Assumptions 1-3 hold. Then*

$$\hat{\sigma}_{i,T}^2 - \omega_{i,T}^2 = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{S.10}$$

$$\frac{1}{n}\sum_{i=1}^{n} b_{in}\left(\hat{\sigma}_{i,T}^2 - \omega_{i,T}^2\right) = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{S.11}$$

$$\hat{\sigma}_{i,T} - \omega_{i,T} = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{S.12}$$

$$\frac{1}{n}\sum_{i=1}^{n} b_{in}\left(\hat{\sigma}_{i,T} - \omega_{i,T}\right) = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{S.13}$$

$$\frac{1}{\hat{\sigma}_{i,T}} - \frac{1}{\omega_{i,T}} = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{S.14}$$

$$\frac{1}{n}\sum_{i=1}^{n} b_{in}\left(\frac{1}{\hat{\sigma}_{i,T}} - \frac{1}{\omega_{i,T}}\right) = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{S.15}$$

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

Proof. We first note that

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

which yields the following error variance decomposition

$$\begin{aligned}\hat{\sigma}^2_{i,T} &= \frac{\mathbf{u}_i^{'}\mathbf{M}_{\mathbf{F}}\mathbf{u}_i}{T} + \frac{\boldsymbol{\gamma}_i^{'}\mathbf{F}^{'}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{F}\boldsymbol{\gamma}_i}{T} + \frac{\mathbf{u}_i^{'}\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{u}_i}{T} \\ &+ \frac{2\mathbf{u}_i^{'}\mathbf{M}_{\mathbf{F}}\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{u}_i}{T} + \frac{2\mathbf{u}_i^{'}\mathbf{M}_{\mathbf{F}}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{F}\boldsymbol{\gamma}_i}{T} + \frac{2\mathbf{u}_i^{'}\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{F}\boldsymbol{\gamma}_i}{T} \\ &= \sum_{j=1}^{6} B_{j,iT}. \end{aligned} \tag{S.16}$$

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 2 and $\|\mathbf{M}_{\hat{\mathbf{F}}}\| = 1$. Then using (S.1) it follows that $\|B_{2,iT}\| = O_p\left(\delta_{nT}^{-2}\right)$. Before establishing the probability order of the remaining terms $B_{3,iT}$ we first observe that

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

and

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

Then using results (S.1) and (S.5) it follows that

$$\frac{\hat{\mathbf{F}}^{'}\hat{\mathbf{F}}}{T} = \frac{\mathbf{F}^{'}\mathbf{F}}{T} + O_p\left(\frac{1}{\delta_{nT}^2}\right), \text{ and } \left(\frac{\hat{\mathbf{F}}^{'}\hat{\mathbf{F}}}{T}\right)^{-1} = \left(\frac{\mathbf{F}^{'}\mathbf{F}}{T}\right)^{-1} + O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{S.17}$$

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

$$\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T} = O_p(1), \text{ and } \left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1} = O_p(1). \tag{S.18}$$

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

$$\begin{aligned} B_{3,iT} &= T^{-1}\mathbf{u}_i'\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{u}_i \\ &= \frac{1}{T}\mathbf{u}_i'\left[\hat{\mathbf{F}}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}' - \mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}'\right]\mathbf{u}_i. \end{aligned} \tag{S.19}$$

and

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

which allows us to write $B_{3,iT}$ as $B_{3,iT} = \sum_{s=1}^{4} C_{s,iT}$, with the first term satisfying

$$\begin{aligned} \left\|C_{1,iT}\right\| &= \left\|\frac{1}{T}\mathbf{u}_i'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\left(\hat{\mathbf{F}} - \mathbf{F}\right)'\mathbf{u}_i\right\| \\ &\leq \frac{\left\|\mathbf{u}_i\right\|^2}{T}\left\|\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\right\|\left(\frac{\left\|\hat{\mathbf{F}} - \mathbf{F}\right\|^2}{T}\right). \end{aligned}$$

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 (S.18). Using (S.1) 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

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

and by taking spectral norms

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

Under Assumption 3, $T^{-1/2}\mathbf{u}_i'\mathbf{F} = O_p(1)$, and using results (S.1) and (S.5) it follows that $\|C_{2,iT}\| = O_p\left(\frac{1}{\delta_{nT}^2}\right)$. Finally, by result (S.8) $\frac{\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{u}_i}{\sqrt{T}} = O_p\left(\frac{\sqrt{T}}{\delta_{nT}^2}\right)$, and it follows that

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

and overall we $B_{3,iT} = O_p\left(\frac{1}{\delta_{nT}^2}\right)$. Consider now the fourth term of (S.16),

$$\begin{aligned} B_{4,iT} &= \frac{\mathbf{u}_i'\left(\mathbf{I}_m - \mathbf{P}_{\mathbf{F}}\right)\left(\mathbf{P}_{\mathbf{F}} - \mathbf{P}_{\hat{\mathbf{F}}}\right)\mathbf{u}_i}{T} = \frac{\mathbf{u}_i'\left(\mathbf{P}_{\mathbf{F}} - \mathbf{P}_{\hat{\mathbf{F}}} + \mathbf{P}_{\mathbf{F}}\mathbf{P}_{\hat{\mathbf{F}}}\right)\mathbf{u}_i}{T} \\ &= \frac{\mathbf{u}_i'\left(\mathbf{P}_{\mathbf{F}} - \mathbf{P}_{\hat{\mathbf{F}}}\right)\mathbf{u}_i}{T} - \frac{\mathbf{u}_i'\mathbf{P}_{\mathbf{F}}\mathbf{u}_i}{T} + \frac{\mathbf{u}_i'\mathbf{P}_{\mathbf{F}}\mathbf{P}_{\hat{\mathbf{F}}}\mathbf{u}_i}{T}, \end{aligned} \tag{S.20}$$

where $\mathbf{P}_{\mathbf{F}} = \mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}'$, and $\mathbf{P}_{\hat{\mathbf{F}}} = \hat{\mathbf{F}}\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)\hat{\mathbf{F}}'$. The order of the first term of (S.20) is the same as that of (S.19), namely $O_p\left(\delta_{nT}^{-2}\right)$. Since $\mathbf{F}$ is distributed independently of $\mathbf{u}_i$, using (S.18), then it readily follows that the second term is $O_p\left(T^{-1}\right)$. The third term of (S.20) can be written as

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

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

$$
\begin{aligned}
\frac{\boldsymbol{\gamma}_i'\mathbf{F}'\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{M}_{\mathbf{F}}\mathbf{u}_i}{T}&=-\frac{\boldsymbol{\gamma}_i'\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{u}_i}{T}-\frac{\boldsymbol{\gamma}_i'\mathbf{F}'\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}\mathbf{F}'\mathbf{u}_i}{T} \\
&=-\frac{\boldsymbol{\gamma}_i'\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{u}_i}{T}+\frac{\boldsymbol{\gamma}_i'\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\hat{\mathbf{F}}}{T}\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\frac{\hat{\mathbf{F}}'\mathbf{u}_i}{T} \\
&\quad+\boldsymbol{\gamma}_i'\frac{\mathbf{F}'\mathbf{M}_{\hat{\mathbf{F}}}\left(\hat{\mathbf{F}}-\mathbf{F}\right)}{T}\left(\frac{\mathbf{F}'\mathbf{F}}{T}\right)^{-1}\frac{\mathbf{F}'\mathbf{u}_i}{T}. \qquad \text{(S.21)}
\end{aligned}
$$

But, using results in Lemma S.1, we have

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

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

and

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

$$= O_p\left(\frac{1}{\delta_{nT}^2\sqrt{T}}\right).$$

Thus, $B_{5,iT} = T^{-1}\boldsymbol{\gamma}_i'\mathbf{F}'\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{M}_{\mathbf{F}}\mathbf{u}_i = O_p\left(\delta_{nT}^{-2}\right)$. Similarly $B_{6,iT} = O_p\left(\delta_{nT}^{-2}\right)$. The above results now establish (S.10). Result (S.11) can be obtained similarly, either by directly considering the weighted average of (S.10) 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 (S.12) and (S.14) follow from (S.10), noting that under Assumption 3 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 (S.12) 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 (S.10) we have $|\hat{\sigma}_{i,T} - \omega_{i,T}| = O_p\left(\delta_{nT}^{-2}\right)$. Similarly, result (S.14) is established noting that

$$\left|\frac{1}{\hat{\sigma}_{i,T}} - \frac{1}{\omega_{i,T}}\right| \leq \frac{|\hat{\sigma}_{i,T} - \omega_{i,T}|}{\omega_{i,T}\hat{\sigma}_{i,T}} \leq c^{-1}\left|\hat{\sigma}_{i,T} - \omega_{i,T}\right| = O_p\left(\frac{1}{\delta_{nT}^2}\right). \tag{S.22}$$

Finally, results (S.13) and (S.15) can be obtained, respectively, in a similar way to the proof of (S.11), since under Assumption 3, $\hat{\sigma}_{i,T}$, $\omega_{i,T}$, $\hat{\sigma}_{i,T}^{-1}$ and $\omega_{i,T}^{-1}$ are also integrable processes. ■

Lemma S.3 *Suppose that the latent factors, $\mathbf{f}_t$, and their loadings, $\boldsymbol{\gamma}_i$, in model (1) are estimated by principle components, $\hat{\mathbf{F}}$ and $\hat{\boldsymbol{\gamma}}_i$, given by (5). Then under Assumptions 1-3 with n* *and $T \to \infty$, such that $n/T \to \kappa$, for $0 < \kappa < \infty$, we have*

$$\mathbf{d}_{1,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} b_{in} (\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) = O_p\left(\frac{1}{\delta_{nT}}\right), \tag{S.23}$$

$$\mathbf{d}_{2,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\omega_{i,T} - \sigma_i) (\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) = O_p\left(\frac{1}{\delta_{nT}}\right), \tag{S.24}$$

$$\mathbf{d}_{3,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{1}{\omega_{iT}} - \frac{1}{\sigma_i}\right) (\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) = O_p\left(\frac{1}{\delta_{nT}}\right), \tag{S.25}$$

$$\mathbf{d}_{4,nT} = \frac{1}{n} \sum_{i=1}^{n} (\hat{\sigma}_{i,T} - \omega_{i,T}) (\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{S.26}$$

$$\mathbf{d}_{5,nT} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{\hat{\sigma}_{i,T}} - \frac{1}{\omega_{i,T}}\right) (\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) = O_p\left(\frac{1}{\delta_{nT}^2}\right) \tag{S.27}$$

$$\mathbf{d}_{6,nT} = \frac{1}{n} \sum_{i=1}^{n} (\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}) = O_p\left(\frac{1}{\delta_{nT}}\right), \tag{S.28}$$

$$\mathbf{d}_{7,nT} = \frac{1}{n} \sum_{i=1}^{n} b_{in} (\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) \boldsymbol{\gamma}_i' = O_p\left(\frac{1}{\delta_{nT}^2}\right), \tag{S.29}$$

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

$$\begin{aligned}
\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i &= \left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1} \left(\frac{\hat{\mathbf{F}}'\mathbf{F}\boldsymbol{\gamma}_i}{T} + \frac{\hat{\mathbf{F}}'\mathbf{u}_i}{T}\right) - \left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1} \left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right) \boldsymbol{\gamma}_i \\
&= -\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1} \left[\frac{\hat{\mathbf{F}}'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\boldsymbol{\gamma}_i}{T}\right] + \left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1} \left(\frac{\hat{\mathbf{F}}'\mathbf{u}_i}{T}\right),
\end{aligned} \tag{S.30}$$

and we have

$$\begin{aligned}
\mathbf{d}_{1,nT} &= n^{-1/2} \sum_{i=1}^{n} b_{in} (\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) \\
&= -\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1} \left[\frac{\sqrt{n}}{T}\hat{\mathbf{F}}'\left(\hat{\mathbf{F}} - \mathbf{F}\right)\right] \left(\frac{1}{n}\sum_{i=1}^{n} b_{in}\boldsymbol{\gamma}_i\right) + \left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1} T^{-1} \left(n^{-1/2}\sum_{i=1}^{n} b_{in}\hat{\mathbf{F}}'\mathbf{u}_i\right).
\end{aligned} \tag{S.31}$$

By result (S.18) $\left(T^{-1}\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1} = O_p(1)$, and by result (S.5) we have

$$\frac{\sqrt{n}}{T}\hat{\mathbf{F}}'\left(\hat{\mathbf{F}}-\mathbf{F}\right) = O_p\left(\frac{\sqrt{n}T}{T\delta_{nT}^2}\right) = O_p\left(\frac{\sqrt{n}}{\min(n,T)}\right) = O_p\left(\frac{1}{\delta_{nT}}\right). \tag{S.32}$$

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

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

Hence, the first term of (S.31) is $O_p\left(\delta_{nT}^{-1}\right)$. For the second term of (S.31), 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 3, we have

$$\frac{1}{T\sqrt{n}}\sum_{i=1}^{n} b_{in}\mathbf{F}^{'}\mathbf{u}_i = \left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T} b_{in}\mathbf{f}_t u_{it}\right)\frac{1}{\sqrt{T}} = O_p\left(\frac{1}{\sqrt{T}}\right). \tag{S.33}$$

Result (S.23) now follows using (S.32) and (S.33) in (S.31), and noting that by assumption n and T are of the same order. To prove (S.24) 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 q_{iT} are independently distributed across i. Using results in Lemma S.7 it is easily seen that $E\left(q_{iT}\right) = O_p(T^{-1/2})$ 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 (S.2) $n^{-1/2}\left\|\hat{\boldsymbol{\Gamma}}-\boldsymbol{\Gamma}\right\| = O_p(\delta_{nT}^{-1})$, and (S.24) is established. Result (S.25) follows similarly, with q_{iT} defined as $q_{iT} = \sigma_i^{-1}\sqrt{T}\left[\left(\frac{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i}{T}\right)^{-1/2} - 1\right]$. Note that $sup_i(1/\sigma_i^2) < K$, and using results in Lemma S.7 it is again easily established that $E\left(q_{iT}'\right) = O(T^{-1/2})$, and $Var\left(q_{iT}'\right) = O_p(1)$. Result (S.26) is also obtained using Cauchy-Schwarz inequality, namely[S1]

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

where by (S.2) $n^{-1/2}\left\|\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}\right\| = O_p\left(\delta_{nT}^{-1}\right)$, and by (S.12) $n^{-1}\sum_{i=1}^{n}\left(\hat{\sigma}_{i,T} - \omega_{i,T}\right)^2 = O_p(\delta_{nT}^{-2})$. Similarly by (S.2) and (S.14) 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 (S.28), 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 (S.23), and becomes identical to it if we replace a_i in (S.23) 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 $1 - \left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{-1} = O_p(T^{-1/2})$. Therefore (S.28) is established as required. Finally, consider result (S.29) and note

[S1] The proofs of (S.24) and (S.25) are different from that of (S.26) and (S.27) due to the fact that $\omega_{iT}^2 - \sigma_i^2 = O_p\left(T^{-1/2}\right)$, but $\hat{\sigma}_{iT}^2 - \omega_{iT}^2 = O_p\left(n^{-1}\right) + O_p\left(T^{-1}\right)$.

that

$$n^{-1/2}\sum_{i=1}^{n} b_{in}\left(\hat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i' = n^{-1/2}\sum_{i=1}^{n} b_{in}\left[-\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}'\left(\hat{\mathbf{F}}-\mathbf{F}\right)\boldsymbol{\gamma}_i+\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}'\mathbf{u}_i\right]\boldsymbol{\gamma}_i'$$

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

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

$$+\sqrt{n}\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\left(n^{-1}\sum_{i=1}^{n} T^{-1}b_{in}\left(\hat{\mathbf{F}}-\mathbf{F}\right)'\mathbf{u}_i\boldsymbol{\gamma}_i'\right)+\left(\frac{\hat{\mathbf{F}}'\hat{\mathbf{F}}}{T}\right)^{-1}\left(n^{-1/2}\sum_{i=1}^{n} T^{-1}b_{in}\mathbf{F}'\mathbf{u}_i\boldsymbol{\gamma}_i'\right). \tag{S.34}$$

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 (S.7) 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 (S.8)

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

Finally, the last term of (S.34) 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 (S.29) is established, as required. ■

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

$$p_{nT} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} s_{t,nT}^2 = o_p(1), \tag{S.35}$$

$$q_{nT} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} \psi_{t,nT}s_{t,nT} = o_p(1), \tag{S.36}$$

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

Proof. Using (A.7), recall that

$$s_{t,nT} = \boldsymbol{\varphi}_{nT}' \left[n^{-1/2} \sum_{i=1}^{n} (\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) u_{it} \right] + \boldsymbol{\varphi}_{nT}' \left[n^{-1/2} \sum_{i=1}^{n} (\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) \boldsymbol{\gamma}_i' \right] \mathbf{f}_t + \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T} \right)' \right] \mathbf{f}_t \tag{S.37}$$

$$+ \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T} \right)' \right] \left(\hat{\mathbf{f}}_t - \mathbf{f}_t \right).$$

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

$$\psi_{t,nT} = \xi_{t,n} - (\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n)' \kappa_{t,n} + \upsilon_{t,nT} \tag{S.38}$$

where

$$\xi_{t,n} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} a_{i,n} \varepsilon_{it}, \;\; a_{i,n} = 1 - \sigma_i \boldsymbol{\varphi}_n' \boldsymbol{\gamma}_i, \tag{S.39}$$

$$\kappa_{t,n} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{\gamma}_i \sigma_i \varepsilon_{it}, \tag{S.40}$$

$$\upsilon_{t,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{1}{(\boldsymbol{\varepsilon}_i' \mathbf{M}_F \boldsymbol{\varepsilon}_i / T)^{1/2}} - 1 \right) \varepsilon_{it}. \tag{S.41}$$

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

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

where $\mathbf{V}_T = T^{-1} \sum_{t=1}^{T} \mathbf{u}_t \mathbf{u}_t'$. But by Assumption 3,$\|\mathbf{V}_T\| = \lambda_{max}(\mathbf{V}_T) = O_p(1)$, and using (S.2) $\left\| \left(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma} \right) \right\| = O_p\left(\frac{\sqrt{n}}{\delta_{nT}} \right)$. 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

$$A_{2,nT} = \sqrt{T} \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T} \right)' \right] \left(T^{-1} \mathbf{F}' \mathbf{F} \right) \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T} \right) \right].$$

By assumption $T^{-1}\mathbf{F}'\mathbf{F} = O_p(1)$, and using (S.28) $n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right) = O_p\left(\delta_{nT}^{-1}\right)$. Hence, $A_{2,nT} = O_p\left(\sqrt{T}\delta_{nT}^{-2}\right) = o_p(1)$. Similarly,

$$\begin{aligned}
A_{3,nT} &= \sqrt{T}\boldsymbol{\varphi}_{nT}'\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i'\right]\left(T^{-1}\sum_{t=1}^{T}\mathbf{f}_t\mathbf{f}_t'\right)\left[n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\left(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_i\right)'\right]\boldsymbol{\varphi}_{nT} \\
&= \sqrt{T}\boldsymbol{\varphi}_{nT}'\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i'\right]\left(T^{-1}\mathbf{F}'\mathbf{F}\right)\left[n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\left(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_i\right)'\right]\boldsymbol{\varphi}_{nT},
\end{aligned}$$

where $\|\boldsymbol{\varphi}_{nT}\|$ is bounded, and by (S.29) $n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i' = O_p\left(\delta_{nT}^{-1}\right)$. Hence

$$\begin{aligned}
A_{3,nT} &= \sqrt{T}\boldsymbol{\varphi}_{nT}'O_p\left(\frac{1}{\min(\sqrt{n},\sqrt{T})}\right)\left(T^{-1}\mathbf{F}'\mathbf{F}\right)O_p\left(\frac{1}{\min(\sqrt{n},\sqrt{T})}\right)\boldsymbol{\varphi}_{nT} \\
&= O_p\left(\frac{\sqrt{T}}{\min\left(n,T\right)}\right) = o_p\left(1\right).
\end{aligned}$$

Next

$$A_{4,nT} = \sqrt{T}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right)\right]'\left(\frac{\left\|\hat{\mathbf{F}} - \mathbf{F}\right\|^2}{T}\right)\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right)\right],$$

where by (S.1) $T^{-1}\left\|\mathbf{F} - \hat{\mathbf{F}}\right\|_F^2 = O_p\left(\delta_{nT}^{-2}\right)$, and by (S.28) $n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right) = O_p(1)$. Hence, $A_{4,nT} = \sqrt{T}O_p\left(\delta_{nT}^{-2}\right) = o_p(1)$. Consider now the cross product terms of p_{nT}, starting with

$$\begin{aligned}
A_{5,nT} &= 2\boldsymbol{\varphi}_{nT}'\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)u_{it}\right]\mathbf{f}_t'\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right)\right] \\
&= 2\boldsymbol{\varphi}_{nT}'\left(\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma})'\mathbf{u}_t\mathbf{f}_t'\right)\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right)\right],
\end{aligned}$$

where $\boldsymbol{\varphi}_{nT}$ and $n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T} - \boldsymbol{\delta}_{i,T}\right)$ are bounded in probability and by (S.9) $n^{-1/2}(\hat{\boldsymbol{\Gamma}} - \boldsymbol{\Gamma})'\mathbf{u}_t = O_p\left(\frac{\sqrt{n}}{\delta_{nT}^2}\right) = o_p(1)$, and we have $A_{5,nT} = o_p(1)$, as well (since $\mathbf{u}_t$ and $\mathbf{f}_t$ are distributed independently). Similarly we have

$$
\begin{aligned}
A_{6,nT} &= \frac{2}{\sqrt{T}} \sum_{t=1}^{T} \boldsymbol{\varphi}_{nT}' \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i \right) u_{it} \right] \mathbf{f}_t' \left[n^{-1/2} \sum_{i=1}^{n} \boldsymbol{\gamma}_i \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i \right)' \right] \boldsymbol{\varphi}_{nT} \\
&= \boldsymbol{\varphi}_{nT}' \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i \right) \left(\frac{2}{\sqrt{T}} \sum_{t=1}^{T} u_{it} \mathbf{f}_t' \right) \right] \left[n^{-1/2} \sum_{i=1}^{n} \boldsymbol{\gamma}_i \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i \right)' \right] \boldsymbol{\varphi}_{nT} \\
&= O_p \left(\frac{1}{\min\left(n, T\right)} \right) = o_p\left(1\right),
\end{aligned}
$$

and

$$
\begin{aligned}
A_{7,nT} &= \frac{2}{\sqrt{T}} \sum_{t=1}^{T} \boldsymbol{\varphi}_{nT}' \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i \right) u_{it} \right] \left(\hat{\mathbf{f}}_t - \mathbf{f}_t \right)' \left[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\delta}}_i - \boldsymbol{\delta}_i \right) \right] \\
&= \boldsymbol{\varphi}_{nT}' \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i \right) \left(\frac{2}{\sqrt{T}} \sum_{t=1}^{T} u_{it} \left(\hat{\mathbf{f}}_t - \mathbf{f}_t \right)' \right) \right] \left[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\delta}}_i - \boldsymbol{\delta}_i \right) \right] \\
&= O_p \left(\frac{1}{\sqrt{T}} \right) O_p \left(\frac{1}{\min(n, T)} \right) = o_p\left(1\right).
\end{aligned}
$$

Also

$$
\begin{aligned}
A_{8,nT} &= 2\sqrt{T} \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\delta}}_i - \boldsymbol{\delta}_i \right)' \right] \left(T^{-1} \sum_{t=1}^{T} \mathbf{f}_t \mathbf{f}_t' \right) \left[n^{-1/2} \sum_{i=1}^{n} \boldsymbol{\gamma}_i \left(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_i \right)' \right] \boldsymbol{\varphi}_{nT} \\
&= 2\sqrt{T} O_p \left(\frac{1}{\min(\sqrt{n}, \sqrt{T})} \right) \left(\frac{\mathbf{F}'\mathbf{F}}{T} \right) O_p \left(\frac{1}{\min(\sqrt{n}, \sqrt{T})} \right) \\
&= O_p \left(\frac{\sqrt{T}}{\min(n, T)} \right) = o_p(1),
\end{aligned}
$$

$$
\begin{aligned}
A_{9,nT} &= 2\sqrt{T} \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\delta}}_i - \boldsymbol{\delta}_i \right)' \right] \left(T^{-1} \sum_{t=1}^{T} \mathbf{f}_t \left(\mathbf{f}_t - \hat{\mathbf{f}}_t \right)' \right) \left[n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\delta}}_i - \boldsymbol{\delta}_i \right) \right] \\
&= 2\sqrt{T} O_p \left(\frac{1}{\min(\sqrt{n}, \sqrt{T})} \right) \left(\frac{\mathbf{F}' \left(\hat{\mathbf{F}} - \mathbf{F} \right)}{T} \right) O_p \left(\frac{1}{\min(\sqrt{n}, \sqrt{T})} \right) \\
&= O_p \left(\frac{\sqrt{T}}{\min(n^2, T^2)} \right) = o_p(1),
\end{aligned}
$$

and

$$
\begin{aligned}
A_{10,nT} &= 2\sqrt{T}\boldsymbol{\varphi}_{nT}'\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i'\right]\left(T^{-1}\sum_{t=1}^{T}\mathbf{f}_t\left(\mathbf{f}_t-\hat{\mathbf{f}}_t\right)'\right)\left[n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\delta}}_{i,T}-\boldsymbol{\delta}_{i,T}\right)\right] \\
&= 2\sqrt{T}\boldsymbol{\varphi}_{nT}'O_p\left(\frac{1}{\min(\sqrt{n},\sqrt{T})}\right)\left(\frac{\mathbf{F}'\left(\hat{\mathbf{F}}-\mathbf{F}\right)}{T}\right)O_p\left(\frac{1}{\min(\sqrt{n},\sqrt{T})}\right) \\
&= O_p\left(\frac{\sqrt{T}}{\min(n^2,T^2)}\right)=o_p(1).
\end{aligned}
$$

Overall, we have $p_{nT}=o_p(1)$, as required. Consider now q_{nT} and note that it can be written as (using (S.38) in (S.36))

$$
q_{nT}=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}s_{t,nT}\xi_{t,n}-\left(\boldsymbol{\varphi}_{nT}-\boldsymbol{\varphi}_n\right)'\frac{1}{\sqrt{T}}\sum_{t=1}^{T}s_{t,nT}\kappa_{t,n}+\frac{1}{\sqrt{T}}\sum_{t=1}^{T}s_{t,nT}\upsilon_{t,nT},
$$

where $\xi_{t,n}$, $\upsilon_{t,nT}$ and $\kappa_{t,n}$ are given by (S.39), (S.40) and (S.41), respectively. Consider the first term of the above and using (S.39) write it as

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

Using (S.39), $B_{1,nT}$ 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 (S.23) 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 (S.39), $B_{2,nT}$ 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 (S.29) $n^{-1/2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i' = O_p(\delta_{nT}^{-1})$. Also, under Assumption 3 $\frac{1}{\sqrt{nT}}\sum_{j=1}^{n}\sum_{t=1}^{T} a_{j,n}\mathbf{f}_t\varepsilon_{jt} = O_p(1)$. Then it follows that $B_{2,nT} = o_p(1)$. Similarly, it is established that $B_{3,nT} = o_p(1)$, noting that by (S.28) 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 (S.42) $\boldsymbol{\varphi}_{nT} - \boldsymbol{\varphi}_n = O_p(n^{-1/2}T^{-1/2}) + O_p(T^{-1})$. 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 (S.41) we have

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

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

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

and

$$\frac{1}{T} \sum_{t=1}^{T} \sigma_i \left(\frac{1}{\left(\boldsymbol{\varepsilon}_j' \mathbf{M}_F \boldsymbol{\varepsilon}_j / T\right)^{1/2}} - 1 \right) \varepsilon_{it} \varepsilon_{jt}$$
$$\to_p lim_T \frac{1}{T} \sum_{t=1}^{T} \sigma_i \left\{ E\left[\varepsilon_{it}^2 \left(\frac{\boldsymbol{\varepsilon}_i' \mathbf{M}_{\mathbf{F}} \boldsymbol{\varepsilon}_i}{T} \right)^{-1/2} \right] - 1 \right\}, \text{ if } i = j$$

Also by (S.55), $E\left[\varepsilon_{it}^2 \left(\frac{\boldsymbol{\varepsilon}_i' \mathbf{M}_{\mathbf{F}} \boldsymbol{\varepsilon}_i}{T} \right)^{-1/2} \right] = 1 + O\left(\frac{1}{T}\right)$, n and T being of the same order, and by (S.23) $n^{-1/2} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right) = O_p(\delta_{nT}^{-1})$. It then follows that $C_{1,nT} = o_p(1)$. Similarly to $B_{2,nT}$, we have

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

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

$$\sqrt{T} \left(\boldsymbol{\varphi}_n - \boldsymbol{\varphi}_{nT} \right) = O_p\left(n^{-1/2}\right) + O_p\left(T^{-1/2}\right) = o_p(1), \tag{S.42}$$

$$\sqrt{T} \left(\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}_n \right) = o_p(1) \tag{S.43}$$

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

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

where

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

Since $\sigma_i/\omega_{i,T}=\left(T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i\right)^{-1/2}$, $\|\boldsymbol{\gamma}_i\|<K$, then using result (S.50) in Lemma S.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 $\mathbf{d}_{1,nT}=n^{-1}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\chi_{i,T}$, where $\chi_{i,T}=-\sqrt{T}\left[\sigma_i/\omega_{i,T}-E\left(\sigma_i/\omega_{i,T}\right)\right]$. It is clear that $\chi_{i,T}$ are distributed independently over i with mean zero and bounded variances:[S2]

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

Hence, $\mathbf{d}_{1,nT}=O_p\left(n^{-1/2}\right)$, and the desired result (S.42) follows. Consider now (S.43) and note that it can be decomposed as

$$\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT}-\boldsymbol{\varphi}_n\right)=\sqrt{T}\left(\boldsymbol{\varphi}_{nT}-\boldsymbol{\varphi}_n\right)+\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{nT}-\boldsymbol{\varphi}_{nT}\right), \tag{S.44}$$

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

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

Now using (S.15) of Lemma S.2 we have

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

[S2] When ε_{it} are normally distributed we have the exact result $E\left(\frac{T}{\boldsymbol{\varepsilon}_i'\mathbf{M}_F\boldsymbol{\varepsilon}_i}\right)=T/(T-m-2)$.

Also by Cauchy–Schwarz inequality using (S.2) and (S.12) we have

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

Using the above results, we have $\sqrt{T}\,(\hat{\boldsymbol{\varphi}}_{nT} - \boldsymbol{\varphi}_{nT}) = o_p(1)$, which in turn establishes (S.43), as required. ■

Lemma S.6 *Suppose that $\boldsymbol{\varepsilon} \sim IID(\mathbf{0}, \mathbf{I}_T)$, where $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_T)'$, $\kappa_1 = E\left(\varepsilon_t^3\right)$, $\kappa_2 = E\left(\varepsilon_t^4\right) - 3$, 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*

$$E\left(\boldsymbol{\varepsilon}'\mathbf{A}\boldsymbol{\varepsilon}\right) = tr(\mathbf{A}) \tag{S.45}$$

$$E\left[\left(\boldsymbol{\varepsilon}'\mathbf{A}\boldsymbol{\varepsilon}\right)\left(\boldsymbol{\varepsilon}'\mathbf{A}\boldsymbol{\varepsilon}\right)\right] = \kappa_2 Tr\left[\left(\mathbf{A} \odot \mathbf{B}\right)\right] + Tr\left(\mathbf{A}\right) Tr\left(\mathbf{B}\right) + 2Tr\left(\mathbf{AB}\right) \tag{S.46}$$

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 S.7 *Suppose that $\boldsymbol{\varepsilon}\,|\mathbf{F} \sim IID(\mathbf{0}, \mathbf{I}_T)$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_T)'$, $sup_t E\left(\left|\varepsilon_t\right|^{4+\epsilon}\right)$, for some small $\epsilon > 0$ and let $\mathbf{M}_F = \mathbf{I}_T - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$, where $\mathbf{F}$ is $T \times m$ matrix such that $\mathbf{F}'\mathbf{F}$ is non-singular. Also let $z_T = \left(T^{-1}\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}\right)^{1/2} - 1$. Then there exists T_0 such that for all $T > T_0$ we* *have*

$$E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right) = 1 - \frac{m}{T} = 1 + O\left(\frac{1}{T}\right) \tag{S.47}$$

$$E\left[\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^2\right] = 1 + O\left(\frac{1}{T}\right) \tag{S.48}$$

$$E\left[\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{1/2}\right] = 1 + O\left(\frac{1}{T}\right) \tag{S.49}$$

$$E\left[\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{-1/2}\right] = 1 + O\left(\frac{1}{T}\right) \tag{S.50}$$

$$Var\left[\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{1/2}\right] = O\left(\frac{1}{T}\right) \tag{S.51}$$

$$E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)\right] = 1 + O\left(\frac{1}{T}\right) \tag{S.52}$$

$$E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{1/2}\right] = 1 + O\left(\frac{1}{T}\right) \tag{S.53}$$

$$E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{-1}\right] = 1 + O\left(\frac{1}{T}\right) \tag{S.54}$$

$$E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{-1/2}\right] = 1 + O\left(\frac{1}{T}\right) \tag{S.55}$$

$$E(z_T) = O\left(\frac{1}{T}\right), \quad \textit{and } E(z_T^2) = O\left(\frac{1}{T}\right) \tag{S.56}$$

$$z_T = O_p\left(T^{-1/2}\right) \tag{S.57}$$

Proof. Result (S.47) follows immediately from (S.45) in Lemma S.6, noting that $tr(\mathbf{M}_F) = T - m$. Result (S.48) follows using (S.46) in Lemma S.6, by setting $\mathbf{A} = \mathbf{B} = \mathbf{M}_F$, and noting that $tr\left[(\mathbf{M}_F \odot \mathbf{M}_F)\right] = \sum_{t=1}^{T} m^2 = O(T)$, $tr\left(\mathbf{M}_F^2\right) = tr\left(\mathbf{M}_F\right) = T - m$. To establish (S.49) note that since $\sqrt{x}$ is a concave function of x, then by Jensen inequality we have

$$E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{T}\right)^{1/2} \leq \left[E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}{T}\right)\right]^{1/2} = \left[1 + O\left(\frac{1}{T}\right)\right]^{1/2} = 1 + O\left(\frac{1}{T}\right).$$

Similarly, by Jensen inequality

$$E\left[\left(\frac{T}{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}\right)^{1/2}\right] \leq \left[E\left(\frac{T}{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}\right)\right]^{1/2}. \tag{S.58}$$

But using a result due to Lieberman (1994) (see Lemmas 5 and 21 of Pesaran and Yamagata

(2017)) we have

$$E\left(\frac{T}{\boldsymbol{\varepsilon}'\mathbf{M}_F\boldsymbol{\varepsilon}}\right) = 1 + O\left(\frac{1}{T}\right).$$

Result (S.50) now follows using the above in (S.58). As for Result (S.51), it follows using (S.47) and (S.49),

$$\begin{aligned} Var\left[\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{1/2}\right] &= E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right) - \left[E\left[\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{1/2}\right]\right]^2 \\ &= 1 + O\left(\frac{1}{T}\right) - \left[1 + O\left(\frac{1}{T}\right)\right]^2 \\ &= O\left(\frac{1}{T}\right) \end{aligned}$$

Result (S.52) follows by writing $\varepsilon_t^2 = \boldsymbol{\varepsilon}'\mathbf{A}\boldsymbol{\varepsilon}$ where $\mathbf{A}$ has only one non-zero element on its diagonal, and then using result (S.46). Results (S.53) follows since by Cauchy–Schwarz inequality we have

$$\begin{aligned} E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{1/2}\right] &= E\left[\varepsilon_t\varepsilon_t\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{1/2}\right] \\ &\leq \left[E\left(\varepsilon_t^2\right)\right]^{1/2}\left\{E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)\right]\right\}^{1/2} = \left\{E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)\right]\right\}^{1/2}, \end{aligned}$$

and using (S.52) it follows that

$$E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{1/2}\right] \leq \left[1 + O\left(\frac{1}{T}\right)\right]^{1/2} = 1 + O\left(\frac{1}{T}\right).$$

The last equality follows by using Maclaurian's expansion of $\sqrt{1+x}$, where x is small. To establish (S.54) we note that (using results in Lieberman (1994))

$$E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{-1}\right] = E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{A}\boldsymbol{\varepsilon}}{T^{-1}\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}\right) = \frac{E\left(\boldsymbol{\varepsilon}'\mathbf{A}\boldsymbol{\varepsilon}\right)}{E\left(T^{-1}\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}\right)} + O\left(\frac{1}{T}\right) = \frac{T}{T-m} + O\left(\frac{1}{T}\right) = 1 + O\left(\frac{1}{T}\right).$$

Similarly, result (S.55) follows noting that

$$\begin{aligned} E\left[\varepsilon_t\varepsilon_t\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{-1/2}\right] &\leq \left[E\left(\varepsilon_t^2\right)\right]^{1/2}\left\{E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{-1}\right]\right\}^{1/2} = \left\{E\left[\varepsilon_t^2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M_F}\boldsymbol{\varepsilon}}{T}\right)^{-1}\right]\right\}^{1/2} \\ &= \left[1 + O\left(\frac{1}{T}\right)\right]^{1/2} = 1 + O\left(\frac{1}{T}\right). \end{aligned}$$

To establish (S.56), using (S.49) we first note that

$$E\left(z_T\right)=E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_{\mathbf{F}}\boldsymbol{\varepsilon}}{T}\right)^{1/2}-1=O\left(\frac{1}{T}\right). \tag{S.59}$$

Similarly,

$$\begin{aligned}
E\left(z_T^2\right) &= E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_{\mathbf{F}}\boldsymbol{\varepsilon}}{T}\right)+1-2\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_{\mathbf{F}}\boldsymbol{\varepsilon}}{T}\right)^{1/2} \\
&= \left[E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_{\mathbf{F}}\boldsymbol{\varepsilon}}{T}\right)-1\right]-2\left[E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_{\mathbf{F}}\boldsymbol{\varepsilon}}{T}\right)^{1/2}-1\right] \\
&= \left[E\left(\frac{\boldsymbol{\varepsilon}'\mathbf{M}_{\mathbf{F}}\boldsymbol{\varepsilon}}{T}\right)-1\right]-2E\left(z_T\right).
\end{aligned}$$

The desired results now follows using (S.47) and (S.59). Finally, to establish (S.57), we first note that since $\mathbf{M}_{\mathbf{F}}$ is an idempotent matrix with rank $T-m$, and $\mathbf{F}$ is distributed independently of $\boldsymbol{\varepsilon}$, then

$$z_T=\left(\frac{\sum_{t=1}^{T-m}\left(\eta_t^2-1\right)}{T}+\frac{T-m}{T}\right)^{1/2}-1,$$

where $\eta_t \sim IID(0,1)$. It also follows that (note that η_t^2-1 is independent over t and has a zero mean and a finite variance (since by assumption ε_t has fourth order moments) then $T^{-1}\sum_{t=1}^{T-m}\left(\eta_t^2-1\right)=O_p(T^{-1/2})$.Hence, $z_T=\left[1+O_p(T^{-1/2})\right]^{1/2}-1=O_p\left(T^{-1/2}\right)$,as required. ■

Lemma S.8 *The CD statistic defined by (7) can be written equivalently as,*

$$CD=\left(\sqrt{\frac{n}{n-1}}\right)\frac{1}{\sqrt{2T}}\sum_{t=1}^{T}\left[\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{e_{it}}{\hat{\sigma}_{i,T}}\right)^2-1\right]. \tag{S.60}$$

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 (7) we have:

$$CD=\sqrt{\frac{2T}{n(n-1)}}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{\frac{1}{T}\sum_{t=1}^{T}e_{it}e_{jt}}{\hat{\sigma}_{i,T}\hat{\sigma}_{j,T}}=\sqrt{\frac{2T}{n(n-1)}}\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(\frac{e_{it}}{\hat{\sigma}_{i,T}}\right)\left(\frac{e_{jt}}{\hat{\sigma}_{j,T}}\right)\right). \tag{S.61}$$

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 (S.61), and after some algebra, we have

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

as required. ■

Lemma S.9 *Consider the CD and $\widetilde{CD}$ statistics defined by (16) and (17), respective 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*

$$
CD=\widetilde{CD}+o_p(1). \tag{S.62}
$$

Proof. Using (16) and (17) we first note that

$$
\left(\sqrt{\frac{2(n-1)}{n}}\right)\left(CD-\widetilde{CD}\right)=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{e_{it}}{\hat{\sigma}_{i,T}}\right)^2-\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{e_{it}}{\omega_{i,T}}\right)^2\right]. \tag{S.63}
$$

Also note that

$$
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{e_{it}}{\hat{\sigma}_{i,T}}=h_{t,nT}+g_{t,nT} \tag{S.64}
$$

where (see also (A.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 (S.64) and using the result in (S.63) we have

$$
\begin{aligned}
\left(\sqrt{\frac{n-1}{2n}}\right)\left(CD-\widetilde{CD}\right) &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}g_{t,nT}^2+\frac{2}{\sqrt{T}}\sum_{t=1}^{T}h_{t,nT}\,g_{t,nT} \\
&= \sqrt{\frac{T}{n}}\left(\frac{1}{\sqrt{n}}\mathbf{d}_{nT}'\mathbf{V}_{eT}\mathbf{d}_{nT}+\frac{1}{\sqrt{n}}\mathbf{c}_{nT}'\mathbf{V}_{eT}\mathbf{d}_{nT}\right),
\end{aligned} \tag{S.65}
$$

where $\mathbf{V}_{eT} = T^{-1}\sum_{t=1}^{T}\mathbf{e}_t\mathbf{e}_t'$. Now using (A.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

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

or in matrix forms

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

where $\mathbf{V}_T = T^{-1}\sum_{t=1}^{T}\mathbf{u}_t\mathbf{u}_t'$ and by Assumption 3 $\|\mathbf{V}_T\| = O_p(1)$. Also by results in Lemma S.1 all other terms of the above are either $O_p(1)$ or of lower order, and we also have $\|\mathbf{V}_{eT}\| = O_p(1)$. Consider now the terms in (S.65) and note that

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

But

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

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

By assumption $\omega_{iT} > c > 0$, and $\omega_{iT}^{-2} < c^{-1} < \infty$, and hence $n^{-1/2}\left\|\mathbf{c}_{nT}\right\| = O_p(1)$. Also, using (S.22) 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 $\left\|\mathbf{d}_{nT}\right\| = O_p\left(\frac{\sqrt{n}}{\delta_{nT}^{2}}\right) = o_p(1)$, recalling that n and T are of the same order. Hence, $\left|CD - \widetilde{CD}\right| = o_p(1)$, as required. ■

Lemma S.10 *Consider the latent factor loadings, $\boldsymbol{\gamma}_i$, in model (1) and their estimates $\hat{\boldsymbol{\gamma}}_i$ given by (5). Then under Assumptions 1-3 with n and $T \to \infty$, such that $n/T \to \kappa$, for $0 < \kappa < \infty$, we have*

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i}{\sigma_i} = O_p\left(\frac{1}{\delta_{nT}^{2}}\right), \tag{S.66}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right)\sigma_i = O_p\left(\frac{1}{\delta_{nT}^{2}}\right), \tag{S.67}$$

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\sigma}_{i,T}\hat{\boldsymbol{\gamma}}_i - \frac{1}{n}\sum_{i=1}^{n}\sigma_i\boldsymbol{\gamma}_i = O_p\left(\frac{1}{\delta_{nT}^{2}}\right), \tag{S.68}$$

$$\frac{1}{n}\sum_{i=1}^{n}\sigma_i^{2}\left(\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}_i'-\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\right) = O_p\left(\frac{1}{\delta_{nT}^{2}}\right). \tag{S.69}$$

Proof. Results (S.66) and (S.67) follow directly from (S.23) by setting $b_{in} = \sigma_i^{-1}$ and $b_{in} = \sigma_i$, respectively. To prove (S.68) note that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\hat{\sigma}_{i,T}\hat{\boldsymbol{\gamma}}_i - \frac{1}{n}\sum_{i=1}^{n}\sigma_i\boldsymbol{\gamma}_i \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\left(\hat{\sigma}_{i,T}-\omega_{i,T}\right)+\omega_{i,T}\right]\left[\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i+\boldsymbol{\gamma}_i\right] - \frac{1}{n}\sum_{i=1}^{n}\sigma_i\boldsymbol{\gamma}_i \\
&= \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\left(\omega_{i,T}-\sigma_i\right) + \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\gamma}_i\left(\hat{\sigma}_{i,T}-\omega_{i,T}\right) + \frac{1}{n}\sum_{i=1}^{n}\sigma_i\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right) \\
&+ \frac{1}{n}\sum_{i=1}^{n}\left(\omega_{i,T}-\sigma_i\right)\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right) + \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_{i,T}-\omega_{i,T}\right)\left(\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\right) \\
&= \mathbf{A}_{1,nT}+\mathbf{A}_{2,nT}+\mathbf{A}_{3,nT}+\mathbf{A}_{4,nT}+\mathbf{A}_{5,nT}.
\end{aligned} \tag{S.70}$$

Recall also that under Assumptions 2 and 3 σ_i and $\boldsymbol{\gamma}_i$ are bounded and $\omega_{i,T} = T^{-1}\boldsymbol{\varepsilon}_i'\mathbf{M}_{\mathbf{F}}\boldsymbol{\varepsilon}_i$, 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 (S.49) we have

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

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 (S.51)

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

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 (S.13) 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 (S.67), which is already established, then $\mathbf{A}_{3,nT} = O_p\left(\delta_{nT}^{-2}\right)$. Using result (S.24) 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 (S.26) 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 (S.68) now follows since $\mathbf{A}_{j,nT} = O_p(\delta_{nT}^{-2})$, for $j = 1, 2, ..., 5$. Finally, consider (S.69) and note that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\gamma}}_i' - \boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\right) &= \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)' \\ &+ \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)\boldsymbol{\gamma}_i' + \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\boldsymbol{\gamma}_i\left(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\right)'. \qquad \text{(S.71)}\end{aligned}$$

Since σ_i^2 is bounded, then

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

and using (S.2) 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 (S.29), 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 (S.69) follows. ■

References

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

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.