2011

Essays on Testing Hypotheses When Non-stationarity Exists in Panel Data Models

Sang Gon Na
Syracuse University

Abstract

This dissertation consists of two essays on testing hypotheses in panel data models when non-stationarity exists in the model. This is done under the high-dimensional framework where both n (cross-section dimension) and T (time series dimension) are large. In the first essay, I discuss the limiting distribution of the t-statistic for $H_0 : \beta = \beta_0$ using different panel data estimators and propose using the t-statistic based on Feasible GLS estimator. In the second essay, I develop the bootstrap F-statistic for cross-sectional independence in a panel data model with factor structure.

The first essay considers the problem of hypotheses testing in a simple panel data regression model with random individual effects and serially correlated disturbances. Following Baltagi, Kao and Liu (2008), I allow for the possibility of non-stationarity in the regressor and/or the disturbance term. While Baltagi *et al.* (2008) focus on the asymptotic properties and distributions of the standard panel data estimators, this essay focuses on test of hypotheses in this setting. One important finding, is that unlike the time series case, one does not necessarily need to rely on the "super-efficient" type AR estimator by Perron and Yabu (2009) to make inference in panel data. In fact, I show that the simple t-ratio always converges to the standard normal distribution regardless of whether the disturbances and/or the regressor are stationary. One caveat is that this may not be robust to heteroskedasticity of the error terms, but it is robust to heterogenous AR parameters across individu-

als. The Monte Carlo simulations in support of all the results are also provided in this essay.

The second essay discusses testing hypotheses of cross-sectional dependence in a panel data model with an introduction of factor structure. Following Bai (2003, 2004, 2009) and Bai, Kao and Ng (2009), I again allow for the possibility of non-stationarity in the regressor and the factor. I give attention to test of hypotheses using F-tests in this setting. The limiting distribution of F-statistics under the high-dimensional framework has not been derived yet in the literature perhaps because of its theoretical complexity. To circumvent this difficulty, this essay suggests the use of wild bootstrap F-tests based on simulation results under various cases where both regressors and factors can be stationary or non-stationary. The Monte Carlo results show that the bootstrap F-tests perform well in testing cross-sectional independence and are recommended in practice. They have the advantage of being feasible even when we do not observe the factors and do not require for formal theoretical approximations. It is also shown that the bootstrap F-tests are robust to heteroskedasticity but sensitive to serial correlation.

Essays on Testing Hypotheses When Non-Stationarity Exists in Panel Data Models

Sang Gon Na

B.A. Yonsei University, 1999

E.M.P.A. Syracuse University, 2009

M.A. Syracuse University, 2010

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Graduate School of Syracuse University

May 2011

Table of Contents

List of Tables and Figures

TABLES

Essay I: Test of Hypotheses in Panel Data Models When the Regressor and Disturbances are Possibly Nonstationary

Essay II: Testing Cross-sectional Dependence Using Bootstrap F-tests

FIGURES

Essay I: Test of Hypotheses in Panel Data Models When the Regressor and Disturbances are Possibly Nonstationary

Essay II: Testing Cross-sectional Dependence Using Bootstrap F-tests

Acknowledgements

It is my pleasure to thank the many people who made this dissertation possible. My deepest gratitude goes to all of whom helped me complete this dissertation and contributed to my intellectual progress at Syracuse University.

In particular, I would like to thank my two advisors, Professor Chihwa Kao and Professor Badi Baltagi for all the encouragement they gave me especially when I thought I could not do any econometric research at all. They have always stimulated me, and enriched my life not as my strict advisors but as my dear friends, even when I was struggling with something other than research. Throughout my dissertation-writing period, each of them provided precious advice, nice teaching, good company, and lots of great ideas. I would have been completely lost without them.

I also wish to thank Professor Donald Dutkowsky, Professor William Horrace and Professor Jan Ondrich who not only helped me build knowledge and techniques for economic research but also participated in the oral examination committee. In addition, I acknowledge the contribution of Professor David Popp for chairing my defense. I thank also the many people who have taught me economics at Syracuse University. Moreover, I would like to acknowledge the assistance of Laura Sauta and Sue Lewis in the Department of Economics at Syracuse University.

I am indebted to my fellow students for providing a fun and encouraging environment. I am especially grateful to Jie Dai, Anshula Johri, Jing Li, Allison Marier, Wael Moussa, Hyun Joon Shin, Liu Tian, and Yin-Fang Yen at Syracuse University. Ying Deng was particularly helpful in patiently teaching me how to program using

GAUSS as well as sharing her interest in the econometrics field. I am also sincerely grateful to Qu Feng who provided advice in each step of my progress.

I thank also my parents, my mother Jong Soon Lee and my father Kang Soo Na. They raised me, supported me, and loved me. To them I wish to dedicate this dissertation.

Last but most important, a big thanks to my wife, Jeong Eun Shin. Without her, I would be a very different person now, and it must have been impossible to finish a Ph.D. She has warmly and devotedly helped me keep focused on my research even when her heart was full of trouble through a fact of life.

Essay I: Test of Hypotheses in Panel Data Models When the Regressor and Disturbances Are Possibly Nonstationary

1 Introduction

In the time series literature, estimation and test of hypotheses of the deterministic time trend model with serially correlated disturbances have been studied by Canjels and Watson (1997), Vogelsang (1998) and Perron and Yabu (2009) to mention a few. For the panel data model, Baltagi and Krämer (1997) and Kao and Emerson (2004, 2005) study the corresponding time trend model with unobservable individual effects and autoregressive remainder disturbances. Baltagi, Kao and Liu (2008) extend this analysis to the case of a panel data regression model with possible non-stationarity in the regressor and/or the disturbance term. They derive the asymptotic distributions of the standard panel data estimators including ordinary least squares (OLS), fixed effects (FE), first-difference (FD), and generalized least squares (GLS) estimators when both the time-series length (T) and the number of cross-sections (n) are large. They show that these estimators have asymptotic normal distributions and have different convergence rates dependent on the non-stationarity of the regressor and the remainder disturbances. Some of their important findings include the following: (i) When the disturbance term is $I(0)$ and the regressor is $I(1)$, the FE estimator is asymptotically equivalent to the GLS estimator and OLS is less efficient than GLS; (ii) When the disturbance term and the regressor are $I(1)$, GLS is more efficient than the FE estimator since GLS is $\sqrt{nT}$ consistent, while FE is $\sqrt{n}$ consistent. As a result, they recommend the GLS estimator as the preferred estimator, and they show using Monte Carlo experiments that the loss in efficiency of the OLS, FE, and FD estimators relative to true GLS can be substantial. This paper is a follow up paper which is concerned with *test of hypotheses* using these standard panel data estimators. One important finding, is that unlike the time series setting, one does not necessarily need to rely on the "*super-efficient*" type AR estimator by Perron and Yabu (2009) to make inference in panel data. In fact, we show that the simple t-ratio based on the FGLS estimator of Baltagi and Li (1991), will always converge to the standard normal distribution regardless of whether the disturbances and/or the regressor are stationary or not. We also show using Monte Carlo experiments that inference based on the OLS, FE, and FD estimators could be misleading relative to that based on feasible GLS. The outline of the paper is as follows: Section 2 considers a simple panel data regression model with unobserved individual effects and AR(1) remainder disturbances and derives the asymptotic distributions of the t statistics of the standard FE and FD estimators, respectively. This is done for four cases, corresponding to whether the remainder disturbances and/or the regressor are stationary or not. In Section 3, we derive the corresponding asymptotic distributions of the t statistic for the FGLS estimator under these four cases. Section 4 reports the finite sample properties of the proposed tests using Monte Carlo experiments. Section 5 concludes. All proofs are given in the appendix.

Unless otherwise specified, for all the asymptotic results in this paper, we let n and T go to infinity simultaneously (i.e., $(n, T) \to \infty$), see Phillips and Moon (1999). We require $\frac{n}{T} \to 0$ in some cases. We write the integral $\int_0^1 W(s)ds$ as $\int W$ and $\bar{W}$ as $W - \int W$ when there is no ambiguity over limits. We use $\xrightarrow{p}$ to denote convergence in probability, $\xrightarrow{d}$ to denote convergence in distribution, $\otimes$ to denote Kronecker product, and $[x]$ to denote the largest integer $\leq x$.

2 The Model and Assumptions

Consider the following panel data regression model:

$$y_{it} = \gamma + \beta x_{it} + u_{it}, \quad i = 1, \ldots, n, \quad t = 1, \ldots, T \tag{1}$$

where $u_{it} = \mu_i + \nu_{it}$, and γ and β are scalars.[1] We assume that the individual effect μ_i is random with $\mu_i \sim iid(0, \sigma_\mu^2)$ and $\{\nu_{it}\}$ is an $AR(1)$

$$\nu_{it} = \rho\nu_{it-1} + e_{it}, \quad |\rho| \leq 1 \tag{2}$$

where e_{it} is a white noise process with variance σ_e^2. The μ_i is independent of the ν_{it} for all i and t.[2] Let $\{x_{it}\}$ be also an AR(1) such that

$$x_{it} = \lambda x_{it-1} + \varepsilon_{it}, \quad |\lambda| \leq 1 \tag{3}$$

where ε_{it} is a white noise process with variance σ_ε^2. In this paper we assume that

$$E(\mu_i \mid x_{it}) = 0. \tag{4}$$

The initialization of this system is $y_{i1} = x_{i1} = O_p(1)$ for all i. Baltagi *et al.* (2008) derive the asymptotic distributions of the standard panel data estimators including ordinary least squares (OLS), fixed effects (FE), first-difference (FD), and generalized least squares (GLS) estimators of β when both T and n are large. They find that, when ν_{it} is $I(0)$ (i.e., $\rho < 1$), the FE[3] and the GLS estimators are both $\sqrt{n}T$ consistent and (asymptotically) equivalent. However, this asymptotic equivalence breaks down when ν_{it} is $I(1)$ (i.e., $\rho = 1$). In this case, the GLS and the FD[4] estimators are both

[1] For simplicity, we consider the case of one regressor, but our results can be extended to the multiple regressors case. In fact, we assume that for the multiple regressors case, $X'X$ is of full rank to avoid the complexity from possible cointegration.

[2] This model was studied by Baltagi and Li (1991) under stationarity of the regressors and the disturbances.

[3] The fixed effects estimator of β is given by,

$$\hat{\beta}_{FE} = \frac{\sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_{i.})(y_{it} - \bar{y}_{i.})}{\sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_{i.})^2}$$

where $\bar{x}_{i.} = \frac{1}{T}\sum_{t=1}^T x_{it}$ and $\bar{y}_{i.} = \frac{1}{T}\sum_{t=1}^T y_{it}$, see Hsiao (2003).

[4] The FD estimator is the OLS estimator of a first-differenced regression, see Hsiao (2003). That is,

$$\Delta y_{it} = \beta \Delta x_{it} + \Delta \nu_{it}.$$

$\sqrt{nT}$ consistent and more efficient than the FE estimator (which is $\sqrt{n}$ consistent).

Define the innovation vector $\mathbf{w}_{it} = (e_{it}, \varepsilon_{it})'$. We assume that $\mathbf{w}_{it}$ is a linear process that satisfies the following assumptions:

Assumption 1 *For each i, we assume:*

1. $\mathbf{w}_{it} = \Pi(L)\boldsymbol{\eta}_{it} = \sum_{j=0}^{\infty} \Pi_j \boldsymbol{\eta}_{it-j}$, $\sum_{j=0}^{\infty} j^a \|\Pi_j\| < \infty$, $|\Pi(1)| \neq 0$ *for some* $a > 1$.

2. *For a given i, $\boldsymbol{\eta}_{it}$ is i.i.d. with zero mean and variance-covariance matrix Ξ, and finite fourth order cumulants.*

Assumption 2 *We assume $\boldsymbol{\eta}_{it}$ and $\boldsymbol{\eta}_{jt}$ are independent for $i \neq j$. That is, we assume cross-sectional independence.*

Assumption 3 *We also assume $E(e_{it}\varepsilon_{i(t+k)}) = 0$ for all i and k and ε_{it} and e_{it} are independent.*

Assumption 1 implies that the partial sum process $\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \mathbf{w}_{it}$ satisfies the following properties:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \mathbf{w}_{it} \xrightarrow{d} \mathbf{B}_i(r) = \mathbf{BM}_i(\Omega) \text{ as } T \to \infty \text{ for all } i \tag{5}$$

where

$$\mathbf{B}_i = \begin{bmatrix} B_{ei} \\ B_{\epsilon i} \end{bmatrix}.$$

The long-run variance covariance matrix of $\{\mathbf{w}_{it}\}$ with Assumption 3 is given by

$$\begin{aligned} \Omega &= \sum_{j=-\infty}^{\infty} E(\mathbf{w}_{ij}\mathbf{w}_{i0}') \\ &= \Pi(1)\Xi\Pi(1)' \\ &= \begin{bmatrix} \sigma_e^2 & 0 \\ 0 & \sigma_\varepsilon^2 \end{bmatrix}. \end{aligned}$$

Then $\mathbf{B}_i$ can be rewritten as

$$\mathbf{B}_i = \begin{bmatrix} B_{ei} \\ B_{\epsilon i} \end{bmatrix} = \begin{bmatrix} \sigma_e & 0 \\ 0 & \sigma_\varepsilon \end{bmatrix} \begin{bmatrix} V_i \\ W_i \end{bmatrix} \tag{6}$$

where $\begin{bmatrix} V_i \\ W_i \end{bmatrix}$ is a standard Brownian motion.

2.1 The Fixed Effects and the First Difference Estimators

In this paper, we focus on testing the common slope β,

$$H_0 : \beta = \beta_0.$$

We start by investigating the asymptotic distributions of the t-statistics for H_0 based on the FE and FD estimators. Let us denote these by t_{FE} and t_{FD}, respectively. We derive these asymptotic distributions under four scenarios where the disturbances and the regressor are allowed to be $I(0)$ or $I(1)$.

If v_{it} is known to be $I(0)$,[5] the corresponding t-test for H_0 using the FE estimator $\widehat{\beta}_{FE}$, is given by:

$$t_{FE} = \frac{\widehat{\beta}_{FE} - \beta_0}{s_{FE}} \tag{7}$$

where $s_{FE} = \sqrt{\frac{\widehat{\sigma}_\nu^2}{\sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_{i.})^2}}$ with $\widehat{\sigma}_\nu^2 = \frac{\hat{\sigma}_e^2}{1-\hat{\rho}^2}$ and $\widehat{\sigma}_e^2 = \frac{1}{nT} \sum_{i=1}^n \sum_{t=2}^T (\widehat{\nu}_{it} - \widehat{\rho\nu}_{it-1})^2$. Here $\widehat{\nu}_{it} = (y_{it} - \overline{y}_{i.}) - \widehat{\beta}_{FE}(x_{it} - \overline{x}_{i.})$ denotes the FE residuals from equation (1), and $\widehat{\rho} = \frac{\sum_{i=1}^n \sum_{t=2}^T \hat{\nu}_{it}\hat{\nu}_{it-1}}{\sum_{i=1}^n \sum_{t=2}^T \hat{\nu}_{it-1}^2}$ is the estimator of ρ suggested by Baltagi and Li (1991). Next, we derive the limiting distribution of $\widehat{\rho}$ when $|\rho| < 1$ as well as when $\rho = 1$.

Lemma 1 *Assume* $(n, T) \to \infty$.

[5] Note that the FE and the GLS estimators are asymptotically equivalent for this case, see Baltagi *et al.* (2008).

1. *If* $|\rho| < 1$ *with* $\frac{n}{T} \to 0$,

$$\sqrt{nT}\left(\widehat{\rho} - \rho\right) \xrightarrow{d} N\left(0, 1 - \rho^2\right)$$

and

$$\widehat{\sigma}_e^2 \xrightarrow{p} \sigma_e^2.$$

2. *If* $\rho = 1$,

$$T\left(\widehat{\rho} - 1\right) \xrightarrow{p} -3$$

and

$$\widehat{\sigma}_e^2 \xrightarrow{p} \sigma_e^2.$$

Theorem 1 derives the corresponding asymptotic distribution of the t statistic based on the FE estimator (t_{FE}) under various scenarios involving the stationarity or non-stationarity of the regressor and the disturbances.

Theorem 1 *Assume* $(n, T) \to \infty$.

1. *If* $|\rho| < 1$, $|\lambda| < 1$ *with* $\frac{n}{T} \to 0$,

$$t_{FE} \xrightarrow{d} N\left(0, \frac{1 + \rho\lambda}{1 - \rho\lambda}\right).$$

2. *If* $\rho = 1$, $|\lambda| < 1$,

$$t_{FE} \text{ cannot be obtained.}$$

3. *If* $|\rho| < 1$, $\lambda = 1$ *with* $\frac{n}{T} \to 0$,

$$t_{FE} \xrightarrow{d} N\left(0, \frac{1 + \rho}{1 - \rho}\right).$$

4. *If* $\rho = 1,\ \lambda = 1,$

$$t_{FE} \text{ cannot be obtained.}$$

The results of Theorem 1 show that, under the null, t_{FE} has a normal distribution if the disturbance term is $I(0)$ regardless of the stationarity or non-stationarity of the regressor. Note that we cannot even implement the t test when the error term is $I(1)$. In fact, one cannot compute the standard error s_{FE} in this setting as shown in the Appendix.

Next, we turn to the case of the FD estimator, $\widehat{\beta}_{FD}$.[6] The corresponding t-test for H_0 using the FD estimator $\widehat{\beta}_{FD}$, is given by:

$$t_{FD} = \frac{\widehat{\beta}_{FD} - \beta_0}{s_{FD}} \tag{8}$$

where

$$s_{FD} = \sqrt{\frac{\widehat{\sigma}^2_{\triangle\nu}}{\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it} - x_{it-1})^2}}$$

with $\widehat{\sigma}^2_{\triangle\nu} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\Delta y_{it} - \widehat{\beta}_{FD} \triangle x_{it}\right)^2$.

Theorem 2 *Assume* $(n,T) \to \infty$ *and* $\frac{n}{T} \to 0$.

1. *If* $|\rho| < 1,\ |\lambda| < 1,$

$$t_{FD} \xrightarrow{d} N\left(0, \frac{(1+\rho)(1+\lambda)[(2-\rho-\lambda)^2 + \frac{(1-\rho)^3}{1+\rho} + \frac{(1-\lambda)^3}{1+\lambda}]}{4(1-\rho\lambda)^2}\right).$$

2. *If* $\rho = 1,\ |\lambda| < 1,$

$$t_{FD} \xrightarrow{d} N(0,1).$$

[6] Note that if v_{it} is known to be $I(1)$, the FD and the GLS estimators are asymptotically equivalent, see Baltagi *et al.* (2008).

3. If $|\rho| < 1,\ \lambda = 1,$

$$t_{FD} \xrightarrow{d} N(0,1).$$

4. If $\rho = 1,\ \lambda = 1,$

$$t_{FD} \xrightarrow{d} N(0,1).$$

The results of Theorem 2 show that, under the null, t_{FD} has a normal distribution regardless of the stationarity or non-stationarity of the regressor and/or the disturbance term.

3 The Feasible GLS Estimator

We rewrite equation (1) in vector form

$$\mathbf{y} = \gamma \boldsymbol{\iota}_{nT} + \mathbf{x}\beta + \mathbf{u} = \gamma \boldsymbol{\iota}_{nT} + \mathbf{x}\beta + \mathbf{Z}_{\boldsymbol{\mu}}\boldsymbol{\mu} + \boldsymbol{\nu} \tag{9}$$

where $\mathbf{y}$ is $nT \times 1$, $\mathbf{x}$ is a vector of x_{it} of dimension $nT \times 1$, $\boldsymbol{\iota}_{nT}$ is a vector of ones of dimension nT, $\mathbf{u}$ is $nT \times 1$, $\boldsymbol{\mu}$ is a vector of μ_i, $\boldsymbol{\nu}$ is a vector of ν_{it} and $\mathbf{Z}_{\mu} = I_n \otimes \boldsymbol{\iota}_T$.

By the partitioned inverse rule, it can be shown, see Baltagi *et al.* (2008), that

$$\begin{aligned} \widehat{\beta}_{GLS} &= \left[\mathbf{x}'\Phi^{-1}\mathbf{x} - \mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\left(\boldsymbol{\iota}_{nT}'\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\boldsymbol{\iota}_{nT}'\Phi^{-1}\mathbf{x}\right]^{-1} \\ &\quad \times \left[\mathbf{x}'\Phi^{-1}\mathbf{y} - \mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\left(\boldsymbol{\iota}_{nT}'\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\boldsymbol{\iota}_{nT}'\Phi^{-1}\mathbf{y}\right] \end{aligned} \tag{10}$$

Substituting (9), one gets:

$$\begin{aligned} \widehat{\beta}_{GLS} - \beta &= \left[\mathbf{x}'\Phi^{-1}\mathbf{x} - \mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\left(\boldsymbol{\iota}_{nT}'\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\boldsymbol{\iota}_{nT}'\Phi^{-1}\mathbf{x}\right]^{-1} \\ &\quad \times \left[\mathbf{x}'\Phi^{-1}\mathbf{u} - \mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\left(\boldsymbol{\iota}_{nT}'\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\boldsymbol{\iota}_{nT}'\Phi^{-1}\mathbf{u}\right] = G_1^{-1}G_2 \end{aligned} \tag{11}$$

where G_1 and G_2 are defined accordingly, see also the Appendix. The variance-covariance matrix is given by:

$$\Phi = E\left(\mathbf{uu}'\right) = \sigma_\mu^2 \left(I_n \otimes \boldsymbol{\iota}_T \boldsymbol{\iota}_T'\right) + \sigma_e^2 \left(I_n \otimes \mathbf{A}\right) \tag{12}$$

where $\boldsymbol{\iota}_T$ is a vector of ones of dimension T. $\mathbf{A}$ is the variance-covariance matrix of ν_{it}, which for the $AR(1)$ is given by:

$$\mathbf{A} = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{bmatrix} \tag{13}$$

when $|\rho| < 1$, and

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & T \end{bmatrix}$$

when $\rho = 1$. Thus, it can be shown, see Baltagi *et al.* (2008), that

$$\Phi^{-1} = I_n \otimes \left[\frac{1}{\sigma_e^2}\left(\mathbf{A}^{-1} - \frac{\sigma_\mu^2}{\sigma_e^2 + \theta\sigma_\mu^2}\mathbf{A}^{-1}\boldsymbol{\iota}_T\boldsymbol{\iota}_T'\mathbf{A}^{-1}\right)\right] \tag{14}$$

where $\theta = \boldsymbol{\iota}_T'\mathbf{A}^{-1}\boldsymbol{\iota}_T$. When $|\rho| < 1$, this estimation is equivalent to the Prais-Winsten transformation method suggested by Baltagi and Li (1991) for the panel data model.

One can easily verify that $\mathbf{A}^{-1} = \mathbf{C}'\mathbf{C}$, where

$$\mathbf{C} = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & -\rho & 1 & 0 \\ 0 & 0 & 0 & 0 & -\rho & 1 \end{bmatrix} \tag{15}$$

is the well known Prais-Winsten transformation for the $AR(1)$ model. Baltagi and Li (1991) suggest pre-multiplying the panel model (9) by $(I_n \otimes \mathbf{C})$ to get rid of serial correlation in the remainder term, and then performing a Fuller and Battese (1973) transformation in the second step to take care of the random effects.

In order to obtain the FGLS estimator, $\widehat{\beta}_{FGLS}$, we use an estimate of ρ suggested by Baltagi and Li (1991) based on FE residuals given below equation (7). The asymptotic distribution of $\widehat{\rho}$ was derived in Lemma 1. Define $\hat{\alpha} = \sqrt{(1+\widehat{\rho})/(1-\widehat{\rho})}$ and $\hat{\boldsymbol{\iota}}_T^{\alpha\prime} = \left(\hat{\alpha}, \boldsymbol{\iota}_{T-1}'\right)$, where $\boldsymbol{\iota}_{T-1}$ is a vector of ones of dimension $T-1$. Using a trick by Wansbeek and Kapteyn (1983), define $\widehat{\overline{J}}_T^{\alpha} = \hat{\boldsymbol{\iota}}_T^{\alpha}\hat{\boldsymbol{\iota}}_T^{\alpha\prime}/\hat{d}^2$, where $\hat{d}^2 = \hat{\boldsymbol{\iota}}_T^{\alpha\prime}\hat{\boldsymbol{\iota}}_T^{\alpha} = \frac{2\widehat{\rho}}{1-\widehat{\rho}} + T$. Then, $\hat{E}_T^{\alpha} = I_T - \widehat{\overline{J}}_T^{\alpha}$. Also let $\sigma_\alpha^2 = \theta\sigma_\mu^2 + \sigma_e^2$ where $\theta = (1-\rho)^2 d^2$. Estimates for σ_e^2 and σ_μ^2 can be obtained from

$$\hat{\sigma}_e^2 = \frac{1}{n(T-1)}\hat{\mathbf{u}}^{*\prime}\left(I_n \otimes \hat{E}_T^{\alpha}\right)\hat{\mathbf{u}}^*$$

and

$$\hat{\sigma}_\alpha^2 = \frac{1}{n}\hat{\mathbf{u}}^{*\prime}\left(I_n \otimes \widehat{\overline{J}}_T^{\alpha}\right)\hat{\mathbf{u}}^*$$

where $\hat{\mathbf{u}}^*$ are the Prais-Winsten transformed residuals (see Baltagi and Li (1991) for more details). Hence, $\hat{\sigma}_\mu^2$ can be estimated as

$$\hat{\sigma}_\mu^2 = \frac{1}{\hat{\theta}} \left(\hat{\sigma}_\alpha^2 - \hat{\sigma}_e^2 \right).$$

Substituting $\hat{\sigma}_e^2$, $\hat{\sigma}_\mu^2$, and $\widehat{\rho}$ into equation (14), one obtains $\widehat{\beta}_{FGLS}$. The corresponding t-test for H_0 using the FGLS estimator $\widehat{\beta}_{FGLS}$, is given by:

$$t_{FGLS} = \frac{\hat{\beta}_{FGLS} - \beta}{\sqrt{var(\hat{\beta}_{FGLS})}} = \frac{\widehat{G}_1^{-1}\widehat{G}_2}{\sqrt{\widehat{G}_1^{-1}}} = \widehat{G}_1^{-1/2}\widehat{G}_2 \tag{16}$$

where $\widehat{G}_1$ and $\widehat{G}_2$ are given as equation (11) with the replacement of Φ by $\hat{\Phi}$.

3.1 Case 1: Without Individual Effects

We begin with a simple case where $\mu_i = 0$. That is, the individual effects are not included in the true model, but there is first order serial correlation. This is not realistic in panel data economic models, but we study it as a base case. The variance-covariance matrix given in (12) reduces to

$$\Phi = E(\mathbf{u}\mathbf{u}') = \sigma_e^2 (I_n \otimes \mathbf{A}) \tag{17}$$

with

$$\Phi^{-1} = \frac{1}{\sigma_e^2}(I_n \otimes \mathbf{A}^{-1}).$$

In this case, the FGLS estimator, $\widehat{\beta}_{FGLS}$, will be based on $\widetilde{\rho}$ and $\widetilde{\sigma}_e^2$ given by,

$$\widetilde{\rho} = \frac{\sum_{i=1}^n \sum_{t=2}^T \widehat{u}_{it}\widehat{u}_{it-1}}{\sum_{i=1}^n \sum_{t=2}^T \widehat{u}_{it-1}^2} \tag{18}$$

and

$$\widetilde{\sigma}_e^2 = \frac{1}{n(T-1)} \widehat{\mathbf{u}}^{*\prime}\widehat{\mathbf{u}}^*$$

where $\widehat{u}_{it}$ denotes the OLS residual.[7]

Lemma 2 *Assume* $(n,T) \to \infty$ *and* $\frac{n}{T} \to 0$.

1. *If* $|\rho| < 1$, $|\lambda| < 1$,

$$\sqrt{nT}\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, \frac{(1+\rho\lambda)(1-\lambda^2)\sigma_e^2}{(1-\rho\lambda)(1-\rho^2)\sigma_\varepsilon^2}\right).$$

2. *If* $\rho = 1$, $|\lambda| < 1$,

$$\sqrt{n}\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, \frac{(1+\lambda)^2\sigma_e^2}{2\sigma_\varepsilon^2}\right).$$

3. *If* $|\rho| < 1$, $\lambda = 1$,

$$\sqrt{n}T\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, \frac{2\sigma_e^2}{(1-\rho)^2\sigma_\varepsilon^2}\right).$$

4. *If* $\rho = 1$, $\lambda = 1$,

$$\sqrt{n}\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, \frac{2\sigma_e^2}{3\sigma_\varepsilon^2}\right).$$

As shown above, we have the same rate of converging speed as that assuming individual effects except for case (3). That is, in the panel cointegration case, we have the convergence rate $\sqrt{n}T$ which is the same as that of the GLS estimator and the FE estimator. However, note that once we add the individual effects, the OLS estimator has the slower convergence rate $\sqrt{nT}$ rather than $\sqrt{n}T$ because μ_i dominates v_i.[8]

[7] Note that we use the OLS residuals instead of the FE residuals in this case. That is, $\widehat{u}_{it} = y_{it} - \widehat{\gamma}_{OLS} - \hat{\beta}_{OLS}x_{it} = \left(y_{it} - \overline{\overline{y}}\right) - \hat{\beta}_{OLS}(x_{it} - \overline{\overline{x}})$ with $\overline{\overline{y}} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} y_{it}$ and $\overline{\overline{x}} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} x_{it}$.

[8] The limiting distribution of the OLS estimator with individual effects is given by

$$\sqrt{nT}\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, \frac{4\sigma_\mu^2}{3\sigma_\varepsilon^2}\right)$$

e.g., see Baltagi *et al.* (2008) for details.

Lemma 3 *Assume* $(n, T) \to \infty$.

1. *If* $|\rho| < 1$ *with* $\frac{n}{T} \to 0$,

$$\sqrt{nT}\left(\widetilde{\rho} - \rho\right) \xrightarrow{d} N\left(0, 1 - \rho^2\right)$$

and

$$\widetilde{\sigma}_e^2 \xrightarrow{p} \sigma_e^2.$$

2. *If* $\rho = 1$,

$$T\left(\widetilde{\rho} - 1\right) \xrightarrow{p} 0$$

and

$$\widetilde{\sigma}_e^2 \xrightarrow{p} \sigma_e^2.$$

As can be seen in Lemma 3, we find that the limiting distribution of $\widetilde{\rho}$ is the same as that of $\widehat{\rho}$ using the FE residuals, when $|\rho| < 1$ with $\frac{n}{T} \to 0$. However, this limiting distribution is different when $\rho = 1$. Compare, $T\left(\widetilde{\rho} - 1\right) \xrightarrow{p} 0$ without individual effects with $T(\widehat{\rho} - 1) \xrightarrow{p} -3$ with individual effects. We also find that, in both cases, the consistency of $\widetilde{\sigma}_e^2$ can be achieved. Based on the above results, one can derive the asymptotic distribution of the t-ratio for each case.

Theorem 3 *Assume* $(n, T) \to \infty$ *and* $\frac{n}{T} \to 0$. *Without individual effects,* $\widetilde{\rho}$ *always leads to* $t_{FGLS} \xrightarrow{d} N(0, 1)$ *regardless of the stationarity or non-stationarity of the regressor and/or the disturbance term.*

Theorem 3 shows that t_{FGLS} always converges to the standard normal case whether the disturbance term is $I(0)$ or $I(1)$ and whether the regressor is $I(0)$ or $I(1)$. That is, without individual effects, the t-ratio based on the FGLS, can be used for inference using the standard normal distribution. Hence, in this case, one does not have to consider the "*super-efficient*" type estimator by Perron and Yabu (2009) which is designed to bridge the gap between $I(0)$ and $I(1)$.[9]

3.2 Case 2: With Individual Effects

This section derives the asymptotic distribution of t_{FGLS} given in (16) and discussed in Section 3.

Theorem 4 *Assume* $(n,T)\to\infty$ *and* $\frac{n}{T}\to 0$.

1. *If* $|\rho|<1,\ |\lambda|<1$

   (a)
   $$\hat{\sigma}_e^2 \xrightarrow{p} \sigma_e^2,$$

   (b)
   $$\hat{\sigma}_\mu^2 \xrightarrow{p} \sigma_\mu^2,$$

   (c)
   $$t_{FGLS} \xrightarrow{d} N(0,1).$$

2. *If* $\rho=1,\ |\lambda|<1$

   (a)
   $$\hat{\sigma}_e^2 \xrightarrow{p} \sigma_e^2,$$

   (b)
   $$(1-\hat{\rho})\,\hat{\sigma}_\mu^2 \xrightarrow{p} \frac{16}{25}\sigma_e^2,$$

[9] One can define the "super-efficient" estimator $\widehat{\rho}_s$ as

$$\widehat{\rho}_s=\begin{cases}\widehat{\rho} & \text{if } |\widehat{\rho}-1|>\frac{\varepsilon}{T^\delta}\\ 1 & \text{if } |\widehat{\rho}-1|\le\frac{\varepsilon}{T^\delta}\end{cases}$$

for some $\delta\in(0,1)$ and $\varepsilon>0$. Hence, when $\widehat{\rho}$ is in a $T^{-\delta}$ neighborhood of 1, it is assigned a value of 1. For details, see Perron and Yabu (2009).

(c)

$$t_{FGLS} \xrightarrow{d} N\left(0,1\right).$$

3. If $|\rho| < 1,\ \lambda = 1$

(a)

$$\hat{\sigma}_e^2 \xrightarrow{p} \sigma_e^2,$$

(b)

$$\hat{\sigma}_\mu^2 \xrightarrow{p} \sigma_\mu^2,$$

(c)

$$t_{FGLS} \xrightarrow{d} N\left(0,1\right).$$

4. If $\rho = 1,\ \lambda = 1$

(a)

$$\hat{\sigma}_e^2 \xrightarrow{p} \sigma_e^2,$$

(b)

$$\left(1-\hat{\rho}\right)\hat{\sigma}_\mu^2 \xrightarrow{p} \frac{16}{25}\sigma_e^2,$$

(c)

$$t_{FGLS} \xrightarrow{d} N\left(0,1\right).$$

Theorem 4 implies that the t-ratio based on $\widehat{\rho}$ by Baltagi and Li (1991) asymptotically leads to the standard normal distribution regardless of the stationarity or non-stationarity of the regressor and/or the disturbance term. This is an interesting finding because despite the fact that we do not have a consistent estimate of σ_μ^2 when $\rho = 1$, we can still obtain t_{FGLS} converging to $N(0,1)$. Accordingly, we have a similar result to that of Theorem 3 except that one cannot expect consistent estimates for all the variance components when $\rho = 1$.

4 Monte Carlo Results

This section runs Monte Carlo experiments in order to study the finite sample properties of the t-statistics for $H_0 : \beta = \beta_0$; based on OLS, FE, FD, GLS, FGLS using Cochrane-Orcutt (GLS-CO), and FGLS using Prais-Winsten (GLS-PW) estimators. We denote these t-statistics by tOLS, tFE, tFD, tGLS, tGLSCO, and tGLSPW, respectively. The model is generated by

$$y_{it} = x_{it}\beta + \mu_i + \nu_{it}, \quad i = 1, \ldots, n, \quad t = 1, \ldots, T \tag{19}$$

with $\beta = 10$. ν_{it} and x_{it} follow an AR(1) process as in (2) and (3) respectively with ρ and λ varying over the range $(0, 0.2, 0.4, 0.6, 0.8, 1)$. We set the variance from signal, see (3), at $\sigma_\varepsilon^2 = 5$. We also control the total variance from noise across experiments, see (2), to be $\sigma_\mu^2 + \sigma_e^2 = 10$. Hence, we have a fixed signal to noise ratio $\frac{\sigma_\varepsilon^2}{\sigma_\mu^2 + \sigma_e^2} = \frac{1}{2}$ across experiments.[10] Next, we vary $\xi = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_e^2}$ over the range $(0, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8)$, respectively. The sample sizes n and T are varied over the range $(20, 40, 60, 120, 240)$. In our experiments, ρ is estimated as the sample correlation coefficient between $\widehat{\nu}_{it}$ and $\widehat{\nu}_{it-1}$, i.e.,

$$\widehat{\rho} = \frac{\sum_{i=1}^n \sum_{t=2}^T \left(\widehat{\nu}_{it} - \overline{\widehat{\nu}}\right)\left(\widehat{\nu}_{it-1} - \overline{\widehat{\nu}}\right)}{\sqrt{\sum_{i=1}^n \sum_{t=2}^T \left(\widehat{\nu}_{it-1} - \overline{\widehat{\nu}}\right)^2}\sqrt{\sum_{i=1}^n \sum_{t=2}^T \left(\widehat{\nu}_{it-1} - \overline{\widehat{\nu}}\right)^2}}$$

where $\overline{\widehat{\nu}}$ is the sample average of $\widehat{\nu}_{it}$. We choose the correlation coefficient estimator because it ensures that $\widehat{\rho}$ is always between 0 and 1.

For each experiment, we perform 10,000 replications and compute the t-statistics using OLS, FE, FD, GLS-CO, GLS-PW, and true GLS. With this design we have 900 experiments. GAUSS 7.0.6 is used to perform the simulations. Random numbers

[10] Note that Baltagi and Li (1997) fix $\sigma_\mu^2 + \sigma_\nu^2$ across experiments. Here, one cannot obtain σ_ν^2 in the nonstationary case. Instead we fix $\sigma_\mu^2 + \sigma_e^2$ and our results are not sensitive to the choice of this sum. In fact, we tried 5, 10, and 20, and the results are similar.

for e_{it}, μ_i, and ε_{it} are generated by the GAUSS procedure RNDNS. We generate $n(T+1000)$ random numbers and then split them into n series so that each series has the same mean and variance. The first $1,000$ observations are discarded for each series.

Tables 1 to 4 report the empirical size of these various t-statistics, when the true size is 5%, for $(\rho, \lambda) = (0.4, 0.4), (1, 0.4), (0.4, 1), (1, 1)$, respectively. Note that $(\rho, \lambda) = (0.4, 1)$ is the *panel cointegration* case and $(\rho, \lambda) = (1, 1)$ is the *spurious regression* case. Some of our findings are the following: (i) As expected, tOLS and tFE perform badly and their performance deteriorate as ρ or λ increase. For Table 1, the size of tOLS varies between 10 and 18%, while the size of tFE varies between 9 and 11%. This gets worse for the non-stationary disturbances case in Table 2, where the size of tOLS and tFE varies between 18 and 20%. For the non-stationary regressor case in Table 3, the size of tOLS varies between 24 and 80%, while the size of tFE varies between 17 and 20%. The spurious regression case in Table 4 gives the worst performance for tOLS with size varying between 59 and 83%. The size for tFE is also bad varying between 51 and 78%. (ii) In all cases, except case 1, tFD performs well with empirical size close to 5%. For case 1, tFD is slightly over-sized at 7 to 9%. (iii) tGLS gives the best performance, with empirical size not statistically different from 5%, for all cases considered. (iv) Both tGLSPW and tGLSCO perform well across experiments. In fact, for small sample sizes such as $(n, T) = (20, 20)$, they are undersized in case 2, and oversized in cases 3 and 4. However, as n and/or T increase, the empirical size of tGLSPW and tGLSCO improves considerably. For example, in case 4, tGLSPW and tGLSCO are oversized at about 10 to 12% for $(n, T) = (20, 20)$, but their empirical size improves to around 6% for $(n, T) = (120, 120)$.

We also note that the size of tOLS gets worse as the percentage of heterogeneity across individuals (ξ) increases. However, this heterogeneity measure does not affect the performance of tFE and tFD, since both estimators wipe out the individual effects.

Theorems 3 and 4 also imply that the t-ratio using FGLS should converge to $N(0,1)$ whether or not the individual effects are included in the model. In fact, Figures 1 to 5 show the overlap of the $N(0,1)$ distribution and the distribution of tGLSPW for various sample sizes (fixing $\xi = 0.4$).

In conclusion, we note that tGLS gives the best performance, but it is infeasible. We recommend tFGLS for testing $H_0 : \beta = \beta_0$ when the researcher has no perfect foresight on stationarity of the regressor and/or the error term. tFD is a viable alternative to tFGLS if either the regressor or the error is nonstationary. tOLS and tFE are not recommended in these cases.

4.1 Robustness to Heterogeneous AR Parameters and Heteroskedasticity

In this section we check the robustness of our results to (i) heterogeneity in the AR parameters in both the regressor and the error term and also to (ii) heteroskedasticity in the error terms. To accomplish this we run two sets of Monte Carlo experiments. The first set of experiments allow the AR parameters to vary across individuals. More specifically, λ_i (for the regressor) and ρ_i (for the error term) are allowed to be uniformly distributed, i.e., $IIDU(0,1)$. The estimation and test procedure are the same as before while the Data Generating Process is different. Table 5 reports the empirical size of these new experiments. Interestingly, the t-statistics using FGLS turn out to be robust across these experiments. In fact, tGLSPW and tGLSCO have empirical size that varies between $4-5\%$. tOLS and tFE perform badly again. In fact, tOLS has empirical size that varies between 19% and 67%, while tFE has empirical size that varies between 16% and 34%. tFD is slightly oversized with empirical size that varies between 6% and 7%.

As for the presence of heteroskedasticity in the error terms, we generate the error terms using the following design:

$$\begin{aligned} e_{it} &= \sigma_i \zeta_{it}^1, \text{ and} \\ \varepsilon_{it} &= \sigma_i \zeta_{it}^2 \end{aligned}$$

where ζ_{it}^1 and ζ_{it}^2 are generated from $N(0,1)$, respectively. To incorporate heteroskedasticity, σ_i are generated as follows:

$$\sigma_i \begin{cases} = 1 \text{ for } i = 1, \ldots, \frac{4n}{5} \\ = c \text{ for } i = \frac{4n}{5} + 1, \ldots, n \end{cases}$$

where $c = \sqrt{2}$ or 10. The simulation results are reported in Table 6 and 7. For Case 1, i.e., $(\rho, \lambda) = (0.4, 0.4)$, we find the following: (i) Table 6 reports the results under relatively low degree of heteroskedasticity ($c = \sqrt{2}$). tFGLS are slightly oversized. In fact, the size for tGLSPW and tGLSCO varies between 6 and 7% for various sample sizes. tOLS and tFE are bad with size varying between 12 to 18% and 11 to 12%, respectively. tFD is also oversized at 9-10%. (ii) Table 7 presents the results under a higher degree of heteroskedasticity ($c = 10$). In this case all the t-statistics are way oversized. The size for tGLSPW and tGLSCO varies between 36 and 40%.[11] Hence, we conclude that tFGLS is robust to heterogeneous AR parameters, but not to heteroskedasticity in the error terms.

5 Conclusion

This paper derived the limiting distribution of the t-statistic for $H_0 : \beta = \beta_0$; using different panel data estimators including FE, FD, and FGLS. This is done in the context of a linear panel data regression model with possible nonstationarity in the

[11] The case of heteroskedastic error terms remains to be studied in the future. For possible ideas on how to handle this problem, see, Baltagi and Kao (2000).

regressor and/or the error term. We showed that one can use t statistics based on the FGLS estimator regardless of the nonstationarity of the regressor and/or the disturbance term. This is unlike the time-series case, where one has to consider a "super-efficient" type AR estimator of Perron and Yabu (2009) to achieve the normal limiting distribution of the t-ratio. One caveat is that this may not be robust to heteroskedasticity of the error terms, but it is robust to heterogenous AR parameters across individuals.

References

[1] Baltagi, B. H., and Kao, C. (2000), "Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey," *Advances in Econometrics*, 15, 7-51.

[2] Baltagi, B. H., Kao, C., and Liu, L. (2008), "Asymptotic Properties of Estimators for the Linear Panel Regression Model with Random Individual Effects and Serially Correlated Errors: The Case of Stationary and Non-Stationary Regressors and Residuals," *Econometrics Journal*, 11, 554-572.

[3] Baltagi, B. H., and Krämer, W. (1997), "A Simple Linear Trend Model with Error Components," *Econometric Theory*, 13, 463-463.

[4] Baltagi, B. H., and Li, Q. (1991), "A Transformation That Will Circumvent the Problem of Autocorrelation in an Error Component Model," *Journal of Econometrics*, 48, 385-393.

[5] Baltagi, B. H., and Li, Q. (1997), "Monte Carlo Results on Pure and Pretest Estimators of an Error Component Model with Autocorrelated Disturbances," *Annales D'Économie et de Statistique*, 48, 69-82.

[6] Canjels, E., and Watson, M. W. (1997), "Estimating Deterministic Trends in the Presence of Serially Correlated Errors," *Review of Economics and Statistics*, 79, 184-200.

[7] Fuller, W. A., and Battese, G. E. (1973), "Transformations for Estimation of Linear Models with Nested Error Structure," *Journal of the American Statistical Association*, 68, 626–632.

[8] Hsiao, C. (2003), *Analysis of Panel Data*, Cambridge, Cambridge University Press.

[9] Kao, C. (1999), "Spurious Regression and Residual-based Tests for Cointegration in Panel Data," *Journal of Econometrics*, 90, 1-44.

[10] Kao, C., and Emerson, J. (2004), "On the Estimation of a Linear Time Trend Regression with a One-Way Error Component Model in the Presence of Serially Correlated Errors: Part I," *Journal of Probability and Statistical Science*, 2, 213-243.

[11] Kao, C., and Emerson, J. (2005), "On the Estimation of a Linear Time Trend Regression with a One-Way Error Component Model in the Presence of Serially Correlated Errors: Part II," *Journal of Probability and Statistical Science*, 3, 59-96.

[12] Perron, P., and Yabu, T. (2009), "Estimating Deterministic Trends with an Integrated or Stationary Noise Component," *Journal of Econometrics*, 151, 56-69.

[13] Phillips, P. C. B., and Moon, H. R. (1999), "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica*, 67, 1057-1111.

[14] Vogelsang, T. J. (1998), "Trend Function Hypothesis Testing in the Presence of Serial Correlation," *Econometrica*, 66, 123-148.

[15] Wansbeek, T. J., and Kapteyn, A. (1983), "A Note on Spectral Decomposition and Maximum likelihood Estimation of ANOVA Models with Balanced Data," *Statistics and Probability Letters,* 1, 213–215.

Figure 1: The Histogram of tFGLSPW for $(n, T) = (20, 20)$

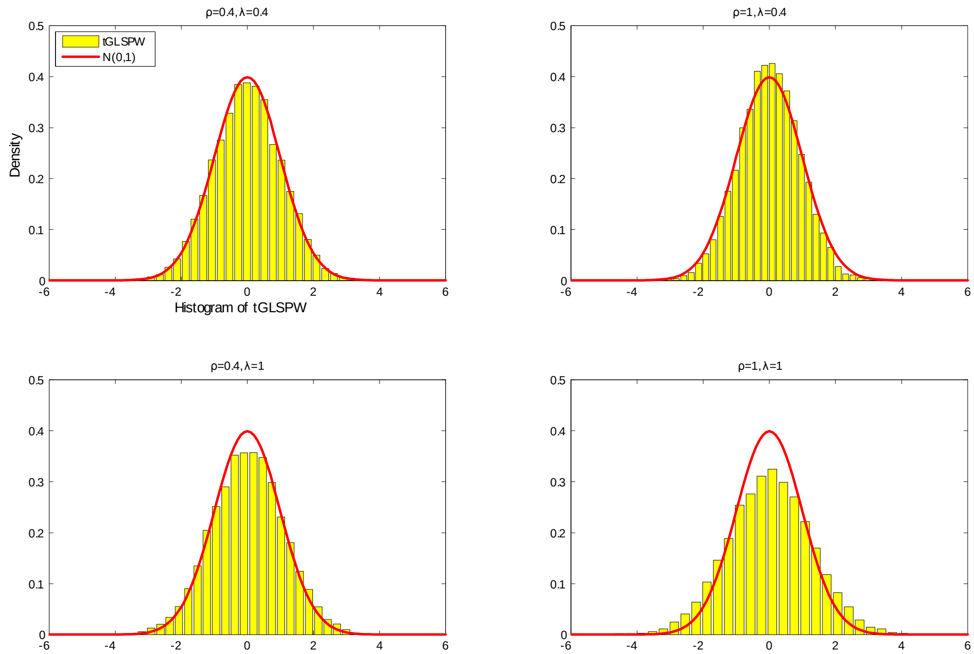

Figure 2: The Histogram of tFGLSPW for $(n, T) = (40, 40)$

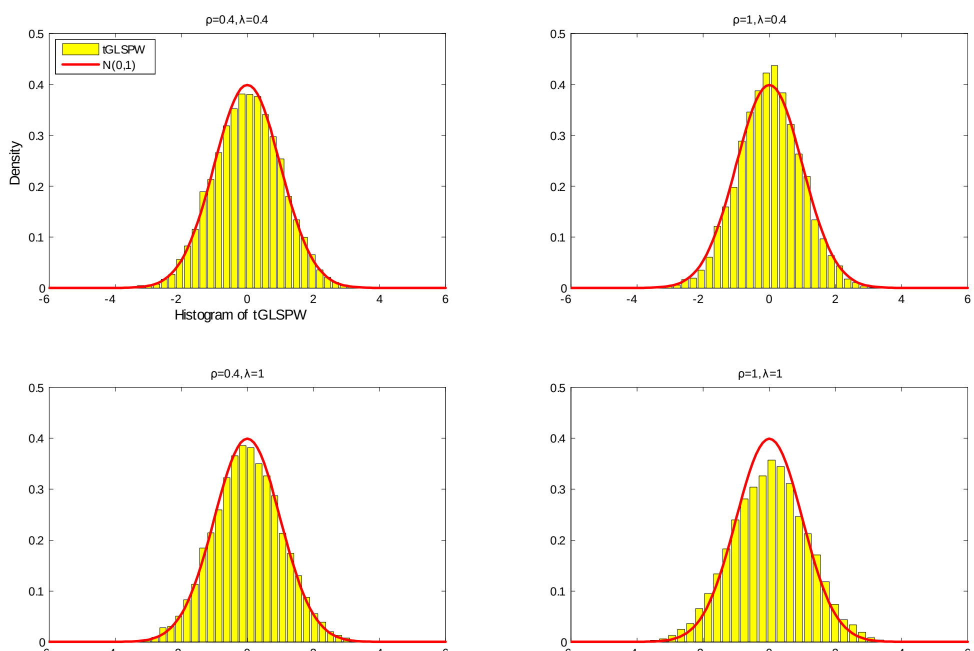

Figure 3: The Histogram of tFGLSPW for $(n, T) = (40, 120)$

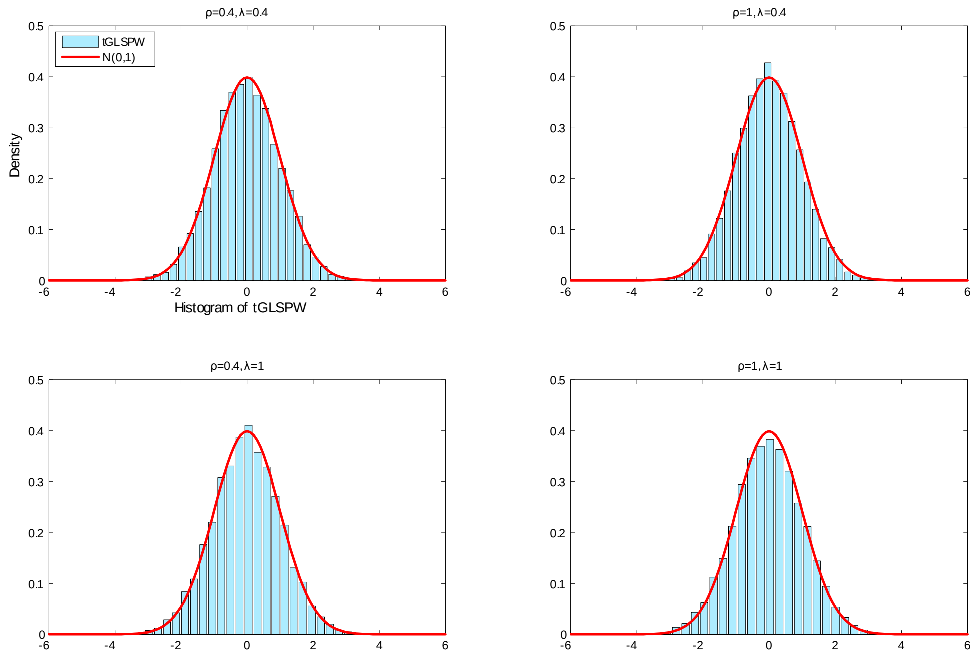

Figure 4: The Histogram of tFGLSPW for $(n, T) = (120, 40)$

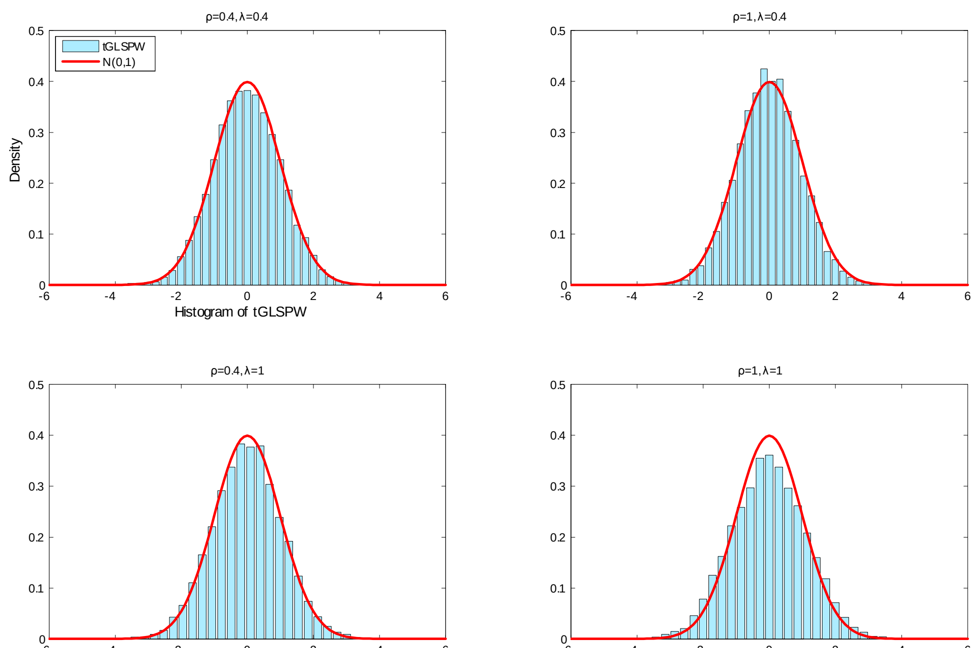

Figure 5: The Histogram of tFGLSPW for $(n, T) = (120, 120)$

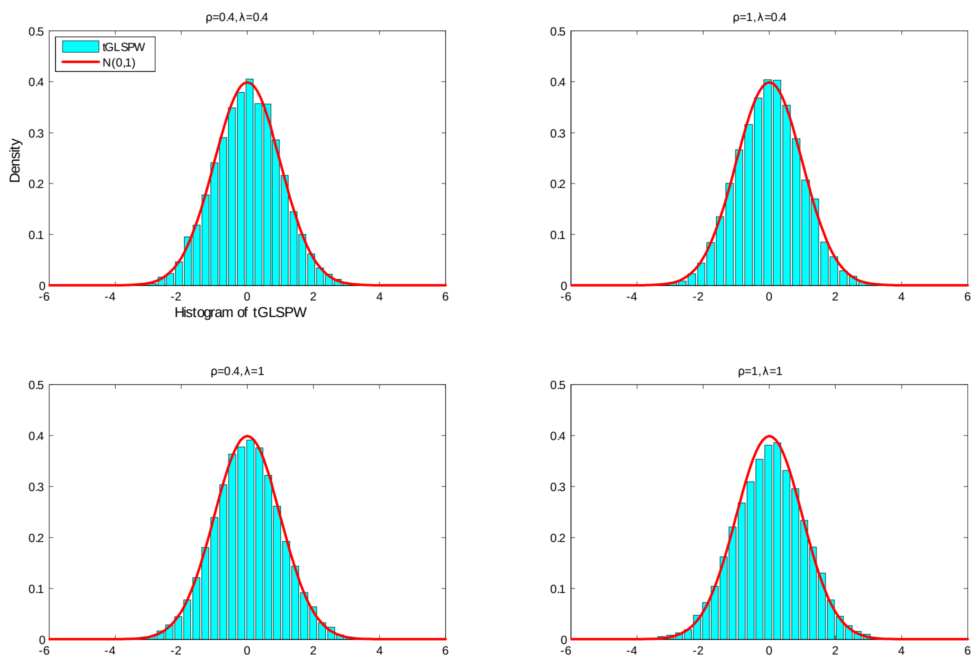

Table 1: The Empirical Size (%) of Case 1 ($\rho = 0.4$, $\lambda = 0.4$) with True Size 5%

(n, T)	ξ	tGLS	tGLSPW	tGLSCO	tOLS	tFE	tFD
(20, 20)	0.05	5.11	5.63	5.64	10.77		
	0.1	5.02	5.51	5.54	11.07		
	0.2	5.02	5.44	5.54	11.97	10.65	8.60
	0.4	4.96	5.34	5.40	13.73		
	0.6	4.92	5.23	5.23	15.63		
	0.8	4.90	5.13	5.07	17.28		
(40, 40)	0.05	5.12	5.15	5.34	10.20		
	0.1	5.06	5.20	5.35	10.74		
	0.2	4.99	5.14	5.36	11.67	10.23	7.84
	0.4	4.94	5.16	5.27	13.32		
	0.6	5.03	5.18	5.30	15.21		
	0.8	5.03	5.14	5.25	17.16		
(40, 120)	0.05	5.04	5.05	4.98	10.12		
	0.1	5.03	5.05	4.99	10.53		
	0.2	5.07	5.04	5.04	11.18	9.89	7.37
	0.4	5.06	5.05	5.00	13.32		
	0.6	5.04	5.02	5.00	15.47		
	0.8	5.04	5.02	4.96	17.78		
(120, 40)	0.05	4.90	4.96	5.06	10.63		
	0.1	4.84	4.98	5.10	11.04		
	0.2	4.87	5.03	5.03	11.54	10.33	8.01
	0.4	4.94	5.04	5.08	13.62		
	0.6	4.92	5.16	5.08	15.63		
	0.8	4.92	5.15	5.00	18.02		
(120, 120)	0.05	4.92	4.96	5.00	10.06		
	0.1	4.94	4.95	4.98	10.53		
	0.2	4.90	4.91	4.97	11.48	9.68	7.72
	0.4	4.94	4.93	5.00	13.59		
	0.6	4.94	4.97	4.98	15.35		
	0.8	4.96	4.95	4.96	17.39		

Table 2: The Empirical Size (%) of Case 2 ($\rho = 1.0$, $\lambda = 0.4$) with True Size 5%

(n, T)	ξ	tGLS	tGLSPW	tGLSCO	tOLS	tFE	tFD
(20, 20)	0.05	5.23	3.28	3.46	18.91		
	0.1	5.21	3.28	3.45	18.90		
	0.2	5.24	3.27	3.43	18.84	19.99	5.78
	0.4	5.11	3.27	3.39	19.01		
	0.6	5.15	3.24	3.37	18.88		
	0.8	5.16	3.19	3.25	18.89		
(40, 40)	0.05	5.25	4.17	4.14	18.37		
	0.1	5.27	4.17	4.14	18.38		
	0.2	5.17	4.17	4.14	18.42	19.78	5.48
	0.4	5.08	4.17	4.14	18.51		
	0.6	5.13	4.15	4.13	18.39		
	0.8	5.08	4.09	4.08	18.73		
(40, 120)	0.05	5.08	4.43	4.21	19.41		
	0.1	5.04	4.42	4.21	19.46		
	0.2	5.03	4.42	4.22	19.52	20.29	5.11
	0.4	5.02	4.41	4.23	19.44		
	0.6	4.94	4.36	4.22	19.53		
	0.8	5.01	4.35	4.22	19.89		
(120, 40)	0.05	5.32	4.36	4.31	19.16		
	0.1	5.33	4.36	4.31	19.05		
	0.2	5.29	4.35	4.31	18.90	19.44	5.55
	0.4	5.29	4.34	4.29	18.93		
	0.6	5.29	4.33	4.30	19.21		
	0.8	5.24	4.34	4.29	19.26		
(120, 120)	0.05	5.18	4.72	4.77	19.30		
	0.1	5.20	4.72	4.77	19.27		
	0.2	5.21	4.71	4.77	19.22	20.11	5.24
	0.4	5.23	4.71	4.78	19.29		
	0.6	5.17	4.71	4.79	19.03		
	0.8	5.16	4.71	4.78	19.20		

Table 3: *The Empirical Size (%) of Case 3 ($\rho = 0.4$, $\lambda = 1.0$) with True Size 5%*

(n, T)	ξ	tGLS	tGLSPW	tGLSCO	tOLS	tFE	tFD
(20, 20)	0.05	5.21	7.79	7.57	24.06		
	0.1	5.09	7.67	7.42	28.40		
	0.2	5.16	7.25	7.44	34.85	17.55	5.83
	0.4	4.90	7.12	7.23	44.86		
	0.6	4.88	7.02	7.05	51.77		
	0.8	4.89	6.67	6.69	57.25		
(40, 40)	0.05	4.92	6.26	6.16	28.53		
	0.1	5.01	6.16	6.03	35.16		
	0.2	4.95	6.14	6.12	44.19	18.47	5.48
	0.4	4.85	6.08	5.94	55.58		
	0.6	4.75	5.89	5.98	62.41		
	0.8	4.77	5.77	5.79	66.73		
(40, 120)	0.05	4.78	5.57	5.61	41.13		
	0.1	4.87	5.55	5.75	51.45		
	0.2	4.99	5.62	5.66	62.38	19.76	4.96
	0.4	5.10	5.76	5.57	72.81		
	0.6	5.30	5.77	5.69	77.72		
	0.8	5.24	5.65	5.63	80.63		
(120, 40)	0.05	4.71	5.56	5.60	28.08		
	0.1	4.78	5.57	5.59	35.28		
	0.2	4.90	5.80	5.82	44.61	19.01	5.72
	0.4	4.90	6.08	5.96	56.25		
	0.6	4.88	5.81	5.91	62.60		
	0.8	4.83	5.73	5.94	67.72		
(120, 120)	0.05	5.06	5.57	5.54	40.54		
	0.1	5.13	5.57	5.73	51.57		
	0.2	5.01	5.51	5.84	62.51	19.62	5.17
	0.4	5.17	5.60	5.78	72.29		
	0.6	5.14	5.65	5.74	77.34		
	0.8	5.23	5.62	5.77	80.61		

Table 4: The Empirical Size (%) of Case 4 ($\rho = 1.0$, $\lambda = 1.0$) with True Size 5%

(n, T)	ξ	tGLS	tGLSPW	tGLSCO	tOLS	tFE	tFD
(20, 20)	0.05	5.15	12.22	11.91	59.95		
	0.1	5.17	12.08	11.93	59.77		
	0.2	5.15	12.11	11.76	59.75	51.43	5.78
	0.4	5.01	11.93	11.59	59.90		
	0.6	5.03	11.55	11.24	60.30		
	0.8	4.97	11.04	10.52	60.95		
(40, 40)	0.05	5.02	8.48	8.50	70.70		
	0.1	5.04	8.44	8.43	70.74		
	0.2	4.99	8.43	8.40	70.84	62.56	5.41
	0.4	4.96	8.41	8.33	70.76		
	0.6	4.93	8.24	8.20	70.55		
	0.8	4.98	7.95	8.04	71.72		
(40, 120)	0.05	4.98	5.99	5.90	82.67		
	0.1	4.94	6.00	5.90	82.74		
	0.2	4.94	5.99	5.96	82.73	77.77	4.96
	0.4	4.92	5.98	5.97	82.67		
	0.6	4.94	6.00	5.92	82.81		
	0.8	4.96	5.92	5.86	82.73		
(120, 40)	0.05	4.84	8.06	8.01	70.69		
	0.1	4.80	8.03	7.97	70.75		
	0.2	4.81	8.06	7.95	70.53	63.09	5.26
	0.4	4.85	7.99	7.92	70.65		
	0.6	4.90	7.90	7.82	70.48		
	0.8	4.81	7.69	7.49	70.46		
(120, 120)	0.05	5.06	6.22	6.19	82.25		
	0.1	5.10	6.21	6.20	82.32		
	0.2	5.13	6.16	6.19	82.49	78.46	5.18
	0.4	5.12	6.14	6.18	82.17		
	0.6	5.15	6.15	6.15	82.26		
	0.8	5.11	6.10	6.08	82.55		

Table 5: *The Empirical Size (%) with Heterogeneous AR Parameters* $\left(\lambda_i, \rho_i \overset{i.i.d.}{\sim} U(0,1)\right)$

(n, T)	ξ	tGLSPW	tGLSCO	tOLS	tFE	tFD
(20, 20)	0.05	5.22	5.30	19.51		
	0.1	5.22	5.33	20.36		
	0.2	5.11	5.29	22.36	16.17	7.70
	0.4	5.13	5.23	26.61		
	0.6	5.00	5.12	30.83		
	0.8	4.73	4.80	35.05		
(40, 40)	0.05	4.73	4.75	21.71		
	0.1	4.72	4.80	23.57		
	0.2	4.77	4.86	27.00	20.00	7.21
	0.4	4.91	4.89	33.78		
	0.6	4.86	4.96	40.15		
	0.8	4.79	4.91	46.78		
(40, 120)	0.05	4.34	4.20	22.97		
	0.1	4.31	4.23	25.83		
	0.2	4.28	4.25	31.38	21.12	6.37
	0.4	4.29	4.25	40.31		
	0.6	4.32	4.27	47.84		
	0.8	4.35	4.29	55.02		
(120, 40)	0.05	5.26	5.23	36.51		
	0.1	5.28	5.28	37.75		
	0.2	5.40	5.33	40.42	26.06	7.53
	0.4	5.38	5.38	44.92		
	0.6	5.47	5.35	49.40		
	0.8	5.37	5.29	54.34		
(120, 120)	0.05	5.51	5.62	42.24		
	0.1	5.51	5.59	43.76		
	0.2	5.50	5.53	47.65	34.91	6.79
	0.4	5.41	5.53	54.71		
	0.6	5.38	5.53	61.22		
	0.8	5.36	5.49	67.46		

Table 6: The Empirical Size (%) under Heteroskedasticity $\left(c = \sqrt{2}\right)$

(n, T)	ξ	tGLSPW	tGLSCO	tOLS	tFE	tFD
(20, 20)	0.05	7.22	7.37	12.24		
	0.1	7.18	7.28	12.59		
	0.2	7.21	7.32	13.00	12.48	10.27
	0.4	7.10	7.25	14.15		
	0.6	7.12	7.25	15.67		
	0.8	7.01	7.15	17.67		
(40, 40)	0.05	6.66	6.80	12.25		
	0.1	6.66	6.74	12.61		
	0.2	6.60	6.71	13.24	11.92	10.35
	0.4	6.59	6.70	14.64		
	0.6	6.57	6.70	16.33		
	0.8	6.58	6.71	17.95		
(40, 120)	0.05	6.28	6.38	11.71		
	0.1	6.31	6.43	12.01		
	0.2	6.33	6.48	12.71	11.33	9.23
	0.4	6.38	6.53	14.14		
	0.6	6.42	6.58	15.99		
	0.8	6.42	6.53	17.63		
(120, 40)	0.05	6.72	6.72	12.31		
	0.1	6.69	6.63	12.77		
	0.2	6.62	6.56	13.49	12.45	9.51
	0.4	6.57	6.67	14.89		
	0.6	6.62	6.69	16.28		
	0.8	6.66	6.68	18.42		

Table 7: The Empirical Size (%) under Heteroskedasticity $(c = 10)$

(n, T)	ξ	tGLSPW	tGLSCO	tOLS	tFE	tFD
(20, 20)	0.05	40.13	40.23	43.94		
	0.1	40.01	40.06	43.91		
	0.2	39.83	39.89	43.76	44.71	42.78
	0.4	39.67	39.74	43.74		
	0.6	39.52	39.67	43.07		
	0.8	39.35	39.43	41.87		
(40, 40)	0.05	37.45	37.51	43.98		
	0.1	37.27	37.32	43.92		
	0.2	37.16	37.21	43.81	44.44	41.92
	0.4	36.97	36.98	43.47		
	0.6	36.71	36.74	42.95		
	0.8	36.76	36.83	41.76		
(40, 120)	0.05	36.53	36.62	43.36		
	0.1	36.61	36.62	43.35		
	0.2	36.51	36.54	43.33	43.67	41.37
	0.4	36.55	36.56	43.24		
	0.6	36.51	36.52	42.88		
	0.8	36.55	36.65	41.67		
(120, 40)	0.05	36.44	36.44	43.67		
	0.1	36.45	36.48	43.49		
	0.2	36.35	36.37	43.40	43.98	41.40
	0.4	36.18	36.21	43.27		
	0.6	36.24	36.22	43.06		
	0.8	36.34	36.37	41.96		

Appendix: Proofs of Lemmas and Theorems

This appendix provides all the proofs for the lemmas and theorems in the text.

A Proof of Lemma 1

Proof. We investigate $|\rho| < 1$ and $\rho = 1$ cases, consecutively.

1. $|\rho| < 1$ case

   (a) $|\rho| < 1,\ |\lambda| < 1$ case

   Consider the limiting distribution of $\hat{\rho}$. Note that

$$\widehat{\rho} - \rho = \frac{\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{\nu}_{it} - \rho\widehat{\nu}_{it-1}\right)\widehat{\nu}_{it-1}}{\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^{2}}$$

   and $\widehat{\nu}_{it} = (y_{it} - \bar{y}_{i.}) - \hat{\beta}_{FE}\left(x_{it} - \bar{x}_{i.}\right) = (\nu_{it} - \overline{\nu}_{i.}) - \left(\hat{\beta}_{FE} - \beta\right)(x_{it} - \bar{x}_{i.})$.

   For the denominator,

$$\begin{aligned}
\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^{2} &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[(\nu_{it-1} - \overline{\nu}_{i.}) - (\hat{\beta}_{FE} - \beta)(x_{it-1} - \bar{x}_{i.})\right]^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1} - \overline{\nu}_{i.})^2 \\
&\quad + \frac{1}{nT}\left\{\sqrt{nT}(\hat{\beta}_{FE} - \beta)\right\}^2\left\{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1} - \bar{x}_{i.})^2\right\} \\
&\quad - \frac{2}{nT}(\hat{\beta}_{FE} - \beta)\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1} - \overline{\nu}_{i.})(x_{it-1} - \bar{x}_{i.}) \\
&= I + II + III.
\end{aligned}$$

   Consider II first. It is easy to see $II = O_p(\frac{1}{nT}) = o_p(1)$ because if $(n, T) \to \infty$ and $\frac{n}{T} \to 0$, then

$$\sqrt{nT}\left(\hat{\beta}_{FE} - \beta\right) = O_p\left(1\right)$$

by, e.g., Baltagi *et al.* (2008). Similarly, we have

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1}-\bar{x}_{i.})^2 = O_p(1).$$

Next we show that $III = O_p(\frac{1}{nT})$. It can be shown that

$$\begin{aligned}
& \frac{2}{nT}(\hat{\beta}_{FE}-\beta)\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})(x_{it-1}-\bar{x}_{i.}) \\
= & \frac{2}{nT}(\hat{\beta}_{FE}-\beta)\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\nu_{it-1}x_{it-1}-\nu_{it-1}\bar{x}_{i.}-\overline{\nu}_{i.}x_{it-1}+\overline{\nu}_{i.}\bar{x}_{i.}\right] \\
= & O_p(\frac{1}{nT}) = o_p(1)
\end{aligned}$$

using

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\nu_{it-1}x_{it-1} = O_p(1)$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\nu_{it-1}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}x_{it}\right) = O_p(1)$$

since ν_{it-1} and x_{it-1} are uncorrelated.

Hence, we have

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^2 = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})^2 + O_p(\frac{1}{nT}) \xrightarrow{p} \frac{\sigma_e^2}{1-\rho^2}$$

as $(n,T)\to\infty$.

For the numerator, it can be shown that $\widehat{\nu}_{it}-\rho\widehat{\nu}_{it-1} = e_{it}-(\hat{\beta}_{FE}-$

$\beta) \{(\lambda - \rho)x_{it-1} + \varepsilon_{it}\} + o_p(1)$ and accordingly we have

$$\begin{aligned}
& \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} (\widehat{\nu}_{it} - \rho\widehat{\nu}_{it-1})\, \widehat{\nu}_{it-1} \\
= & \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \left\{ \begin{array}{c} \left[e_{it} - (\hat{\beta}_{FE} - \beta) \{(\lambda - \rho)x_{it-1} + \varepsilon_{it}\} + o_p(1)\right] \\ \left[(\nu_{it-1} - \overline{\nu}_{i.}) - (\hat{\beta}_{FE} - \beta)(x_{it-1} - \overline{x}_{i.})\right] \end{array} \right\} \\
= & \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} (\nu_{it-1} - \overline{\nu}_{i.})\, e_{it} \\
& -\sqrt{nT}(\hat{\beta}_{FE} - \beta) \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=2}^{T} \{(\lambda - \rho)x_{it-1} + \varepsilon_{it}\} (\nu_{it-1} - \overline{\nu}_{i.}) \\
& -\sqrt{nT}(\hat{\beta}_{FE} - \beta) \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=2}^{T} e_{it}(x_{it-1} - \overline{x}_{i.}) \\
& + \left\{\sqrt{nT}(\hat{\beta}_{FE} - \beta)\right\}^2 \frac{1}{(nT)^{3/2}} \sum_{i=1}^{n} \sum_{t=2}^{T} \{(\lambda - \rho)x_{it-1} + \varepsilon_{it}\} (x_{it-1} - \overline{x}_{i.}) + o_p(1) \\
= & \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} [(\nu_{it-1} - \overline{\nu}_{i.})\, e_{it}] + O_p(\frac{1}{\sqrt{nT}})
\end{aligned}$$

because $\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} x_{it-1}\nu_{it-1} = O_p(1)$, $\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \varepsilon_{it}\nu_{it-1} = O_p(1)$, $\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=2}^{T} x_{it-1}^2 = O_p(1)$, and $\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} x_{it-1}e_{it} = O_p(1)$ as $(n, T) \to \infty$.

Therefore, we have

$$\begin{aligned}
& \tfrac{1}{\sqrt{nT}} \textstyle\sum_{i=1}^{n} \sum_{t=2}^{T} [(\nu_{it-1} - \overline{\nu}_{i.})\, e_{it}] + o_p\,(1) \\
= & \tfrac{1}{\sqrt{nT}} \textstyle\sum_{i=1}^{n} \sum_{t=2}^{T} \nu_{it-1}e_{it} - \tfrac{1}{\sqrt{nT}} \sum_{i=1}^{n} \left[\left(\tfrac{1}{\sqrt{T}} \sum_{t=2}^{T} \nu_{it}\right)\left(\tfrac{1}{\sqrt{T}} \sum_{t=2}^{T} e_{it}\right)\right] + o_p\,(1) \\
\xrightarrow{d} & \sigma_e^2 N\left(0, \tfrac{1}{1-\rho^2}\right)
\end{aligned}$$

if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$. Note that

$$\begin{aligned}
& \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\left[\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\nu_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}e_{it}\right)\right] \\
= & \sqrt{\frac{n}{T}}\frac{1}{n}\sum_{i=1}^{n}\left[\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\nu_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}e_{it}\right)\right] \\
= & O_p\left(\sqrt{\frac{n}{T}}\right)O_p(1) \\
= & o_p(1)
\end{aligned}$$

with a condition $\frac{n}{T}\to 0$ since

$$\frac{1}{n}\sum_{i=1}^{n}\left[\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\nu_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}e_{it}\right)\right]=O_p(1).$$

Finally, we conclude that if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$\begin{aligned}
& \sqrt{nT}\left(\widehat{\rho}-\rho\right) \\
= & \frac{\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{\nu}_{it}-\rho\widehat{\nu}_{it-1}\right)\widehat{\nu}_{it-1}}{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^{2}} \xrightarrow{d} \frac{\sigma_e^2 N\left(0,\frac{1}{1-\rho^2}\right)}{\frac{\sigma_e^2}{1-\rho^2}} = N\left(0,1-\rho^2\right).
\end{aligned}$$

Next we consider $\widehat{\sigma}_e^2$. Note that

$$
\begin{aligned}
\widehat{\sigma}_e^2 &= \tfrac{1}{nT}\sum_{i=1}^n \sum_{t=2}^T \left(\widehat{\nu}_{it} - \widehat{\rho}\widehat{\nu}_{it-1}\right)^2 \\
&= \tfrac{1}{nT}\sum_{i=1}^n \sum_{t=2}^T \left[\begin{array}{c} \rho\nu_{it-1} + e_{it} - \overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)(\lambda x_{it-1} + \varepsilon_{it} - \bar{x}_{i.}) \\ -\hat{\rho}\left(\nu_{it-1} - \overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)(x_{it-1} - \bar{x}_{i.})\right) \end{array}\right]^2 \\
&= \tfrac{1}{nT}\sum_{i=1}^n \sum_{t=2}^T \left[\begin{array}{c} e_{it} - (\hat{\rho} - \rho)\left\{\nu_{it-1} - (\hat{\beta}_{FE} - \beta)x_{it-1}\right\} \\ +(\hat{\rho} - \rho)\left\{\overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)\bar{x}_{i.}\right\} \\ +(\rho - 1)\left\{\overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)\bar{x}_{i.}\right\} \\ -(\hat{\beta}_{FE} - \beta)\left\{\varepsilon_{it} + (\lambda - \rho)x_{it-1}\right\} \end{array}\right]^2 \\
&= \tfrac{1}{nT}\sum_{i=1}^n \sum_{t=2}^T e_{it}^2 + \tfrac{1}{nT}\sum_{i=1}^n \sum_{t=2}^T \left[(\hat{\rho} - \rho)\left\{\nu_{it-1} - (\hat{\beta}_{FE} - \beta)x_{it-1}\right\}\right]^2 \\
&\quad + \tfrac{1}{nT}\sum_{i=1}^n \sum_{t=2}^T \left[(\hat{\rho} - \rho)\left\{\overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)\bar{x}_{i.}\right\}\right]^2 \\
&\quad + \tfrac{1}{nT}\sum_{i=1}^n \sum_{t=2}^T \left[(\rho - 1)\left\{\overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)\bar{x}_{i.}\right\}\right]^2 \\
&\quad + \tfrac{1}{nT}\sum_{i=1}^n \sum_{t=2}^T \left[(\hat{\beta}_{FE} - \beta)\left\{\varepsilon_{it} + (\lambda - \rho)x_{it-1}\right\}\right]^2 \\
&\quad - \tfrac{2}{nT}\sum_{i=1}^n \sum_{t=2}^T e_{it}(\hat{\rho} - \rho)\left\{\nu_{it-1} - (\hat{\beta}_{FE} - \beta)x_{it-1}\right\} \\
&\quad + \tfrac{2}{nT}\sum_{i=1}^n \sum_{t=2}^T e_{it}(\hat{\rho} - \rho)\left\{\overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)\bar{x}_{i.}\right\} \\
&\quad + \tfrac{2}{nT}\sum_{i=1}^n \sum_{t=2}^T e_{it}(\rho - 1)\left\{\overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)\bar{x}_{i.}\right\} \\
&\quad - \tfrac{2}{nT}\sum_{i=1}^n \sum_{t=2}^T e_{it}(\hat{\beta}_{FE} - \beta)\left\{\varepsilon_{it} + (\lambda - \rho)x_{it-1}\right\} + o_p(1) \\
&= I + II + III + IV + V + VI + VII + VIII + VIIII + o_p(1).
\end{aligned}
$$

One can show that all the rest of terms except I is $o_p(1)$. For example, consider II.

$$
\begin{aligned}
&\left\{\sqrt{nT}(\hat{\rho} - \rho)\right\}^2 \frac{1}{(nT)^2}\sum_{i=1}^n \sum_{t=2}^T \left(\begin{array}{c} \nu_{it-1}^2 - 2(\hat{\beta}_{FE} - \beta)\nu_{it-1}x_{it-1} \\ +(\hat{\beta}_{FE} - \beta)^2 x_{it-1}^2 \end{array}\right) \\
= \;& O_p\left(\frac{1}{nT}\right) + O_p\left(\frac{1}{n^2T^2}\right) + O_p\left(\frac{1}{n^2T^2}\right) = O_p\left(\frac{1}{nT}\right) = o_p(1)
\end{aligned}
$$

using

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\nu_{it-1}^{2} = O_p(1),$$
$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\nu_{it-1}x_{it-1} = O_p(1)$$

and

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}^{2} = O_p(1)$$

as $(n,T)\to\infty$.

Similarly, it can be easily shown that

$$\widehat{\sigma}_e^2 = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}^2 + o_p(1) \xrightarrow{p} \sigma_e^2$$

as $(n,T)\to\infty$.

(b) $|\rho|<1,\ \lambda=1$ case

From

$$\widehat{\rho}-\rho = \frac{\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{\nu}_{it}-\rho\widehat{\nu}_{it-1}\right)\widehat{\nu}_{it-1}}{\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^{2}},$$

we have

$$\begin{aligned}
\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^{2} &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[(\nu_{it-1}-\overline{\nu}_{i.})-(\hat{\beta}_{FE}-\beta)(x_{it-1}-\bar{x}_{i.})\right]^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})^2 \\
&\quad +\frac{1}{nT}\left\{\sqrt{n}T(\hat{\beta}_{FE}-\beta)\right\}^2\left\{\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1}-\bar{x}_{i.})^2\right\} \\
&\quad -\frac{2}{n^{3/2}}\sqrt{n}T(\hat{\beta}_{FE}-\beta)\sum_{i=1}^{n}\frac{1}{T^2}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})(x_{it-1}-\bar{x}_{i.}) \\
&= I+II+III.
\end{aligned}$$

Using a similar argument as in Phillips and Moon (1999) and Baltagi *et al.* (2008), it can be shown that $II = O_p(\frac{1}{nT})$ if $(n,T) \to \infty$ and $\frac{n}{T} \to 0$, which yields

$$\sqrt{n}T\left(\hat{\beta}_{FE} - \beta\right) = O_p\left(1\right).$$

Also note that as $(n,T) \to \infty$, we have

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it-1} - \bar{x}_{i.})^2 = O_p(1),$$

see equation (C.3) in Kao (1999).

Next we show that $III = O_p\left(\frac{1}{nT}\right)$. This follows because

$$\begin{aligned}
& \frac{2}{n^{3/2}}\sum_{i=1}^{n}\frac{1}{T^2}\sum_{t=2}^{T}(\nu_{it-1} - \overline{\nu}_{i.})(x_{it-1} - \bar{x}_{i.}) \\
= & \frac{2}{nT}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1} - \overline{\nu}_{i.})\frac{(x_{it-1} - \bar{x}_{i.})}{\sqrt{T}} \\
= & O_p\left(\frac{1}{nT}\right)
\end{aligned}$$

and

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1} - \overline{\nu}_{i.})\frac{(x_{it-1} - \bar{x}_{i.})}{\sqrt{T}} = O_p\left(1\right)$$

where v_{it} and x_{it} are not correlated. Hence,

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^2 = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1} - \overline{\nu}_{i.})^2 + O_p\left(\frac{1}{nT}\right) \xrightarrow{p} \frac{\sigma_e^2}{1-\rho^2}$$

as $(n,T) \to \infty$.

For the numerator, $\widehat{\nu}_{it} - \rho\widehat{\nu}_{it-1} = e_{it} - (\hat{\beta}_{FE} - \beta)\left\{(1-\rho)x_{it-1} + \varepsilon_{it}\right\} + o_p(1)$

and by using a similar argument, we get

$$
\begin{aligned}
&\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{\nu}_{it}-\rho\widehat{\nu}_{it-1}\right)\widehat{\nu}_{it-1} \\
&=\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\begin{array}{c}\left[e_{it}-(\hat{\beta}_{FE}-\beta)\left\{(1-\rho)x_{it-1}+\varepsilon_{it}\right\}+o_p(1)\right] \\ \cdot\left[(\nu_{it-1}-\overline{\nu}_{i.})-(\hat{\beta}_{FE}-\beta)(x_{it-1}-\bar{x}_{i.})\right]\end{array}\right] \\
&=\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\left(\nu_{it-1}-\overline{\nu}_{i.}\right)e_{it}\right]+o_p\left(1\right) \\
&=\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\nu_{it-1}e_{it}-\frac{1}{\sqrt{T}}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\nu_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}e_{it}\right)\right]+o_p\left(1\right) \\
&\xrightarrow{d}\sigma_e^2 N\left(0,\frac{1}{1-\rho^2}\right)
\end{aligned}
$$

if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$.

We conclude that if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$
\begin{aligned}
&\sqrt{nT}\left(\widehat{\rho}-\rho\right) \\
=\;&\frac{\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{\nu}_{it}-\rho\widehat{\nu}_{it-1}\right)\widehat{\nu}_{it-1}}{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^2}\xrightarrow{d}\frac{\sigma_e^2 N\left(0,\frac{1}{1-\rho^2}\right)}{\frac{\sigma_e^2}{1-\rho^2}}=N\left(0,1-\rho^2\right),
\end{aligned}
$$

which is the same result as in Lemma 1.(1).

Next consider $\widehat{\sigma}_e^2$.

$$\begin{aligned}
\widehat{\sigma}_e^2 &= \tfrac{1}{nT}\textstyle\sum_{i=1}^n \sum_{t=2}^T \left(\widehat{\nu}_{it} - \widehat{\rho}\widehat{\nu}_{it-1}\right)^2 \\
&= \tfrac{1}{nT}\textstyle\sum_{i=1}^n \sum_{t=2}^T [\rho\nu_{it-1} + e_{it} - \overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)(x_{it-1} + \varepsilon_{it} - \bar{x}_{i.}) \\
&\quad -\hat{\rho}(\nu_{it-1} - \overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)(x_{it-1} - \bar{x}_{i.}))]^2 \\
&= \tfrac{1}{nT}\textstyle\sum_{i=1}^n \sum_{t=2}^T [e_{it} - (\hat{\rho} - \rho)\left\{\nu_{it-1} - (\hat{\beta}_{FE} - \beta)x_{it-1}\right\} \\
&\quad +(\hat{\rho} - \rho)\left\{\overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)\bar{x}_{i.}\right\} \\
&\quad +(\rho - 1)\left\{\overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)\bar{x}_{i.}\right\} - (\hat{\beta}_{FE} - \beta)\left\{\varepsilon_{it} + (1-\rho)x_{it-1}\right\}]^2 \\
&= \tfrac{1}{nT}\textstyle\sum_{i=1}^n \sum_{t=2}^T e_{it}^2 + \tfrac{1}{nT}\sum_{i=1}^n \sum_{t=2}^T \left[(\hat{\rho} - \rho)\left\{\nu_{it-1} - (\hat{\beta}_{FE} - \beta)x_{it-1}\right\}\right]^2 \\
&\quad +\tfrac{1}{nT}\textstyle\sum_{i=1}^n \sum_{t=2}^T \left[(\hat{\rho} - \rho)\left\{\overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)\bar{x}_{i.}\right\}\right]^2 \\
&\quad +\tfrac{1}{nT}\textstyle\sum_{i=1}^n \sum_{t=2}^T \left[(\rho - 1)\left\{\overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)\bar{x}_{i.}\right\}\right]^2 \\
&\quad +\tfrac{1}{nT}\textstyle\sum_{i=1}^n \sum_{t=2}^T \left[(\hat{\beta}_{FE} - \beta)\left\{\varepsilon_{it} + (1-\rho)x_{it-1}\right\}\right]^2 \\
&\quad -\tfrac{2}{nT}\textstyle\sum_{i=1}^n \sum_{t=2}^T e_{it}(\hat{\rho} - \rho)\left\{\nu_{it-1} - (\hat{\beta}_{FE} - \beta)x_{it-1}\right\} \\
&\quad +\tfrac{2}{nT}\textstyle\sum_{i=1}^n \sum_{t=2}^T e_{it}(\hat{\rho} - \rho)\left\{\overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)\bar{x}_{i.}\right\} \\
&\quad +\tfrac{2}{nT}\textstyle\sum_{i=1}^n \sum_{t=2}^T e_{it}(\rho - 1)\left\{\overline{\nu}_{i.} - (\hat{\beta}_{FE} - \beta)\bar{x}_{i.}\right\} \\
&\quad -\tfrac{2}{nT}\textstyle\sum_{i=1}^n \sum_{t=2}^T e_{it}(\hat{\beta}_{FE} - \beta)\left\{\varepsilon_{it} + (1-\rho)x_{it-1}\right\} + o_p(1) \\
&= I + II + III + IV + V + VI + VII + VIII + VIIII + o_p(1)
\end{aligned}$$

and it can be shown that

$$\widehat{\sigma}_e^2 = I + o_p(1).$$

Let us consider II, for example.

$$\frac{\left\{\sqrt{nT}(\hat{\rho} - \rho)\right\}^2}{(nT)^2}\sum_{i=1}^n \sum_{t=2}^T (\nu_{it-1}^2 - 2(\hat{\beta}_{FE} - \beta)\nu_{it-1}x_{it-1} + (\hat{\beta}_{FE} - \beta)^2 x_{it-1}^2) = O_p\left(\frac{1}{nT}\right)$$

because

$$\frac{1}{(nT)^2}\sum_{i=1}^n \sum_{t=2}^T \nu_{it}^2 = \frac{1}{nT}\frac{1}{nT}\sum_{i=1}^n \sum_{t=2}^T \nu_{it}^2 = O_p\left(\frac{1}{nT}\right)$$

using $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\nu_{it}^{2}\xrightarrow{p}\frac{\sigma_e^2}{1-\rho^2}=O_p(1)$.

Also

$$\frac{2}{n^{3/2}T^2}\sqrt{n}T(\hat{\beta}_{FE}-\beta)\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=2}^{T}\frac{\nu_{it}}{\sqrt{T}}\frac{x_{it-1}}{\sqrt{T}}\right)=O_p\left(\frac{1}{n^{3/2}T^2}\right)$$

using $\frac{1}{n}\sum_{i=1}^{n}\sum_{t=2}^{T}\frac{\nu_{it}}{\sqrt{T}}\frac{x_{it-1}}{\sqrt{T}}=O_p(1)$

and

$$\begin{aligned}&\frac{1}{(nT)^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\hat{\beta}_{FE}-\beta)^2x_{it-1}^2\\ =\;&\frac{(\sqrt{n}T(\hat{\beta}_{FE}-\beta))^2}{n^2T^2}\left(\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}^2\right)=O_p\left(\frac{1}{n^2T^2}\right)\end{aligned}$$

as $(n,T)\to\infty$ with the joint limit argument.

By a similar process, we have

$$\widehat{\sigma}_e^2=\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}^2+o_p(1)\xrightarrow{p}\sigma_e^2.$$

as $(n,T)\to\infty$.

2. $\rho=1$ case

   (a) $\rho=1$, $|\lambda|<1$ case

   Since we have $\rho=1$,

$$\widehat{\rho}-1=\frac{\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{\nu}_{it}-\widehat{\nu}_{it-1}\right)\widehat{\nu}_{it-1}}{\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^2}.$$

For the denominator,

$$\begin{aligned}
\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^2 &= \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[(\nu_{it-1}-\overline{\nu}_{i.})-(\hat{\beta}_{FE}-\beta)(x_{it-1}-\bar{x}_{i.})\right]^2 \\
&= \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})^2 \\
&\quad +\frac{1}{nT}\left\{\sqrt{n}(\hat{\beta}_{FE}-\beta)\right\}^2\left\{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1}-\bar{x}_{i.})^2\right\} \\
&\quad -\sqrt{n}(\hat{\beta}_{FE}-\beta)\frac{2}{n^{3/2}T^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})(x_{it-1}-\bar{x}_{i.}) \\
&= I+II+III.
\end{aligned}$$

Consider II first. Using a similar argument as in Phillips and Moon (1999) and Baltagi *et al.* (2008), it can be shown that $II = O_p\left(\frac{1}{nT}\right)$ if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, which yields

$$\sqrt{n}\left(\hat{\beta}_{FE}-\beta\right) = O_p\left(1\right).$$

Also as $(n,T)\to\infty$, we have

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it-1}-\bar{x}_{i.})^2 = O_p(1).$$

Consider III. It can be shown that

$$\begin{aligned}
III &= \sqrt{n}(\hat{\beta}_{FE}-\beta)\frac{2}{n^{3/2}T^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})(x_{it-1}-\bar{x}_{i.}) \\
&= \frac{1}{nT}\sqrt{n}(\hat{\beta}_{FE}-\beta)\frac{2}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=2}^{T}\frac{\nu_{it-1}-\overline{\nu}_{i.}}{\sqrt{T}}\frac{x_{it-1}-\bar{x}_{i.}}{\sqrt{T}} \\
&= O_p(\frac{1}{nT})
\end{aligned}$$

as $(n,T)\to\infty$ using

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=2}^{T}\frac{\nu_{it-1}-\overline{\nu}_{i.}}{\sqrt{T}}\frac{x_{it-1}-\bar{x}_{i.}}{\sqrt{T}}=O_p\left(1\right)$$

and

$$\sqrt{n}(\hat{\beta}_{FE}-\beta)=O_p\left(1\right).$$

Hence, we have

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^2=\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})^2+O_p\left(\frac{1}{nT}\right)\xrightarrow{p}\frac{\sigma_e^2}{6}$$

by equation (C.3) in Kao (1999).

For the numerator, $\widehat{\nu}_{it}-\widehat{\nu}_{it-1}=e_{it}-(\hat{\beta}_{FE}-\beta)\left\{(\lambda-1)x_{it-1}+\varepsilon_{it}\right\}$. One can show that

$$\begin{aligned}
&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{\nu}_{it}-\widehat{\nu}_{it-1}\right)\widehat{\nu}_{it-1}\\
&=\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\begin{array}{c}\left[e_{it}-(\hat{\beta}_{FE}-\beta)\left\{(\lambda-1)x_{it-1}+\varepsilon_{it}\right\}\right]\\ \cdot\left[(\nu_{it-1}-\overline{\nu}_{i.})-(\hat{\beta}_{FE}-\beta)(x_{it-1}-\overline{x}_{i.})\right]\end{array}\right]\\
&=\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\begin{array}{c}e_{it}\left(\nu_{it-1}-\overline{\nu}_{i.}\right)-(\hat{\beta}_{FE}-\beta)\left\{(\lambda-1)x_{it-1}+\varepsilon_{it}\right\}\left(\nu_{it-1}-\overline{\nu}_{i.}\right)\\ -(\hat{\beta}_{FE}-\beta)e_{it}(x_{it-1}-\overline{x}_{i.})\\ +(\hat{\beta}_{FE}-\beta)^2\left\{(\lambda-1)x_{it-1}+\varepsilon_{it}\right\}\left(x_{it-1}-\overline{x}_{i.}\right)\end{array}\right]\\
&=I+II+III+IV.
\end{aligned}$$

Consider II. It can be easily shown that

$$\begin{aligned}
II &= \frac{\sqrt{n}(\hat{\beta}_{FE}-\beta)}{n^{3/2}T}\sum_{i=1}^{n}\sum_{t=2}^{T}\left\{(\lambda-1)x_{it-1}+\varepsilon_{it}\right\}\left(\nu_{it-1}-\overline{\nu}_{i.}\right) \\
&= \frac{\sqrt{n}(\hat{\beta}_{FE}-\beta)}{n^{3/2}T}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\begin{array}{c}(\lambda-1)x_{it-1}\nu_{it-1}-(\lambda-1)x_{it-1}\overline{\nu}_{i.} \\ +\varepsilon_{it}\nu_{it-1}-\varepsilon_{it}\overline{\nu}_{i.}\end{array}\right] \\
&= O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{n}\right) \\
&= O_p\left(\frac{1}{n}\right)
\end{aligned}$$

using

$$\begin{aligned}
\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}\nu_{it-1} &= O_p(1), \\
\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=2}^{T}\varepsilon_{it}\nu_{it-1} &= O_p(1),
\end{aligned}$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}x_{it-1}\right)\left(\frac{1}{T^{3/2}}\sum_{t=2}^{T}\nu_{it-1}\right)=O_p(1).$$

Consider III. It is also easy to see that

$$\begin{aligned}
III &= -\frac{\sqrt{n}(\hat{\beta}_{FE}-\beta)}{n^{3/2}T}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}(x_{it-1}-\overline{x}_{i.}) \\
&= O_p\left(\frac{1}{n\sqrt{T}}\right)+O_p\left(\frac{1}{nT^{3/2}}\right)=O_p\left(\frac{1}{n\sqrt{T}}\right).
\end{aligned}$$

Consider IV.

$$
\begin{aligned}
IV &= \frac{\left(\sqrt{n}(\hat{\beta}_{FE}-\beta)\right)^2}{n^2T}\sum_{i=1}^{n}\sum_{t=2}^{T}\left\{(\lambda-1)x_{it-1}+\varepsilon_{it}\right\}(x_{it-1}-\overline{x}_{i.}) \\
&= \frac{\left(\sqrt{n}(\hat{\beta}_{FE}-\beta)\right)^2}{n^2T}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\begin{array}{c}(\lambda-1)x_{it-1}^2-(\lambda-1)x_{it-1}\overline{x}_{i.} \\ +\varepsilon_{it}x_{it-1}-\varepsilon_{it}\overline{x}_{i.}\end{array}\right] \\
&= O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{nT}\right)+O_p\left(\frac{1}{n^{3/2}\sqrt{T}}\right)+O_p\left(\frac{1}{nT}\right) \\
&= O_p\left(\frac{1}{n}\right)
\end{aligned}
$$

using

$$
\begin{aligned}
\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}^2 &= O_p(1), \\
\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}x_{it-1}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}x_{it}\right) &= O_p(1),
\end{aligned}
$$

and

$$
\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\varepsilon_{it-1}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}x_{it}\right)=O_p(1).
$$

We conclude that

$$
\begin{aligned}
&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(\widehat{\nu}_{it}-\widehat{\nu}_{it-1})\widehat{\nu}_{it-1} \\
&= I+O_p\left(\frac{1}{nT}\right)+O_p\left(\frac{1}{n\sqrt{T}}\right)+O_p\left(\frac{1}{n}\right) \\
&= I+O_p\left(\frac{1}{n}\right) \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[(\nu_{it}-\overline{\nu}_{i.})e_{it}\right]+o_p(1)\xrightarrow{p}-\frac{\sigma_e^2}{2}
\end{aligned}
$$

using equation (C.5) in Kao (1999). Hence,

$$T\left(\widehat{\rho}-1\right)=\frac{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\Delta\widehat{\nu}_{it}\widehat{\nu}_{it-1}}{\frac{1}{nT^{2}}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^{2}}\xrightarrow{p}-\frac{\frac{\sigma_{e}^{2}}{2}}{\frac{\sigma_{e}^{2}}{6}}=-3.$$

Next we consider $\widehat{\sigma}_e^2$.

$$\begin{aligned}
\widehat{\sigma}_{e}^{2} &= \tfrac{1}{nT}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{\nu}_{it}-\widehat{\rho}\widehat{\nu}_{it-1}\right)^{2}\\
&= \tfrac{1}{nT}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\begin{array}{c}\nu_{it-1}+e_{it}-\overline{\nu}_{i.}-(\hat{\beta}_{FE}-\beta)(\lambda x_{it-1}+\varepsilon_{it}-\bar{x}_{i.})\\ -\hat{\rho}\left\{\nu_{it-1}-\overline{\nu}_{i.}-(\hat{\beta}_{FE}-\beta)(x_{it-1}-\bar{x}_{i.})\right\}\end{array}\right]^{2}\\
&= \tfrac{1}{nT}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\begin{array}{c}e_{it}-(\hat{\rho}-1)\left\{\nu_{it-1}-(\hat{\beta}_{FE}-\beta)x_{it-1}\right\}\\ +(\hat{\rho}-1)\left\{\overline{\nu}_{i.}-(\hat{\beta}_{FE}-\beta)\bar{x}_{i.}\right\}\\ -(\hat{\beta}_{FE}-\beta)\left\{\varepsilon_{it}+(\lambda-1)x_{it-1}\right\}\end{array}\right]^{2}\\
&= \tfrac{1}{nT}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}^{2}+\tfrac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[(\hat{\rho}-1)\left\{\nu_{it-1}-(\hat{\beta}_{FE}-\beta)x_{it-1}\right\}\right]^{2}\\
&\quad+\tfrac{1}{nT}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}\left[(\hat{\rho}-1)\left\{\overline{\nu}_{i.}-(\hat{\beta}_{FE}-\beta)\bar{x}_{i.}\right\}\right]^{2}\\
&\quad+\tfrac{1}{nT}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}\left[(\hat{\beta}_{FE}-\beta)\left\{\varepsilon_{it}+(\lambda-1)x_{it-1}\right\}\right]^{2}\\
&\quad-\tfrac{2}{nT}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}(\hat{\rho}-1)\left\{\nu_{it-1}-(\hat{\beta}_{FE}-\beta)x_{it-1}\right\}\\
&\quad+\tfrac{2}{nT}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}(\hat{\rho}-1)\left\{\overline{\nu}_{i.}-(\hat{\beta}_{FE}-\beta)\bar{x}_{i.}\right\}\\
&\quad-\tfrac{2}{nT}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}(\hat{\beta}_{FE}-\beta)\left\{\varepsilon_{it}+(\lambda-1)x_{it-1}\right\}+o_{p}(1)\\
&= I+II+III+IV+V+VI+VII+VIII+VIIII+o_{p}(1).
\end{aligned}$$

From above, it can be shown that

$$\widehat{\sigma}_{e}^{2}=\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}^{2}+o_{p}(1)\xrightarrow{p}\sigma_{e}^{2}$$

as $(n,T)\to\infty$. We illustrate II only as an example. It can be shown that

$$\begin{aligned}&\{T(\hat{\rho}-1)\}^2\frac{1}{nT^3}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}^2-2(\hat{\beta}_{FE}-\beta)\nu_{it-1}x_{it-1}+(\hat{\beta}_{FE}-\beta)^2x_{it-1}^2)\\ =\ & O_p(\frac{1}{\sqrt{n}T})\end{aligned}$$

as $(n,T)\to\infty$ because

$$\frac{1}{nT^3}\sum_{i=1}^{n}\sum_{t=2}^{T}\nu_{it-1}^2=O_p(\frac{1}{T})$$

with $\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\nu_{it-1}^2=O_p(1)$ and

$$\left(\sqrt{n}(\hat{\beta}_{FE}-\beta)\right)^2\frac{1}{nT^2}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}^2=O_p(\frac{1}{nT^2})$$

with $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}^2=O_p(1)$.

Also note that

$$-2\sqrt{n}(\hat{\beta}_{FE}-\beta)\frac{1}{nT^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=2}^{T}\frac{\nu_{it-1}}{\sqrt{T}}\frac{x_{it-1}}{\sqrt{T}}=O_p(\frac{1}{nT^2})$$

using $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=2}^{T}\frac{\nu_{it-1}}{\sqrt{T}}\frac{x_{it-1}}{\sqrt{T}}=O_p(1)$.

(b) $\rho=1$, $\lambda=1$ case

Recall that

$$\widehat{\rho}-1=\frac{\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{\nu}_{it}-\widehat{\nu}_{it-1}\right)\widehat{\nu}_{it-1}}{\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^2}.$$

We have

$$\begin{aligned}
\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^2 &= \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[(\nu_{it-1}-\overline{\nu}_{i.})-(\hat{\beta}_{FE}-\beta)(x_{it-1}-\bar{x}_{i.})\right]^2 \\
&= \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})^2 \\
&\quad +\frac{1}{n}\left\{\sqrt{n}(\hat{\beta}_{FE}-\beta)\right\}^2\left\{\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1}-\bar{x}_{i.})^2\right\} \\
&\quad -\sqrt{n}(\hat{\beta}_{FE}-\beta)\frac{2}{n^{3/2}T^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})(x_{it-1}-\bar{x}_{i.}) \\
&= I+II+III.
\end{aligned}$$

Consider II first. Using a similar argument as in Phillips and Moon (1999) and Baltagi *et al.* (2008), it can be shown that $II = O_p(\frac{1}{n})$ if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, which yields

$$\sqrt{n}\left(\hat{\beta}_{FE}-\beta\right)\xrightarrow{d} N\left(0,\frac{2\sigma_e^2}{5\sigma_\varepsilon^2}\right) = O_p\left(1\right).$$

Also as $(n,T)\to\infty$, we have

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it-1}-\bar{x}_{i.})^2 \xrightarrow{p} \frac{\sigma_\varepsilon^2}{6} = O_p(1)$$

see equation (C.3) in Kao (1999).

Next consider III.

$$\begin{aligned}
III &= \sqrt{n}(\hat{\beta}_{FE}-\beta)\frac{2}{n^{3/2}T^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})(x_{it-1}-\bar{x}_{i.}) \\
&= \sqrt{n}(\hat{\beta}_{FE}-\beta)\frac{2}{n}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=2}^{T}(\frac{\nu_{it-1}-\overline{\nu}_{i.}}{\sqrt{T}})(\frac{x_{it-1}-\bar{x}_{i.}}{\sqrt{T}})\frac{1}{T} \\
&= O_p(\frac{1}{n})
\end{aligned}$$

as $(n,T)\to\infty$ since

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=2}^{T}(\frac{\nu_{it-1}-\overline{\nu}_{i.}}{\sqrt{T}})(\frac{x_{it-1}-\bar{x}_{i.}}{\sqrt{T}})\frac{1}{T}=O_p\left(1\right)$$

where ν_{it} and x_{it} are not correlated.

Hence, we have

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^2=\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})^2+o_p(1)\xrightarrow{p}\frac{\sigma_e^2}{6}$$

by equation (C.3) in Kao (1999).

For the numerator, $\widehat{\nu}_{it}-\widehat{\nu}_{it-1}=e_{it}-(\hat{\beta}_{FE}-\beta)\varepsilon_{it}$, and it can be shown that

$$\begin{aligned}
&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{\nu}_{it}-\widehat{\nu}_{it-1}\right)\widehat{\nu}_{it-1}\\
=\;&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[e_{it}-(\hat{\beta}_{FE}-\beta)\varepsilon_{it}\right]\left[(\nu_{it-1}-\overline{\nu}_{i.})-(\hat{\beta}_{FE}-\beta)(x_{it-1}-\bar{x}_{i.})\right]\\
=\;&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\begin{array}{c}(\nu_{it-1}-\overline{\nu}_{i.})e_{it}-(\hat{\beta}_{FE}-\beta)(x_{it-1}-\bar{x}_{i.})e_{it}\\ -(\hat{\beta}_{FE}-\beta)(\nu_{it-1}-\overline{\nu}_{i.})\varepsilon_{it}+(\hat{\beta}_{FE}-\beta)^2\varepsilon_{it}(x_{it-1}-\bar{x}_{i.})\end{array}\right]\\
=\;&I+II+III+IV.
\end{aligned}$$

Consider II. It can be shown that

$$\begin{aligned}
II &= -\frac{\sqrt{n}(\hat{\beta}_{FE}-\beta)}{n^{3/2}T}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1}-\bar{x}_{i.})e_{it}\\
&= -\frac{\sqrt{n}(\hat{\beta}_{FE}-\beta)}{n^{3/2}}\frac{1}{T}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}e_{it}\\
&\quad+\frac{\sqrt{n}(\hat{\beta}_{FE}-\beta)}{n^{3/2}}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}e_{it}\right)\left(\frac{1}{T^{3/2}}\sum_{t=2}^{T}x_{it}\right)\\
&= O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{n}\right)=O_p\left(\frac{1}{n}\right)
\end{aligned}$$

using

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}e_{it}=O_p\left(1\right)$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}e_{it}\right)\left(\frac{1}{T^{3/2}}\sum_{t=2}^{T}x_{it}\right)=O_p\left(1\right).$$

Consider III and IV. In a similar vein as II, one can see that

$$III=-\frac{\sqrt{n}(\hat{\beta}_{FE}-\beta)}{n^{3/2}T}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\nu_{it-1}\varepsilon_{it}-\overline{\nu}_{i.}\varepsilon_{it}\right]=O_p\left(\frac{1}{n}\right)$$

and

$$IV=\frac{\left(\sqrt{n}(\hat{\beta}_{FE}-\beta)\right)^2}{n^2T}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\varepsilon_{it}x_{it-1}-\varepsilon_{it}\bar{x}_{i.}\right]=O_p\left(\frac{1}{n}\right).$$

We conclude that

$$\begin{aligned}&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{\nu}_{it}-\widehat{\nu}_{it-1}\right)\widehat{\nu}_{it-1}\\&=\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\nu}_{i.})e_{it}+O_p\left(\frac{1}{n}\right)\xrightarrow{p}-\frac{\sigma_e^2}{2}\end{aligned}$$

using equation (C.5) in Kao (1999). Combining these results, we get

$$T\left(\widehat{\rho}-1\right)=\frac{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\Delta\widehat{\nu}_{it}\widehat{\nu}_{it-1}}{\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^2}\xrightarrow{p}-\frac{\frac{\sigma_e^2}{2}}{\frac{\sigma_e^2}{6}}=-3$$

which is the required result.

Next we consider $\widehat{\sigma}_e^2$. Note that

$$\begin{aligned}
\widehat{\sigma}_e^2 &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{\nu}_{it}-\widehat{\rho}\widehat{\nu}_{it-1}\right)^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left\{\begin{array}{c} \nu_{it-1}+e_{it}-\overline{\nu}_{i.}-(\hat{\beta}_{FE}-\beta)(x_{it-1}+\varepsilon_{it}-\bar{x}_{i.}) \\ -\hat{\rho}(\nu_{it-1}-\overline{\nu}_{i.}-(\hat{\beta}_{FE}-\beta)(x_{it-1}-\bar{x}_{i.})) \end{array}\right\}^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left\{\begin{array}{c} e_{it}-(\hat{\rho}-1)\left\{\nu_{it-1}-(\hat{\beta}_{FE}-\beta)x_{it-1}\right\} \\ +(\hat{\rho}-1)\left\{\overline{\nu}_{i.}-(\hat{\beta}_{FE}-\beta)\bar{x}_{i.}\right\}-(\hat{\beta}_{FE}-\beta)\varepsilon_{it} \end{array}\right\}^2.
\end{aligned}$$

After some tedious algebra, it can be easily shown that

$$\widehat{\sigma}_e^2 = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}^2 + o_p(1) \xrightarrow{p} \sigma_e^2.$$

■

B Proof of Theorem 1

Proof. Now we are ready to prove Theorem 1.

1. $|\rho|<1,\ |\lambda|<1$ case

   Recall

   $$S_{FE} = \sqrt{\frac{\widehat{\sigma}_\nu^2}{\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\bar{x}_{i.}\right)^2}}.$$

   Now note that

   $$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it}-\bar{x}_{i.})^2 \xrightarrow{p} \frac{\sigma_\varepsilon^2}{1-\lambda^2}$$

   and

   $$\widehat{\sigma}_v^2 = \frac{\hat{\sigma}_e^2}{1-\widehat{\rho}^2} \xrightarrow{p} \frac{\sigma_e^2}{(1-\rho^2)}$$

   since $1-\hat{\rho}^2 = (1-\rho^2)+(1-\rho)(\hat{\rho}-\rho)-(\hat{\rho}-\rho)^2 = (1-\rho^2)+o_p(1)$ using

$\hat{\rho} - \rho = O_p\left(\frac{1}{\sqrt{nT}}\right)$.

Therefore, if $(n, T) \to \infty$ and $\frac{n}{T} \to 0$, then

$$t_{FE} = \frac{\sqrt{nT}(\hat{\beta}_{FE} - \beta)}{\sqrt{\widehat{\sigma}_\nu^2 / \left(\frac{1}{nT}\sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_{i.})^2\right)}} \xrightarrow{d} \frac{N\left(0, \frac{(1+\rho\lambda)(1-\lambda^2)\sigma_e^2}{(1-\rho\lambda)(1-\rho^2)\sigma_\varepsilon^2}\right)}{\sqrt{\left(\frac{\sigma_e^2}{1-\rho^2}\right) / \left(\frac{\sigma_\varepsilon^2}{1-\lambda^2}\right)}} = N(0, \frac{1+\rho\lambda}{1-\rho\lambda}).$$

2. $\rho = 1$, $|\lambda| < 1$ case

   Consider $\widehat{\sigma}_\nu^2$.

$$\frac{\widehat{\sigma}_\nu^2}{T} = \frac{\hat{\sigma}_e^2}{T(1-\widehat{\rho}^2)} = \frac{\hat{\sigma}_e^2}{T(1-\hat{\rho})(1+\hat{\rho})} \xrightarrow{p} -\frac{\sigma_e^2}{6}.$$

   From the construction of S_{FE} , it is obvious that we cannot obtain s_{FE} because we have a complex number problem. Accordingly we cannot have t_{FE} either.

3. $|\rho| < 1$, $\lambda = 1$ case

   Because

$$\frac{1}{nT^2}\sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_{i.})^2 \xrightarrow{p} \frac{\sigma_\varepsilon^2}{6}$$

   and

$$\widehat{\sigma}_\nu^2 = \frac{\hat{\sigma}_e^2}{1-\widehat{\rho}^2} \xrightarrow{p} \frac{\sigma_e^2}{(1-\rho^2)}$$

   as shown in Theorem 1.(1), one can easily verify that if $(n, T) \to \infty$ and $\frac{n}{T} \to 0$, then

$$t_{FE} = \frac{\sqrt{n}T(\hat{\beta}_{FE} - \beta)}{\sqrt{\widehat{\sigma}_\nu^2 / \left(\frac{1}{nT^2}\sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_{i.})^2\right)}} \xrightarrow{d} \frac{N\left(0, \frac{6\sigma_e^2}{(1-\rho)^2\sigma_\varepsilon^2}\right)}{\sqrt{\left(\frac{\sigma_e^2}{1-\rho^2}\right) / \left(\frac{\sigma_\varepsilon^2}{6}\right)}} = N(0, \frac{1+\rho}{1-\rho}).$$

4. $\rho = 1$, $\lambda = 1$ case

In a similar vein as Theorem 1.(2), since

$$\frac{\widehat{\sigma}_{\nu}^{2}}{T}=\frac{\widehat{\sigma}_{e}^{2}}{T(1-\hat{\rho}^{2})}=\frac{\widehat{\sigma}_{e}^{2}}{T(1-\hat{\rho})(1+\hat{\rho})}\xrightarrow{p}-\frac{\sigma_{e}^{2}}{6}$$

we cannot obtain s_{FE} or t_{FE} either.

■

C Proof of Theorem 2

Proof. Now we prove Theorem 2.

1. $|\rho|<1,\ |\lambda|<1$ case

   Recall

$$\Delta y_{it}=\beta\Delta x_{it}+\Delta\nu_{it}$$

   with $\Delta y_{it}-\hat{\beta}_{FD}\Delta x_{it}=\Delta\nu_{it}-(\hat{\beta}_{FD}-\beta)\triangle x_{it}$.

   Let us look at $\hat{\sigma}_{\triangle\nu}^{2}$ first.

$$\begin{aligned}\hat{\sigma}_{\triangle\nu}^{2} &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left\{\Delta\nu_{it}-(\hat{\beta}_{FD}-\beta)\triangle x_{it}\right\}^{2}\\ &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\{(\rho-1)\nu_{it-1}+e_{it}\}-(\hat{\beta}_{FD}-\beta)\{(\lambda-1)x_{it-1}+\varepsilon_{it}\}\right]^{2}\\ &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}[(\rho-1)\nu_{it-1}+e_{it}]^{2}\\ &\quad+\frac{1}{n^{2}T^{2}}[\sqrt{nT}(\hat{\beta}_{FD}-\beta)]^{2}\sum_{i=1}^{n}\sum_{t=2}^{T}[(\lambda-1)x_{it-1}+\varepsilon_{it}]^{2}\\ &\quad-\frac{2}{nT}\left[\sqrt{nT}(\hat{\beta}_{FD}-\beta)\right]\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}[(\rho-1)\nu_{it-1}+e_{it}][(\lambda-1)x_{it-1}+\varepsilon_{it}]\\ &= I+II+III.\end{aligned}$$

Consider I. It can be shown that

$$\begin{aligned} I &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[(\rho-1)\nu_{it-1}+e_{it}\right]^2 \\ &= \frac{(\rho-1)^2}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\nu_{it-1}^2+\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}^2+\frac{2(\rho-1)}{\sqrt{nT}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\nu_{it-1}e_{it} \\ &= (\rho-1)^2\frac{\sigma_e^2}{1-\rho^2}+\sigma_e^2+O_p(\frac{1}{\sqrt{nT}})=\frac{2\sigma_e^2}{1+\rho} \end{aligned}$$

as $(n,T)\to\infty$ using

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\nu_{it-1}e_{it}=O_p(1).$$

For II and III, one can verify that $II=O_p\left(\frac{1}{nT}\right)$ and $III=O_p\left(\frac{1}{nT}\right)$ using the fact that if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$\sqrt{nT}(\hat{\beta}_{FD}-\beta)\xrightarrow{d}N(0,\frac{(1+\lambda)^2\left[(2-\rho-\lambda)^2+\frac{(1-\rho)^3}{1+\rho}+\frac{(1-\lambda)^3}{1+\lambda}\right]\sigma_e^2}{4(1-\rho\lambda)^2\sigma_\varepsilon^2})=O_p(1).$$

This uses a similar argument as in Phillips and Moon (1999), also Corollary 5.1 in Baltagi *et al.* (2008).

Hence, we have

$$\begin{aligned} \hat{\sigma}_{\triangle\nu}^2 &= (\rho-1)^2\frac{\sigma_e^2}{1-\rho^2}+\sigma_e^2+O_p(\frac{1}{\sqrt{nT}})+O_p\left(\frac{1}{nT}\right)+O_p\left(\frac{1}{nT}\right) \\ &= (\rho-1)^2\frac{\sigma_e^2}{1-\rho^2}+\sigma_e^2+O_p(\frac{1}{\sqrt{nT}}) \\ &\xrightarrow{p}\frac{2\sigma_e^2}{1+\rho}. \end{aligned}$$

Now recall that

$$t_{FD}=\frac{\widehat{\beta}_{FD}-\beta_0}{s_{FD}}$$

with $s_{FD} = \sqrt{\frac{\hat{\sigma}^2_{\triangle\nu}}{\sum_{i=1}^n \sum_{t=1}^T (x_{it}-x_{it-1})^2}}$. Here one can easily see that

$$\begin{aligned}
&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it}-x_{it-1})^2 \\
= \; &\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it}^2 + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}^2 - \frac{2}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it}x_{it-1} \xrightarrow{p} \frac{2\sigma_\varepsilon^2}{1+\lambda}
\end{aligned}$$

as $(n,T) \to \infty$. We conclude that if $(n,T) \to \infty$ and $\frac{n}{T} \to 0$, then

$$\begin{aligned}
t_{FD} &= \frac{\sqrt{nT}(\hat{\beta}_{FD}-\beta)}{\sqrt{\hat{\sigma}^2_{\triangle\nu}/\left(\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T (x_{it}-x_{it-1})^2\right)}} \\
&\xrightarrow{d} \frac{N(0, \frac{(1+\lambda)^2\left[(2-\rho-\lambda)^2+\frac{(1-\rho)^3}{1+\rho}+\frac{(1-\lambda)^3}{1+\lambda}\right]\sigma_e^2}{4(1-\rho\lambda)^2\sigma_\varepsilon^2})}{\sqrt{\left(\frac{2\sigma_e^2}{1+\rho}\right)/\left(\frac{2\sigma_\varepsilon^2}{1+\lambda}\right)}} \\
&= N\left(0, \frac{(1+\lambda)(1+\rho)\left[(2-\rho-\lambda)^2+\frac{(1-\rho)^3}{1+\rho}+\frac{(1-\lambda)^3}{1+\lambda}\right]}{4(1-\rho\lambda)^2}\right).
\end{aligned}$$

2. $\rho = 1$, $|\lambda| < 1$ case

From $\Delta y_{it} - \hat{\beta}_{FD}\Delta x_{it} = \Delta\nu_{it} - (\hat{\beta}_{FD}-\beta)\triangle x_{it}$, one can show that

$$\begin{aligned}
\hat{\sigma}^2_{\triangle v} &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left\{\Delta\nu_{it}-(\hat{\beta}_{FD}-\beta)\triangle x_{it}\right\}^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[e_{it}-(\hat{\beta}_{FD}-\beta)\left\{(\lambda-1)x_{it-1}+\varepsilon_{it}\right\}\right]^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}^2 + \frac{1}{n^2T^2}[\sqrt{nT}(\hat{\beta}_{FD}-\beta)]^2\sum_{i=1}^{n}\sum_{t=2}^{T}[(\lambda-1)x_{it-1}+\varepsilon_{it}]^2 \\
&\quad -\frac{2}{\sqrt{nT}}\left[\sqrt{nT}(\hat{\beta}_{FD}-\beta)\right]\frac{1}{n\sqrt{T}}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}\left[(\lambda-1)x_{it-1}+\varepsilon_{it}\right] \\
&= I + II + III.
\end{aligned}$$

Obviously, $I = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T} e_{it}^2 \xrightarrow{p} \sigma_e^2$.

For II, using a similar argument as in Phillips and Moon (1999) and Baltagi *et al.* (2008), it can be shown that $II = O_p\left(\frac{1}{nT}\right)$ because if $(n,T) \to \infty$ and $\frac{n}{T} \to 0$, which yields

$$\sqrt{nT}(\hat{\beta}_{FD} - \beta) \xrightarrow{d} N(0, \frac{(1+\lambda)\sigma_e^2}{2\sigma_\varepsilon^2}) = O_p(1).$$

Also note that

$$\frac{(\lambda-1)^2}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T} x_{it-1}^2 = O_p(1),$$
$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T} \varepsilon_{it}^2 = O_p(1),$$

and

$$\frac{2(\lambda-1)}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T} x_{it-1}\varepsilon_{it} = O_p(1).$$

For III, it is easy to see that

$$III = \frac{2}{nT}\left[\sqrt{nT}(\hat{\beta}_{FD} - \beta)\right]\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T} e_{it}\left[(\lambda-1)x_{it-1} + \varepsilon_{it}\right] = O_p(\frac{1}{nT})$$

using

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T} e_{it}x_{it-1} = O_p(1)$$

and

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T} e_{it}\varepsilon_{it} = O_p(1).$$

Then we have,

$$\begin{aligned}\hat{\sigma}^2_{\triangle v} &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T} e_{it}^2 + O_p\left(\frac{1}{nT}\right) + O_p(\frac{1}{nT}) \\ &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T} e_{it}^2 + O_p(\frac{1}{\sqrt{nT}}) \xrightarrow{p} \sigma_e^2.\end{aligned}$$

Also recall that

$$\begin{aligned}&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it}-x_{it-1})^2 \\ =\ &\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it}^2 + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}^2 - \frac{2}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it}x_{it-1} \xrightarrow{p} \frac{2\sigma_\varepsilon^2}{1+\lambda}\end{aligned}$$

as $(n,T)\to\infty$. Hence, we conclude that if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$t_{FD} = \frac{\sqrt{nT}(\hat{\beta}_{FD}-\beta)}{\sqrt{\hat{\sigma}^2_{\triangle\nu}/\left(\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it}-x_{it-1})^2\right)}} \xrightarrow{d} \frac{N(0,\frac{(1+\lambda)\sigma_e^2}{2\sigma_\varepsilon^2})}{\sqrt{\sigma_e^2/\left(\frac{2\sigma_\varepsilon^2}{1+\lambda}\right)}} = N(0,1).$$

3. $|\rho|<1,\ \lambda=1$ case

By a similar argument as above, it can be easily shown that

$$\begin{aligned}\hat{\sigma}^2_{\triangle v} &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left\{\Delta\nu_{it}-(\hat{\beta}_{FD}-\beta)\triangle x_{it}\right\}^2 \\ &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\{(\rho-1)\nu_{it-1}+e_{it}\}-(\hat{\beta}_{FD}-\beta)\varepsilon_{it}\right]^2 \\ &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}[(\rho-1)\nu_{it-1}+e_{it}]^2 + \frac{1}{n^2T^2}[\sqrt{nT}(\hat{\beta}_{FD}-\beta)]^2\sum_{i=1}^{n}\sum_{t=2}^{T}\varepsilon_{it}^2 \\ &\quad -\frac{2}{nT}\left[\sqrt{nT}(\hat{\beta}_{FD}-\beta)\right]\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}[(\rho-1)\nu_{it-1}+e_{it}]\,\varepsilon_{it} \\ &= \frac{2\sigma_e^2}{1+\rho}+O_p\left(\frac{1}{nT}\right)+O_p\left(\frac{1}{nT}\right) = \frac{2\sigma_e^2}{1+\rho}+O_p\left(\frac{1}{nT}\right).\end{aligned}$$

This is because if $(n,T) \to \infty$ and $\frac{n}{T} \to 0$, then

$$\sqrt{nT}(\hat{\beta}_{FD} - \beta) \xrightarrow{d} N\left(0, \frac{2\sigma_e^2}{(1+\rho)\sigma_\varepsilon^2}\right) = O_p(1)$$

using a similar argument as in Phillips and Moon (1999) and Baltagi *et al.* (2008).

Also recall that

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it} - x_{it-1})^2 = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\varepsilon_{it}^2 \xrightarrow{p} \sigma_\varepsilon^2.$$

Hence, if $(n,T) \to \infty$ and $\frac{n}{T} \to 0$, we have

$$t_{FD} = \frac{\sqrt{nT}(\hat{\beta}_{FD} - \beta)}{\sqrt{\hat{\sigma}_{\triangle\nu}^2 / \left(\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - x_{it-1})^2\right)}} \xrightarrow{d} \frac{N(0, \frac{2\sigma_e^2}{(1+\rho)\sigma_\varepsilon^2})}{\sqrt{\left(\frac{2\sigma_e^2}{1+\rho}\right)/\sigma_\varepsilon^2}} = N(0,1).$$

4. $\rho = 1$, $\lambda = 1$ case

Again, it can be easily shown that

$$\begin{aligned}
\hat{\sigma}_{\triangle\nu}^2 &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\Delta\nu_{it} - (\hat{\beta}_{FD} - \beta)\triangle x_{it}\right]^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}[e_{it} - (\hat{\beta}_{FD} - \beta)\varepsilon_{it}]^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}^2 + \frac{1}{n^2T^2}[\sqrt{nT}(\hat{\beta}_{FD} - \beta)]\sum_{i=1}^{n}\sum_{t=2}^{T}\varepsilon_{it}^2 \\
&\quad - \frac{2}{nT}[\sqrt{nT}(\hat{\beta}_{FD} - \beta)]\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}\varepsilon_{it} \\
&= \sigma_e^2 + O_p\left(\frac{1}{nT}\right) + O_p\left(\frac{1}{nT}\right) \\
&= \sigma_e^2 + O_p\left(\frac{1}{nT}\right) \xrightarrow{p} \sigma_e^2
\end{aligned}$$

because if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$\sqrt{nT}(\hat{\beta}_{FD}-\beta)\xrightarrow{d}N(0,\frac{\sigma_e^2}{\sigma_\varepsilon^2})=O_p(1)$$

using a similar argument as in Phillips and Moon (1999) and Baltagi *et al.* (2008).

Hence, we conclude that if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$t_{FD}=\frac{\sqrt{nT}(\hat{\beta}_{FD}-\beta)}{\sqrt{\hat{\sigma}^2_{\triangle\nu}/\left(\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(x_{it}-x_{it-1})^2\right)}}\xrightarrow{d}\frac{N(0,\sigma_e^2/\sigma_\varepsilon^2)}{\sqrt{\sigma_e^2/\sigma_\varepsilon^2}}=N(0,1).$$

■

D Proof of Lemma 2

The OLS estimator of β is given by

$$\widehat{\beta}_{OLS}=\frac{\sum_{i=1}^n\sum_{t=1}^T\left(x_{it}-\overline{\overline{x}}\right)\left(y_{it}-\overline{\overline{y}}\right)}{\sum_{i=1}^n\sum_{t=1}^T\left(x_{it}-\overline{\overline{x}}\right)^2}$$

where $\overline{\overline{x}}=\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T x_{it}$ and $\overline{\overline{y}}=\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T y_{it}$. Rewriting the equation, we have

$$\widehat{\beta}_{OLS}-\beta=\frac{\sum_{i=1}^n\sum_{t=1}^T\left(x_{it}-\overline{\overline{x}}\right)\left(v_{it}-\overline{\overline{v}}\right)}{\sum_{i=1}^n\sum_{t=1}^T\left(x_{it}-\overline{\overline{x}}\right)^2}.$$

Proof. We consider the denominator first and then move to the numerator to prove Lemma 2.[1]

1. The denominator

[1] Note that μ_i is not included in error term here.

(a) when $|\lambda| < 1$,

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\overline{\overline{x}}\right)^2 = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}^2 - \left(\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\right)^2 \xrightarrow{p} \frac{\sigma_\varepsilon^2}{1-\lambda^2}$$

as $(n,T)\to\infty$ using $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}^2 = \frac{\sigma_\varepsilon^2}{1-\lambda^2}$ and $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it} = O_p(1)$.

(b) When $\lambda = 1$,

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\overline{\overline{x}}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\frac{x_{it}}{\sqrt{T}}\right)^2\frac{1}{T} - \left[\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it}}{\sqrt{T}}\frac{1}{T}\right]^2 \xrightarrow{p} \frac{\sigma_\varepsilon^2}{2}$$

as $(n,T)\to\infty$.

2. The numerator

(a) If $|\rho| < 1$, $|\lambda| < 1$, it can be shown that

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\overline{\overline{x}}\right)\left(\nu_{it}-\overline{\overline{\nu}}\right) = \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\nu_{it} + O_p\left(\frac{1}{\sqrt{nT}}\right)$$

and

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\nu_{it} \xrightarrow{d} N\left(0, \frac{\sigma_\varepsilon^2}{(1-\lambda^2)}\frac{\sigma_e^2}{(1-\rho^2)}\frac{(1+\rho\lambda)}{(1-\rho\lambda)}\right)$$

as $(n,T)\to\infty$.

(b) If $\rho = 1$, $|\lambda| < 1$, it can be shown that

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\overline{\overline{x}}\right)\left(\nu_{it}-\overline{\overline{\nu}}\right) = \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\nu_{it} + O_p\left(\frac{1}{\sqrt{n}}\right)$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it}}{\sqrt{T}}\frac{\nu_{it}}{\sqrt{T}} \xrightarrow{d} N\left(0, \frac{\sigma_\varepsilon^2\sigma_e^2}{2(1-\lambda)^2}\right)$$

as $(n,T) \to \infty$.

(c) If $|\rho| < 1$, $\lambda = 1$, again we have

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\overline{\overline{x}}\right)\left(\nu_{it}-\overline{\overline{\nu}}\right) = \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\nu_{it} + O_p(\frac{1}{\sqrt{n}})$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it}}{\sqrt{T}}\frac{\nu_{it}}{\sqrt{T}} \xrightarrow{d} N(0, \frac{\sigma_e^2\sigma_\varepsilon^2}{2(1-\rho)^2})$$

as $(n,T) \to \infty$.

(d) If $\rho = 1$, $\lambda = 1$,

$$\frac{1}{\sqrt{n}T^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\overline{\overline{x}}\right)\left(\nu_{it}-\overline{\overline{\nu}}\right) = \frac{1}{\sqrt{n}T^2}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\nu_{it} + O_p(\frac{1}{\sqrt{n}}).$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it}}{\sqrt{T}}\frac{\nu_{it}}{\sqrt{T}}\frac{1}{T} \xrightarrow{d} N(0, \frac{\sigma_e^2\sigma_\varepsilon^2}{6})$$

as $(n,T) \to \infty$.

■

Using the results above, the proof of Lemma 2 is straightforward

E Proof of Lemma 3

In this section, we consider the limiting distribution of ρ using OLS residuals and we check the consistency of σ_e^2 under nonstationarity of both the error term and the regressor.

Proof. Assume $(n,T) \to \infty$.

1. $|\rho| < 1$ case

(a) $|\rho| < 1$, $|\lambda| < 1$ case

Recall

$$\widetilde{\rho} - \rho = \frac{\sum_{i=1}^{n}\sum_{t=2}^{T}(\widehat{u}_{it} - \rho\widehat{u}_{it-1})\widehat{u}_{it-1}}{\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it-1}^{2}}$$

where $\widehat{u}_{it} = (y_{it} - \bar{\bar{y}}) - \hat{\beta}_{OLS}(x_{it} - \bar{\bar{x}}) = \left(\nu_{it} - \bar{\bar{\nu}}\right) - \left(\hat{\beta}_{OLS} - \beta\right)(x_{it} - \bar{\bar{x}})$.

For the denominator,

$$\begin{aligned}
& \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it-1}^{2} \\
= \; & \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[(\nu_{it-1} - \bar{\bar{\nu}}) - (\hat{\beta}_{OLS} - \beta)(x_{it-1} - \bar{\bar{x}})\right]^{2} \\
= \; & \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1} - \bar{\bar{\nu}})^{2} \\
& + \frac{1}{nT}\left(\sqrt{nT}(\hat{\beta}_{OLS} - \beta)\right)^{2}\left\{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1} - \bar{\bar{x}})^{2}\right\} \\
& - \frac{2}{\sqrt{nT}}\left(\sqrt{nT}(\hat{\beta}_{OLS} - \beta)\right)\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1} - \bar{\bar{\nu}})(x_{it-1} - \bar{\bar{x}}) \\
= \; & I + II + III.
\end{aligned}$$

Consider II first. It is easy to see that $II = O_p\left(\frac{1}{nT}\right)$ because if $(n, T) \to \infty$ and $\frac{n}{T} \to 0$, then

$$\sqrt{nT}\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, \frac{(1+\rho\lambda)(1-\lambda^{2})\sigma_{e}^{2}}{(1-\rho\lambda)(1-\rho^{2})\sigma_{\varepsilon}^{2}}\right).$$

Also, by Lemma 2.(1), we have

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1} - \bar{\bar{x}})^{2} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it}^{2} - \left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it}\right]^{2} \xrightarrow{p} \frac{\sigma_{\varepsilon}^{2}}{1-\lambda^{2}}.$$

$III = O_p\left(\frac{1}{nT}\right)$ because

$$\frac{1}{nT}\left(\sqrt{nT}(\hat{\beta}_{OLS}-\beta)\right)\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\overline{\nu}})(x_{it-1}-\bar{\bar{x}}) = O_p\left(\frac{1}{nT}\right)$$

as $(n,T)\to\infty$ using

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\overline{\nu}})(x_{it-1}-\bar{\bar{x}}) = O_p(1)$$

since ν_{it} and x_{it} are not correlated.

Hence,

$$\begin{aligned}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it-1}^{2} &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(v_{it-1}-\bar{\bar{v}})^2 + O_p\left(\frac{1}{nT}\right) + O_p\left(\frac{1}{nT}\right)\\ &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(v_{it-1}-\bar{\bar{v}})^2 + O_p\left(\frac{1}{nT}\right)\xrightarrow{p}\frac{\sigma_e^2}{1-\rho^2}.\end{aligned}$$

Let us look at the numerator. Because $\widehat{u}_{it}-\rho\widehat{u}_{it-1} = e_{it}-(\hat{\beta}_{OLS}-\beta)\{(\lambda-\rho)x_{it-1}+\varepsilon_{it}\}+o_p(1)$, we have

$$\begin{aligned}&\tfrac{1}{\sqrt{nT}}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{u}_{it}-\rho\widehat{u}_{it-1}\right)\widehat{u}_{it-1}\\ =\ &\tfrac{1}{\sqrt{nT}}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\begin{array}{l}\left[e_{it}-(\hat{\beta}_{OLS}-\beta)\{(\lambda-\rho)x_{it-1}+\varepsilon_{it}\}\right]\\ \cdot\left[\left(\nu_{it-1}-\overline{\overline{\nu}}\right)-\left(\hat{\beta}_{OLS}-\beta\right)(x_{it-1}-\bar{\bar{x}})\right]\end{array}\right]\\ =\ &\begin{array}{l}\tfrac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\nu_{it-1}-\overline{\overline{\nu}}\right)e_{it}\\ -\sqrt{nT}(\hat{\beta}_{OLS}-\beta)\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\{(\lambda-\rho)x_{it-1}+\varepsilon_{it}\}\left(\nu_{it-1}-\overline{\overline{\nu}}\right)\\ -\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}(x_{it-1}-\bar{\bar{x}})e_{it}\\ +\left(\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2\frac{1}{n^{3/2}}\sum_{i=1}^{n}\frac{1}{T^{3/2}}\sum_{t=2}^{T}\{(\lambda-\rho)x_{it-1}+\varepsilon_{it}\}(x_{it-1}-\bar{\bar{x}})\end{array}\\ =\ &I+II+III+IV.\end{aligned}$$

Consider I. One can see that

$$
\begin{aligned}
I &= \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \left[\left(\nu_{it-1} - \overline{\overline{\nu}} \right) e_{it} \right] \\
&= \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \nu_{it-1} e_{it} - \frac{1}{\sqrt{nT}} \left[\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \nu_{it} \left(\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} e_{it} \right) \right] \\
&= \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \nu_{it-1} e_{it} + O_p \left(\frac{1}{\sqrt{nT}} \right) \xrightarrow{d} \sigma_e^2 N \left(0, \frac{1}{1-\rho^2} \right).
\end{aligned}
$$

Consider II.

$$
\begin{aligned}
II &= -\sqrt{nT} (\hat{\beta}_{OLS} - \beta) \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=2}^{T} \{ (\lambda - \rho) x_{it-1} + \varepsilon_{it} \} \left(\nu_{it-1} - \overline{\overline{\nu}} \right) \\
&= -\sqrt{nT} (\hat{\beta}_{OLS} - \beta) \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=2}^{T} \left[\begin{array}{c} (\lambda - \rho) x_{it-1} \nu_{it-1} - (\lambda - \rho) x_{it-1} \overline{\overline{\nu}} \\ + \varepsilon_{it} \nu_{it-1} - \varepsilon_{it} \overline{\overline{\nu}} \end{array} \right] \\
&= O_p \left(\frac{1}{\sqrt{nT}} \right) + O_p \left(\frac{1}{nT} \right) + O_p \left(\frac{1}{\sqrt{nT}} \right) + O_p \left(\frac{1}{nT} \right) = O_p \left(\frac{1}{\sqrt{nT}} \right)
\end{aligned}
$$

as $(n, T) \to \infty$ using

$$
\begin{gathered}
\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} x_{it-1} \nu_{it-1} = O_p(1), \\
\left(\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} x_{it-1} \right) \left(\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \nu_{it-1} \right) = O_p(1), \\
\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \varepsilon_{it} \nu_{it-1} = O_p(1),
\end{gathered}
$$

and

$$
\left(\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \varepsilon_{it} \right) \left(\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \nu_{it-1} \right) = O_p(1).
$$

Consider III. Using a similar argument, it can be shown that

$$
III = -\sqrt{nT} \left(\hat{\beta}_{OLS} - \beta \right) \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=2}^{T} (x_{it-1} - \bar{\bar{x}}) e_{it} = O_p \left(\frac{1}{\sqrt{nT}} \right).
$$

Consider IV.

$$\begin{aligned}
& \left(\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2 \frac{1}{n^{3/2}T^{3/2}} \sum_{i=1}^{n}\sum_{t=2}^{T} \left\{(\lambda-\rho)x_{it-1}+\varepsilon_{it}\right\}\left(x_{it-1}-\bar{\bar{x}}\right) \\
= & \left(\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2 \frac{1}{\sqrt{nT}}\frac{1}{nT} \sum_{i=1}^{n}\sum_{t=2}^{T} \left[\begin{array}{c} (\lambda-\rho)x_{it-1}^2-(\lambda-\rho)x_{it-1}\bar{\bar{x}} \\ +\varepsilon_{it}x_{it-1}-\varepsilon_{it}\bar{\bar{x}} \end{array}\right] \\
= & O_p\left(\frac{1}{\sqrt{nT}}\right)+O_p\left(\frac{1}{n^{3/2}T^{3/2}}\right)+O_p\left(\frac{1}{nT}\right)+O_p\left(\frac{1}{n^{3/2}T^{3/2}}\right) \\
= & O_p\left(\frac{1}{\sqrt{nT}}\right)
\end{aligned}$$

as $(n,T)\to\infty$.

Hence, we have

$$\begin{aligned}
& \frac{1}{\sqrt{nT}} \sum_{i=1}^{n}\sum_{t=2}^{T} \left(\widehat{u}_{it}-\rho\widehat{u}_{it-1}\right)\widehat{u}_{it-1}+O_p\left(\frac{1}{\sqrt{nT}}\right) \\
& \xrightarrow{d} \sigma_e^2 N\left(0,\frac{1}{1-\rho^2}\right)
\end{aligned}$$

as $(n,T)\to\infty$.

Therefore, we conclude that if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$\begin{aligned}
& \sqrt{nT}\left(\widetilde{\rho}-\rho\right) \\
= & \frac{\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{u}_{it}-\rho\widehat{u}_{it-1}\right)\widehat{u}_{it-1}}{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it-1}^2} \xrightarrow{d} \frac{\sigma_e^2 N\left(0,\frac{1}{1-\rho^2}\right)}{\frac{\sigma_e^2}{1-\rho^2}} = N\left(0,1-\rho^2\right).
\end{aligned}$$

Next we show $\widetilde{\sigma}_e^2$ is a consistent estimator. Note that

$$\widehat{\mathbf{u}}=\mathbf{y}-\widehat{\gamma}_{OLS}\boldsymbol{\iota}_{nT}-\mathbf{x}\hat{\beta}_{OLS}=E_{nT}\left[\boldsymbol{\nu}+\mathbf{x}\left(\hat{\beta}_{OLS}-\beta\right)\right]$$

where $E_{nT} = I_{nT} - \bar{J}_{nT}$ and $\bar{J}_{nT} = \boldsymbol{\iota}_{nT}\boldsymbol{\iota}_{nT}'/nT$. Hence,

$$\begin{aligned}
\widetilde{\sigma}_e^2 &= \frac{1}{n(T-1)}\widehat{\mathbf{u}}^{*\prime}\widehat{\mathbf{u}}^* \approx \frac{1}{nT}\widehat{\mathbf{u}}^{*\prime}\widehat{\mathbf{u}}^* \\
&= \frac{1}{nT}\widehat{\mathbf{u}}'\left(I_n \otimes \hat{C}'\right)\left(I_n \otimes \hat{C}\right)\widehat{\mathbf{u}} \\
&= \frac{1}{nT}\left[\boldsymbol{\nu} + \mathbf{x}\left(\hat{\beta}_{OLS} - \beta\right)\right]' E_{nT}\left(I_n \otimes \hat{C}'\hat{C}\right)E_{nT}\left[\boldsymbol{\nu} + \mathbf{x}\left(\hat{\beta}_{OLS} - \beta\right)\right] \\
&= \frac{1}{nT}\boldsymbol{\nu}' E_{nT}\left(I_n \otimes \hat{C}'\hat{C}\right)E_{nT}\boldsymbol{\nu} \\
&\quad + \frac{1}{nT}\left(\hat{\beta}_{OLS} - \beta\right)^2 \mathbf{x}' E_{nT}\left(I_n \otimes \hat{C}'\hat{C}\right)E_{nT}\mathbf{x} \\
&\quad + \frac{2}{nT}\left(\hat{\beta}_{OLS} - \beta\right)\mathbf{x}' E_{nT}\left(I_n \otimes \hat{C}'\hat{C}\right)E_{nT}\boldsymbol{\nu} \\
&= I + II + III
\end{aligned}$$

using $\widehat{\mathbf{u}}^* = \left(I_n \otimes \hat{C}\right)\widehat{\mathbf{u}}$ and $\widehat{\mathbf{u}} = E_{nT}\left[\boldsymbol{\nu} + \mathbf{x}\left(\hat{\beta}_{OLS} - \beta\right)\right]$.

To rearrange the terms, note that

$$\begin{aligned}
E_{nT} &= I_{nT} - \bar{J}_{nT} \\
&= E_n \otimes I_T + \bar{J}_n \otimes I_T - \bar{J}_n \otimes \bar{J}_T \\
&= E_n \otimes I_T + \bar{J}_n \otimes E_T
\end{aligned}$$

and accordingly

$$\begin{aligned}
& E_{nT}\left(I_n \otimes \hat{C}'\hat{C}\right)E_{nT} \\
&= \left(E_n \otimes I_T + \bar{J}_n \otimes E_T\right)\left(I_n \otimes \hat{C}'\hat{C}\right)\left(E_n \otimes I_T + \bar{J}_n \otimes E_T\right) \\
&= E_n \otimes \hat{C}'\hat{C} + \bar{J}_n \otimes E_T\hat{C}'\hat{C}E_T.
\end{aligned}$$

Consider I.

$$\begin{aligned}
I &= \frac{1}{nT}\boldsymbol{\nu}' E_{nT}\left(I_n \otimes \hat{C}'\hat{C}\right) E_{nT}\boldsymbol{\nu} \\
&= \frac{1}{nT}\boldsymbol{\nu}'\left(E_n \otimes \hat{C}'\hat{C}\right)\boldsymbol{\nu} + \frac{1}{nT}\boldsymbol{\nu}'\left(\bar{J}_n \otimes E_T\hat{C}'\hat{C}E_T\right)\boldsymbol{\nu} \\
&= \frac{1}{nT}\boldsymbol{\nu}'\left(I_n \otimes \hat{C}'\hat{C}\right)\boldsymbol{\nu} - \frac{1}{nT}\boldsymbol{\nu}'\left(\bar{J}_n \otimes \hat{C}'\hat{C}\right)\boldsymbol{\nu} + \frac{1}{nT}\boldsymbol{\nu}'\left(\bar{J}_n \otimes E_T\hat{C}'\hat{C}E_T\right)\boldsymbol{\nu} \\
&\approx \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\nu_{it} - \widetilde{\rho}\nu_{it-1}\right)^2 - \left\{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\nu_{it} - \widetilde{\rho}\nu_{it-1}\right)\right\}^2 \\
&\quad + \left\{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\left(\nu_{it} - \overline{\nu}_{i.}\right) - \widetilde{\rho}\left(\nu_{it-1} - \overline{\nu}_{i.}\right)\right)\right\}^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it}^2 - 2\left(\widetilde{\rho} - \rho\right)\frac{1}{T}\sum_{t=1}^{T}e_{it}\nu_{it-1} + \left(\widetilde{\rho} - \rho\right)^2\frac{1}{T}\sum_{t=1}^{T}\nu_{it-1}^2\right] \\
&\quad - \left\{\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it} - \left(\widetilde{\rho} - \rho\right)\frac{1}{T}\sum_{t=1}^{T}\nu_{it-1}\right]\right\}^2 \\
&\quad + \left\{\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it} - \frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{\rho} - \rho\right)\nu_{it-1} - \frac{1}{T}\sum_{t=1}^{T}(1 - \widetilde{\rho})\overline{\nu}_{i.}\right]\right\}^2.
\end{aligned}$$

Now it is easy to see that

$$I = \sigma_e^2 + o_p\left(1\right)$$

using

$$\begin{aligned}
\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}^2 &\xrightarrow{p} \sigma_e^2, \\
\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it-1}^2 &\xrightarrow{p} \frac{\sigma_e^2}{(1-\rho^2)}, \\
\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it} &= O_p\left(1\right),
\end{aligned}$$

and

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}(1-\widetilde{\rho})\overline{\nu}_{i.}=O_p\left(1\right)$$

as $(n,T)\to\infty$ with $\sqrt{nT}\left(\widetilde{\rho}-\rho\right)=O_p\left(1\right)$.

Consider II. In a similar vein as I, it is easy to see that

$$\begin{aligned}
& \frac{1}{nT}\mathbf{x}'E_{nT}\left(I_n\otimes\hat{C}'\hat{C}\right)E_{nT}\mathbf{x} \\
= & \frac{1}{nT}\mathbf{x}'\left(E_n\otimes\hat{C}'\hat{C}\right)\mathbf{x}+\frac{1}{nT}\mathbf{x}'\left(\bar{J}_n\otimes E_T\hat{C}'\hat{C}E_T\right)\mathbf{x} \\
= & \frac{1}{nT}\mathbf{x}'\left(I_n\otimes\hat{C}'\hat{C}\right)\mathbf{x}-\frac{1}{nT}\mathbf{x}'\left(\bar{J}_n\otimes\hat{C}'\hat{C}\right)\mathbf{x}+\frac{1}{nT}\mathbf{x}'\left(\bar{J}_n\otimes E_T\hat{C}'\hat{C}E_T\right)\mathbf{x}.
\end{aligned}$$

Expanding this equation, one can show that

$$\begin{aligned}
& \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)^2-\left\{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)\right\}^2 \\
& +\left\{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\left(x_{it}-\overline{x}_{i.}\right)-\widetilde{\rho}\left(x_{it-1}-\overline{x}_{i.}\right)\right)\right\}^2 \\
= & \frac{1}{nT}\sum_{i=1}^{n}\left\{\begin{array}{c}\sum_{t=1}^{T}\varepsilon_{it}^2+\left(\widetilde{\rho}-\rho\right)^2\sum_{t=1}^{T}x_{it-1}^2 \\ +(\lambda-\rho)^2\sum_{t=1}^{T}x_{it-1}^2-2\left(\widetilde{\rho}-\rho\right)\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} \\ +2(\lambda-\rho)\sum_{t=1}^{T}\varepsilon_{it}x_{it-1}-2\left(\widetilde{\rho}-\rho\right)\left(\lambda-\rho\right)\sum_{t=1}^{T}x_{it-1}^2\end{array}\right\} \\
& -\left[\frac{1}{nT}\sum_{i=1}^{n}\left\{\sum_{t=1}^{T}\varepsilon_{it}-\left(\widetilde{\rho}-\rho\right)\sum_{t=1}^{T}x_{it-1}+\left(\lambda-\rho\right)\sum_{t=1}^{T}x_{it-1}\right\}\right]^2 \\
& +\left[\frac{1}{nT}\sum_{i=1}^{n}\left\{\begin{array}{c}\sum_{t=1}^{T}\varepsilon_{it}-\sum_{t=1}^{T}\left(\widetilde{\rho}-\rho\right)x_{it-1} \\ +(\lambda-\rho)\sum_{t=1}^{T}x_{it-1}-\sum_{t=1}^{T}(1-\widetilde{\rho})\overline{x}_{i.}\end{array}\right\}\right]^2 \\
= & \sigma_\varepsilon^2+\frac{\sigma_\varepsilon^2(\lambda-\rho)^2}{\left(1-\lambda^2\right)}+o_p(1)
\end{aligned}$$

as $(n,T)\to\infty$. This is because in the first term

$$\begin{aligned}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^{2} &\xrightarrow{p}\sigma_{\varepsilon}^{2},\\ \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} &= O_{p}\left(1\right),\\ \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^{2} &\xrightarrow{p}\sigma_{\varepsilon}^{2}/\left(1-\lambda^{2}\right),\end{aligned}$$

and $\sqrt{nT}\left(\widetilde{\rho}-\rho\right)=O_{p}\left(1\right)$. Also note that $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}=O_{p}\left(1\right)$, $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}=O_{p}\left(1\right)$, and $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}(1-\widetilde{\rho})\overline{x}_{i.}=O_{p}\left(1\right)$ as $(n,T)\to\infty$. Now we get

$$\begin{aligned} II &\approx \frac{1}{nT}\left(\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\right)^{2}\left[\frac{1}{nT}\mathbf{x}'E_{nT}\left(I_{n}\otimes\hat{C}'\hat{C}\right)E_{nT}\mathbf{x}\right]\\ &= O_{p}\left(\frac{1}{nT}\right)=o_{p}\left(1\right)\end{aligned}$$

because if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\xrightarrow{d}N\left(0,\frac{(1+\rho\lambda)(1-\lambda^{2})\sigma_{e}^{2}}{(1-\rho\lambda)(1-\rho^{2})\sigma_{\varepsilon}^{2}}\right)$$

and accordingly

$$\frac{1}{nT}\mathbf{x}'E_{nT}\left(I_{n}\otimes\hat{C}'\hat{C}\right)E_{nT}\mathbf{x}=O_{p}\left(1\right).$$

Consider III.

$$\begin{aligned}
III &= \frac{2}{nT}\left(\hat{\beta}_{OLS}-\beta\right)\mathbf{x}'E_{nT}\left(I_n\otimes\hat{C}'\hat{C}\right)E_{nT}\boldsymbol{\nu} \\
&\leq 2\sqrt{\left[\begin{array}{c}\left[\frac{1}{nT}\boldsymbol{\nu}'E_{nT}\left(I_n\otimes\hat{C}'\hat{C}\right)E_{nT}\boldsymbol{\nu}\right] \\ \cdot\left[\frac{1}{nT}\left(\hat{\beta}_{OLS}-\beta\right)^2\mathbf{x}'E_{nT}\left(I_n\otimes\hat{C}'\hat{C}\right)E_{nT}\mathbf{x}\right]\end{array}\right]} \\
&= 2\sqrt{I\times II}=2\sqrt{\sigma_e^2\times 0}=0
\end{aligned}$$

as $(n,T)\to\infty$ using the Cauchy-Schwarz inequality.

Summarizing, we proved that

$$\widetilde{\sigma}_e^2\xrightarrow{p}\sigma_e^2.$$

(b) $|\rho|<1$, $\lambda=1$ case

This is the *panel cointegration* case. Consider

$$\widetilde{\rho}-\rho=\frac{\sum_{i=1}^n\sum_{t=2}^T(\widehat{u}_{it}-\rho\widehat{u}_{it-1})\widehat{u}_{it-1}}{\sum_{i=1}^n\sum_{t=2}^T\widehat{u}_{it-1}^2}.$$

For the denominator,

$$\begin{aligned}
&\frac{1}{nT}\sum_{i=1}^n\sum_{t=2}^T\widehat{u}_{it-1}^2 \\
=& \frac{1}{nT}\sum_{i=1}^n\sum_{t=2}^T\left[(\nu_{it-1}-\overline{\overline{\nu}})-(\hat{\beta}_{OLS}-\beta)(x_{it-1}-\bar{\bar{x}})\right]^2 \\
=& \frac{1}{nT}\sum_{i=1}^n\sum_{t=2}^T(\nu_{it-1}-\overline{\overline{\nu}})^2 \\
&+\frac{1}{nT}\left\{\sqrt{n}T(\hat{\beta}_{OLS}-\beta)\right\}^2\left\{\frac{1}{nT^2}\sum_{i=1}^n\sum_{t=2}^T(x_{it-1}-\bar{\bar{x}})^2\right\} \\
&-2\sqrt{n}T(\hat{\beta}_{OLS}-\beta)\frac{1}{n^{3/2}T^2}\sum_{i=1}^n\sum_{t=2}^T(\nu_{it-1}-\overline{\overline{\nu}})(x_{it-1}-\bar{\bar{x}}) \\
=& I+II+III.
\end{aligned}$$

Consider II first. With the joint limit, one can verify $II = O_p\left(\frac{1}{nT}\right)$ using that if $(n,T) \to \infty$ and $\frac{n}{T} \to 0$, then

$$\sqrt{n}T\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, \frac{2\sigma_e^2}{(1-\rho)^2\sigma_\varepsilon^2}\right)$$

by Lemma 2.(3). Also note that as $(n,T) \to \infty$, we have

$$\begin{aligned} & \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1} - \bar{\bar{x}})^2 \\ = & \left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^2}\sum_{t=2}^{T}x_{it-1}^2\right] - \left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^{3/2}}\sum_{t=2}^{T}x_{it}\right]^2 \xrightarrow{p} \frac{\sigma_\varepsilon^2}{2} = O_p(1). \end{aligned}$$

Next, $III = O_p\left(\frac{1}{nT}\right)$ using

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\nu_{it-1} - \bar{\bar{\nu}}\right)\left(\frac{x_{it-1} - \bar{\bar{x}}}{\sqrt{T}}\right) = O_p\left(1\right)$$

since ν_{it} and x_{it} are not correlated.

Hence, we have

$$\begin{aligned} \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{\nu}_{it-1}^2 &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1} - \bar{\bar{\nu}})^2 + O_p\left(\frac{1}{nT}\right) + O_p\left(\frac{1}{nT}\right) \\ &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1} - \bar{\bar{\nu}})^2 + O_p\left(\frac{1}{nT}\right) \\ &\xrightarrow{p} \frac{\sigma_e^2}{1-\rho^2} \end{aligned}$$

as $(n,T) \to \infty$.

For the numerator, $\widehat{u}_{it} - \rho\widehat{u}_{it-1} = e_{it} - (\hat{\beta}_{OLS} - \beta)\left\{(1-\rho)x_{it-1} + \varepsilon_{it}\right\} + o_p(1)$

and

$$
\begin{aligned}
& \tfrac{1}{\sqrt{nT}} \textstyle\sum_{i=1}^{n} \sum_{t=2}^{T} \left(\widehat{u}_{it} - \rho \widehat{u}_{it-1}\right) \widehat{u}_{it-1} \\
= \quad & \tfrac{1}{\sqrt{nT}} \textstyle\sum_{i=1}^{n} \sum_{t=2}^{T} \left[\begin{array}{c} \left[e_{it} - (\hat{\beta}_{OLS} - \beta)\left\{(1-\rho)x_{it-1} + \varepsilon_{it}\right\} + o_p(1)\right] \\ \cdot \left[\left(\nu_{it-1} - \overline{\overline{\nu}}\right) - \left(\hat{\beta}_{OLS} - \beta\right)\left(x_{it-1} - \bar{\bar{x}}\right)\right] \end{array} \right] \\
= \quad & \tfrac{1}{\sqrt{nT}} \textstyle\sum_{i=1}^{n} \sum_{t=2}^{T} \left(\nu_{it-1} - \overline{\overline{\nu}}\right) e_{it} \\
& - \tfrac{\sqrt{nT}\left(\hat{\beta}_{OLS} - \beta\right)}{nT^{3/2}} \textstyle\sum_{i=1}^{n} \sum_{t=2}^{T} \left\{(1-\rho)x_{it-1} + \varepsilon_{it}\right\} \left(\nu_{it-1} - \overline{\overline{\nu}}\right) \\
& - \tfrac{\sqrt{nT}\left(\hat{\beta}_{OLS} - \beta\right)}{nT^{3/2}} \textstyle\sum_{i=1}^{n} \sum_{t=2}^{T} \left(x_{it-1} - \bar{\bar{x}}\right) e_{it} \\
& + \tfrac{\left\{\sqrt{nT}(\hat{\beta}_{OLS} - \beta)\right\}^2}{n^{3/2}T^{5/2}} \textstyle\sum_{i=1}^{n} \sum_{t=2}^{T} \left\{(1-\rho)x_{it-1} + \varepsilon_{it}\right\} \left(x_{it-1} - \bar{\bar{x}}\right) \\
= \quad & I + II + III + IV.
\end{aligned}
$$

Consider II first. Again, it can be shown that $II = O_p\left(\frac{1}{\sqrt{nT}}\right)$ because

$$
\begin{aligned}
& \frac{1}{nT^{3/2}} \sum_{i=1}^{n} \sum_{t=2}^{T} \left\{(1-\rho)x_{it-1} + \varepsilon_{it}\right\} \left(\nu_{it-1} - \overline{\overline{\nu}}\right) \\
= \quad & \frac{(1-\rho)}{\sqrt{nT}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=2}^{T} \frac{x_{it-1}}{\sqrt{T}} \frac{\nu_{it-1}}{\sqrt{T}} \\
& - \frac{(1-\rho)}{n\sqrt{T}} \left(\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \nu_{it-1} \right) \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=2}^{T} \frac{x_{it-1}}{\sqrt{T}} \frac{1}{T} \right) \\
& + \frac{1}{\sqrt{nT}} \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \varepsilon_{it} \nu_{it-1} \\
& - \frac{1}{nT^{3/2}} \left(\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \varepsilon_{it} \right) \left(\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \nu_{it-1} \right) \\
= \quad & O_p\left(\frac{1}{\sqrt{nT}}\right) + O_p\left(\frac{1}{n\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{nT}}\right) + O_p\left(\frac{1}{nT^{3/2}}\right) \\
= \quad & O_p\left(\frac{1}{\sqrt{nT}}\right)
\end{aligned}
$$

as $(n, T) \to \infty$.

Consider III.

$$\begin{aligned} III &= -\frac{\sqrt{n}T\left(\hat{\beta}_{OLS}-\beta\right)}{nT^{3/2}}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(x_{it-1}-\bar{\bar{x}}\right)e_{it} \\ &= -\frac{\sqrt{n}T\left(\hat{\beta}_{OLS}-\beta\right)}{nT^{3/2}}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}e_{it}+\frac{\sqrt{n}T\left(\hat{\beta}_{OLS}-\beta\right)}{nT^{3/2}}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}\bar{\bar{x}} \\ &= -\frac{\sqrt{n}T\left(\hat{\beta}_{OLS}-\beta\right)}{\sqrt{n}T}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=2}^{T}\frac{x_{it-1}}{\sqrt{T}}\frac{e_{it}}{\sqrt{T}} \\ &\quad +\frac{\sqrt{n}T\left(\hat{\beta}_{OLS}-\beta\right)}{n\sqrt{T}}\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}\right)\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it}}{\sqrt{T}}\frac{1}{T}\right) \\ &= O_p\left(\frac{1}{\sqrt{nT}}\right)+O_p\left(\frac{1}{n\sqrt{T}}\right)=O_p\left(\frac{1}{\sqrt{nT}}\right). \end{aligned}$$

Consider IV.

$$\begin{aligned} IV &= \frac{\left\{\sqrt{n}T(\hat{\beta}_{OLS}-\beta)\right\}^2}{n^{3/2}T^{5/2}}\sum_{i=1}^{n}\sum_{t=2}^{T}\left\{(1-\rho)x_{it-1}+\varepsilon_{it}\right\}\left(x_{it-1}-\bar{\bar{x}}\right) \\ &= \frac{\left\{\sqrt{n}T(\hat{\beta}_{OLS}-\beta)\right\}^2}{n^{3/2}}\sum_{i=1}^{n}\frac{1}{T^{5/2}}\sum_{t=2}^{T}\left[\begin{array}{c}(1-\rho)x_{it-1}^2-(1-\rho)x_{it-1}\bar{\bar{x}} \\ +\varepsilon_{it}x_{it-1}-\varepsilon_{it}\bar{\bar{x}}\end{array}\right] \\ &= O_p\left(\frac{1}{\sqrt{nT}}\right) \end{aligned}$$

using

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^2}\sum_{t=2}^{T}x_{it-1}^2 &= O_p(1), \\ \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{T^{3/2}}\sum_{t=2}^{T}x_{it-1} &= O_p(1), \end{aligned}$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\varepsilon_{it}x_{it-1}=O_p(1).$$

Hence, one can see that

$$\begin{aligned}
&\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \left(\widehat{u}_{it} - \rho \widehat{u}_{it-1}\right) \widehat{u}_{it-1} \\
&= \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=2}^{T} \left(\nu_{it-1} - \overline{\overline{\nu}}\right) e_{it} + o_p\left(1\right)
\end{aligned}$$

and we conclude that if $(n, T) \to \infty$ and $\frac{n}{T} \to 0$, then

$$\sqrt{nT}\left(\widetilde{\rho} - \rho\right) \xrightarrow{d} \frac{\sigma_e^2 N\left(0, \frac{1}{1-\rho^2}\right)}{\frac{\sigma_e^2}{1-\rho^2}} = N\left(0, 1-\rho^2\right).$$

We check the consistency of $\widetilde{\sigma}_e^2$ next. From Lemma 3.1.(a), we know

$$I \to \sigma_e^2.$$

Moreover, it can be easily shown that $II = o_p(1)$ because if $(n, T) \to \infty$ and $\frac{n}{T} \to 0$, then

$$\sqrt{n}T\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, \frac{2\sigma_e^2}{(1-\rho)^2 \sigma_\varepsilon^2}\right)$$

and

$$\sqrt{nT}\left(\widetilde{\rho} - \rho\right) \xrightarrow{d} N\left(0, 1-\rho^2\right).$$

Also note that $III = o_p(1)$ with Cauchy-Schwarz inequality.

We proved

$$\widetilde{\sigma}_e^2 \xrightarrow{p} \sigma_e^2.$$

2. $\rho = 1$ case

   (a) $\rho = 1$, $|\lambda| < 1$ case

Here we have

$$\widetilde{\rho}-1=\frac{\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{u}_{it}-\widehat{u}_{it-1}\right)\widehat{u}_{it-1}}{\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it-1}^{2}}.$$

For the denominator,

$$\begin{aligned}
& \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it-1}^{2} \\
= & \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[(\nu_{it-1}-\overline{\overline{\nu}})-(\hat{\beta}_{OLS}-\beta)(x_{it-1}-\overline{\overline{x}})\right]^2 \\
= & \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\overline{\nu}})^2 \\
& +\left\{\sqrt{n}(\hat{\beta}_{OLS}-\beta)\right\}^2\frac{1}{n^2T^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1}-\overline{\overline{x}})^2 \\
& -\sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{2}{n^{3/2}T^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\overline{\nu}})(x_{it-1}-\overline{\overline{x}}) \\
= & I+II+III.
\end{aligned}$$

Consider II first. It is easy to see $II=O_p\left(\frac{1}{nT}\right)$ because if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\xrightarrow{d}N\left(0,\frac{(1+\lambda)^2\sigma_e^2}{2\sigma_\varepsilon^2}\right)=O_p\left(1\right)$$

by Lemma 2.(2) and because as $(n,T)\to\infty$, we have

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1}-\bar{\bar{x}})^2\xrightarrow{p}\frac{\sigma_\varepsilon^2}{1-\lambda^2}.$$

Next consider III. One can show that

$$
\begin{aligned}
III &= -\sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{2}{n^{3/2}T^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it}-\overline{\overline{\nu}})(x_{it}-\overline{\overline{x}}) \\
&= -\sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{1}{nT}\frac{2}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it}-\overline{\overline{x}})(\frac{\nu_{it}-\overline{\overline{\nu}}}{\sqrt{T}}) \\
&= O_p\left(\frac{1}{nT}\right)
\end{aligned}
$$

since ν_{it} and x_{it} are not correlated and accordingly $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it}-\overline{\overline{x}})(\frac{\nu_{it}-\overline{\overline{\nu}}}{\sqrt{T}})=O_p(1)$.

Hence, we have

$$
\begin{aligned}
\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it-1}^2 &= \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\overline{\nu}})^2+O_p\left(\frac{1}{nT}\right)+O_p\left(\frac{1}{nT}\right) \\
&= \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\overline{\nu}})^2+O_p\left(\frac{1}{nT}\right)\xrightarrow{p}\frac{\sigma_e^2}{6}
\end{aligned}
$$

as $(n,T)\to\infty$ by, e.g., equation (C.3) in Kao (1999).

For the numerator, $\widehat{u}_{it}-\widehat{u}_{it-1}=e_{it}-(\hat{\beta}_{OLS}-\beta)\{(\lambda-1)x_{it-1}+\varepsilon_{it}\}$ and it can be shown that

$$
\begin{aligned}
& \tfrac{1}{nT}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{u}_{it}-\widehat{u}_{it-1}\right)\widehat{u}_{it-1} \\
= \; & \tfrac{1}{nT}\textstyle\sum_{i=1}^{n}\sum_{t=2}^{T}\left[\begin{array}{l}\left[e_{it}-(\hat{\beta}_{OLS}-\beta)\{(\lambda-1)x_{it-1}+\varepsilon_{it}\}\right] \\ \cdot\left[\left(\nu_{it-1}-\overline{\overline{\nu}}\right)-\left(\hat{\beta}_{OLS}-\beta\right)(x_{it-1}-\bar{\bar{x}})\right]\end{array}\right] \\
= \; & \begin{array}{l} \tfrac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\nu_{it-1}-\overline{\overline{\nu}}\right)e_{it} \\ -\sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{1}{n^{3/2}}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\{(\lambda-1)x_{it-1}+\varepsilon_{it}\}\left(\nu_{it-1}-\overline{\overline{\nu}}\right) \\ -\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\frac{1}{n^{3/2}}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}(x_{it-1}-\bar{\bar{x}})\,e_{it} \\ +\left(\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2\frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\{(\lambda-1)x_{it-1}+\varepsilon_{it}\}(x_{it-1}-\bar{\bar{x}}) \end{array} \\
= \; & I+II+III+IV.
\end{aligned}
$$

Consider II.

$$
\begin{aligned}
II &= -\sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{1}{n^{3/2}}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\{(\lambda-1)x_{it-1}+\varepsilon_{it}\}\left(\nu_{it-1}-\overline{\overline{\nu}}\right) \\
&= -\sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{1}{n^{3/2}}\sum_{i=1}^{n}\left\{\begin{array}{c} \frac{1}{T}\sum_{t=2}^{T}(\lambda-1)x_{it-1}\nu_{it-1} \\ -\frac{1}{T}\sum_{t=2}^{T}(\lambda-1)x_{it-1}\overline{\overline{\nu}} \\ +\frac{1}{T}\sum_{t=2}^{T}\varepsilon_{it}\nu_{it-1}-\frac{1}{T}\sum_{t=2}^{T}\varepsilon_{it}\overline{\overline{\nu}} \end{array}\right\} \\
&= -\sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{(\lambda-1)}{n}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=2}^{T}\frac{x_{it-1}}{\sqrt{T}}\frac{\nu_{it-1}}{\sqrt{T}}\right) \\
&\quad +\sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{(\lambda-1)}{n^{3/2}}\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}\right)\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{\nu_{it}}{\sqrt{T}}\frac{1}{T}\right) \\
&\quad -\sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{1}{n}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=2}^{T}\frac{\varepsilon_{it}}{\sqrt{T}}\frac{\nu_{it-1}}{\sqrt{T}}\right) \\
&\quad +\sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{1}{n^{3/2}}\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\varepsilon_{it}\right)\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{\nu_{it}}{\sqrt{T}}\frac{1}{T}\right) \\
&= O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{n^{3/2}}\right)+O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{n^{3/2}}\right)=O_p\left(\frac{1}{n}\right)
\end{aligned}
$$

using

$$
\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)=O_p\left(1\right)
$$

as $(n,T)\to\infty$ with the joint CLT.

Consider III.

$$
\begin{aligned}
III &= -\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\frac{1}{n^{3/2}}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\left(x_{it-1}e_{it}-\overline{\overline{x}}e_{it}\right) \\
&= -\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\frac{1}{n\sqrt{T}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}x_{it-1}e_{it} \\
&\quad +\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\frac{1}{n^{3/2}T}\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}\right)\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\right) \\
&= O_p\left(\frac{1}{n\sqrt{T}}\right)+O_p\left(\frac{1}{n^{3/2}T}\right)=O_p\left(\frac{1}{n\sqrt{T}}\right).
\end{aligned}
$$

Consider IV. After some algebra, it can be shown that

$$\begin{aligned}
IV &= \left(\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2 \frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\left\{(\lambda-1)x_{it-1}+\varepsilon_{it}\right\}\left(x_{it-1}-\bar{\bar{x}}\right) \\
&= \left(\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2 \frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\left[\begin{array}{c}(\lambda-1)x_{it-1}^2-(\lambda-1)x_{it-1}\bar{\bar{x}} \\ +\varepsilon_{it}x_{it-1}-\varepsilon_{it}\bar{\bar{x}}\end{array}\right] \\
&= O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{n^2T}\right)+O_p\left(\frac{1}{n^{3/2}\sqrt{T}}\right)+O_p\left(\frac{1}{n^2T}\right)=O_p\left(\frac{1}{n}\right).
\end{aligned}$$

Lastly, consider I.

$$\begin{aligned}
I &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\nu_{it-1}-\overline{\overline{\nu}}\right)e_{it} \\
&= \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=2}^{T}\frac{\nu_{it-1}}{\sqrt{T}}\frac{e_{it}}{\sqrt{T}} \\
&\quad -\frac{1}{n}\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}\right)\left(\frac{1}{\sqrt{n}T^{3/2}}\sum_{i=1}^{n}\sum_{t=2}^{T}\nu_{it-1}\right) \\
&= O_p\left(\frac{1}{\sqrt{n}}\right)+O_p\left(\frac{1}{n}\right)=O_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Therefore,

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{u}_{it}-\widehat{u}_{it-1}\right)\widehat{u}_{it-1}=o_p(1).$$

We finally have

$$T\left(\widetilde{\rho}-1\right)=\frac{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\Delta\widehat{u}_{it}\widehat{u}_{it-1}}{\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it-1}^2}\xrightarrow{p}\frac{0}{\frac{\sigma_e^2}{6}}=0.$$

Next we show $\widetilde{\sigma}_e^2$ is a consistent estimator. Again we have

$$\begin{aligned}
\widetilde{\sigma}_e^2 &= \frac{1}{nT}\boldsymbol{\nu}' E_{nT}\left(I_n \otimes \hat{C}'\hat{C}\right) E_{nT}\boldsymbol{\nu} \\
&\quad + \frac{1}{nT}\left(\hat{\beta}_{OLS} - \beta\right)^2 \mathbf{x}' E_{nT}\left(I_n \otimes \hat{C}'\hat{C}\right) E_{nT}\mathbf{x} \\
&\quad + \frac{2}{nT}\left(\hat{\beta}_{OLS} - \beta\right) \mathbf{x}' E_{nT}\left(I_n \otimes \hat{C}'\hat{C}\right) E_{nT}\boldsymbol{\nu} \\
&= I + II + III.
\end{aligned}$$

Consider I.

$$\begin{aligned}
I &= \frac{1}{nT}\boldsymbol{\nu}' E_{nT}\left(I_n \otimes \hat{C}'\hat{C}\right) E_{nT}\boldsymbol{\nu} \\
&\approx \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\nu_{it} - \widetilde{\rho}\nu_{it-1}\right)^2 - \left\{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\nu_{it} - \widetilde{\rho}\nu_{it-1}\right)\right\}^2 \\
&\quad + \left\{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}\left(\left(\nu_{it} - \overline{\nu}_{i.}\right) - \widetilde{\rho}\left(\nu_{it-1} - \overline{\nu}_{i.}\right)\right)\right\}^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\begin{array}{c} \frac{1}{T}\sum_{t=1}^{T} e_{it}^2 - 2T\left(\widetilde{\rho} - 1\right)\frac{1}{T^2}\sum_{t=1}^{T} e_{it}\nu_{it-1} \\ + \left\{T\left(\widetilde{\rho} - 1\right)\right\}^2 \frac{1}{T^3}\sum_{t=1}^{T}\nu_{it-1}^2 \end{array}\right] \\
&\quad - \left\{\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T} e_{it} - T\left(\widetilde{\rho} - 1\right)\frac{1}{T^2}\sum_{t=1}^{T}\nu_{it-1}\right]\right\}^2 \\
&\quad + \left\{\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T} e_{it} - \frac{T\left(\widetilde{\rho} - 1\right)}{T^2}\sum_{t=1}^{T}\nu_{it-1} + \frac{T\left(\widetilde{\rho} - 1\right)}{T^2}\sum_{t=1}^{T}\overline{\nu}_{i.}\right]\right\}^2.
\end{aligned}$$

Now it is easy to see that

$$I \to \sigma_e^2$$

as $(n, T) \to \infty$ using $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} e_{it}^2 \xrightarrow{p} \sigma_e^2$, $\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T} e_{it}\nu_{it-1} = O_p(1)$, $\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it-1}^2 = O_p(1)$, and $T\left(\widetilde{\rho} - 1\right) = o_p\left(1\right)$ with the joint limit.

Consider II. After some tedious algebra, it can be shown that

$$\begin{aligned}
&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)^{2}-\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)\right]^{2}\\
&+\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\left(x_{it}-\overline{x}_{i.}\right)-\widetilde{\rho}\left(x_{it-1}-\overline{x}_{i.}\right)\right)\right]^{2}\\
= \;\; & \sigma_{\varepsilon}^{2}+\frac{\sigma_{\varepsilon}^{2}(\lambda-1)^{2}}{\left(1-\lambda^{2}\right)}+o_{p}(1)
\end{aligned}$$

as $(n,T)\rightarrow\infty$. Now one can see that

$$II\approx\frac{\left(\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\right)^{2}}{n}\left[\frac{1}{nT}\mathbf{x}'E_{nT}\left(I_{n}\otimes\hat{C}'\hat{C}\right)E_{nT}\mathbf{x}\right]=O_{p}\left(\frac{1}{n}\right)=o_{p}(1)$$

using the fact that if $(n,T)\rightarrow\infty$ and $\frac{n}{T}\rightarrow 0$, then

$$\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\xrightarrow{d}N\left(0,\frac{(1+\lambda)^{2}\sigma_{e}^{2}}{2\sigma_{\varepsilon}^{2}}\right).$$

Because $III=o_{p}(1)$ by the Cauchy-Schwarz inequality, we get

$$\widetilde{\sigma}_{e}^{2}\xrightarrow{p}\sigma_{e}^{2}.$$

(b) $\rho=1$, $\lambda=1$ case

Note that

$$\widetilde{\rho}-1=\frac{\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{u}_{it}-\widehat{u}_{it-1}\right)\widehat{u}_{it-1}}{\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it-1}^{2}}.$$

For the denominator,

$$
\begin{aligned}
& \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it-1}^2 \\
= \; & \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[(\nu_{it-1}-\overline{\overline{\nu}})-(\hat{\beta}_{OLS}-\beta)(x_{it-1}-\overline{\overline{x}})\right]^2 \\
= \; & \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\overline{\nu}})^2 \\
& +\left(\sqrt{n}(\hat{\beta}_{OLS}-\beta)\right)^2\frac{1}{n^2T^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1}-\overline{\overline{x}})^2 \\
& -\left(\sqrt{n}(\hat{\beta}_{OLS}-\beta)\right)\frac{2}{n^{3/2}T^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\nu_{it-1}-\overline{\overline{\nu}})(x_{it-1}-\overline{\overline{x}}) \\
= \; & I+II+III.
\end{aligned}
$$

Consider II first. With the joint limit, one can see that

$$
\begin{aligned}
& \left(\sqrt{n}(\hat{\beta}_{OLS}-\beta)\right)^2\frac{1}{n^2T^2}\sum_{i=1}^{n}\sum_{t=2}^{T}(x_{it-1}-\overline{\overline{x}})^2 \\
= \; & \left(\sqrt{n}(\hat{\beta}_{OLS}-\beta)\right)^2\frac{1}{n}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\frac{x_{it-1}}{\sqrt{T}}\right)^2\frac{1}{T} \\
& -\left(\sqrt{n}(\hat{\beta}_{OLS}-\beta)\right)^2\frac{1}{n^{3/2}}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=2}^{T}\frac{x_{it-1}}{\sqrt{T}}\frac{1}{T}\right)^2 \\
= \; & O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{n^{3/2}}\right)=O_p\left(\frac{1}{n}\right)
\end{aligned}
$$

as $(n,T)\to\infty$ using the fact that if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$
\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\xrightarrow{d}N\left(0,\frac{2\sigma_e^2}{3\sigma_\varepsilon^2}\right)=O_p(1)
$$

by Lemma 2.(4).

Next $III = O_p(\frac{1}{n})$ because

$$\begin{aligned}
& \sqrt{n}(\hat{\beta}_{OLS} - \beta)\frac{2}{n^{3/2}} \sum_{i=1}^{n} \frac{1}{T^2} \sum_{t=2}^{T} (\nu_{it-1} - \overline{\overline{\nu}})(x_{it-1} - \overline{\overline{x}}) \\
= \ & \sqrt{n}(\hat{\beta}_{OLS} - \beta)\frac{2}{n}\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=2}^{T} (\frac{\nu_{it-1} - \overline{\overline{\nu}}}{\sqrt{T}})(\frac{x_{it-1} - \overline{\overline{x}}}{\sqrt{T}})\frac{1}{T} \\
= \ & O_p(\frac{1}{n})
\end{aligned}$$

using

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=2}^{T} (\frac{\nu_{it-1} - \overline{\overline{\nu}}}{\sqrt{T}})(\frac{x_{it-1} - \overline{\overline{x}}}{\sqrt{T}})\frac{1}{T} = O_p(1)$$

where ν_{it} and x_{it} are not correlated.

Accordingly, we conclude that

$$\begin{aligned}
\frac{1}{nT^2} \sum_{i=1}^{n} \sum_{t=2}^{T} \widehat{u}_{it-1}^2 &= \frac{1}{nT^2} \sum_{i=1}^{n} \sum_{t=2}^{T} (\nu_{it-1} - \overline{\overline{\nu}})^2 + O_p\left(\frac{1}{n}\right) + O_p(\frac{1}{n}) \\
&= \frac{1}{nT^2} \sum_{i=1}^{n} \sum_{t=2}^{T} (\nu_{it-1} - \overline{\overline{\nu}})^2 + O_p(\frac{1}{n}) \xrightarrow{p} \frac{\sigma_e^2}{6}
\end{aligned}$$

by equation (C.3) in Kao (1999).

For the numerator, $\widehat{u}_{it} - \widehat{u}_{it-1} = e_{it} - (\hat{\beta}_{OLS} - \beta)\varepsilon_{it} + o_p(1)$. Hence

$$
\begin{aligned}
& \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{u}_{it} - \widehat{u}_{it-1}\right)\widehat{u}_{it-1} \\
= \; & \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left[e_{it} - (\hat{\beta}_{OLS} - \beta)\varepsilon_{it}\right]\left[\begin{array}{c}\left(\nu_{it-1} - \overline{\overline{\nu}}\right) \\ -\left(\hat{\beta}_{OLS} - \beta\right)\left(x_{it-1} - \bar{\bar{x}}\right)\end{array}\right] \\
= \; & \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\nu_{it-1} - \overline{\overline{\nu}}\right)e_{it} \\
& -\sqrt{n}(\hat{\beta}_{OLS} - \beta)\frac{1}{n^{3/2}}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\varepsilon_{it}\left(\nu_{it-1} - \overline{\overline{\nu}}\right) \\
& -\sqrt{n}\left(\hat{\beta}_{OLS} - \beta\right)\frac{1}{n^{3/2}}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\left(x_{it-1} - \bar{\bar{x}}\right)e_{it} \\
& +\left(\sqrt{n}\left(\hat{\beta}_{OLS} - \beta\right)\right)^2\frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\varepsilon_{it}\left(x_{it-1} - \bar{\bar{x}}\right) \\
= \; & I + II + III + IV.
\end{aligned}
$$

Consider I. One can verify that

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\nu_{it-1} - \overline{\overline{\nu}}\right)e_{it} = O_p\left(\frac{1}{\sqrt{n}}\right)$$

using $\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=2}^{T}\nu_{it-1}e_{it} = O_p(1)$ and $\frac{1}{T}\sum_{i=1}^{n}\sum_{t=2}^{T}e_{it}\overline{\overline{\nu}} = O_p(1)$ as $(n,T) \to \infty$.

Consider II.

$$\begin{aligned}
II &= \sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{1}{n^{3/2}}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\varepsilon_{it}\left(\nu_{it-1}-\overline{\overline{\nu}}\right) \\
&= \sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{1}{n}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=2}^{T}\frac{\varepsilon_{it}}{\sqrt{T}}\frac{\nu_{it-1}}{\sqrt{T}}\right) \\
&\quad -\sqrt{n}(\hat{\beta}_{OLS}-\beta)\frac{1}{n^{3/2}}\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=2}^{T}\varepsilon_{it}\right)\left(\frac{1}{\sqrt{n}T^{3/2}}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it}\right) \\
&= O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{n^{3/2}}\right)=O_p\left(\frac{1}{n}\right)
\end{aligned}$$

as $(n,T)\to\infty$.

Consider III and IV. In a similar vein as II, it is easy to see that

$$III=-\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\frac{1}{n^{3/2}}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\left(x_{it-1}-\bar{\bar{x}}\right)e_{it}=O_p\left(\frac{1}{n}\right)$$

and

$$IV=\left(\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2\frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=2}^{T}\varepsilon_{it}\left(x_{it-1}-\bar{\bar{x}}\right)=O_p\left(\frac{1}{n^{3/2}}\right).$$

Therefore,

$$\begin{aligned}
&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\widehat{u}_{it}-\widehat{u}_{it-1}\right)\widehat{u}_{it-1} \\
&= O_p(\frac{1}{\sqrt{n}})+O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{n^{3/2}}\right) \\
&= O_p(\frac{1}{\sqrt{n}})=o_p\left(1\right).
\end{aligned}$$

Summarizing, we have

$$T\left(\widetilde{\rho}-1\right)=\frac{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=2}^{T}\Delta\widehat{u}_{it}\widehat{u}_{it-1}}{\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it-1}^2}\xrightarrow{p}\frac{0}{\frac{\sigma_e^2}{6}}=0.$$

Next we show $\widetilde{\sigma}_e^2$ is a consistent estimator. It is clear that $I \rightarrow \sigma_e^2$ as $(n, T) \rightarrow \infty$ as shown already.

Consider II. In a similar process as above, one can show that

$$
\begin{aligned}
& \tfrac{1}{nT} \textstyle\sum_{i=1}^{n} \sum_{t=1}^{T} \left(x_{it} - \widehat{\rho} x_{it-1}\right)^2 - \left\{ \tfrac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(x_{it} - \widehat{\rho} x_{it-1}\right) \right\}^2 \\
& + \left\{ \tfrac{1}{n} \textstyle\sum_{i=1}^{n} \tfrac{1}{T} \sum_{t=1}^{T} \left(\left(x_{it} - \overline{x}_{i.}\right) - \widehat{\rho}\left(x_{it-1} - \overline{x}_{i.}\right)\right) \right\}^2 \\
= \; & \tfrac{1}{n} \textstyle\sum_{i=1}^{n} \left[\tfrac{1}{T} \sum_{t=1}^{T} \varepsilon_{it}^2 + \tfrac{\{T(\widehat{\rho}-1)\}^2}{T^3} \sum_{t=1}^{T} x_{it-1}^2 - \tfrac{2T(\widehat{\rho}-1)}{T^2} \sum_{t=1}^{T} \varepsilon_{it} x_{it-1} \right] \\
& - \left[\tfrac{1}{n} \textstyle\sum_{i=1}^{n} \left\{ \tfrac{1}{T} \sum_{t=1}^{T} \varepsilon_{it} - \tfrac{T(\widehat{\rho}-1)}{T^2} \sum_{t=1}^{T} x_{it-1} \right\} \right]^2 \\
& + \left\{ \tfrac{1}{n} \textstyle\sum_{i=1}^{n} \left[\tfrac{1}{T} \sum_{t=1}^{T} \varepsilon_{it} - \tfrac{T(\widehat{\rho}-1)}{T^2} \sum_{t=1}^{T} x_{it-1} + \tfrac{T(\widehat{\rho}-1)}{T^2} \sum_{t=1}^{T} \overline{x}_{i.} \right] \right\}^2 \\
= \; & \sigma_\varepsilon^2 + o_p(1)
\end{aligned}
$$

as $(n, T) \rightarrow \infty$. Hence we have

$$
II \approx \frac{\left(\sqrt{n}\left(\hat{\beta}_{OLS} - \beta\right)\right)^2}{n} \left[\frac{1}{nT} \mathbf{x}' E_{nT} \left(I_n \otimes \hat{C}'\hat{C} \right) E_{nT} \mathbf{x} \right] = O_p\left(\frac{1}{n}\right) = o_p\left(1\right)
$$

since if $(n, T) \rightarrow \infty$ and $\frac{n}{T} \rightarrow 0$, then

$$
\sqrt{n}\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, \frac{2\sigma_e^2}{3\sigma_\varepsilon^2}\right).
$$

Also because $III = o_p(1)$ by the Cauchy-Schwarz inequality, we get

$$
\widetilde{\sigma}_e^2 \xrightarrow{p} \sigma_e^2.
$$

■

F Proof of Theorem 3

Preparation: Note that from equation (9), we have

$$\mathbf{y} = \gamma \boldsymbol{\iota}_{nT} + \mathbf{x}\beta + \mathbf{u} = \gamma \boldsymbol{\iota}_{nT} + \mathbf{x}\beta + \mathbf{Z}_{\mu}\boldsymbol{\mu} + \boldsymbol{\nu}$$

where $\mathbf{y}$ is $nT \times 1$, $\mathbf{x}$ is a vector of x_{it} of dimension $nT \times 1$, $\boldsymbol{\iota}_{nT}$ is a vector of ones of dimension nT, $\mathbf{u}$ is $nT \times 1$, $\boldsymbol{\mu}$ is a vector of μ_i, $\boldsymbol{\nu}$ is a vector of ν_{it} and $\mathbf{Z}_{\mu} = I_n \otimes \boldsymbol{\iota}_T$. Also recall from equation (13) that

$$\Phi^{-1} = I_n \otimes \left[\frac{1}{\sigma_e^2} \left(\mathbf{A}^{-1} - \frac{\sigma_\mu^2}{\sigma_e^2 + \theta \sigma_\mu^2} \mathbf{A}^{-1} \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \mathbf{A}^{-1} \right) \right].$$

Here we define $\mathbf{z} = [\boldsymbol{\iota}_{nT}, \mathbf{x}]$, then

$$\begin{aligned}
\begin{pmatrix} \widehat{\gamma}_{GLS} \\ \widehat{\beta}_{GLS} \end{pmatrix} &= \left(\mathbf{z}'\Phi^{-1}\mathbf{z}\right)^{-1} \left(\mathbf{z}'\Phi^{-1}\mathbf{y}\right) \\
&= \left(\begin{bmatrix} \boldsymbol{\iota}_{nT}' \\ \mathbf{x}' \end{bmatrix} \Phi^{-1} \begin{bmatrix} \boldsymbol{\iota}_{nT} & \mathbf{x} \end{bmatrix} \right)^{-1} \left(\begin{bmatrix} \boldsymbol{\iota}_{nT}' \\ \mathbf{x}' \end{bmatrix} \Phi^{-1}\mathbf{y} \right) \\
&= \begin{bmatrix} \boldsymbol{\iota}_{nT}'\Phi^{-1}\boldsymbol{\iota}_{nT} & \boldsymbol{\iota}_{nT}'\Phi^{-1}\mathbf{x} \\ \mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT} & \mathbf{x}'\Phi^{-1}\mathbf{x} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{\iota}_{nT}'\Phi^{-1}\mathbf{y} \\ \mathbf{x}'\Phi^{-1}\mathbf{y} \end{bmatrix} \\
&= \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{\iota}_{nT}'\Phi^{-1}\mathbf{y} \\ \mathbf{x}'\Phi^{-1}\mathbf{y} \end{bmatrix} \\
&= \begin{bmatrix} F_{11}\boldsymbol{\iota}_{nT}'\Phi^{-1}\mathbf{y} + F_{12}\mathbf{x}'\Phi^{-1}\mathbf{y} \\ F_{21}\boldsymbol{\iota}_{nT}'\Phi^{-1}\mathbf{y} + F_{22}\mathbf{x}'\Phi^{-1}\mathbf{y} \end{bmatrix}
\end{aligned}$$

where

$$\begin{aligned}
F_{11} &= \left[\boldsymbol{\iota}'_{nT}\Phi^{-1}\boldsymbol{\iota}_{nT} - \boldsymbol{\iota}'_{nT}\Phi^{-1}\mathbf{x}\left(\mathbf{x}'\Phi^{-1}\mathbf{x}\right)^{-1}\mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\right]^{-1}, \\
F_{12} &= -\left(\boldsymbol{\iota}'_{nT}\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\boldsymbol{\iota}'_{nT}\Phi^{-1}\mathbf{x}\left[\mathbf{x}'\Phi^{-1}\mathbf{x} - \mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\left(\boldsymbol{\iota}'_{nT}\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\boldsymbol{\iota}'_{nT}\Phi^{-1}\mathbf{x}\right]^{-1}, \\
F_{21} &= -\left[\mathbf{x}'\Phi^{-1}\mathbf{x} - \mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\left(\boldsymbol{\iota}'_{nT}\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\boldsymbol{\iota}'_{nT}\Phi^{-1}\mathbf{x}\right]^{-1}\mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\left(\boldsymbol{\iota}'_{nT}\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1},
\end{aligned}$$

and

$$F_{22} = \left[\mathbf{x}'\Phi^{-1}\mathbf{x} - \mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\left(\boldsymbol{\iota}'_{nT}\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\boldsymbol{\iota}'_{nT}\Phi^{-1}\mathbf{x}\right]^{-1}.$$

Hence, we have

$$\begin{aligned}
\widehat{\beta}_{GLS} &= F_{21}\boldsymbol{\iota}'_{nT}\Phi^{-1}\mathbf{y} + F'_{22}\mathbf{x}'\Phi^{-1}\mathbf{y} \\
&= \left[\mathbf{x}'\Phi^{-1}\mathbf{x} - \mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\left(\boldsymbol{\iota}'_{nT}\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\boldsymbol{\iota}'_{nT}\Phi^{-1}\mathbf{x}\right]^{-1} \\
&\quad \times \left[\mathbf{x}'\Phi^{-1}\mathbf{y} - \mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\left(\boldsymbol{\iota}'_{nT}\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\boldsymbol{\iota}'_{nT}\Phi^{-1}\mathbf{y}\right]
\end{aligned}$$

and

$$\widehat{\beta}_{GLS} - \beta = G_1^{-1}G_2$$

where

$$G_1 = \mathbf{x}'\Phi^{-1}\mathbf{x} - \mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\left(\boldsymbol{\iota}'_{nT}\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\boldsymbol{\iota}'_{nT}\Phi^{-1}\mathbf{x}$$

and

$$G_2 = \mathbf{x}'\Phi^{-1}\mathbf{u} - \mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\left(\boldsymbol{\iota}'_{nT}\Phi^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\boldsymbol{\iota}'_{nT}\Phi^{-1}\mathbf{u}$$

respectively.

Proof. Following Baltagi *et al.* (2008), we first define matrices $\hat{\mathbf{A}}$ and $\hat{\mathbf{C}}$ which replace ρ in the matrix A and C in equation (12) and (14) with $\widetilde{\rho}$ given by,

$$\widetilde{\rho} = \frac{\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it}\widehat{u}_{it-1}}{\sum_{i=1}^{n}\sum_{t=2}^{T}\widehat{u}_{it-1}^2}$$

where $\widehat{u}_{it}$ denotes the it-th OLS residual. Using the definition of Φ^{-1} in equation (13) and $\widetilde{\sigma}_e^2$ given by,

$$\widetilde{\sigma}_e^2 = \frac{1}{n(T-1)}\widehat{\mathbf{u}}^{*\prime}\widehat{\mathbf{u}}^*$$

where $\widehat{\mathbf{u}}^* = (I_n \otimes \hat{C})\widehat{\mathbf{u}}$ and $\widehat{\mathbf{u}}$ denotes the $nT \times 1$ vector of the OLS residuals, one obtains:

$$\mathbf{x}'\hat{\Phi}^{-1}\mathbf{x} = \frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\left(\mathbf{x}_i'\hat{\mathbf{A}}_i^{-1}\mathbf{x}_i\right),$$

$$\mathbf{x}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT} = \frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\left(\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T\right),$$

$$\boldsymbol{\iota}_{nT}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT} = \frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\left(\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T\right),$$

$$\mathbf{x}'\hat{\Phi}^{-1}\boldsymbol{\nu} = \frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\left(\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i\right),$$

and

$$\boldsymbol{\iota}_{nT}'\hat{\Phi}^{-1}\boldsymbol{\nu} = \frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\left(\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i\right)$$

where

$$\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\mathbf{x}_i = \mathbf{x}_i'\hat{\mathbf{C}}'\hat{\mathbf{C}}\mathbf{x}_i \approx \sum_{t=1}^{T}\left(x_{it} - \widetilde{\rho}x_{it-1}\right)^2,$$

$$\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i = \mathbf{x}_i'\hat{\mathbf{C}}'\hat{\mathbf{C}}\boldsymbol{\nu}_i \approx \sum_{t=1}^{T}\left(x_{it} - \widetilde{\rho}x_{it-1}\right)\left(\nu_{it} - \widetilde{\rho}\nu_{it-1}\right),$$

$$\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T = \mathbf{x}_i'\hat{\mathbf{C}}'\hat{\mathbf{C}}\boldsymbol{\iota}_T \approx (1-\widetilde{\rho})\sum_{t=1}^{T}\left(x_{it} - \widetilde{\rho}x_{it-1}\right),$$

$$\boldsymbol{\iota}_T'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i = \boldsymbol{\iota}_T'\widehat{\mathbf{C}}'\widehat{\mathbf{C}}\boldsymbol{\nu}_i \approx (1-\widetilde{\rho})\sum_{t=1}^{T}\left(\nu_{it} - \widetilde{\rho}\nu_{it-1}\right),$$

and

$$\theta = \boldsymbol{\iota}_T'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T = \boldsymbol{\iota}_T'\widehat{\mathbf{C}}'\widehat{\mathbf{C}}\boldsymbol{\iota}_T = (1-\widetilde{\rho}^2) + (T-1)(1-\widetilde{\rho})^2 \approx \sum_{t=1}^{T}(1-\widetilde{\rho})^2 = T\left(1-\widetilde{\rho}\right)^2.$$

In this section, we assume that $(n,T)\to\infty$ and $\frac{n}{T}\to 0$ unless otherwise specified.

1. $|\rho|<1$, $|\lambda|<1$ case

   (a) Define

$$\frac{1}{nT}\widehat{G}_1=\frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x}-\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{x}}{nT}.$$

First we consider

$$\begin{aligned}&\frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x}\\ =\;&\frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\frac{1}{T}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i.\end{aligned}$$

Expanding this equation, we will show that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i &\approx \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)^2\\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}\left(\varepsilon_{it}+\lambda x_{it-1}-\widetilde{\rho}x_{it-1}\right)^2\\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{\begin{array}{c}\frac{1}{T}\sum_{t=2}^{T}\varepsilon_{it}^2+\left(\widetilde{\rho}-\rho\right)^2\frac{1}{T}\sum_{t=1}^{T}x_{it-1}^2\\ +(\lambda-\rho)^2\frac{1}{T}\sum_{t=1}^{T}x_{it-1}^2\\ -\left(\widetilde{\rho}-\rho\right)\frac{2}{T}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1}\\ +(\lambda-\rho)\frac{2}{T}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1}\\ -\left(\widetilde{\rho}-\rho\right)\left(\lambda-\rho\right)\frac{2}{T}\sum_{t=1}^{T}x_{it-1}^2\end{array}\right\}\\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2+\frac{(\lambda-\rho)^2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2\\
&\quad +I+II+III+IV\\
&= \frac{(1-2\rho\lambda+\rho^2)}{\left(1-\lambda^2\right)}\sigma_\varepsilon^2+o_p(1)\end{aligned}$$

as $(n,T)\to\infty$.

Consider I.

$$I=\left(\sqrt{nT}\left(\widetilde{\rho}-\rho\right)\right)^{2}\frac{1}{nT}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^{2}=O_{p}(\frac{1}{nT})$$

using

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^{2}=O_{p}\left(1\right).$$

Consider II.

$$II=-2\left(\sqrt{nT}\left(\widetilde{\rho}-\rho\right)\right)\frac{1}{nT}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1}=O_{p}(\frac{1}{nT})$$

using

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1}=O_{p}\left(1\right)$$

and

$$\sqrt{nT}\left(\widetilde{\rho}-\rho\right)=O_{p}\left(1\right).$$

Consider III.

$$2(\lambda-\rho)\frac{1}{\sqrt{nT}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1}=O_{p}(\frac{1}{\sqrt{nT}}).$$

Consider IV.

$$-2\left(\sqrt{nT}\left(\widetilde{\rho}-\rho\right)\right)(\lambda-\rho)\frac{1}{\sqrt{nT}}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^{2}=O_{p}(\frac{1}{\sqrt{nT}}).$$

Hence, we have

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}(x_{it}-\widetilde{\rho}x_{it-1})^2 \\
=\;& \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2+\frac{(\lambda-\rho)^2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2 \\
&+O_p(\frac{1}{nT})+O_p(\frac{1}{nT})+O_p(\frac{1}{\sqrt{nT}})+O_p(\frac{1}{\sqrt{nT}}) \\
=\;& \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2+\frac{(\lambda-\rho)^2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2+O_p(\frac{1}{\sqrt{nT}}).
\end{aligned}$$

Because

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2\xrightarrow{p}\sigma_\varepsilon^2$$

and

$$\frac{(\lambda-\rho)^2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2\xrightarrow{p}\frac{(\lambda-\rho)^2\sigma_\varepsilon^2}{(1-\lambda^2)},$$

one concludes that

$$\frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x}=\frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\frac{1}{T}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i\xrightarrow{p}\frac{(1-2\rho\lambda+\rho^2)\sigma_\varepsilon^2}{\left(1-\lambda^2\right)\sigma_e^2}.$$

Next consider

$$\begin{aligned}
\frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT} &= \frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\left(\frac{1}{T}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T\right) \\
&\approx (1-\widetilde{\rho})\frac{1}{\widetilde{\sigma}_e^2}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}[\varepsilon_{it}+(\lambda-\rho)x_{it-1}-(\widetilde{\rho}-\rho)x_{it-1}] \\
&= O_p\left(\frac{1}{\sqrt{nT}}\right)+O_p\left(\frac{1}{\sqrt{nT}}\right)+O_p\left(\frac{1}{nT}\right)=O_p\left(\frac{1}{\sqrt{nT}}\right)
\end{aligned}$$

as $(n,T)\rightarrow\infty$.

Also note

$$\frac{1}{nT}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT} \approx \frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\frac{1}{T}T\left(1-\widetilde{\rho}\right)^2 \xrightarrow{p} \frac{\left(1-\rho\right)^2}{\sigma_e^2} = O_p\left(1\right).$$

Hence,

$$\begin{aligned}
\frac{1}{nT}\widehat{G}_1 &= \frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} - \frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{x}}{nT} \\
&\xrightarrow{p} \frac{(1-2\rho\lambda+\rho^2)\sigma_\varepsilon^2}{\left(1-\lambda^2\right)\sigma_e^2}
\end{aligned}$$

as $(n,T)\rightarrow\infty$.

(b) Now we investigate $\widehat{G}_2$.

$$\frac{1}{\sqrt{nT}}\widehat{G}_2 = \frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\nu} - \frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\nu}}{\sqrt{nT}}$$

Consider first

$$\begin{aligned}
&\frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\nu} \\
=& \frac{1}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}.
\end{aligned}$$

Here

$$
\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} &\approx \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)\left(\nu_{it}-\widetilde{\rho}\nu_{it-1}\right) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\begin{array}{c}\left[\varepsilon_{it}+(\lambda-\rho)x_{it-1}-(\widetilde{\rho}-\rho)x_{it-1}\right] \\ \cdot\left[e_{it}-(\widetilde{\rho}-\rho)\nu_{it-1}\right]\end{array}\right] \\
&= \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\varepsilon_{it}+(\lambda-\rho)x_{it-1}\right]e_{it} \\
&\quad -\frac{\sqrt{nT}\left(\widetilde{\rho}-\rho\right)}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\varepsilon_{it}+(\lambda-\rho)x_{it-1}\right]\nu_{it-1} \\
&\quad -\frac{\sqrt{nT}\left(\widetilde{\rho}-\rho\right)}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}e_{it} \\
&\quad +\frac{\left(\sqrt{nT}\left(\widetilde{\rho}-\rho\right)\right)^2}{n^{3/2}T^{3/2}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}\nu_{it-1} \\
&= \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\varepsilon_{it}+(\lambda-\rho)x_{it-1}\right]e_{it}+I+II+III.
\end{aligned}
$$

Consider I. Note that we have

$$
\begin{aligned}
I &= -\frac{\sqrt{nT}\left(\widetilde{\rho}-\rho\right)}{\sqrt{nT}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}\nu_{it-1} \\
&\quad -(\lambda-\rho)\frac{\sqrt{nT}\left(\widetilde{\rho}-\rho\right)}{\sqrt{nT}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}\nu_{it-1} \\
&= O_p\left(\frac{1}{\sqrt{nT}}\right)+O_p\left(\frac{1}{\sqrt{nT}}\right)=O_p\left(\frac{1}{\sqrt{nT}}\right)
\end{aligned}
$$

using $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}\nu_{it-1} = O_p\left(1\right)$ and $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}\nu_{it-1} = O_p\left(1\right)$.

By a similar process, it can be shown that

$$
II=-\frac{\sqrt{nT}\left(\widetilde{\rho}-\rho\right)}{\sqrt{nT}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}e_{it}=O_p\left(\frac{1}{\sqrt{nT}}\right)
$$

and

$$III = \frac{(\sqrt{nT}\,(\widetilde{\rho}-\rho))^2}{nT}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T} x_{it-1}\nu_{it-1} = O_p\left(\frac{1}{nT}\right).$$

Hence,

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} \quad &\approx \quad \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\varepsilon_{it}+(\lambda-\rho)x_{it-1}\right]e_{it} \\
&\quad +O_p\left(\frac{1}{\sqrt{nT}}\right)+O_p\left(\frac{1}{\sqrt{nT}}\right)+O_p\left(\frac{1}{nT}\right) \\
&= \quad \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\begin{array}{c}\varepsilon_{it}+(\lambda-\rho)\varepsilon_{it-1} \\ +\lambda(\lambda-\rho)\varepsilon_{it-2}+\cdots\end{array}\right]e_{it} \\
&\quad +O_p\left(\frac{1}{\sqrt{nT}}\right).
\end{aligned}$$

Because

$$\begin{aligned}
&\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\varepsilon_{it}+(\lambda-\rho)\varepsilon_{it-1}+\lambda(\lambda-\rho)\varepsilon_{it-2}+\cdots\right]e_{it} \\
&\xrightarrow{d} N(0,\frac{(1-2\rho\lambda+\rho^2)\sigma_e^2\sigma_\varepsilon^2}{1-\lambda^2}),
\end{aligned}$$

we get

$$\frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\nu} = \frac{1}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} \xrightarrow{d} N(0,\frac{(1-2\rho\lambda+\rho^2)\sigma_\varepsilon^2}{\left(1-\lambda^2\right)\sigma_e^2})$$

Next consider

$$\frac{1}{\sqrt{nT}}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\nu} = \frac{1}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\boldsymbol{\iota}_T'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}.$$

From

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\boldsymbol{\iota}_T'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} &\approx \frac{(1-\widetilde{\rho})}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}(\nu_{it}-\widetilde{\rho}\nu_{it-1}) \\ &= (1-\widetilde{\rho})\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}(e_{it}-(\widetilde{\rho}-\rho)\nu_{it-1}) \\ &= (1-\widetilde{\rho})\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it} \\ &\quad -(1-\widetilde{\rho})\frac{\sqrt{nT}(\widetilde{\rho}-\rho)}{\sqrt{nT}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it-1},\end{aligned}$$

it is easy to see that

$$(1-\widetilde{\rho})\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it}\xrightarrow{d}(1-\rho)N(0,\sigma_e^2)$$

and

$$(1-\widetilde{\rho})\frac{\sqrt{nT}(\widetilde{\rho}-\rho)}{\sqrt{nT}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it-1}=O_p\left(\frac{1}{\sqrt{nT}}\right).$$

Accordingly, we have

$$\frac{1}{\sqrt{nT}}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\nu}\xrightarrow{d}(1-\rho)N(0,\frac{1}{\sigma_e^2}).$$

Also recall that $\frac{1}{nT}\mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}=O_p(\frac{1}{\sqrt{nT}})$; $\frac{1}{nT}\boldsymbol{\iota}_{nT}'\Phi^{-1}\boldsymbol{\iota}_{nT}\xrightarrow{p}\frac{(1-\rho)^2}{\sigma_e^2}$ as shown above.

Hence, we have

$$\begin{aligned}\frac{1}{\sqrt{nT}}\widehat{G}_2 &= \frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\nu}-\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\nu}}{\sqrt{nT}} \\ &\xrightarrow{d} N(0,\frac{(1-2\rho\lambda+\rho^2)\sigma_\varepsilon^2}{(1-\lambda^2)\sigma_e^2})\end{aligned}$$

as $(n,T)\to\infty$.

(c) We conclude that

$$t_{FGLS} = \left[\frac{(1-2\rho\lambda+\rho^2)\sigma_\varepsilon^2}{\left(1-\lambda^2\right)\sigma_e^2} \right]^{-1/2} N\left(0, \frac{(1-2\rho\lambda+\rho^2)\sigma_\varepsilon^2}{(1-\lambda^2)\sigma_e^2}\right) = N(0,1).$$

2. $\rho = 1$, $|\lambda| < 1$ case

(a) Let

$$\frac{1}{nT}\widehat{G}_1 = \frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} - T\left(1-\widetilde{\rho}\right)\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT\left(1-\widetilde{\rho}\right)}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n\left(1-\widetilde{\rho}\right)}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{x}}{nT\left(1-\widetilde{\rho}\right)}.$$

Using a similar argument as above, we first consider

$$\frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} = \frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\frac{1}{T}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i.$$

Expanding this equation, we get

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i &\approx \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\varepsilon_{it}+\lambda x_{it-1}-\widetilde{\rho}x_{it-1}\right)^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2+\frac{1}{T^2}\left\{T\left(\widetilde{\rho}-1\right)\right\}^2\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2 \\
&\quad +(\lambda-1)^2\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2 \\
&\quad -\frac{2}{T}T\left(\widetilde{\rho}-1\right)\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} \\
&\quad +(\lambda-1)\frac{2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} \\
&\quad -\frac{2}{T}T\left(\widetilde{\rho}-1\right)(\lambda-1)\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2+(\lambda-1)^2\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2 \\
&\quad +I+II+III+IV.
\end{aligned}$$

Consider I. With the joint limit, we have

$$I=\left(\widetilde{\rho}-1\right)^2\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2=o_p\left(\frac{1}{T^2}\right)$$

using

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2=O_p\left(1\right)$$

and

$$\left(\widetilde{\rho}-1\right)=o_p\left(\frac{1}{T}\right).$$

Consider II. In a similar vein as I,

$$II=-\frac{2}{\sqrt{nT}}\left(\widetilde{\rho}-1\right)\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1}=o_p\left(\frac{1}{\sqrt{n}T^{3/2}}\right)$$

using $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1}=O_p\left(1\right)$.

Consider III.

$$III=(\lambda-1)\frac{1}{\sqrt{nT}}\frac{2}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1}=O_p\left(\frac{1}{\sqrt{nT}}\right)$$

as $(n,T)\to\infty$.

Consider IV.

$$IV=-2\left(\widetilde{\rho}-1\right)\left(\lambda-1\right)\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2=o_p\left(\frac{1}{T}\right).$$

Finally, because we know

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2\xrightarrow{p}\sigma_\varepsilon^2$$

and

$$(\lambda-1)^2\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2\xrightarrow{p}\frac{(\lambda-1)^2\sigma_\varepsilon^2}{1-\lambda^2},$$

we get

$$\frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x}=\frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\frac{1}{T}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i\xrightarrow{p}\frac{2\sigma_\varepsilon^2}{(1+\lambda)\sigma_e^2}$$

as $(n,T)\to\infty$.

Next, it can be shown that

$$\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT\left(1-\widetilde{\rho}\right)}=\frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\left(\frac{1}{T\left(1-\widetilde{\rho}\right)}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T\right)=o_p\left(1\right)$$

as $(n,T)\to\infty$ because

$$\begin{aligned}\frac{1}{T\left(1-\widetilde{\rho}\right)}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T &\approx \frac{1}{T}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)\\ &= \frac{1}{T}\sum_{t=1}^{T}\left[\varepsilon_{it}+(\lambda-1)x_{it-1}-\left(\widetilde{\rho}-1\right)x_{it-1}\right]\end{aligned}$$

and accordingly

$$\begin{aligned}\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT\left(1-\widehat{\rho}\right)} &= \frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\left(\frac{1}{T\left(1-\widetilde{\rho}\right)}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T\right)\\ &\approx \frac{1}{\widetilde{\sigma}_e^2}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\varepsilon_{it}+(\lambda-1)x_{it-1}-\left(\widetilde{\rho}-1\right)x_{it-1}\right]\\ &= O_p\left(\frac{1}{\sqrt{nT}}\right)+O_p\left(\frac{1}{\sqrt{nT}}\right)+o_p\left(\frac{1}{\sqrt{n}T^{3/2}}\right).\end{aligned}$$

We also know that

$$\frac{1}{n\left(1-\widetilde{\rho}\right)}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}=\frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\left[\frac{2\widetilde{\rho}(1-\widetilde{\rho})}{(1-\widetilde{\rho})}+\frac{T\left(1-\widetilde{\rho}\right)^2}{(1-\widetilde{\rho})}\right]\xrightarrow{p}\frac{2}{\sigma_e^2}.$$ [2]

Hence, we get

$$\frac{1}{nT}\widehat{G}_1\xrightarrow{p}\frac{2\sigma_\varepsilon^2}{(1+\lambda)\sigma_e^2}$$

as $(n,T)\to\infty$.

(b) Now we investigate $\widehat{G}_2$. Recall that

$$\frac{1}{\sqrt{nT}}\widehat{G}_2=\frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\nu}-T\left(1-\widetilde{\rho}\right)\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT\left(1-\widetilde{\rho}\right)}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n\left(1-\widetilde{\rho}\right)}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\nu}}{\sqrt{nT}\left(1-\widetilde{\rho}\right)}.$$

[2] Note that we are using the entire form including the first observation, not the abbreviated form. That is,

$$\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}=\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\left[2\widetilde{\rho}(1-\widetilde{\rho})+T\left(1-\widetilde{\rho}\right)^2\right].$$

Firstly, consider

$$\frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\nu} = \frac{1}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}.$$

Here

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} &\approx \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)\left(\nu_{it}-\widetilde{\rho}\nu_{it-1}\right) \\
&= \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\begin{array}{c} \left[\varepsilon_{it}+(\lambda-1)x_{it-1}-(\widetilde{\rho}-1)x_{it-1}\right] \\ \cdot\left[e_{it}-(\widetilde{\rho}-1)\nu_{it-1}\right] \end{array}\right] \\
&= \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\begin{array}{c} \left[\varepsilon_{it}+(\lambda-1)x_{it-1}\right]e_{it} \\ -\left(\widetilde{\rho}-1\right)\left[\varepsilon_{it}+(\lambda-1)x_{it-1}\right]\nu_{it-1} \\ -\left(\widetilde{\rho}-1\right)x_{it-1}e_{it} \\ +\left(\widetilde{\rho}-1\right)^2 x_{it-1}\nu_{it-1} \end{array}\right\} \\
&= \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\varepsilon_{it}+(\lambda-1)x_{it-1}\right]e_{it}+I+II+III.
\end{aligned}$$

Consider I. It can be shown that

$$\begin{aligned}
I &= \left(\widetilde{\rho}-1\right)\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\varepsilon_{it}+(\lambda-1)x_{it-1}\right]\nu_{it-1} \\
&= \frac{T\left(\widetilde{\rho}-1\right)}{\sqrt{T}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{\varepsilon_{it}}{\sqrt{T}}\frac{\nu_{it-1}}{\sqrt{T}}+\frac{T\left(\widetilde{\rho}-1\right)}{\sqrt{T}}\frac{(\lambda-1)}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it-1}}{\sqrt{T}}\frac{\nu_{it-1}}{\sqrt{T}} \\
&= o_p\left(\frac{1}{\sqrt{T}}\right)+o_p\left(\frac{1}{\sqrt{T}}\right)=o_p\left(\frac{1}{\sqrt{T}}\right)
\end{aligned}$$

using $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{\varepsilon_{it}}{\sqrt{T}}\frac{\nu_{it-1}}{\sqrt{T}}=O_p(1)$ and $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it-1}}{\sqrt{T}}\frac{\nu_{it-1}}{\sqrt{T}}=O_p(1)$ with $\left(\widetilde{\rho}-1\right)=o_p\left(\frac{1}{T}\right)$.

Consider II and III. One can easily verify that

$$II=\left(T\left(\widetilde{\rho}-1\right)\right)\frac{1}{T}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}e_{it}=o_p\left(\frac{1}{T}\right)$$

and

$$III = \left(T\left(\widetilde{\rho}-1\right)\right)^2 \frac{1}{T^{3/2}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{x_{it-1}}{\sqrt{T}} \frac{\nu_{it-1}}{\sqrt{T}} = o_p\left(\frac{1}{T^{3/2}}\right).$$

We conclude that

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\mathbf{x}_i' \widehat{\mathbf{A}}^{-1} \boldsymbol{\nu}_i}{\sqrt{T}} &\approx \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[\varepsilon_{it} + (\lambda-1)x_{it-1}\right] e_{it} \\ &\quad + o_p\left(\frac{1}{\sqrt{T}}\right) + o_p\left(\frac{1}{T}\right) + o_p\left(\frac{1}{T^{3/2}}\right) \\ &= \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[\varepsilon_{it} + (\lambda-1)x_{it-1}\right] e_{it} + o_p\left(\frac{1}{\sqrt{T}}\right) \\ &\xrightarrow{d} N(0, \frac{2\sigma_e^2\sigma_\varepsilon^2}{1+\lambda}). \end{aligned}$$

Accordingly, we have

$$\frac{1}{\sqrt{nT}} \mathbf{x}' \widehat{\Phi}^{-1} \boldsymbol{\nu} = \frac{1}{\widetilde{\sigma}_e^2} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\mathbf{x}_i' \widehat{\mathbf{A}}^{-1} \boldsymbol{\nu}_i}{\sqrt{T}} \xrightarrow{d} \frac{1}{\sigma_e^2} N(0, \frac{2\sigma_e^2\sigma_\varepsilon^2}{1+\lambda}).$$

Next consider

$$\frac{1}{\sqrt{nT}\left(1-\widetilde{\rho}\right)} \boldsymbol{\iota}_{nT}' \widehat{\Phi}^{-1} \boldsymbol{\nu} = \frac{1}{\widetilde{\sigma}_e^2} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\boldsymbol{\iota}_T' \widehat{\mathbf{A}}^{-1} \boldsymbol{\nu}_i}{\sqrt{T}\left(1-\widetilde{\rho}\right)}.$$

One can see that

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\boldsymbol{\iota}_T' \widehat{\mathbf{A}}^{-1} \boldsymbol{\nu}_i}{\sqrt{T}\left(1-\widetilde{\rho}\right)} &\approx \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} (e_{it} + (1-\widetilde{\rho})\nu_{it-1}) \\ &= \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} e_{it} + \frac{T(1-\widetilde{\rho})}{\sqrt{n}T^{3/2}} \sum_{i=1}^{n} \sum_{t=1}^{T} \nu_{it-1} \\ &= \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} e_{it} + o_p\left(1\right) \xrightarrow{d} \sigma_e N(0,1) \end{aligned}$$

using $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{\nu_{it-1}}{\sqrt{T}} \frac{1}{T} = O_p\left(1\right)$.

Therefore,

$$\frac{1}{\sqrt{nT}\,(1-\widetilde{\rho})}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\nu} \xrightarrow{d} N(0, \frac{1}{\sigma_e^2}).$$

Also note that

$$\frac{1}{nT\,(1-\widetilde{\rho})}\mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT} \xrightarrow{p} 0; \; \frac{1}{n\,(1-\widetilde{\rho})}\boldsymbol{\iota}_{nT}'\Phi^{-1}\boldsymbol{\iota}_{nT} = O_p\,(1) \xrightarrow{p} \frac{2}{\sigma_e^2},$$

as has been shown already.

Therefore,

$$\frac{1}{\sqrt{nT}}\widehat{G}_2 \xrightarrow{d} \frac{1}{\sigma_e^2}N\left(0, \frac{2\sigma_e^2\sigma_\varepsilon^2}{1+\lambda}\right) = N\left(0, \frac{2\sigma_\varepsilon^2}{(1+\lambda)\sigma_e^2}\right)$$

as $(n, T) \to \infty$.

(c) Summarizing, we have

$$t_{FGLS} = \left[\frac{2\sigma_\varepsilon^2}{(1+\lambda)\sigma_e^2}\right]^{-1/2} N\left(0, \frac{2\sigma_\varepsilon^2}{(1+\lambda)\sigma_e^2}\right) = N(0,1).$$

3. $|\rho| < 1$, $\lambda = 1$ case

(a) This is the *panel cointegration* case. Note that

$$\frac{1}{nT^2}\widehat{G}_1 = \frac{1}{nT^2}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} - \frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT^{3/2}}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{x}}{nT^{3/2}}.$$

First we consider

$$\frac{1}{nT^2}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} = \frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\frac{1}{T^2}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i.$$

Expanding this equation, we get

$$
\begin{aligned}
& \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^2}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i \\
\approx\ & \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^2}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)^2 \\
=\ & \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^2}\sum_{t=1}^{T}\left[\varepsilon_{it}+(1-\rho)x_{it-1}-(\widetilde{\rho}-\rho)x_{it-1}\right]^2 \\
=\ & \frac{1}{n}\sum_{i=1}^{n}\left\{\begin{array}{c}
\frac{1}{T^2}\sum_{t=1}^{T}\varepsilon_{it}^2+\left\{\sqrt{nT}\left(\widetilde{\rho}-\rho\right)\right\}^2\frac{1}{nT^3}\sum_{t=1}^{T}x_{it-1}^2 \\
+(1-\rho)^2\frac{1}{T^2}\sum_{t=1}^{T}x_{it-1}^2 \\
-\sqrt{nT}\left(\widetilde{\rho}-\rho\right)\frac{2}{\sqrt{n}T^{5/2}}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} \\
+(1-\rho)\frac{2}{T^2}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} \\
-\sqrt{nT}\left(\widetilde{\rho}-\rho\right)(1-\rho)\frac{2}{\sqrt{n}T^{5/2}}\sum_{t=1}^{T}x_{it-1}^2
\end{array}\right\} \\
=\ & (1-\rho)^2\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2+I+II+III+IV+V.
\end{aligned}
$$

Consider I.

$$I=\frac{1}{T}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2=O_p\left(\frac{1}{T}\right)$$

using $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2=O_p(1)$.

Consider II.

$$II=\left\{\sqrt{nT}\left(\widetilde{\rho}-\rho\right)\right\}^2\frac{1}{nT}\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^2}\sum_{t=1}^{T}x_{it-1}^2=O_p\left(\frac{1}{nT}\right)$$

using $\sqrt{nT}\left(\widetilde{\rho}-\rho\right)=O_p(1)$ and $\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\frac{x_{it-1}}{\sqrt{T}}\right)^2\frac{1}{T}=O_p(1)$.

Consider III.

$$III=-2\left(\sqrt{nT}\left(\widetilde{\rho}-\rho\right)\right)\frac{1}{nT^{3/2}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{\varepsilon_{it}}{\sqrt{T}}\frac{x_{it-1}}{\sqrt{T}}=O_p\left(\frac{1}{nT^{3/2}}\right)$$

using $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{\varepsilon_{it}}{\sqrt{T}}\frac{x_{it-1}}{\sqrt{T}}=O_p(1)$.

Consider IV and V. It is easy to see that

$$IV = (1-\rho)\frac{2}{\sqrt{nT}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{\varepsilon_{it}}{\sqrt{T}}\frac{x_{it-1}}{\sqrt{T}} = O_p\left(\frac{1}{\sqrt{nT}}\right)$$

and

$$V = -(1-\rho)\left(\sqrt{nT}\left(\widetilde{\rho}-\rho\right)\right)\frac{1}{\sqrt{nT}}\frac{1}{n}\sum_{i=1}^{n}\frac{2}{T^2}\sum_{t=1}^{T}x_{it-1}^2 = O_p\left(\frac{1}{\sqrt{nT}}\right).$$

Therefore,

$$\begin{aligned}
& \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^2}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i \\
\approx\ & (1-\rho)^2\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^2}\sum_{t=1}^{T}x_{it-1}^2 \\
& +O_p\left(\frac{1}{T}\right)+O_p\left(\frac{1}{nT}\right)+O_p\left(\frac{1}{nT^{3/2}}\right)+O_p\left(\frac{1}{\sqrt{nT}}\right)+O_p\left(\frac{1}{\sqrt{nT}}\right) \\
=\ & (1-\rho)^2\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^2}\sum_{t=1}^{T}x_{it-1}^2+\max\left\{O_p\left(\frac{1}{T}\right),O_p\left(\frac{1}{\sqrt{nT}}\right)\right\}.
\end{aligned}$$

Finally, because we know

$$(1-\rho)^2\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^2}\sum_{t=1}^{T}x_{it-1}^2 \xrightarrow{p} \frac{\sigma_\varepsilon^2(1-\rho)^2}{2},$$

we have

$$\frac{1}{nT^2}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} = \frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\frac{1}{T^2}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i \xrightarrow{p} \frac{\sigma_\varepsilon^2(1-\rho)^2}{2\sigma_e^2}$$

as $(n,T)\to\infty$.

Next note that

$$
\begin{aligned}
\frac{1}{nT^{3/2}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT} &= \frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\left(\frac{1}{T^{3/2}}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T\right) \\
&\approx \frac{(1-\widetilde{\rho})}{\widetilde{\sigma}_e^2}\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^{3/2}}\sum_{t=1}^{T}(x_{it}-\widetilde{\rho}x_{it-1}) \\
&= \frac{(1-\widetilde{\rho})}{\widetilde{\sigma}_e^2}\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}[\varepsilon_{it}-(1-\widetilde{\rho})x_{it-1}] \\
&= I+II.
\end{aligned}
$$

For I, one can see that

$$
\frac{(1-\widetilde{\rho})}{\widehat{\sigma}_e^2}\frac{1}{\sqrt{n}T}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it} = O_p\left(\frac{1}{\sqrt{n}T}\right).
$$

For II,

$$
\frac{(1-\widetilde{\rho})^2}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_{it-1} = O_p\left(\frac{1}{\sqrt{n}}\right).
$$

Therefore, we have

$$
\frac{1}{nT^{3/2}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT} = O_p\left(\frac{1}{\sqrt{n}}\right) = o_p(1).
$$

Also note that

$$
\frac{1}{nT}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT} \approx \frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\frac{1}{T}T(1-\widetilde{\rho})^2 \xrightarrow{p} \frac{(1-\rho)^2}{\sigma_e^2}.
$$

Hence, we have

$$
\begin{aligned}
\frac{1}{nT^2}\widehat{G}_1 &= \frac{1}{nT^2}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} - \frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT^{3/2}}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{x}}{nT^{3/2}} \\
&\xrightarrow{p} \frac{(1-\rho)^2\sigma_\varepsilon^2}{2\sigma_e^2}
\end{aligned}
$$

as $(n,T)\to\infty$.

(b) Now we investigate $\widehat{G}_2$. Note that

$$\frac{1}{\sqrt{nT}}\widehat{G}_2=\frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\nu}-\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT^{3/2}}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\nu}}{\sqrt{nT}}.$$

We first consider

$$\frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\nu}=\frac{1}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{T}.$$

Here

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{T}\\
\approx\ &\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)\left(\nu_{it}-\widetilde{\rho}\nu_{it-1}\right)\\
=\ &\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\frac{1}{T}\sum_{t=1}^{T}\left[\varepsilon_{it}+(1-\rho)x_{it-1}-(\widetilde{\rho}-\rho)x_{it-1}\right]\left[e_{it}-(\widetilde{\rho}-\rho)\nu_{it-1}\right]\right\}\\
=\ &\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\begin{array}{c}\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}e_{it}+\frac{1}{T}\sum_{t=1}^{T}(1-\rho)x_{it-1}e_{it}\\ -\frac{\sqrt{nT}(\widetilde{\rho}-\rho)}{\sqrt{n}}\frac{1}{T^{3/2}}\sum_{t=1}^{T}\varepsilon_{it}\nu_{it-1}\\ -(1-\rho)\frac{\sqrt{nT}(\widetilde{\rho}-\rho)}{\sqrt{n}}\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_{it-1}\nu_{it-1}\\ -\frac{\sqrt{nT}(\widetilde{\rho}-\rho)}{\sqrt{n}}\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_{it-1}e_{it}\\ +\frac{(\sqrt{nT}(\widetilde{\rho}-\rho))^2}{n}\frac{1}{T^2}\sum_{t=1}^{T}x_{it-1}\nu_{it-1}\end{array}\right\}\\
=\ &\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}(1-\rho)x_{it-1}e_{it}+I+II+III+IV+V.
\end{aligned}$$

Consider I. One can see that

$$I=\frac{1}{\sqrt{T}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}e_{it}=O_p\left(\frac{1}{\sqrt{T}}\right)$$

using $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}e_{it}=O_p(1)$.

Consider II.

$$II = \frac{1}{\sqrt{nT}} \left(\sqrt{nT} \left(\widetilde{\rho} - \rho \right) \right) \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \varepsilon_{it} \nu_{it-1} = O_p \left(\frac{1}{\sqrt{nT}} \right)$$

using $\sqrt{nT} \left(\widetilde{\rho} - \rho \right) = O_p(1)$ and $\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \varepsilon_{it} \nu_{it-1} = O_p(1)$.

Consider III.

$$III = -(1-\rho)\sqrt{nT} \left(\widetilde{\rho} - \rho \right) \frac{1}{\sqrt{nT}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{x_{it-1}}{\sqrt{T}} \frac{\nu_{it-1}}{\sqrt{T}} = O_p \left(\frac{1}{\sqrt{nT}} \right)$$

using $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{x_{it-1}}{\sqrt{T}} \frac{\nu_{it-1}}{\sqrt{T}} = O_p(1)$.

Consider IV and V. In a similar vein as above, it is easy to see that

$$IV = -\sqrt{nT} \left(\widetilde{\rho} - \rho \right) \frac{1}{\sqrt{nT}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{x_{it-1}}{\sqrt{T}} \frac{e_{it}}{\sqrt{T}} = O_p \left(\frac{1}{\sqrt{nT}} \right)$$

and

$$V = (\sqrt{nT} \left(\widetilde{\rho} - \rho \right))^2 \frac{1}{nT} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{x_{it-1}}{\sqrt{T}} \frac{\nu_{it-1}}{\sqrt{T}} = O_p \left(\frac{1}{nT} \right).$$

Therefore,

$$\begin{aligned}
& \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\mathbf{x}_i' \widehat{\mathbf{A}}^{-1} \boldsymbol{\nu}_i}{T} \\
\approx \; & \frac{(1-\rho)}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{x_{it-1}}{\sqrt{T}} \frac{e_{it}}{\sqrt{T}} \\
& + O_p \left(\frac{1}{\sqrt{T}} \right) + O_p \left(\frac{1}{\sqrt{nT}} \right) + O_p \left(\frac{1}{\sqrt{nT}} \right) + O_p \left(\frac{1}{\sqrt{nT}} \right) + O_p \left(\frac{1}{nT} \right) \\
= \; & \frac{(1-\rho)}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{x_{it-1}}{\sqrt{T}} \frac{e_{it}}{\sqrt{T}} + O_p \left(\frac{1}{\sqrt{T}} \right).
\end{aligned}$$

Because it can be shown that

$$\frac{(1-\rho)}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it-1}}{\sqrt{T}}\frac{e_{it}}{\sqrt{T}}\xrightarrow{d}(1-\rho)N(0,\frac{\sigma_\varepsilon^2\sigma_e^2}{2}),$$

one concludes that

$$\frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\nu}=\frac{1}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{T}\xrightarrow{d}(1-\rho)N(0,\frac{\sigma_\varepsilon^2}{2\sigma_e^2})$$

as $(n,T)\to\infty$.

Next, recall that

$$\frac{1}{\sqrt{nT}}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\nu}\xrightarrow{d}(1-\rho)N(0,\frac{1}{\sigma_e^2})$$

as shown in 1.(2). Also, $\frac{1}{nT^{3/2}}\mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT}\xrightarrow{p}0$; $\frac{1}{nT}\boldsymbol{\iota}_{nT}'\Phi^{-1}\boldsymbol{\iota}_{nT}\xrightarrow{p}\frac{(1-\rho)^2}{\sigma_e^2}$ as shown above.

Hence, we get

$$\frac{1}{\sqrt{nT}}\widehat{G}_2\xrightarrow{d}N(0,\frac{(1-\rho)^2\sigma_\varepsilon^2}{2\sigma_e^2})$$

as $(n,T)\to\infty$.

(c) Finally, we conclude that

$$t_{FGLS}=\left[\frac{(1-\rho)^2\sigma_\varepsilon^2}{2\sigma_e^2}\right]^{-1/2}N(0,\frac{(1-\rho)^2\sigma_\varepsilon^2}{2\sigma_e^2})=N(0,1).$$

4. $\rho=1$, $\lambda=1$ case

(a) Recall

$$\frac{1}{nT}\widehat{G}_1=\frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x}-T\left(1-\widetilde{\rho}\right)\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT\left(1-\widetilde{\rho}\right)}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n\left(1-\widetilde{\rho}\right)}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{x}}{nT\left(1-\widetilde{\rho}\right)}.$$

In a similar process as above, we consider first

$$\frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} = \frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\frac{1}{T}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i.$$

Expanding this equation, we get

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i &\approx \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}\left(\varepsilon_{it}+x_{it-1}-\widetilde{\rho}x_{it-1}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{\begin{array}{c} \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}^2+\frac{1}{T^2}\left\{T\left(\widetilde{\rho}-1\right)\right\}^2\frac{1}{T}\sum_{t=1}^{T}x_{it-1}^2 \\ -\frac{2}{T}T\left(\widetilde{\rho}-1\right)\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} \end{array}\right\} \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2+I+II.
\end{aligned}$$

Consider I. It is easy to see that

$$I = \left\{T\left(\widetilde{\rho}-1\right)\right\}^2\frac{1}{T}\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2 = o_p\left(\frac{1}{T}\right)$$

using

$$T\left(\widetilde{\rho}-1\right) = o_p\left(1\right)$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\frac{x_{it-1}}{\sqrt{T}}\right)^2\frac{1}{T} = O_p\left(1\right).$$

For II,

$$II = -2T\left(\widetilde{\rho}-1\right)\frac{1}{\sqrt{nT}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} = o_p\left(\frac{1}{\sqrt{nT}}\right)$$

because $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{\varepsilon_{it}}{\sqrt{T}}\frac{x_{it-1}}{\sqrt{T}}=O_p(1)$. Hence,

$$\begin{aligned}
\frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} &= \frac{1}{n}\frac{1}{\widetilde{\sigma}_e^2}\sum_{i=1}^{n}\frac{1}{T}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i \\
&\approx \frac{1}{nT}\frac{1}{\hat{\sigma}_e^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2+o_p\left(\frac{1}{T}\right)+o_p\left(\frac{1}{\sqrt{n}T}\right) \\
&= \frac{1}{nT}\frac{1}{\hat{\sigma}_e^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2+o_p\left(\frac{1}{T}\right) \\
&\xrightarrow{p} \frac{\sigma_\varepsilon^2}{\sigma_e^2}
\end{aligned}$$

as $(n,T)\to\infty$.

Next note that

$$\begin{aligned}
\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT(1-\widehat{\rho})} &= \frac{1}{\widetilde{\sigma}_e^2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T(1-\widetilde{\rho})}\right) \\
&\approx \frac{1}{\widetilde{\sigma}_e^2}\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}-T(\widetilde{\rho}-1)\frac{1}{T^2}\sum_{t=1}^{T}x_{it-1}\right\} \\
&= \frac{1}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{nT}}\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}\right) \\
&\quad -\frac{T(\widetilde{\rho}-1)}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{nT}}\left(\frac{1}{\sqrt{n}T^{3/2}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}\right) \\
&= O_p\left(\frac{1}{\sqrt{nT}}\right)+o_p\left(\frac{1}{\sqrt{nT}}\right)\xrightarrow{p}0
\end{aligned}$$

because $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}=O_p(1)$ and $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it-1}}{\sqrt{T}}\frac{1}{T}=O_p(1)$ as $(n,T)\to\infty$.

We also know

$$\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n(1-\widetilde{\rho})}\xrightarrow{p}\frac{2}{\sigma_e^2}$$

as shown in 2.(a). Hence,

$$\frac{1}{nT}\widehat{G}_1\xrightarrow{p}\frac{\sigma_\varepsilon^2}{\sigma_e^2}$$

as $(n,T) \to \infty$.

(b) Now we turn to $\widehat{G}_2$.

$$\begin{aligned}
&\frac{1}{\sqrt{nT}}\widehat{G}_2 \\
&= \frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\nu} - T\left(1-\widetilde{\rho}\right)\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT\left(1-\widetilde{\rho}\right)}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n\left(1-\widetilde{\rho}\right)}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\nu}}{\sqrt{nT}\left(1-\widetilde{\rho}\right)}.
\end{aligned}$$

First, we consider

$$\frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\nu} = \frac{1}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}.$$

Here

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} &\approx \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(x_{it}-\widetilde{\rho}x_{it-1}\right)\left(\nu_{it}-\widetilde{\rho}\nu_{it-1}\right)\right\} \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\varepsilon_{it}-(\widetilde{\rho}-1)x_{it-1}\right]\left[e_{it}-(\widetilde{\rho}-1)\nu_{it-1}\right]\right\} \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\begin{array}{c}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}e_{it} \\ -\frac{T(\widetilde{\rho}-1)}{T^{3/2}}\sum_{t=1}^{T}\varepsilon_{it}\nu_{it-1} \\ -\frac{T(\widetilde{\rho}-1)}{T^{3/2}}\sum_{t=1}^{T}x_{it-1}e_{it} \\ +\frac{\{T(\widetilde{\rho}-1)\}^2}{T^{5/2}}\sum_{t=1}^{T}x_{it-1}\nu_{it-1}\end{array}\right\} \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}e_{it} + I + II + III.
\end{aligned}$$

Consider I.

$$I = -T\left(\widetilde{\rho}-1\right)\frac{1}{\sqrt{T}}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{\varepsilon_{it}}{\sqrt{T}}\frac{\nu_{it-1}}{\sqrt{T}}\right) = o_p\left(\frac{1}{\sqrt{T}}\right)$$

using

$$T\left(\widetilde{\rho}-1\right) = o_p\left(1\right)$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{\varepsilon_{it}}{\sqrt{T}}\frac{\nu_{it-1}}{\sqrt{T}}=O_p\left(1\right).$$

Consider II. In a similar vein as I, one can also verify that

$$II=-T\left(\widetilde{\rho}-1\right)\frac{1}{\sqrt{T}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it-1}}{\sqrt{T}}\frac{e_{it}}{\sqrt{T}}=o_p\left(\frac{1}{\sqrt{T}}\right)$$

using

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it-1}}{\sqrt{T}}\frac{e_{it}}{\sqrt{T}}=O_p\left(1\right).$$

Consider III.

$$III=\left\{T\left(\widetilde{\rho}-1\right)\right\}^2\frac{1}{\sqrt{T}}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it-1}}{\sqrt{T}}\frac{\nu_{it-1}}{\sqrt{T}}\frac{1}{T}\right)=o_p\left(\frac{1}{\sqrt{T}}\right)$$

using $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{x_{it-1}}{\sqrt{T}}\frac{\nu_{it-1}}{\sqrt{T}}\frac{1}{T}=O_p\left(1\right)$.

We conclude that

$$\begin{aligned}\frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\nu} &= \frac{1}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} \\ &\approx \frac{1}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}e_{it}+o_p\left(\frac{1}{\sqrt{T}}\right)+o_p\left(\frac{1}{\sqrt{T}}\right)+o_p\left(\frac{1}{\sqrt{T}}\right) \\ &= \frac{1}{\widetilde{\sigma}_e^2}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}e_{it}+o_p\left(\frac{1}{\sqrt{T}}\right) \\ &\xrightarrow{d} N(0,\frac{\sigma_\varepsilon^2}{\sigma_e^2})\end{aligned}$$

as $(n,T)\to\infty$.

Also recall that $\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{nT(1-\widetilde{\rho})}\xrightarrow{p}0$, $\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n(1-\widetilde{\rho})}\xrightarrow{p}\frac{2}{\sigma_e^2}$ from above and $\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\nu}}{\sqrt{nT}(1-\widetilde{\rho})}\xrightarrow{d}$ $N(0,\frac{1}{\sigma_e^2})$ from 2.(b).

Hence, the second term of $\frac{1}{\sqrt{nT}}\widehat{G}_2$ is $o_p(1)$ and we conclude that

$$\frac{1}{\sqrt{nT}}\widehat{G}_2 \xrightarrow{d} N(0, \frac{\sigma_\varepsilon^2}{\sigma_e^2})$$

as $(n, T) \to \infty$.

(c) Finally, we have

$$t_{FGLS} = \left[\frac{\sigma_\varepsilon^2}{\sigma_e^2}\right]^{-1/2} N(0, \frac{\sigma_\varepsilon^2}{\sigma_e^2}) = N(0, 1).$$

■

G Proof of Theorem 4

We study the following lemmas before proving Theorem 4.

Lemma 1 (B)

$$\hat{\sigma}_e^2 = I + II + III$$

where

$$\begin{aligned} I &= \tfrac{1}{n}\textstyle\sum_{i=1}^n \left[\tfrac{1}{T}\sum_{t=1}^T e_{it}^2 - 2\left(\widehat{\rho}-\rho\right)\tfrac{1}{T}\sum_{t=1}^T e_{it}\nu_{it-1} + \left(\widehat{\rho}-\rho\right)^2 \tfrac{1}{T}\sum_{t=1}^T \nu_{it-1}^2\right] \\ &\quad -\tfrac{T}{\hat{d}^2}\tfrac{1}{n}\textstyle\sum_{i=1}^n \left[\tfrac{1}{T}\sum_{t=1}^T e_{it} - \left(\widehat{\rho}-\rho\right)\tfrac{1}{T}\sum_{t=1}^T \nu_{it-1}\right]^2, \end{aligned}$$

$$\begin{aligned} II &= \tfrac{\left(\hat{\beta}_{OLS}-\beta\right)^2}{n}\textstyle\sum_{i=1}^n [\tfrac{1}{T}\sum_{t=1}^T \varepsilon_{it}^2 + \tfrac{(\widehat{\rho}-\rho)^2}{T}\sum_{t=1}^T x_{it-1}^2 + \tfrac{(\lambda-\rho)^2}{T}\sum_{t=1}^T x_{it-1}^2 \\ &\quad -\tfrac{2(\widehat{\rho}-\rho)}{T}\textstyle\sum_{t=1}^T \varepsilon_{it}x_{it-1} + \tfrac{2(\lambda-\rho)}{T}\sum_{t=1}^T \varepsilon_{it}x_{it-1} - \tfrac{2(\widehat{\rho}-\rho)(\lambda-\rho)}{T}\sum_{t=1}^T x_{it-1}^2] \\ &\quad -\tfrac{\left(\hat{\beta}_{OLS}-\beta\right)^2}{n}\left(\tfrac{T}{\hat{d}^2}\right)\textstyle\sum_{i=1}^n \left[\tfrac{1}{T}\sum_{t=1}^T \varepsilon_{it} - \left(\widehat{\rho}-\rho\right)\tfrac{1}{T}\sum_{t=1}^T x_{it-1} + (\lambda-\rho)\tfrac{1}{T}\sum_{t=1}^T x_{it-1}\right]^2, \end{aligned}$$

and $III \leq \sqrt{I \times II}$.

Proof. Note that

$$\hat{\mathbf{u}} = \mathbf{y} - \widehat{\gamma}_{OLS}\boldsymbol{\iota}_{nT} - \mathbf{x}\hat{\beta}_{OLS} = E_{nT}\left[\mathbf{u} + \mathbf{x}\left(\hat{\beta}_{OLS} - \beta\right)\right].$$

Because

$$\begin{aligned}
\left(I_n \otimes \hat{E}_T^{\alpha}\right)\hat{\mathbf{u}}^* &= \left(I_n \otimes \hat{E}_T^{\alpha}\hat{C}\right)\hat{\mathbf{u}} \\
&= \left(I_n \otimes \hat{E}_T^{\alpha}\hat{C}\right)\left(I_{nT} - \bar{J}_{nT}\right)\left[\mathbf{u} + \mathbf{x}\left(\hat{\beta}_{OLS} - \beta\right)\right] \\
&= \left(I_n \otimes \hat{E}_T^{\alpha}\hat{C}\right)\left[\mathbf{u} + \mathbf{x}\left(\hat{\beta}_{OLS} - \beta\right)\right] \\
&= \left(I_n \otimes \hat{E}_T^{\alpha}\hat{C}\right)\left[\left(I_n \otimes \boldsymbol{\iota}_T\right)\boldsymbol{\mu} + \boldsymbol{\nu} + \mathbf{x}\left(\hat{\beta}_{OLS} - \beta\right)\right] \\
&= \left(I_n \otimes \hat{E}_T^{\alpha}\hat{C}\right)\left[\boldsymbol{\nu} + \mathbf{x}\left(\hat{\beta}_{OLS} - \beta\right)\right]
\end{aligned}$$

using

$$\hat{E}_T^{\alpha}\hat{C}\boldsymbol{\iota}_T = (1 - \widehat{\rho})\hat{E}_T^{\alpha}\hat{\boldsymbol{\iota}}_T^{\alpha} = 0,$$

one can show that

$$\begin{aligned}
\hat{\sigma}_e^2 &= \frac{1}{n\left(T-1\right)}\hat{\mathbf{u}}^{*\prime}\left(I_n \otimes \hat{E}_T^{\alpha}\right)\hat{\mathbf{u}}^* \\
&= \frac{1}{n\left(T-1\right)}\hat{\mathbf{u}}^{\prime}\left(I_n \otimes \hat{C}^{\prime}\hat{E}_T^{\alpha}\hat{C}\right)\hat{\mathbf{u}} \\
&= \frac{1}{n\left(T-1\right)}\left[\boldsymbol{\nu} + \mathbf{x}\left(\hat{\beta}_{OLS} - \beta\right)\right]^{\prime}\left(I_n \otimes \hat{C}\hat{E}_T^{\alpha}\hat{C}\right)\left[\boldsymbol{\nu} + \mathbf{x}\left(\hat{\beta}_{OLS} - \beta\right)\right] \\
&= \frac{1}{n\left(T-1\right)}\boldsymbol{\nu}^{\prime}\left(I_n \otimes \hat{C}\hat{E}_T^{\alpha}\hat{C}\right)\boldsymbol{\nu} + \frac{1}{n\left(T-1\right)}\left(\hat{\beta}_{OLS} - \beta\right)^2\mathbf{x}^{\prime}\left(I_n \otimes \hat{C}\hat{E}_T^{\alpha}\hat{C}\right)\mathbf{x} \\
&\quad + \frac{2}{n\left(T-1\right)}\left(\hat{\beta}_{OLS} - \beta\right)\mathbf{x}^{\prime}\left(I_n \otimes \hat{C}\hat{E}_T^{\alpha}\hat{C}\right)\boldsymbol{\nu} \\
&= I + II + III.
\end{aligned}$$

Consider I.

$$\begin{aligned} I &= \frac{1}{n(T-1)}\boldsymbol{\nu}'\left(I_n\otimes\hat{C}'\hat{E}_T^{\alpha}\hat{C}\right)\boldsymbol{\nu} \\ &= \frac{1}{n(T-1)}\boldsymbol{\nu}'\left(I_n\otimes\hat{C}'\hat{C}\right)\boldsymbol{\nu}-\frac{1}{n(T-1)\hat{d}^2}\boldsymbol{\nu}'\left(I_n\otimes\hat{C}'\hat{\boldsymbol{\iota}}_T^{\alpha}\hat{\boldsymbol{\iota}}_T^{\alpha\prime}\hat{C}\right)\boldsymbol{\nu}. \end{aligned}$$

The first term in I is

$$\begin{aligned} &\frac{1}{n(T-1)}\boldsymbol{\nu}'\left(I_n\otimes\hat{C}'\hat{C}\right)\boldsymbol{\nu} \\ \approx\;& \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\left(\nu_{it}-\widehat{\rho}\nu_{it-1}\right)^2\right] \\ =\;& \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it}^2-2\left(\widehat{\rho}-\rho\right)\frac{1}{T}\sum_{t=1}^{T}e_{it}\nu_{it-1}+\left(\widehat{\rho}-\rho\right)^2\frac{1}{T}\sum_{t=1}^{T}\nu_{it-1}^2\right]. \end{aligned}$$

The second term in I is

$$\begin{aligned} \frac{1}{n(T-1)\hat{d}^2}\boldsymbol{\nu}'\left(I_n\otimes\hat{C}'\hat{\boldsymbol{\iota}}_T^{\alpha}\hat{\boldsymbol{\iota}}_T^{\alpha\prime}\hat{C}\right)\boldsymbol{\nu} &\approx \frac{T}{n\hat{d}^2}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\left(\nu_{it}-\widehat{\rho}\nu_{it-1}\right)\right]^2 \\ &= \frac{T}{n\hat{d}^2}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it}-\left(\widehat{\rho}-\rho\right)\frac{1}{T}\sum_{t=1}^{T}\nu_{it-1}\right]^2. \end{aligned}$$

Hence, we have

$$\begin{aligned} I \approx\;& \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it}^2-2\left(\widehat{\rho}-\rho\right)\frac{1}{T}\sum_{t=1}^{T}e_{it}\nu_{it-1}+\left(\widehat{\rho}-\rho\right)^2\frac{1}{T}\sum_{t=1}^{T}\nu_{it-1}^2\right] \\ &-\frac{T}{\hat{d}^2}\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it}-\left(\widehat{\rho}-\rho\right)\frac{1}{T}\sum_{t=1}^{T}\nu_{it-1}\right]^2. \end{aligned}$$

Consider II. In a similar vein as I, we get

$$\begin{aligned}
II &= \frac{1}{n(T-1)}\left(\hat{\beta}_{OLS}-\beta\right)^2 \mathbf{x}'\left(I_n \otimes \hat{C}\hat{E}_T^{\alpha}\hat{C}\right)\mathbf{x} \\
&\approx \frac{\left(\hat{\beta}_{OLS}-\beta\right)^2}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\left(x_{it}-\widehat{\rho}x_{it-1}\right)^2\right] \\
&\quad -\frac{\left(\hat{\beta}_{OLS}-\beta\right)^2 T}{n\hat{d}^2}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\left(x_{it}-\widehat{\rho}x_{it-1}\right)\right]^2 \\
&= \frac{\left(\hat{\beta}_{OLS}-\beta\right)^2}{n}\sum_{i=1}^{n}\left[\begin{array}{c}\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}^2+\frac{(\widehat{\rho}-\rho)^2}{T}\sum_{t=1}^{T}x_{it-1}^2 \\ +\frac{(\lambda-\rho)^2}{T}\sum_{t=1}^{T}x_{it-1}^2-\frac{2(\widehat{\rho}-\rho)}{T}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} \\ +\frac{2(\lambda-\rho)}{T}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1}-\frac{2(\widehat{\rho}-\rho)(\lambda-\rho)}{T}\sum_{t=1}^{T}x_{it-1}^2\end{array}\right] \\
&\quad -\frac{\left(\hat{\beta}_{OLS}-\beta\right)^2 T}{n\hat{d}^2}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}-\left(\widehat{\rho}-\rho\right)\frac{1}{T}\sum_{t=1}^{T}x_{it-1}+(\lambda-\rho)\frac{1}{T}\sum_{t=1}^{T}x_{it-1}\right]^2.
\end{aligned}$$

Consider III. By the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
III &= \frac{2}{n(T-1)}\left(\hat{\beta}_{OLS}-\beta\right)\mathbf{x}'\left(I_n \otimes \hat{C}\hat{E}_T^{\alpha}\hat{C}\right)\boldsymbol{\nu} \\
&\leq \sqrt{\left[\frac{1}{n(T-1)}\boldsymbol{\nu}'\left(I_n \otimes \hat{C}\hat{E}_T^{\alpha}\hat{C}\right)\boldsymbol{\nu}\right]\left[\frac{1}{n(T-1)}\left(\hat{\beta}_{OLS}-\beta\right)^2\mathbf{x}'\left(I_n \otimes \hat{C}\hat{E}_T^{\alpha}\hat{C}\right)\mathbf{x}\right]} \\
&= \sqrt{I \times II}.
\end{aligned}$$

■

Lemma 2 (B)

$$\frac{1}{T}\hat{\sigma}_{\alpha}^2 = I + II + III + IV + V + VI$$

where

$$I = (1-\hat{\rho})^2\frac{\hat{d}^2}{T}\left[\frac{1}{n}\sum_{i=1}^{n}\mu_i^2-\left(\frac{1}{n}\sum_{i=1}^{n}\mu_i\right)^2\right],$$

$$
\begin{aligned}
II &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{T}{\hat{d}^2}\right)\left[\frac{1}{T}\sum_{t=1}^{T}\left(\nu_{it}-\hat{\rho}\nu_{it-1}\right)\right]^2+\frac{(1-\hat{\rho})^2}{nT}\frac{\hat{d}^2}{T}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it}\right]^2 \\
&\quad -\frac{2(1-\hat{\rho})}{n}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\nu_{it}-\hat{\rho}\nu_{it-1}\right)\right]\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it}\right],
\end{aligned}
$$

$$
\begin{aligned}
III &= \frac{\left(\hat{\beta}_{OLS}-\beta\right)^2}{n}\sum_{i=1}^{n}\left(\frac{T}{\hat{d}^2}\right)\left[\frac{1}{T}\sum_{t=1}^{T}\left(x_{it}-\hat{\rho}x_{it-1}\right)\right]^2 \\
&\quad +\frac{\left(\hat{\beta}_{OLS}-\beta\right)^2(1-\hat{\rho})^2}{nT}\frac{\hat{d}^2}{T}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\right]^2 \\
&\quad -\frac{2\left(\hat{\beta}_{OLS}-\beta\right)^2(1-\hat{\rho})}{n}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\hat{\rho}x_{it-1}\right)\right]\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\right],
\end{aligned}
$$

and $IV \leq \sqrt{I \times II}$, $V \leq \sqrt{I \times III}$, $VI \leq \sqrt{II \times III}$.

Proof. It can be shown that

$$
\begin{aligned}
\left(I_n \otimes \widehat{\bar{J}}_T^{\alpha}\right)\hat{\mathbf{u}}^* &= \left(I_n \otimes \widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\hat{\mathbf{u}} \\
&= \left(I_n \otimes \widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\left[\left(I_n\otimes\boldsymbol{\iota}_T\right)\boldsymbol{\mu}+\boldsymbol{\nu}+\mathbf{x}\left(\hat{\beta}_{OLS}-\beta\right)\right] \\
&= (1-\hat{\rho})\left(E_n\otimes\hat{\boldsymbol{\iota}}_T^{\alpha}\right)\boldsymbol{\mu}+\left(I_n \otimes \widehat{\bar{J}}_T^{\alpha}\hat{C}\right)E_{nT}\boldsymbol{\nu} \\
&\quad +\left(I_n \otimes \widehat{\bar{J}}_T^{\alpha}\hat{C}\right)E_{nT}\mathbf{x}\left(\hat{\beta}_{OLS}-\beta\right).
\end{aligned}
$$

using

$$
\begin{aligned}
&\left(I_n \otimes \widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\left(I_n\otimes\boldsymbol{\iota}_T\right)\boldsymbol{\mu} \\
&= \left(I_n \otimes \widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_n\otimes\boldsymbol{\iota}_T-\bar{J}_n\otimes\boldsymbol{\iota}_T\right)\boldsymbol{\mu} \\
&= \left(E_n \otimes \widehat{\bar{J}}_T^{\alpha}\hat{C}\boldsymbol{\iota}_T\right)\boldsymbol{\mu} \\
&= (1-\hat{\rho})\left(E_n\otimes\hat{\boldsymbol{\iota}}_T^{\alpha}\right)\boldsymbol{\mu}
\end{aligned}
$$

where $\hat{C}\iota_T = (1-\hat{\rho})\,\hat{\boldsymbol{\iota}}_T^{\alpha}$ and $\widehat{\bar{J}}_T^{\alpha}\hat{\boldsymbol{\iota}}_T^{\alpha} = \hat{\boldsymbol{\iota}}_T^{\alpha}$. Therefore,

$$\begin{aligned}
& \frac{1}{T}\hat{\sigma}_\alpha^2 \\
= & \frac{1}{nT}\hat{\mathbf{u}}^{*\prime}\left(I_n \otimes \widehat{\bar{J}}_T^{\alpha}\right)\hat{\mathbf{u}}^* \\
= & \frac{1}{nT}(1-\hat{\rho})^2\hat{d}^2\boldsymbol{\mu}' E_n\boldsymbol{\mu} + \frac{1}{nT}\boldsymbol{\nu}'\left(I_{nT}-\bar{J}_{nT}\right)\left(I_n \otimes \hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\boldsymbol{\nu} \\
& +\frac{1}{nT}\left(\hat{\beta}_{OLS}-\beta\right)^2\mathbf{x}'\left(I_{nT}-\bar{J}_{nT}\right)\left(I_n \otimes \hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\mathbf{x} \\
& +\frac{2}{nT}(1-\hat{\rho})\,\boldsymbol{\mu}'\left(E_n \otimes \hat{\boldsymbol{\iota}}_T^{\alpha\prime}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\boldsymbol{\nu} \\
& +\frac{2}{nT}(1-\hat{\rho})\left(\hat{\beta}_{OLS}-\beta\right)\boldsymbol{\mu}'\left(E_n \otimes \hat{\boldsymbol{\iota}}_T^{\alpha\prime}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\mathbf{x} \\
& +\frac{2}{nT}\left(\hat{\beta}_{OLS}-\beta\right)\mathbf{v}'\left(I_{nT}-\bar{J}_{nT}\right)\left(I_n \otimes \hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\mathbf{x} \\
= & I + II + III + IV + V + VI.
\end{aligned}$$

Consider I.

$$\begin{aligned}
I &= \frac{1}{nT}(1-\hat{\rho})^2\hat{d}^2\boldsymbol{\mu}' E_n\boldsymbol{\mu} \\
&= (1-\hat{\rho})^2\frac{\hat{d}^2}{T}\left(\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{n} - \frac{\boldsymbol{\mu}'\boldsymbol{\iota}_n\boldsymbol{\iota}_n'\boldsymbol{\mu}}{n^2}\right) \\
&= (1-\hat{\rho})^2\frac{\hat{d}^2}{T}\left[\frac{1}{n}\sum_{i=1}^n \mu_i^2 - \left(\frac{1}{n}\sum_{i=1}^n \mu_i\right)^2\right].
\end{aligned}$$

Consider II.

$$\begin{aligned}
II &= \frac{1}{nT}\boldsymbol{\nu}'\left(I_{nT}-\bar{J}_{nT}\right)\left(I_n \otimes \hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\boldsymbol{\nu} \\
&= \frac{1}{nT}\boldsymbol{\nu}'\left(I_n \otimes \hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\boldsymbol{\nu} + \frac{1}{nT}\boldsymbol{\nu}'\bar{J}_{nT}\left(I_n \otimes \hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\bar{J}_{nT}\boldsymbol{\nu} \\
&\quad -\frac{2}{nT}\boldsymbol{\nu}'\left(I_n \otimes \hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\bar{J}_{nT}\boldsymbol{\nu} \\
&= \frac{1}{nT}\boldsymbol{\nu}'\left(I_n \otimes \hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\boldsymbol{\nu} + \frac{(1-\hat{\rho})^2\hat{d}^2}{nT^2}\boldsymbol{\nu}'\bar{J}_{nT}\boldsymbol{\nu} - \frac{2(1-\hat{\rho})}{nT^2}\boldsymbol{\nu}'\left(\bar{J}_n \otimes \hat{C}'\hat{\boldsymbol{\iota}}_T^{\alpha}\boldsymbol{\iota}_T'\right)\boldsymbol{\nu}.
\end{aligned}$$

This can be simplified as follows:

$$\begin{aligned}
II &\approx \frac{1}{nT}\frac{1}{\hat{d}^2}\sum_{i=1}^{n}\left[\sum_{t=1}^{T}(\nu_{it}-\hat{\rho}\nu_{it-1})\right]^2+\frac{(1-\hat{\rho})^2\hat{d}^2}{nT^2}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it}\right]^2 \\
&\quad -\frac{2(1-\hat{\rho})}{nT^2}\left[\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}(\nu_{it}-\hat{\rho}\nu_{it-1})\right]\left[\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{T}{\hat{d}^2}\right)\left[\frac{1}{T}\sum_{t=1}^{T}(\nu_{it}-\hat{\rho}\nu_{it-1})\right]^2+\frac{(1-\hat{\rho})^2}{nT}\frac{\hat{d}^2}{T}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it}\right]^2 \\
&\quad -\frac{2(1-\hat{\rho})}{n}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\nu_{it}-\hat{\rho}\nu_{it-1})\right]\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it}\right].
\end{aligned}$$

Consider IV.

$$\begin{aligned}
IV &= \frac{2}{nT}(1-\hat{\rho})\boldsymbol{\mu}'\left(E_n\otimes\hat{\imath}_T^{\alpha\prime}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\boldsymbol{\nu} \\
&\leq \sqrt{\left[\frac{1}{nT}(1-\hat{\rho})^2\hat{d}^2\boldsymbol{\mu}'E_n\boldsymbol{\mu}\right]\left[\frac{1}{nT}\boldsymbol{\nu}'\left(I_{nT}-\bar{J}_{nT}\right)\left(I_n\otimes\hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\boldsymbol{\nu}\right]} \\
&= \sqrt{I\times II}
\end{aligned}$$

by the Cauchy-Schwarz inequality. Consider III next. In a similar process as II, one can easily verify that

$$\begin{aligned}
III &\approx \frac{\left(\hat{\beta}_{OLS}-\beta\right)^2}{n}\sum_{i=1}^{n}\left(\frac{T}{\hat{d}^2}\right)\left[\frac{1}{T}\sum_{t=1}^{T}(x_{it}-\hat{\rho}x_{it-1})\right]^2 \\
&\quad +\frac{\left(\hat{\beta}_{OLS}-\beta\right)^2(1-\hat{\rho})^2}{nT}\frac{\hat{d}^2}{T}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\right]^2 \\
&\quad -\frac{2\left(\hat{\beta}_{OLS}-\beta\right)^2(1-\hat{\rho})}{n}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it}-\hat{\rho}x_{it-1})\right]\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\right].
\end{aligned}$$

We can obtain V and VI as well by using the Cauchy-Schwarz inequality. ■

Lemma 3 (B)

$$
\begin{aligned}
\mathbf{x}'\hat{\Phi}^{-1}\mathbf{x} &= \frac{1}{\hat{\sigma}_e^2}\sum_{i=1}^{n}\left(\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\mathbf{x}_i - \frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2}\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\mathbf{x}_i\right), \\
\mathbf{x}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT} &= \frac{1}{\hat{\sigma}_\alpha^2}\sum_{i=1}^{n}\left(\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T\right), \\
\boldsymbol{\iota}_{nT}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT} &= \frac{n\hat{\theta}}{\hat{\sigma}_\alpha^2}, \\
\mathbf{x}'\hat{\Phi}^{-1}\mathbf{u} &= \frac{1}{\hat{\sigma}_e^2}\sum_{i=1}^{n}\left[\frac{\mu_i\hat{\sigma}_e^2}{\hat{\sigma}_\alpha^2}\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T + \left(\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i - \frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2}\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i\right)\right], \\
\boldsymbol{\iota}_{nT}'\hat{\Phi}^{-1}\mathbf{u} &= \sum_{i=1}^{n}\left[\frac{1}{\hat{\sigma}_\alpha^2}\left(\hat{\theta}\mu_i + \boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i\right)\right].
\end{aligned}
$$

Proof. See Baltagi *et al.* (2008). ■

Now we are ready to prove Theorem 4.

Proof. Assume $(n,T)\to\infty$ and $\frac{n}{T}\to 0$.

1. When $|\rho|<1$, $|\lambda|<1$, if $\widehat{\rho}\xrightarrow{p}\rho$

   (a) First, let us show that $\hat{\sigma}_e^2$ is a consistent estimator.

   From Lemma 1 (B), we consider I first. It can be shown that

$$
\begin{aligned}
I &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it}^2 - 2\left(\widehat{\rho}-\rho\right)\frac{1}{T}\sum_{t=1}^{T}e_{it}\nu_{it-1} + \left(\widehat{\rho}-\rho\right)^2\frac{1}{T}\sum_{t=1}^{T}\nu_{it-1}^2\right] \\
&\quad - \left(\frac{T}{\hat{d}^2}\right)\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it} - \left(\widehat{\rho}-\rho\right)\frac{1}{T}\sum_{t=1}^{T}\nu_{it-1}\right]^2 \\
&\xrightarrow{p}\sigma_e^2
\end{aligned}
$$

as $(n,T)\to\infty$. This is because in the first term,

$$
\begin{aligned}
& \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it}^{2}-2\left(\widehat{\rho}-\rho\right)\frac{1}{T}\sum_{t=1}^{T}e_{it}\nu_{it-1}+\left(\widehat{\rho}-\rho\right)^{2}\frac{1}{T}\sum_{t=1}^{T}\nu_{it-1}^{2}\right] \\
= & \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it}^{2}-2\sqrt{nT}\left(\widehat{\rho}-\rho\right)\frac{1}{nT}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it}\nu_{it-1} \\
& +\left(\sqrt{nT}\left(\widehat{\rho}-\rho\right)\right)^{2}\frac{1}{nT}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it-1}^{2} \\
= & \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it}^{2}+O_{p}\left(\frac{1}{nT}\right)+O_{p}\left(\frac{1}{nT}\right) \\
= & \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it}^{2}+O_{p}\left(\frac{1}{nT}\right)\xrightarrow{p}\sigma_{e}^{2}
\end{aligned}
$$

using $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it}\nu_{it-1}=O_p(1)$, $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it-1}^{2}=O_p(1)$, and $\sqrt{nT}\left(\widehat{\rho}-\rho\right)=O_p(1)$.

Let us look at the second term. One can verify that

$$
\begin{aligned}
& \left(\frac{T}{\hat{d}^{2}}\right)\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it}-\left(\widehat{\rho}-\rho\right)\frac{1}{T}\sum_{t=1}^{T}\nu_{it-1}\right]^{2} \\
= & \left(\frac{T}{\hat{d}^{2}}\right)\frac{1}{T}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right)^{2} \\
& +\left(\frac{T}{\hat{d}^{2}}\right)\left(\sqrt{nT}\left(\widehat{\rho}-\rho\right)\right)^{2}\frac{1}{nT^{2}}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\nu_{it-1}\right)^{2} \\
& -2\left(\frac{T}{\hat{d}^{2}}\right)\frac{\sqrt{nT}\left(\widehat{\rho}-\rho\right)}{\sqrt{n}T^{3/2}}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\nu_{it-1}\right) \\
= & O_{p}(\frac{1}{T})+O_{p}(\frac{1}{nT^{2}})+O_{p}(\frac{1}{\sqrt{n}T^{3/2}})=O_{p}(\frac{1}{T})
\end{aligned}
$$

using

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right)^2 &= O_p(1),\\
\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\nu_{it-1}\right)^2 &= O_p(1),
\end{aligned}
$$

and

$$
\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\nu_{it-1}\right) = O_p(1).
$$

Also note that $\frac{T}{\hat{d}^2} = \frac{T}{\frac{2\hat{\rho}}{1-\hat{\rho}}+T} \xrightarrow{p} 1$ as $T \to \infty$.

Consider II. It can be also shown that

$$
\begin{aligned}
II &= \frac{\left(\hat{\beta}_{OLS}-\beta\right)^2}{n}\textstyle\sum_{i=1}^{n}[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}^2 + \frac{(\widehat{\rho}-\rho)^2}{T}\sum_{t=1}^{T}x_{it-1}^2 + \frac{(\lambda-\rho)^2}{T}\sum_{t=1}^{T}x_{it-1}^2\\
&\quad -\tfrac{2(\widehat{\rho}-\rho)}{T}\textstyle\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} + \frac{2(\lambda-\rho)}{T}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} - \frac{2(\widehat{\rho}-\rho)(\lambda-\rho)}{T}\sum_{t=1}^{T}x_{it-1}^2]\\
&\quad -\frac{\left(\hat{\beta}_{OLS}-\beta\right)^2 T}{n\hat{d}^2}\textstyle\sum_{i=1}^{n}\left[\begin{array}{c}\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it} - (\widehat{\rho}-\rho)\frac{1}{T}\sum_{t=1}^{T}x_{it-1}\\ +(\lambda-\rho)\frac{1}{T}\sum_{t=1}^{T}x_{it-1}\end{array}\right]^2\\
&= O_p(\tfrac{1}{nT}).
\end{aligned}
$$

This follows because if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$
\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right) \xrightarrow{d} N\left(0, \frac{\sigma_\mu^2}{\sigma_\varepsilon^2} + \frac{(1+\rho\lambda)(1-\lambda^2)\sigma_e^2}{(1-\rho\lambda)(1-\rho^2)\sigma_\varepsilon^2}\right)
$$

using a similar argument as in Phillips and Moon (1999) and Baltagi *et al.* (2008).

Consider III. From Lemma 1 (B), we know that

$$
III \leq \sqrt{I \times II} \xrightarrow{p} 0.
$$

Hence,

$$\hat{\sigma}_e^2 \xrightarrow{p} \sigma_e^2.$$

(b) Next, we show that $\hat{\sigma}_\mu^2$ is a consistent estimator of σ_μ^2. From Lemma 2 (B), one can see that

$$I = (1-\hat{\rho})^2 \frac{\hat{d}^2}{T} \left[\frac{1}{n} \sum_{i=1}^{n} \mu_i^2 - \left(\frac{1}{n} \sum_{i=1}^{n} \mu_i \right)^2 \right] \xrightarrow{p} (1-\rho)^2 \sigma_\mu^2.$$

Consider II next. It can be shown that

$$\begin{aligned} II &= \frac{1}{n} \sum_{i=1}^{n} \left(\frac{T}{\hat{d}^2} \right) \left[\frac{1}{T} \sum_{t=1}^{T} (\nu_{it} - \hat{\rho}\nu_{it-1}) \right]^2 + \frac{(1-\hat{\rho})^2}{nT} \frac{\hat{d}^2}{T} \left[\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \nu_{it} \right]^2 \\ &\quad - \frac{2(1-\hat{\rho})}{n} \left[\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (\nu_{it} - \hat{\rho}\nu_{it-1}) \right] \left[\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \nu_{it} \right] \\ &= O_p(\frac{1}{T}) + O_p(\frac{1}{n^{3/2}T^{3/2}}) + O_p(\frac{1}{n^2 T}) = O_p(\frac{1}{T}). \end{aligned}$$

For the first term, from (a), we know that

$$\frac{1}{n} \sum_{i=1}^{n} \left(\frac{T}{\hat{d}^2} \right) \left[\frac{1}{T} \sum_{t=1}^{T} (\nu_{it} - \hat{\rho}\nu_{it-1}) \right]^2 = O_p(\frac{1}{T}).$$

For the second term,

$$\begin{aligned} & \frac{(1-\hat{\rho})^2}{nT} \frac{\hat{d}^2}{T} \left[\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \nu_{it} \right]^2 \\ = & \frac{(1-\hat{\rho})^2}{n^{3/2}T^{3/2}} \frac{\hat{d}^2}{T} \left[\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \nu_{it} \right]^2 = O_p(\frac{1}{n^{3/2}T^{3/2}}) \end{aligned}$$

using

$$\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \nu_{it} = O_p(1).$$

Let us look at the last term. We have

$$
\begin{aligned}
& \frac{2\left(1-\hat{\rho}\right)}{n}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\nu_{it}-\hat{\rho}\nu_{it-1}\right)\right]\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it}\right] \\
= \; & \frac{2\left(1-\hat{\rho}\right)}{n}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it}-\left(\widehat{\rho}-\rho\right)\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it-1}\right]\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it}\right] \\
= \; & \frac{2\left(1-\hat{\rho}\right)}{n^2T}\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it}\right)\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it}\right) \\
& -\frac{2\left(1-\hat{\rho}\right)}{n^{5/2}T^{3/2}}\sqrt{nT}\left(\widehat{\rho}-\rho\right)\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it-1}\right)\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it}\right) \\
= \; & O_p(\frac{1}{n^2T})+O_p(\frac{1}{n^{5/2}T^{3/2}})=O_p(\frac{1}{n^2T}).
\end{aligned}
$$

Accordingly, we have

$$
IV \leq \sqrt{I \times II} \xrightarrow{p} 0.
$$

Finally, consider III.

$$
\begin{aligned}
III \;=\; & \frac{\left(\hat{\beta}_{OLS}-\beta\right)^2}{n}\sum_{i=1}^{n}\left(\frac{T}{\hat{d}^2}\right)\left[\frac{1}{T}\sum_{t=1}^{T}\left(x_{it}-\hat{\rho}x_{it-1}\right)\right]^2 \\
& +\frac{\left(\hat{\beta}_{OLS}-\beta\right)^2\left(1-\hat{\rho}\right)^2}{nT}\frac{\hat{d}^2}{T}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\right]^2 \\
& -\frac{2\left(\hat{\beta}_{OLS}-\beta\right)^2\left(1-\hat{\rho}\right)}{n}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\hat{\rho}x_{it-1}\right)\right]\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\right].
\end{aligned}
$$

It can be easily shown that $III = o_p(1)$ as $(n,T)\to\infty$ in a similar way as above. This follows because if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$
\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\xrightarrow{d} N\left(0, \frac{\sigma_\mu^2}{\sigma_\varepsilon^2}+\frac{(1+\rho\lambda)(1-\lambda^2)\sigma_e^2}{(1-\rho\lambda)(1-\rho^2)\sigma_\varepsilon^2}\right).
$$

Hence, by the fact that $V \leq \sqrt{I \times III} \xrightarrow{p} 0$ and $VI \leq \sqrt{II \times III} \xrightarrow{p} 0$, we get

$$\frac{1}{T}\hat{\sigma}_{\alpha}^{2} \xrightarrow{p} (1-\rho)^{2}\sigma_{\mu}^{2}.$$

One concludes that

$$\hat{\sigma}_{\mu}^{2} = \frac{1}{\hat{\theta}}\left(\hat{\sigma}_{\alpha}^{2} - \hat{\sigma}_{e}^{2}\right) = \frac{T}{\hat{\theta}}\left(\frac{\hat{\sigma}_{\alpha}^{2}}{T} - \frac{\hat{\sigma}_{e}^{2}}{T}\right) \xrightarrow{p} \frac{1}{(1-\rho)^{2}}\left[(1-\rho)^{2}\sigma_{\mu}^{2} - 0\right] = \sigma_{\mu}^{2}$$

using

$$\begin{aligned}
\frac{1}{T}\hat{\theta} &= \frac{1}{T}(1-\hat{\rho})^{2}\hat{d}^{2} \\
&= \frac{1}{T}(1-\hat{\rho})^{2}\left[\left(\frac{2\hat{\rho}}{1-\hat{\rho}}\right) + T\right] \\
&= (1-\hat{\rho})^{2}\left[\frac{1}{T}\left(\frac{2\hat{\rho}}{1-\hat{\rho}}\right) + 1\right] \xrightarrow{p} (1-\rho)^{2}.
\end{aligned}$$

(c) Let us calculate the term $\widehat{G}_1$ in equation (14) first.

$$\frac{1}{nT}\widehat{G}_{1} = \frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} - \frac{1}{T}\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{x}}{n}.$$

We investigate

$$\frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} = \frac{1}{n\hat{\sigma}_{e}^{2}}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_{i}'\hat{\mathbf{A}}^{-1}\mathbf{x}_{i}}{T} - \frac{\hat{\sigma}_{\mu}^{2}}{\hat{\sigma}_{\alpha}^{2}/T}\frac{\mathbf{x}_{i}'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_{T}}{T}\frac{\boldsymbol{\iota}_{T}'\hat{\mathbf{A}}^{-1}\mathbf{x}_{i}}{T}\right)$$

from Lemma 3 (B). As shown in Theorem 3.1.(a), one can see that

$$\frac{1}{n\hat{\sigma}_{e}^{2}}\sum_{i=1}^{n}\frac{1}{T}\mathbf{x}_{i}'\widehat{\mathbf{A}}^{-1}\mathbf{x}_{i} \xrightarrow{p} \frac{(1-2\rho\lambda+\rho^{2})\sigma_{\varepsilon}^{2}}{\left(1-\lambda^{2}\right)\sigma_{e}^{2}}.$$

Next, it can be shown that

$$
\begin{aligned}
& \frac{1}{n}\frac{1}{\widehat{\sigma}_e^2}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T}\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\mathbf{x}_i}{T}\right) \\
= & \frac{(1-\widehat{\rho})^2}{\widehat{\sigma}_e^2}\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}+\frac{(\lambda-\rho)}{T}\sum_{t=1}^{T}x_{it-1}-\frac{(\widehat{\rho}-\rho)}{T}\sum_{t=1}^{T}x_{it-1}\right]^2 \\
= & \frac{(1-\widehat{\rho})^2}{\widehat{\sigma}_e^2}\frac{1}{n}\sum_{i=1}^{n}\left[\begin{array}{c}
\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}\right)^2+\left(\frac{(\lambda-\rho)}{T}\sum_{t=1}^{T}x_{it-1}\right)^2 \\
+\left(\frac{(\widehat{\rho}-\rho)}{T}\sum_{t=1}^{T}x_{it-1}\right)^2 \\
+2\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}\right)\left(\frac{(\lambda-\rho)}{T}\sum_{t=1}^{T}x_{it-1}\right) \\
-2\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}\right)\left(\frac{(\widehat{\rho}-\rho)}{T}\sum_{t=1}^{T}x_{it-1}\right) \\
-2\left(\frac{(\lambda-\rho)}{T}\sum_{t=1}^{T}x_{it-1}\right)\left(\frac{(\widehat{\rho}-\rho)}{T}\sum_{t=1}^{T}x_{it-1}\right)
\end{array}\right] \\
= & I+II+III+IV+V+VI.
\end{aligned}
$$

Consider I and II. One can see that

$$
I=\frac{(1-\widehat{\rho})^2}{\widehat{\sigma}_e^2}\frac{1}{T}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}\right)^2=O_p\left(\frac{1}{T}\right)
$$

and

$$
II=\frac{(1-\widehat{\rho})^2}{\widehat{\sigma}_e^2}\frac{1}{T}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{(\lambda-\rho)}{\sqrt{T}}\sum_{t=1}^{T}x_{it-1}\right)^2=O_p\left(\frac{1}{T}\right).
$$

Consider III.

$$
III=\frac{(1-\widehat{\rho})^2}{\widehat{\sigma}_e^2}\left(\sqrt{nT}(\widehat{\rho}-\rho)\right)^2\frac{1}{nT^2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}x_{it-1}\right)^2=O_p\left(\frac{1}{nT^2}\right).
$$

Consider IV.

$$
IV=\frac{2\,(1-\widehat{\rho})^2}{\widehat{\sigma}_e^2}\frac{1}{T}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}\right)\left(\frac{(\lambda-\rho)}{\sqrt{T}}\sum_{t=1}^{T}x_{it-1}\right)=O_p\left(\frac{1}{T}\right)
$$

using
$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}\right)\left(\frac{(\lambda-\rho)}{\sqrt{T}}\sum_{t=1}^{T}x_{it-1}\right)=O_p\left(1\right).$$

Lastly, consider V and VI. It can be shown that

$$\begin{aligned}
V &= -\frac{2\left(1-\widehat{\rho}\right)^2}{\widehat{\sigma}_e^2}\frac{\sqrt{nT}(\widehat{\rho}-\rho)}{\sqrt{n}T^{3/2}}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}x_{it-1}\right)\\
&= O_p\left(\frac{1}{\sqrt{n}T^{3/2}}\right)
\end{aligned}$$

and

$$\begin{aligned}
VI &= -\frac{2\left(1-\widehat{\rho}\right)^2}{\widehat{\sigma}_e^2}\frac{\sqrt{nT}(\widehat{\rho}-\rho)}{\sqrt{n}T^{3/2}}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{(\lambda-\rho)}{\sqrt{T}}\sum_{t=1}^{T}x_{it-1}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}x_{it-1}\right)\\
&= O_p\left(\frac{1}{\sqrt{n}T^{3/2}}\right).
\end{aligned}$$

Therefore, we conclude that

$$\begin{aligned}
&\frac{1}{n}\frac{1}{\widehat{\sigma}_e^2}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T}\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\mathbf{x}_i}{T}\right)\\
&= O_p\left(\frac{1}{T}\right)+O_p\left(\frac{1}{T}\right)+O_p\left(\frac{1}{nT^2}\right)\\
&\quad +O_p\left(\frac{1}{T}\right)+O_p\left(\frac{1}{\sqrt{n}T^{3/2}}\right)+O_p\left(\frac{1}{\sqrt{n}T^{3/2}}\right)\\
&= O_p\left(\frac{1}{T}\right).
\end{aligned}$$

Then we have
$$\frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x}\xrightarrow{p}\frac{(1-2\rho\lambda+\rho^2)\sigma_\varepsilon^2}{\left(1-\lambda^2\right)\sigma_e^2}.$$

One can also verify that

$$\frac{1}{n}\mathbf{x}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT}=\frac{1}{\sqrt{nT}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_\alpha^2/T}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{\sqrt{T}}\right)=O_p\left(\frac{1}{\sqrt{nT}}\right)$$

and

$$\frac{1}{n}\boldsymbol{\iota}_{nT}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT} = \frac{\hat{\theta}/T}{\hat{\sigma}_\alpha^2/T} \xrightarrow{p} \frac{(1-\rho)^2}{(1-\rho)^2\sigma_\mu^2} = \frac{1}{\sigma_\mu^2}.$$

Finally, we get

$$\frac{1}{nT}\widehat{G}_1 \xrightarrow{p} \frac{(1-2\rho\lambda+\rho^2)\sigma_\varepsilon^2}{\left(1-\lambda^2\right)\sigma_e^2}$$

as $(n,T) \to \infty$.

Now we turn to $\widehat{G}_2$. Note that

$$\frac{1}{\sqrt{nT}}\widehat{G}_2 = \frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{u} - \frac{1}{\sqrt{T}}\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n}\left(\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{u}}{\sqrt{n}}.$$

Consider first

$$\begin{aligned} & \frac{1}{\sqrt{nT}}\mathbf{x}'\hat{\Phi}^{-1}\mathbf{u} \\ = \; & \frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\begin{array}{c} \frac{1}{\sqrt{T}}\frac{\hat{\sigma}_e^2}{\hat{\sigma}_\alpha^2/T}\mu_i\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T} \\ +\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} - \frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2/T}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T}\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}\right) \end{array}\right]. \end{aligned}$$

For the first term, one can show that

$$\begin{aligned} & \frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{T}}\left(\frac{\hat{\sigma}_e^2}{\hat{\sigma}_\alpha^2/T}\mu_i\right)\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T}\right) \\ = \; & (1-\hat{\rho})\left(\frac{1}{\hat{\sigma}_\alpha^2/T}\right)\frac{1}{T}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\mu_i\left[x_{it}-\hat{\rho}x_{it-1}\right] \\ = \; & (1-\hat{\rho})\left(\frac{1}{\hat{\sigma}_\alpha^2/T}\right)\frac{1}{T}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\mu_i\left[\begin{array}{c} \varepsilon_{it}+(\lambda-\rho)x_{it-1} \\ -\sqrt{nT}(\hat{\rho}-\rho)\frac{x_{it-1}}{\sqrt{nT}} \end{array}\right] \\ = \; & O_p\left(\frac{1}{T}\right)+O_p\left(\frac{1}{T}\right)+O_p\left(\frac{1}{\sqrt{n}T^{3/2}}\right) = O_p\left(\frac{1}{T}\right) \end{aligned}$$

using

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\mu_i\varepsilon_{it} = O_p\left(1\right)$$

and

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\mu_i x_{it-1} = O_p\left(1\right).$$

Also recall from Theorem 3.1.(b) that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} \xrightarrow{d} N(0, \frac{(1-2\rho\lambda+\rho^2)\sigma_e^2\sigma_\varepsilon^2}{1-\lambda^2})$$

as $(n,T)\to\infty$.

For the last term, it can be shown that

$$\frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{T}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2/T}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{\sqrt{T}}\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} = O_p\left(\frac{1}{\sqrt{T}}\right)$$

in a similar way as above.

Therefore, we conclude that

$$\begin{aligned}
\frac{1}{\sqrt{nT}}\mathbf{x}'\hat{\Phi}^{-1}\mathbf{u} &= \frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} + O_p\left(\frac{1}{T}\right) + O_p\left(\frac{1}{\sqrt{T}}\right) \\
&= \frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} + O_p\left(\frac{1}{\sqrt{T}}\right) \\
&\xrightarrow{d} N(0, \frac{(1-2\rho\lambda+\rho^2)\sigma_\varepsilon^2}{\left(1-\lambda^2\right)\sigma_e^2}).
\end{aligned}$$

Next, recall that $\frac{1}{n}\mathbf{x}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT} \xrightarrow{p} 0$ and $\frac{1}{n}\boldsymbol{\iota}_{nT}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT} \xrightarrow{p} \frac{1}{\sigma_\mu^2}$ from above.

Finally, it can be shown that

$$\begin{aligned}
\frac{1}{\sqrt{n}}\boldsymbol{\iota}_{nT}'\hat{\Phi}^{-1}\mathbf{u} =& \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\hat{\sigma}_e^2+\hat{\theta}\hat{\sigma}_\mu^2}\left(\hat{\theta}\mu_i + \iota_T'\hat{\mathbf{A}}^{-1}\nu_i\right) \\
= & \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\hat{\sigma}_\alpha^2/T}\left[\left(\frac{\hat{\theta}}{T}\right)\mu_i + \frac{1}{\sqrt{T}}\left(\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}\right)\right] \\
= & I + II.
\end{aligned}$$

Consider I.

$$\begin{aligned} I &= \left(\frac{\hat{\theta}}{T}\right)\left(\frac{1}{\hat{\sigma}_{\alpha}^2/T}\right)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mu_i \\ &\xrightarrow{d} \frac{1}{(1-\rho)^2\sigma_{\mu}^2}(1-\rho)^2 N\left(0,\sigma_{\mu}^2\right) \end{aligned}$$

using $\frac{\hat{\theta}}{T} \to (1-\rho)^2$ and $\hat{\sigma}_{\alpha}^2/T \to (1-\rho)^2\sigma_{\mu}^2$.

Consider II.

$$\begin{aligned} II &= \frac{1}{\hat{\sigma}_{\alpha}^2/T}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i = \frac{(1-\widehat{\rho})}{\hat{\sigma}_{\alpha}^2/T}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}(\nu_{it}-\widehat{\rho}\nu_{it-1}) \\ &= \frac{(1-\widehat{\rho})}{\hat{\sigma}_{\alpha}^2/T}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[e_{it}-(\widehat{\rho}-\rho)\nu_{it-1}\right] \\ &= \frac{(1-\widehat{\rho})}{\hat{\sigma}_{\alpha}^2/T}\frac{1}{\sqrt{T}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it} - \frac{(1-\widehat{\rho})}{\hat{\sigma}_{\alpha}^2/T}\frac{\sqrt{nT}(\widehat{\rho}-\rho)}{\sqrt{n}T}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it-1} \\ &= O_p\left(\frac{1}{\sqrt{T}}\right)+O_p\left(\frac{1}{\sqrt{n}T}\right) = O_p\left(\frac{1}{\sqrt{T}}\right). \end{aligned}$$

Hence,

$$\begin{aligned} \frac{1}{\sqrt{n}}\boldsymbol{\iota}_{nT}'\hat{\Phi}^{-1}\mathbf{u} &= \left(\frac{\hat{\theta}}{T}\right)\left(\frac{1}{\hat{\sigma}_{\alpha}^2/T}\right)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mu_i + O_p\left(\frac{1}{\sqrt{T}}\right) \\ &\xrightarrow{d} \frac{1}{(1-\rho)^2\sigma_{\mu}^2}(1-\rho)^2 N\left(0,\sigma_{\mu}^2\right) = N\left(0,\frac{1}{\sigma_{\mu}^2}\right). \end{aligned}$$

Summarizing, we have

$$\frac{1}{\sqrt{nT}}\widehat{G}_2 \xrightarrow{d} N\left(0,\frac{(1-2\rho\lambda+\rho^2)\sigma_{\varepsilon}^2}{(1-\lambda^2)\sigma_e^2}\right)$$

as $(n,T)\to\infty$. Finally,

$$t_{FGLS} = \left(\frac{1}{nT}\widehat{G}_1\right)^{-1/2}\left(\frac{1}{\sqrt{nT}}\widehat{G}_2\right) \xrightarrow{d} N\left(0,1\right).$$

2. When $\rho = 1$, $|\lambda| < 1$, if $T(\widehat{\rho} - 1) \xrightarrow{p} \kappa$

   (a) First, let us show that $\hat{\sigma}_e^2$ is a consistent estimator of σ_e^2.

   From Lemma 1 (B), it can be shown, in a similar way as 1.(a) that

$$
\begin{aligned}
I &= \frac{1}{n}\sum_{i=1}^{n}\left[\begin{array}{c}\frac{1}{T}\sum_{t=1}^{T} e_{it}^2 - \frac{2[T(\widehat{\rho}-1)]}{T}\left(\frac{1}{T}\sum_{t=1}^{T} e_{it}\nu_{it-1}\right) \\ +\frac{[T(\widehat{\rho}-1)]^2}{T}\left(\frac{1}{T^2}\sum_{t=1}^{T}\nu_{it-1}^2\right)\end{array}\right] \\
&\quad -\left(\frac{T}{\hat{d}^2}\right)\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T} e_{it} - \frac{1}{\sqrt{T}}\left[T(\widehat{\rho}-1)\right]\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}\nu_{it-1}\right)\right]^2 \\
&\quad \xrightarrow{p} \sigma_e^2
\end{aligned}
$$

   as $(n, T) \to \infty$ using $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} e_{it}^2 \xrightarrow{p} \sigma_e^2$, $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T} e_{it}\nu_{it-1} = O_p(1)$, $\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it-1}^2 = O_p(1)$, and $\frac{1}{\sqrt{n}T^{3/2}}\sum_{i=1}^{n}\sum_{t=1}^{T}\nu_{it-1} = O_p(1)$. Also note that $T(\widehat{\rho} - 1) \xrightarrow{p} \kappa$ and

$$
\frac{T}{\hat{d}^2} = \frac{T}{\frac{1+\widehat{\rho}}{1-\widehat{\rho}} + T - 1} = \frac{T(1-\widehat{\rho})}{2\widehat{\rho} + T(1-\widehat{\rho})} \xrightarrow{p} \frac{-\kappa}{2-\kappa}.
$$

   Consider II. From Lemma 1 (B), we have

$$
\begin{aligned}
&\tfrac{1}{n}\textstyle\sum_{i=1}^{n}[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}^2 + \frac{(\widehat{\rho}-1)^2}{T}\sum_{t=1}^{T}x_{it-1}^2 + \frac{(\lambda-1)^2}{T}\sum_{t=1}^{T}x_{it-1}^2 \\
&-\tfrac{2(\widehat{\rho}-1)}{T}\textstyle\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} + \frac{2(\lambda-1)}{T}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} - \frac{2(\widehat{\rho}-1)(\lambda-1)}{T}\sum_{t=1}^{T}x_{it-1}^2] \\
&-\tfrac{T}{n\hat{d}^2}\textstyle\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it} - (\widehat{\rho}-1)\frac{1}{T}\sum_{t=1}^{T}x_{it-1} + (\lambda-1)\frac{1}{T}\sum_{t=1}^{T}x_{it-1}\right]^2 \\
= \;& 2\sigma_\varepsilon^2/(1+\lambda) + o_p(1)
\end{aligned}
$$

   as $(n, T) \to \infty$. This is because $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2 \xrightarrow{p} \sigma_\varepsilon^2$, $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} = O_p(1)$, $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2 \xrightarrow{p} \sigma_\varepsilon^2/\left(1-\lambda^2\right)$, $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it} = O_p(1)$, and $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1} = O_p(1)$.

Hence,

$$II \approx \frac{\left[\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\right]^2}{n}\frac{1}{nT}\mathbf{x}'\left(I_n\otimes\hat{C}\hat{E}_T^{\alpha}\hat{C}\right)\mathbf{x} = O_p(\frac{1}{n}) = o_p(1)$$

using the fact that if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\xrightarrow{d}N\left(0,\frac{(1-\lambda)^2\sigma_e^2}{2\sigma_\varepsilon^2}\right).$$

This follows from a similar argument as in Phillips and Moon (1999) and Baltagi *et al.* (2008).

Since $III \leq \sqrt{I\times II}\xrightarrow{p}0$, we conclude that

$$\hat{\sigma}_e^2\xrightarrow{p}\sigma_e^2.$$

(b) Let us show that $\hat{\sigma}_\mu^2$ is not a consistent estimator of σ_μ^2.

Using Lemma 2 (B), we have

$$\begin{aligned}
&\frac{1}{T(1-\hat{\rho})}\hat{\sigma}_\alpha^2\\
=\;&\frac{1}{nT(1-\hat{\rho})}\hat{\mathbf{u}}^{*\prime}\left(I_n\otimes\widehat{\bar{J}}_T^{\alpha}\right)\hat{\mathbf{u}}^*\\
=\;&\frac{(1-\hat{\rho})\hat{d}^2}{nT}\boldsymbol{\mu}'E_n\boldsymbol{\mu}\\
&+\frac{1}{nT(1-\hat{\rho})}\boldsymbol{\nu}'\left(I_{nT}-\bar{J}_{nT}\right)\left(I_n\otimes\hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\boldsymbol{\nu}\\
&+\frac{1}{nT(1-\hat{\rho})}\left(\hat{\beta}_{OLS}-\beta\right)^2\mathbf{x}'\left(I_{nT}-\bar{J}_{nT}\right)\left(I_n\otimes\hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\mathbf{x}\\
&+\frac{2}{nT}\boldsymbol{\mu}'\left(E_n\otimes\hat{\boldsymbol{\iota}}_T^{\alpha\prime}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\boldsymbol{\nu}\\
&+\frac{2}{nT}\left(\hat{\beta}_{OLS}-\beta\right)\boldsymbol{\mu}'\left(E_n\otimes\hat{\boldsymbol{\iota}}_T^{\alpha\prime}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\mathbf{x}\\
&+\frac{1}{nT(1-\hat{\rho})}\left(\hat{\beta}_{OLS}-\beta\right)\boldsymbol{\nu}'\left(I_{nT}-\bar{J}_{nT}\right)\left(I_n\otimes\hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT}-\bar{J}_{nT}\right)\mathbf{x}\\
=\;&I+II+III+IV+V+VI.
\end{aligned}$$

Consider I. It is easy to see that

$$I = \frac{(1-\hat{\rho})\,\hat{d}^2}{T}\left[\frac{1}{n}\sum_{i=1}^{n}\mu_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n}\mu_i\right)^2\right] \xrightarrow{p} 0$$

using $(1-\hat{\rho})\,\hat{d}^2 = 2\widehat{\rho} + T\,(1-\widehat{\rho}) \xrightarrow{p} 2-\kappa$.

Consider II.

$$\begin{aligned}
II \;=\; & \tfrac{1}{n}\textstyle\sum_{i=1}^{n}\frac{1}{(1-\hat{\rho})\hat{d}^2}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it} + [T\,(1-\hat{\rho})]\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}\nu_{it-1}\right)\right]^2 \\
& + \tfrac{(1-\hat{\rho})\hat{d}^2}{nT}\left[\tfrac{1}{n}\textstyle\sum_{i=1}^{n}\frac{1}{T^{3/2}}\sum_{t=1}^{T}\nu_{it}\right]^2 \\
& - \tfrac{2T(1-\hat{\rho})}{nT}\left[\tfrac{1}{n}\textstyle\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it} + T\,(1-\hat{\rho})\frac{1}{T^{3/2}}\sum_{t=1}^{T}\nu_{it-1}\right)\right] \\
& \cdot\left[\tfrac{1}{n}\textstyle\sum_{i=1}^{n}\frac{1}{T^{3/2}}\sum_{t=1}^{T}\nu_{it}\right].
\end{aligned}$$

Let us look at the first term. It can be shown that

$$\begin{aligned}
& \frac{1}{(1-\hat{\rho})\,\hat{d}^2}\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it} + [T\,(1-\hat{\rho})]\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}\nu_{it-1}\right)\right]^2 \\
= \; & \frac{1}{(1-\hat{\rho})\,\hat{d}^2}\frac{1}{n}\sum_{i=1}^{n}\left[\begin{array}{c}
\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right)^2 \\
+[T\,(1-\hat{\rho})]^2\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}\nu_{it-1}\right)^2 \\
+2T\,(1-\hat{\rho})\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right)\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}\nu_{it-1}\right)
\end{array}\right] \\
= \; & \frac{1}{(1-\hat{\rho})\,\hat{d}^2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right)^2 \\
& +\frac{(T\,(1-\hat{\rho}))^2}{(1-\hat{\rho})\,\hat{d}^2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}\nu_{it-1}\right)^2 \\
& +\frac{2T\,(1-\hat{\rho})}{(1-\hat{\rho})\,\hat{d}^2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right)\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}\nu_{it-1}\right) \\
& \xrightarrow{p} \frac{1}{2-\kappa}\sigma_e^2\left(\frac{\kappa^2}{3}-\kappa+1\right)
\end{aligned}$$

since

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right)^{2} \xrightarrow{p} \sigma_e^2,$$
$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}\nu_{it-1}\right)^{2} \xrightarrow{p} \frac{\sigma_e^2}{3},$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right)\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}\nu_{it-1}\right) \xrightarrow{p} \frac{\sigma_e^2}{2}.$$

Also note that the second and third terms of II are $o_p(1)$ as $(n,T)\to\infty$. Therefore,

$$II \xrightarrow{p} \frac{\kappa^2-3\kappa+3}{3(2-\kappa)}\sigma_e^2$$

and from Lemma 2 (B),

$$IV \leq \sqrt{I\times II} \xrightarrow{p} 0.$$

Consider III next.

$$\begin{aligned} III \;=\; & \frac{\left(\hat{\beta}_{OLS}-\beta\right)^2}{n}\sum_{i=1}^{n}\left(\frac{1}{(1-\hat{\rho})\hat{d}^2}\right)\left[\begin{array}{c}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}-\frac{T(\hat{\rho}-1)}{T^{3/2}}\sum_{t=1}^{T}x_{it-1}\\ +\frac{(\lambda-1)}{\sqrt{T}}\sum_{t=1}^{T}x_{it-1}\end{array}\right]^2 \\ & +\frac{\left(\hat{\beta}_{OLS}-\beta\right)^2(1-\hat{\rho})\hat{d}^2}{nT^2}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}x_{it}\right]^2 \\ & -\frac{2\left(\hat{\beta}_{OLS}-\beta\right)^2}{n}\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it}-\hat{\rho}x_{i,t-1}\right)\right]\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}\right]. \end{aligned}$$

In a similar process as in II, one can verify that $III = O_p(\frac{1}{n})$ as $(n,T)\to\infty$ using the fact that if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right) \xrightarrow{d} N\left(0,\frac{(1-\lambda)^2\sigma_e^2}{2\sigma_\varepsilon^2}\right)$$

and accordingly that $V \leq \sqrt{I\times III} \xrightarrow{p} 0$, $VI \leq \sqrt{II\times III} \xrightarrow{p} 0$, respectively.

Summarizing, we have

$$\frac{\hat{\sigma}_{\alpha}^{2}}{T\left(1-\hat{\rho}\right)} \xrightarrow{p} \frac{\kappa^{2}-3\kappa+3}{3\left(2-\kappa\right)}\sigma_{e}^{2}.$$

Since $\hat{\sigma}_{\mu}^{2}=\frac{1}{\hat{\theta}}\left(\hat{\sigma}_{\alpha}^{2}-\hat{\sigma}_{e}^{2}\right)$ and $\hat{\theta}=\left(1-\hat{\rho}\right)^{2}\hat{d}^{2}$, we have

$$\begin{aligned}
\left(1-\hat{\rho}\right)\hat{\sigma}_{\mu}^{2} &= \left(1-\hat{\rho}\right)\tfrac{1}{\hat{\theta}}\left(\hat{\sigma}_{\alpha}^{2}-\hat{\sigma}_{e}^{2}\right)\\
&= \tfrac{1}{(1-\hat{\rho})\hat{d}^{2}}\left(\hat{\sigma}_{\alpha}^{2}-\hat{\sigma}_{e}^{2}\right)\\
&= \left(\tfrac{T}{\hat{d}^{2}}\right)\left(\tfrac{\hat{\sigma}_{\alpha}^{2}}{T(1-\hat{\rho})}\right)-\tfrac{\hat{\sigma}_{e}^{2}}{(1-\hat{\rho})\hat{d}^{2}}\\
&\xrightarrow{p} \left(\tfrac{-\kappa}{2-\kappa}\right)\left(\tfrac{\kappa^{2}-3\kappa+3}{3(2-\kappa)}\sigma_{e}^{2}\right)-\left(\tfrac{1}{2-\kappa}\sigma_{e}^{2}\right)=\tfrac{-\kappa^{3}+3\kappa^{2}-6}{3(2-\kappa)^{2}}\sigma_{e}^{2}.
\end{aligned}$$

If we plug $k=-3$ into this equation, we get

$$\left(1-\hat{\rho}\right)\hat{\sigma}_{\mu}^{2} \xrightarrow{p} \frac{48}{75}\sigma_{e}^{2}=\frac{16}{25}\sigma_{e}^{2}.$$

(c) We start from $\widehat{G}_1$ in equation (14). Let us define

$$\frac{1}{nT}\widehat{G}_{1}=\frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x}-\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n}\left(\frac{T}{n}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{x}}{n}.$$

From Lemma 3 (B), we have

$$\frac{1}{nT}\mathbf{x}'\hat{\Phi}^{-1}\mathbf{x}=\frac{1}{n\hat{\sigma}_{e}^{2}}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_{i}'\hat{\mathbf{A}}^{-1}\mathbf{x}_{i}}{T}-T\left(1-\widehat{\rho}\right)^{2}\frac{\hat{\sigma}_{\mu}^{2}}{\hat{\sigma}_{\alpha}^{2}}\frac{\mathbf{x}_{i}'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_{T}}{T\left(1-\widehat{\rho}\right)}\frac{\boldsymbol{\iota}_{T}'\hat{\mathbf{A}}^{-1}\mathbf{x}_{i}}{T\left(1-\widehat{\rho}\right)}\right).$$

Firstly, recall from Theorem 3.2.(a) that

$$\frac{1}{n}\frac{1}{\hat{\sigma}_{e}^{2}}\sum_{i=1}^{n}\frac{1}{T}\mathbf{x}_{i}'\widehat{\mathbf{A}}^{-1}\mathbf{x}_{i}\xrightarrow{p}\frac{2\sigma_{\varepsilon}^{2}}{(1+\lambda)\sigma_{e}^{2}}.$$

Note also that

$$\frac{1}{n\hat{\sigma}_e^2}\sum_{i=1}^{n}T\left(1-\widehat{\rho}\right)^2\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T\left(1-\widehat{\rho}\right)}\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\mathbf{x}_i}{T\left(1-\widehat{\rho}\right)}=O_p(\frac{1}{T})$$

using

$$\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T\left(1-\widehat{\rho}\right)}=O_p(\frac{1}{\sqrt{T}})$$

as shown in 1.(c) and

$$T\left(1-\widehat{\rho}\right)^2\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2}=\frac{\left(1-\widehat{\rho}\right)\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2/T\left(1-\widehat{\rho}\right)}\xrightarrow{p}\frac{\frac{-\kappa^3+3\kappa^2-6}{3(2-\kappa)^2}\sigma_e^2}{\frac{\kappa^2-3\kappa+3}{3(2-\kappa)}\sigma_e^2}=\frac{-k^3+3k^2-6}{(2-k)(k^2-3k+3)}.$$

Hence,

$$\frac{1}{nT}\mathbf{x}'\hat{\Phi}^{-1}\mathbf{x}\xrightarrow{p}\frac{2\sigma_\varepsilon^2}{(1+\lambda)\sigma_e^2}$$

as $(n,T)\to\infty$.

Next, one can shown in a similar way that

$$\begin{aligned}\frac{1}{n}\mathbf{x}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT} &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_\alpha^2/T\left(1-\widehat{\rho}\right)}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T\left(1-\widehat{\rho}\right)}\right)\\ &= \frac{1}{\hat{\sigma}_\alpha^2/T\left(1-\widehat{\rho}\right)}\frac{1}{\sqrt{nT}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{\sqrt{T}\left(1-\widehat{\rho}\right)}\right)\\ &= O_p(\frac{1}{\sqrt{nT}}).\end{aligned}$$

Also

$$\frac{T}{n}\boldsymbol{\iota}_{nT}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT}=\frac{\hat{\theta}/\left(1-\widehat{\rho}\right)}{\hat{\sigma}_\alpha^2/T\left(1-\widehat{\rho}\right)}\xrightarrow{p}\frac{2-k}{\frac{\kappa^2-3\kappa+3}{3(2-\kappa)}\sigma_e^2}=\frac{1}{\sigma_e^2}\frac{3\left(2-\kappa\right)^2}{\kappa^2-3\kappa+3}.$$

Therefore,

$$\frac{1}{nT}\widehat{G}_1\xrightarrow{p}\frac{2\sigma_\varepsilon^2}{(1+\lambda)\sigma_e^2}.$$

Now we turn to $\widehat{G}_2$.

$$\frac{1}{\sqrt{nT}}\widehat{G}_2 = \frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{u} - \frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n}\left(\frac{T}{n}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\frac{\sqrt{T}}{\sqrt{n}}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{u}.$$

We have

$$\begin{aligned}
\frac{1}{\sqrt{nT}}\mathbf{x}'\hat{\Phi}^{-1}\mathbf{u} &= \frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{T}}\left(\frac{\hat{\sigma}_e^2}{\hat{\sigma}_\alpha^2/T\left(1-\widehat{\rho}\right)}\right)\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T\left(1-\widehat{\rho}\right)}\right)\mu_i \\
&\quad+\frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} \\
&\quad-T\left(1-\widehat{\rho}\right)^2\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2}\frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T\left(1-\widehat{\rho}\right)}\right)\left(\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}\left(1-\widehat{\rho}\right)}\right) \\
&= I + II + III.
\end{aligned}$$

For I, with the joint CLT we have

$$\begin{aligned}
I &= \frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{T}}\left(\frac{\hat{\sigma}_e^2}{\hat{\sigma}_\alpha^2/T\left(1-\widehat{\rho}\right)}\right)\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T\left(1-\widehat{\rho}\right)}\right)\mu_i \\
&= \left(\frac{1}{\hat{\sigma}_\alpha^2/T\left(1-\widehat{\rho}\right)}\right)\frac{1}{T}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\mu_i\left[x_{it}-\hat{\rho}x_{it-1}\right] \\
&= \left(\frac{1}{\hat{\sigma}_\alpha^2/T\left(1-\widehat{\rho}\right)}\right)\frac{1}{T}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\mu_i\left[\varepsilon_{it}+(\lambda-1)x_{it-1}-T(\hat{\rho}-1)\frac{x_{it-1}}{T}\right] \\
&= O_p\left(\frac{1}{T}\right)+O_p\left(\frac{1}{T}\right)+O_p\left(\frac{1}{\sqrt{n}T^2}\right)=O_p\left(\frac{1}{T}\right).
\end{aligned}$$

For II, recall from Theorem 3 that

$$\frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}\xrightarrow{d}N(0,\frac{2\sigma_\varepsilon^2}{\left(1+\lambda\right)\sigma_e^2}).$$

For III, it is easy to see $III = O_p\left(\frac{1}{\sqrt{T}}\right)$ using

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{\sqrt{T}\left(1-\widehat{\rho}\right)}\right)\left(\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}\left(1-\widehat{\rho}\right)}\right) = O_p\left(1\right).$$

Also, as shown already, $T\left(1-\widehat{\rho}\right)^2\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2} \xrightarrow{p} \frac{-k^3+3k^2-6}{(2-k)(k^2-3k+3)}$ and $\hat{\sigma}_\alpha^2/T\left(1-\widehat{\rho}\right) \xrightarrow{p} \frac{\kappa^2-3\kappa+3}{3(2-\kappa)}\sigma_e^2$.

Finally, we conclude that

$$\frac{1}{\sqrt{nT}}\mathbf{x}'\hat{\Phi}^{-1}\mathbf{u} \xrightarrow{d} N(0,\frac{2\sigma_\varepsilon^2}{(1+\lambda)\sigma_e^2})$$

as $(n,T)\to\infty$.

Next it can be shown that

$$\begin{aligned}\frac{\sqrt{T}}{\sqrt{n}}\boldsymbol{\iota}_{nT}'\hat{\Phi}^{-1}\mathbf{u} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\begin{array}{c}\left(\frac{1}{\sqrt{T}}\right)\frac{\hat{\theta}/(1-\hat{\rho})}{\hat{\sigma}_\alpha^2/T(1-\hat{\rho})}\mu_i \\ +\frac{1}{\hat{\sigma}_\alpha^2/T(1-\hat{\rho})}\left(\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}(1-\widehat{\rho})}\right)\end{array}\right] \\ &= O_p\left(1\right)\end{aligned}$$

using

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}\left(1-\widehat{\rho}\right)} \\ =\ &\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}+\frac{T\left(1-\widehat{\rho}\right)}{T^{3/2}}\sum_{t=1}^{T}\nu_{it-1}\right] \\ =\ &O_p\left(1\right)\end{aligned}$$

and

$$\begin{aligned}\hat{\theta}/\left(1-\hat{\rho}\right) &\xrightarrow{p} 2-k, \\ \frac{\hat{\sigma}_\alpha^2}{T\left(1-\hat{\rho}\right)} &\xrightarrow{p} \frac{\kappa^2-3\kappa+3}{3\left(2-\kappa\right)}\sigma_e^2,\end{aligned}$$

respectively.

Therefore,

$$\frac{1}{\sqrt{nT}}\widehat{G}_2 \xrightarrow{d} N(0, \frac{2\sigma_\varepsilon^2}{(1+\lambda)\sigma_e^2})$$

as $(n,T) \to \infty$ using $\frac{1}{n}\mathbf{x}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT} \xrightarrow{p} 0$ and $\frac{T}{n}\boldsymbol{\iota}'_{nT}\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT} \xrightarrow{p} \frac{1}{\sigma_e^2}\frac{3(2-\kappa)^2}{\kappa^2-3\kappa+3}$ from above.

We conclude that

$$t_{FGLS} = \left(\frac{1}{nT}\widehat{G}_1\right)^{-1/2}\left(\frac{1}{\sqrt{nT}}\widehat{G}_2\right) \xrightarrow{d} N(0,1).$$

3. When $|\rho| < 1$, $\lambda = 1$, if $\widehat{\rho} \xrightarrow{p} \rho$

   (a) First, let us show that $\hat{\sigma}_e^2$ is a consistent estimator of σ_e^2. From Lemma 1 (B), we have

$$\begin{aligned} I &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it}^2 - 2\left(\widehat{\rho}-\rho\right)\frac{1}{T}\sum_{t=1}^{T}e_{it}\nu_{it-1} + \left(\widehat{\rho}-\rho\right)^2\frac{1}{T}\sum_{t=1}^{T}\nu_{it-1}^2\right] \\ &\quad - \left(\frac{T}{\hat{d}^2}\right)\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}e_{it} - \left(\widehat{\rho}-\rho\right)\frac{1}{T}\sum_{t=1}^{T}\nu_{it-1}\right]^2 \\ &\xrightarrow{p} \sigma_e^2 \end{aligned}$$

   as shown in 1.(a).

Consider II. It can be shown similarly that

$$\begin{aligned} II &= \frac{\left(\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2}{n^2}\sum_{i=1}^{n}\left[\begin{array}{c} \frac{1}{T^2}\sum_{t=1}^{T}\varepsilon_{it}^2 \\ +\frac{\left(\sqrt{nT}(\widehat{\rho}-\rho)\right)^2}{nT^3}\sum_{t=1}^{T}x_{it-1}^2 \\ +\frac{(1-\rho)^2}{T^2}\sum_{t=1}^{T}x_{it-1}^2 \\ -\frac{2\sqrt{nT}(\widehat{\rho}-\rho)}{\sqrt{n}T^{5/2}}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} \\ +\frac{2(1-\rho)}{T^2}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} \\ -\frac{2\sqrt{nT}(\widehat{\rho}-\rho)(1-\rho)}{\sqrt{n}T^{5/2}}\sum_{t=1}^{T}x_{it-1}^2 \end{array}\right] \\ &\quad -\frac{\left(\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2}{n^2}\frac{T}{\hat{d}^2}\sum_{i=1}^{n}\left[\begin{array}{c} \frac{1}{T^{3/2}}\sum_{t=1}^{T}\varepsilon_{it} \\ -\sqrt{nT}\left(\widehat{\rho}-\rho\right)\frac{1}{\sqrt{n}T^2}\sum_{t=1}^{T}x_{it-1} \\ +(1-\rho)\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_{it-1} \end{array}\right]^2 \\ &= o_p(1). \end{aligned}$$

This follows because, if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, we get

$$\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\xrightarrow{d} N\left(0,\frac{4\sigma_\mu^2}{3\sigma_\varepsilon^2}\right)$$

using a similar argument as in Phillips and Moon (1999) and Baltagi *et al.* (2008). Also note that

$$\sqrt{nT}\left(\widehat{\rho}-\rho\right)\xrightarrow{d} N\left(0,1-\rho^2\right)$$

from Lemma 1. This is because $\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2 = O_p(1)$, $\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1}^2 = O_p(1)$, $\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1} = O_p(1)$, and $\frac{1}{\sqrt{n}T^{3/2}}\sum_{i=1}^{n}\sum_{t=1}^{T}x_{it-1} = O_p(1)$. Also note that from Lemma 1 (B),

$$III \le \sqrt{I\times II}\xrightarrow{p} 0.$$

We conclude that

$$\hat{\sigma}_e^2 \xrightarrow{p} \sigma_e^2.$$

(b) Next, let us show that $\hat{\sigma}_\mu^2$ is a consistent estimator of σ_μ^2. From Lemma 2 (B), we know that $I \xrightarrow{p} (1-\rho)^2 \sigma_\mu^2$, $II \xrightarrow{p} 0$, and accordingly $IV \leq \sqrt{I \times II} \xrightarrow{p} 0$ as shown 1.(b).

Let us look at III.

$$\begin{aligned} III &= \frac{\left(\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2}{n^2} \sum_{i=1}^{n} \left(\frac{T}{\hat{d}^2}\right) \left[\frac{1}{T^{3/2}} \sum_{t=1}^{T} (x_{it} - \hat{\rho} x_{it-1})\right]^2 \\ &+ \frac{\left(\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2 (1-\hat{\rho})^2}{n^2} \frac{\hat{d}^2}{T} \left[\frac{1}{nT^2} \sum_{i=1}^{n}\sum_{t=1}^{T} x_{it}\right]^2 \\ &- \frac{2\left(\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2 (1-\hat{\rho})}{n^2} \left[\frac{1}{nT^{3/2}} \sum_{i=1}^{n}\sum_{t=1}^{T} (x_{it} - \hat{\rho} x_{it-1})\right] \\ &\cdot \left[\frac{1}{nT^{3/2}} \sum_{i=1}^{n}\sum_{t=1}^{T} x_{it}\right]. \end{aligned}$$

With a similar process to 2.(b), it can be shown that $III = o_p(1)$ because if $(n,T) \to \infty$ and $\frac{n}{T} \to 0$, then

$$\sqrt{nT}\left(\hat{\beta}_{OLS}-\beta\right) \xrightarrow{d} N\left(0, \frac{4\sigma_\mu^2}{3\sigma_\varepsilon^2}\right).$$

Hence, with $V \leq \sqrt{I \times III} \xrightarrow{p} 0$ and $VI \leq \sqrt{II \times III} \xrightarrow{p} 0$, we finally have

$$\frac{1}{T}\hat{\sigma}_\alpha^2 \xrightarrow{p} (1-\rho)^2 \sigma_\mu^2$$

and accordingly

$$\hat{\sigma}_\mu^2 = \frac{1}{\hat{\theta}}\left(\hat{\sigma}_\alpha^2 - \hat{\sigma}_e^2\right) = \frac{T}{\hat{\theta}}\left(\frac{\hat{\sigma}_\alpha^2}{T} - \frac{\hat{\sigma}_e^2}{T}\right) \xrightarrow{p} \frac{1}{(1-\rho)^2}\left[(1-\rho)^2 \sigma_\mu^2 - 0\right] = \sigma_\mu^2.$$

(c) Let us start from the term $\widehat{G}_1$ in equation (14). Recall

$$\frac{1}{nT^2}\widehat{G}_1 = \frac{1}{nT^2}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} - \frac{1}{T}\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n\sqrt{T}}\left(\frac{\boldsymbol{\iota}'_{nT}\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n}\right)^{-1}\frac{\boldsymbol{\iota}'_{nT}\widehat{\Phi}^{-1}\mathbf{x}}{n\sqrt{T}}.$$

From Lemma 3 (B), we have

$$\frac{1}{nT^2}\mathbf{x}'\hat{\Phi}^{-1}\mathbf{x} = \frac{1}{n\hat{\sigma}_e^2}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\mathbf{x}_i}{T^2} - \frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2/T}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T^{3/2}}\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\mathbf{x}_i}{T^{3/2}}\right).$$

From Theorem 3.3.(a), we have

$$\frac{1}{n\hat{\sigma}_e^2}\sum_{i=1}^{n}\frac{1}{T^2}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i \xrightarrow{p} (1-\rho)^2\frac{\sigma_\varepsilon^2}{2\sigma_e^2}.$$

Next, it can be shown that

$$\frac{1}{n\hat{\sigma}_e^2}\sum_{i=1}^{n}\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2/T}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T^{3/2}}\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\mathbf{x}_i}{T^{3/2}} \xrightarrow{p} (1-\rho)^2\frac{\sigma_\varepsilon^2}{3\sigma_e^2}$$

using the fact that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T^{3/2}}\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\mathbf{x}_i}{T^{3/2}} \\
=\;& \frac{1}{n}\sum_{i=1}^{n}\left[(1-\widehat{\rho})\frac{1}{T^{3/2}}\sum_{t=1}^{T}(x_{it}-\widehat{\rho}x_{it-1})\right]^2 \\
=\;& \frac{1}{n}\sum_{i=1}^{n}\left[(1-\widehat{\rho})\frac{1}{T^{3/2}}\sum_{t=1}^{T}((1-\widehat{\rho})\,x_{it-1}+\varepsilon_{it})\right]^2 \\
=\;& \frac{1}{n}\sum_{i=1}^{n}\left[\begin{array}{c}(1-\widehat{\rho})^4\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_{it-1}\right)^2+(1-\widehat{\rho})^2\frac{1}{T^2}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}\right)^2\\ (1-\widehat{\rho})^3\frac{1}{T}\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_{it-1}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}\right)\end{array}\right] \\
=\;& \frac{1}{n}\sum_{i=1}^{n}(1-\widehat{\rho})^4\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_{it-1}\right)^2+O_p\left(\frac{1}{T^2}\right)+O_p\left(\frac{1}{T}\right) \\
=\;& (1-\widehat{\rho})^4\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_{it-1}\right)^2+O_p\left(\frac{1}{T}\right)\xrightarrow{p}(1-\rho)^4\frac{\sigma_\varepsilon^2}{3}
\end{aligned}$$

and $\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2/T} \xrightarrow{p} \frac{1}{(1-\rho)^2}$. Hence,

$$\begin{aligned} \frac{1}{nT^2}\mathbf{x}'\Phi^{-1}\mathbf{x} &= \frac{1}{n\hat{\sigma}_e^2}\sum_{i=1}^{n}\left(\frac{1}{T^2}\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i - \frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2/T}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T^{3/2}}\frac{\boldsymbol{\iota}_T'\widehat{\mathbf{A}}^{-1}\mathbf{x}_i}{T^{3/2}}\right) \\ &\xrightarrow{p} (1-\rho)^2\frac{\sigma_\varepsilon^2}{2\sigma_e^2} - (1-\rho)^2\frac{\sigma_\varepsilon^2}{3\sigma_e^2} = \frac{(1-\rho)^2\,\sigma_\varepsilon^2}{6\sigma_e^2} \end{aligned}$$

as $(n,T) \to \infty$.

Next, one can verify that

$$\begin{aligned} \frac{1}{n\sqrt{T}}\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT} &= \frac{1}{n}\frac{1}{\hat{\sigma}_\alpha^2/T}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T^{3/2}}\right) \\ &= O_p(\frac{1}{\sqrt{n}}) \end{aligned}$$

as shown in Theorem 3.3.(a). Also recall that

$$\frac{1}{n}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT} \xrightarrow{p} \frac{1}{\sigma_\mu^2}$$

as shown in 1.(c).

Hence, we have

$$\frac{1}{nT^2}\widehat{G}_1 \xrightarrow{p} \frac{(1-\rho)^2\,\sigma_\varepsilon^2}{6\sigma_e^2}$$

as $(n,T) \to \infty$.

Now we investigate $\widehat{G}_2$. Let

$$\frac{1}{\sqrt{n}T}\widehat{G}_2 = \frac{1}{\sqrt{n}T}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{u} - \frac{1}{\sqrt{T}}\frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n\sqrt{T}}\left(\frac{1}{n}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\frac{1}{\sqrt{n}}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{u}.$$

Consider that

$$
\begin{aligned}
& \frac{1}{\sqrt{n}T}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{u} \\
= & \frac{1}{\widehat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\begin{array}{c} \frac{1}{\sqrt{T}}\left(\frac{\widehat{\sigma}_e^2}{\hat{\sigma}_\alpha^2/T}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T^{3/2}}\right)\mu_i \\ +\left(\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{T}-\frac{\widehat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2/T}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T^{3/2}}\frac{\boldsymbol{\iota}_T'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}\right)\end{array}\right].
\end{aligned}
$$

Firstly, in a similar vein as 1.(c) it can be shown that

$$
\frac{1}{\widehat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{T}}\left(\frac{\widehat{\sigma}_e^2}{\hat{\sigma}_\alpha^2/T}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T^{3/2}}\right)\mu_i = O_p\left(\frac{1}{\sqrt{T}}\right)
$$

using

$$
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\widehat{\sigma}_e^2}{\hat{\sigma}_\alpha^2/T}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T^{3/2}}\right)\mu_i = O_p(1).
$$

Next, recall from Theorem 3.3.(b) that

$$
\frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{T}\xrightarrow{d}N(0,\frac{(1-\rho)^2\sigma_\varepsilon^2}{2\sigma_e^2}).
$$

Lastly, we consider

$$
\begin{aligned}
& \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T^{3/2}}\frac{\boldsymbol{\iota}_T'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} \\
\approx & \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{T^2}\left[(1-\widehat{\rho})\sum_{t=1}^{T}(x_{it}-\widehat{\rho}x_{it-1})\right]\left[(1-\widehat{\rho})\sum_{t=1}^{T}(\nu_{it}-\widehat{\rho}\nu_{it-1})\right] \\
= & \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{(1-\widehat{\rho})^2}{T^2}\left[\sum_{t=1}^{T}((1-\widehat{\rho})x_{it-1}+\varepsilon_{it})\right]\left[\sum_{t=1}^{T}(e_{it}-(\widehat{\rho}-\rho)\nu_{it-1})\right] \\
= & \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{(1-\widehat{\rho})^2}{T^2}\left[\begin{array}{c}\sum_{t=1}^{T}(1-\widehat{\rho})x_{it-1}\sum_{t=1}^{T}e_{it} \\ +\sum_{t=1}^{T}\varepsilon_{it}\sum_{t=1}^{T}e_{it} \\ -\sum_{t=1}^{T}(1-\widehat{\rho})x_{it-1}\sum_{t=1}^{T}(\widehat{\rho}-\rho)\nu_{it-1} \\ -\sum_{t=1}^{T}(\widehat{\rho}-\rho)\varepsilon_{it}\sum_{t=1}^{T}\nu_{it-1}\end{array}\right] \\
= & I+II+III+IV.
\end{aligned}
$$

Consider II first.

$$\begin{aligned} II &= \frac{1}{T}\frac{(1-\widehat{\rho})^2}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right) \\ &= O_p\left(\frac{1}{T}\right) \end{aligned}$$

using $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right) = O_p(1)$ where ε_{it} and e_{it} are not correlated.

Consider III and IV. It is easy to see that

$$\begin{aligned} III &= -(1-\widehat{\rho})^3\frac{1}{\sqrt{nT}}\frac{\sqrt{nT}(\widehat{\rho}-\rho)}{\sqrt{n}}\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_{it-1}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\nu_{it-1}\right) \\ &= O_p\left(\frac{1}{\sqrt{nT}}\right) \end{aligned}$$

using

$$\frac{\sqrt{nT}(\widehat{\rho}-\rho)}{\sqrt{n}}\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_{it-1}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\nu_{it-1}\right)$$

and

$$\begin{aligned} IV &= -(1-\widehat{\rho})^2\frac{1}{\sqrt{n}T^{3/2}}\frac{\sqrt{nT}(\widehat{\rho}-\rho)}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\nu_{it-1}\right) \\ &= O_p\left(\frac{1}{\sqrt{n}T^{3/2}}\right). \end{aligned}$$

Lastly, consider I. it can be shown that

$$\begin{aligned} I &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(1-\widehat{\rho})^3\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_{it-1}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right) \\ &\xrightarrow{d}(1-\rho)^3 N(0,\frac{\sigma_\varepsilon^2\sigma_e^2}{3}). \end{aligned}$$

Therefore,

$$\frac{1}{\widehat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\widehat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2/T}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T^{3/2}}\frac{\boldsymbol{\iota}_T'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}} \xrightarrow{d} (1-\rho)\,N(0,\frac{\sigma_\varepsilon^2}{3\sigma_e^2})$$

and accordingly it can be shown that

$$\frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{u} \xrightarrow{d} N\left(0,\frac{(1-\rho)^2\,\sigma_\varepsilon^2}{6\sigma_e^2}\right)$$

by using a similar process as in Phillips and Moon (1999).

Next consider

$$\begin{aligned}\frac{1}{\sqrt{n}}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{u} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\hat{\sigma}_\alpha^2/T}\left[\left(\frac{\hat{\theta}}{T}\right)\mu_i+\frac{1}{\sqrt{T}}\left(\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}\right)\right] \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\hat{\sigma}_\alpha^2/T}\left(\frac{\hat{\theta}}{T}\right)\mu_i+\frac{1}{\sqrt{T}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\hat{\sigma}_\alpha^2/T}\left(\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}\right) \\ &= I+II.\end{aligned}$$

Consider I. Recall from 1.(c) that

$$I=\frac{1}{\hat{\sigma}_\alpha^2/T}\left(\frac{\hat{\theta}}{T}\right)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mu_i \xrightarrow{d} \frac{1}{(1-\rho)^2\,\sigma_\mu^2}(1-\rho)^2\,N\left(0,\sigma_\mu^2\right).$$

Consider II.

$$II=\frac{1}{\sqrt{T}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\hat{\sigma}_\alpha^2/T}\left(\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}\right)=O_p\left(\frac{1}{\sqrt{T}}\right)$$

using $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}\right)=O_p(1)$. We conclude that

$$\frac{1}{\sqrt{n}}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{u} \xrightarrow{d} N\left(0,\frac{1}{\sigma_\mu^2}\right).$$

Because we also know that $\frac{1}{n\sqrt{T}}\mathbf{x}'\Phi^{-1}\boldsymbol{\iota}_{nT} \xrightarrow{p} 0$ and $\frac{1}{n}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT} \xrightarrow{p} \frac{1}{\sigma_\mu^2}$, which are proved above, we have

$$\frac{1}{\sqrt{n}T}\widehat{G}_2 \xrightarrow{d} N\left(0, \frac{(1-\rho)^2\sigma_\varepsilon^2}{6\sigma_e^2}\right).$$

Finally,

$$t_{FGLS} = \left(\frac{1}{nT^2}\widehat{G}_1\right)^{-1/2}\left(\frac{1}{\sqrt{n}T}\widehat{G}_2\right) \xrightarrow{d} N(0,1).$$

4. When $\rho = 1$, $\lambda = 1$ if $T(\widehat{\rho}-1) \xrightarrow{p} \kappa$

   (a) First, let us show that $\hat{\sigma}_e^2$ is a consistent estimator of σ_e^2. From Lemma 1 (B), we have

$$\begin{aligned} I &= \frac{1}{n}\sum_{i=1}^{n}\left[\begin{array}{c} \frac{1}{T}\sum_{t=1}^{T} e_{it}^2 - \frac{2[T(\widehat{\rho}-1)]}{T}\left(\frac{1}{T}\sum_{t=1}^{T} e_{it}v_{it-1}\right) \\ +\frac{[T(\widehat{\rho}-1)]^2}{T}\left(\frac{1}{T^2}\sum_{t=1}^{T} v_{it-1}^2\right)\end{array}\right] \\ &\quad -\left(\frac{T}{\hat{d}^2}\right)\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T} e_{it} - \frac{1}{\sqrt{T}}[T(\widehat{\rho}-1)]\left(\frac{1}{T^{3/2}}\sum_{t=1}^{T} v_{it-1}\right)\right]^2 \\ &\xrightarrow{p} \sigma_e^2 \end{aligned}$$

   as $(n,T)\to\infty$, as shown already in 2.(a).

   Consider II. Using a similar argument, one can easily show that

$$\begin{aligned} II &= \frac{\left(\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2}{n^2}\sum_{i=1}^{n}\left[\begin{array}{c} \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}^2 + \frac{(T(\widehat{\rho}-1))^2}{T}\frac{1}{T^2}\sum_{t=1}^{T} x_{it-1}^2 \\ -\frac{2T(\widehat{\rho}-1)}{T}\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}x_{it-1}\end{array}\right] \\ &\quad -\frac{\left(\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2}{n^2}\left(\frac{T}{\hat{d}^2}\right)\sum_{i=1}^{n}\left[\frac{1}{\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it} - \frac{T(\widehat{\rho}-1)}{\sqrt{T}}\frac{1}{T^{3/2}}\sum_{t=1}^{T} x_{it-1}\right]^2 \\ &= o_p(1) \end{aligned}$$

   because if $(n,T)\to\infty$ and $\frac{n}{T}\to 0$, then

$$\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right) \xrightarrow{d} N\left(0, \frac{2\sigma_e^2}{3\sigma_\varepsilon^2}\right).$$

Consider III. From Lemma 1 (B), we know

$$III \leq \sqrt{I \times II} \xrightarrow{p} 0.$$

We conclude that

$$\hat{\sigma}_e^2 \xrightarrow{p} \sigma_e^2.$$

(b) Next, we investigate $\hat{\sigma}_\mu^2$. From Lemma 2 (B), we have

$$\begin{aligned}
& \frac{1}{T(1-\hat{\rho})}\hat{\sigma}_\alpha^2 \\
= \; & \frac{1}{nT(1-\hat{\rho})}\hat{\mathbf{u}}^{*\prime}\left(I_n \otimes \widehat{\bar{J}}_T^{\alpha}\right)\hat{\mathbf{u}}^{*} \\
= \; & \frac{(1-\hat{\rho})\hat{d}^2}{nT}\boldsymbol{\mu}' E_n \boldsymbol{\mu} \\
& + \frac{1}{nT(1-\hat{\rho})}\boldsymbol{\nu}'\left(I_{nT} - \bar{J}_{nT}\right)\left(I_n \otimes \hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT} - \bar{J}_{nT}\right)\boldsymbol{\nu} \\
& + \frac{1}{nT(1-\hat{\rho})}\left(\hat{\beta}_{OLS} - \beta\right)^2 \mathbf{x}'\left(I_{nT} - \bar{J}_{nT}\right)\left(I_n \otimes \hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT} - \bar{J}_{nT}\right)\mathbf{x} \\
& + \frac{2}{nT}\boldsymbol{\mu}'\left(E_n \otimes \hat{\boldsymbol{\iota}}_T^{\alpha\prime}\hat{C}\right)\left(I_{nT} - \bar{J}_{nT}\right)\boldsymbol{\nu} \\
& + \frac{2}{nT}\left(\hat{\beta}_{OLS} - \beta\right)\boldsymbol{\mu}'\left(E_n \otimes \hat{\boldsymbol{\iota}}_T^{\alpha\prime}\hat{C}\right)\left(I_{nT} - \bar{J}_{nT}\right)\mathbf{x} \\
& + \frac{1}{nT(1-\hat{\rho})}\left(\hat{\beta}_{OLS} - \beta\right)\boldsymbol{\nu}'\left(I_{nT} - \bar{J}_{nT}\right)\left(I_n \otimes \hat{C}'\widehat{\bar{J}}_T^{\alpha}\hat{C}\right)\left(I_{nT} - \bar{J}_{nT}\right)\mathbf{x} \\
= \; & I + II + III + IV + V + VI.
\end{aligned}$$

Consider I.

$$I = \frac{(1-\hat{\rho})\hat{d}^2}{T}\left[\frac{1}{n}\sum_{i=1}^{n}\mu_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n}\mu_i\right)^2\right] \xrightarrow{p} 0$$

as $(n,T) \to \infty$ with $(1-\hat{\rho})\hat{d}^2 = 2\widehat{\rho} + T(1-\widehat{\rho}) \xrightarrow{p} 2-\kappa$.

Consider II. As shown in 2.(b), we have

$$II \xrightarrow{p} \frac{\kappa^2 - 3\kappa + 3}{3(2-\kappa)}\sigma_e^2$$

and accordingly

$$IV \leq \sqrt{I \times II} \xrightarrow{p} 0.$$

Consider III next.

$$\begin{aligned} III &= \frac{\left(\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2}{n^2} \sum_{i=1}^n \left(\frac{1}{(1-\hat{\rho})\hat{d}^2}\right) \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \varepsilon_{it} - \frac{T(\widehat{\rho}-1)}{T^{3/2}} \sum_{t=1}^T x_{it-1}\right]^2 \\ & + \frac{\left(\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2 (1-\hat{\rho})\hat{d}^2}{n^2} \left[\frac{1}{n} \sum_{i=1}^n \frac{1}{T^2} \sum_{t=1}^T x_{it}\right]^2 \\ & - \frac{2\left(\sqrt{n}\left(\hat{\beta}_{OLS}-\beta\right)\right)^2}{n^2} \left[\frac{1}{n} \sum_{i=1}^n \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \varepsilon_{it} - \frac{T(\widehat{\rho}-1)}{T^{3/2}} \sum_{t=1}^T x_{it-1}\right)\right] \\ & \cdot \left[\frac{1}{nT^{3/2}} \sum_{i=1}^n \sum_{t=1}^T x_{it}\right]. \end{aligned}$$

One can show that $III = o_p(1)$ as $(n, T) \to \infty$ using the fact that if $(n, T) \to \infty$ and $\frac{n}{T} \to 0$, then

$$\sqrt{n}\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, \frac{2\sigma_e^2}{3\sigma_\varepsilon^2}\right)$$

and that $V \leq \sqrt{I \times III} \xrightarrow{p} 0$, $VI \leq \sqrt{II \times III} \xrightarrow{p} 0$, respectively.

Summarizing, we have the same result as 2.(b),

$$\frac{\hat{\sigma}_\alpha^2}{T(1-\hat{\rho})} \xrightarrow{p} \frac{\kappa^2 - 3\kappa + 3}{3(2-\kappa)}\sigma_e^2$$

and

$$(1-\hat{\rho})\hat{\sigma}_\mu^2 \xrightarrow{p} \frac{-\kappa^3 + 3\kappa^2 - 6}{3(2-\kappa)^2}\sigma_e^2.$$

With $k = -3$,

$$(1-\hat{\rho})\hat{\sigma}_\mu^2 \xrightarrow{p} \frac{16}{25}\sigma_e^2.$$

(c) Let us first look at $\widehat{G}_1$ in equation (14). Define

$$\frac{1}{nT}\widehat{G}_1 = \frac{1}{nT}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{x} - \frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n}\left(\frac{T}{n}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\frac{\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{x}}{n}.$$

From Lemma 3 (B), we have

$$\frac{1}{nT}\mathbf{x}'\hat{\Phi}^{-1}\mathbf{x} = \frac{1}{n\hat{\sigma}_e^2}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\mathbf{x}_i}{T} - T\left(1-\widehat{\rho}\right)^2\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T\left(1-\widehat{\rho}\right)}\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\mathbf{x}_i}{T\left(1-\widehat{\rho}\right)}\right).$$

Note that

$$\frac{1}{n\hat{\sigma}_e^2}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\mathbf{x}_i}{T} \xrightarrow{p} \frac{\sigma_\varepsilon^2}{\sigma_e^2}$$

as shown in Theorem 3.4.(a).

Next, one can easily see that

$$\frac{1}{n\hat{\sigma}_e^2}\sum_{i=1}^{n}T\left(1-\widehat{\rho}\right)^2\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T\left(1-\widehat{\rho}\right)}\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\mathbf{x}_i}{T\left(1-\widehat{\rho}\right)} = O_p\left(\frac{1}{T}\right)$$

using

$$\begin{aligned}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T\left(1-\widehat{\rho}\right)} &\approx \frac{1}{\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(x_{it}-\hat{\rho}x_{it-1}\right)\\ &= \frac{1}{\sqrt{T}}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it} + T\left(1-\widehat{\rho}\right)\sum_{t=1}^{T}\frac{x_{it-1}}{\sqrt{T}}\frac{1}{T}\right] = O_p\left(\frac{1}{\sqrt{T}}\right)\end{aligned}$$

and $T\left(1-\widehat{\rho}\right)^2\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2} \xrightarrow{p} \frac{-k^3+3k^2-6}{(2-k)(k^2-3k+3)}$ as shown already.

Hence, we have

$$\frac{1}{nT}\mathbf{x}'\hat{\Phi}^{-1}\mathbf{x} = \frac{1}{n\hat{\sigma}_e^2}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\mathbf{x}_i}{T} + O_p\left(\frac{1}{T}\right) \xrightarrow{p} \frac{\sigma_\varepsilon^2}{\sigma_e^2}$$

as $(n,T)\rightarrow\infty$.

Note also that

$$\begin{aligned}\frac{1}{n}\mathbf{x}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT} &= \frac{1}{\sqrt{nT}}\frac{1}{\hat{\sigma}_{\alpha}^{2}/T\left(1-\widehat{\rho}\right)}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{\sqrt{T}\left(1-\widehat{\rho}\right)}\right)\\ &= O_p(\frac{1}{\sqrt{nT}}).\end{aligned}$$

Lastly, recall that

$$\frac{T}{n}\boldsymbol{\iota}_{nT}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT} = \frac{\hat{\theta}/\left(1-\widehat{\rho}\right)}{\hat{\sigma}_{\alpha}^{2}/T\left(1-\widehat{\rho}\right)} \xrightarrow{p} \frac{2-k}{\frac{\kappa^2-3\kappa+3}{3(2-\kappa)}\sigma_e^2} = \frac{1}{\sigma_e^2}\frac{3\left(2-\kappa\right)^2}{\kappa^2-3\kappa+3}.$$

Therefore, we conclude that

$$\frac{1}{nT}\widehat{G}_1 \xrightarrow{p} \frac{\sigma_{\varepsilon}^2}{\sigma_e^2}.$$

Now we turn to $\widehat{G}_2$. Let

$$\frac{1}{\sqrt{nT}}\widehat{G}_2 = \frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{u} - \frac{\mathbf{x}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}}{n}\left(\frac{T}{n}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\boldsymbol{\iota}_{nT}\right)^{-1}\frac{\sqrt{T}}{\sqrt{n}}\boldsymbol{\iota}_{nT}'\widehat{\Phi}^{-1}\mathbf{u}.$$

Consider

$$\frac{1}{\sqrt{nT}}\mathbf{x}'\hat{\Phi}^{-1}\mathbf{u} = \frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\begin{array}{c}\frac{1}{\sqrt{T}}\left(\frac{\hat{\sigma}_e^2}{\hat{\sigma}_{\alpha}^2/T(1-\widehat{\rho})}\right)\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T(1-\widehat{\rho})}\right)\mu_i\\ +\left(\begin{array}{c}\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}\\ -T\left(1-\widehat{\rho}\right)^2\frac{\hat{\sigma}_{\mu}^2}{\hat{\sigma}_{\alpha}^2}\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{T(1-\widehat{\rho})}\right)\left(\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}(1-\widehat{\rho})}\right)\end{array}\right)\end{array}\right].$$

Note that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{T}\left(\frac{1}{\hat{\sigma}_{\alpha}^2/T\left(1-\widehat{\rho}\right)}\right)\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{\sqrt{T}\left(1-\widehat{\rho}\right)}\right)\mu_i = O_p\left(\frac{1}{T}\right)$$

and recall from Theorem 3.4.(b) that

$$\frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\widehat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}}\xrightarrow{d}N(0,\frac{\sigma_\varepsilon^2}{\sigma_e^2}).$$

Lastly, it can be shown that

$$\frac{1}{\hat{\sigma}_e^2}\frac{1}{\sqrt{n}}\frac{1}{\sqrt{T}}\sum_{i=1}^{n}T\left(1-\widehat{\rho}\right)^2\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2}\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{\sqrt{T}\left(1-\widehat{\rho}\right)}\right)\left(\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}\left(1-\widehat{\rho}\right)}\right)=O_p\left(\frac{1}{\sqrt{T}}\right)$$

using $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_i'\hat{\mathbf{A}}^{-1}\boldsymbol{\iota}_T}{\sqrt{T}(1-\widehat{\rho})}\right)\left(\frac{\boldsymbol{\iota}_T'\hat{\mathbf{A}}^{-1}\boldsymbol{\nu}_i}{\sqrt{T}(1-\widehat{\rho})}\right)=O_p(1)$ and $T\left(1-\widehat{\rho}\right)^2\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}_\alpha^2}\xrightarrow{p}\frac{-k^3+3k^2-6}{(2-k)(k^2-3k+3)}$.

Therefore,

$$\frac{1}{\sqrt{nT}}\mathbf{x}'\widehat{\Phi}^{-1}\mathbf{u}\xrightarrow{d}N(0,\frac{\sigma_\varepsilon^2}{\sigma_e^2}).$$

Also, using the results above, $\frac{1}{n}\mathbf{x}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT}=O_p(\frac{1}{\sqrt{nT}})$ and $\frac{T}{n}\boldsymbol{\iota}_{nT}'\hat{\Phi}^{-1}\boldsymbol{\iota}_{nT}\xrightarrow{p}\frac{1}{\sigma_e^2}\frac{3(2-\kappa)^2}{\kappa^2-3\kappa+3}$.

Summarizing, we have

$$\frac{1}{\sqrt{nT}}\widehat{G}_2\xrightarrow{d}N(0,\frac{\sigma_\varepsilon^2}{\sigma_e^2})$$

and accordingly,

$$t_{FGLS}=\left(\frac{1}{nT}\widehat{G}_1\right)^{-1/2}\left(\frac{1}{\sqrt{nT}}\widehat{G}_2\right)\xrightarrow{d}N\left(0,1\right).$$

■

Essay II: Testing Cross-sectional Dependence Using Bootstrap F-tests

1 Introduction

Cross-sectional dependence caused by common shocks can seriously impact inference as well as estimation. Andrews (2005) demonstrates that common shocks can result in inconsistent estimates in cross-sectional regressions and accordingly serious consequences for statistical inference.[1] To deal with the problems of common shocks, Bai (2003, 2004) considers the common factor model, and proposes principal component analysis (PCA) to consistently estimate the factors and factor loadings under stationarity, e.g., Bai (2003), and non-stationarity of the factors, e.g., Bai (2004). In addition, Bai (2009) and Bai, Kao and Ng (2009) extend this analysis to a panel data model that includes regressors as well as factors. Bai (2009) assumes stationary regressors and factors while Bai, *et al.* (2009) allow for non-stationary regressors and factors (i.e., *panel cointegration* case[2]). This paper considers the problem of testing cross-sectional independence in a panel data model using the factor structure proposed by Bai (2003, 2004, 2009) and Bai, *et al.* (2009).

Given this setting, it is natural to consider the simple F-statistic to test the null hypothesis that all the factor loadings are zero (i.e., cross-sectional independence). It is well known that the limiting distribution of the F-statistic can be approximated by a chi-squared distribution, when n is fixed and T is large. From the results of Boos and Brownie (1995) and Akritas and Arnold (2000) one can infer that the asymptotic distribution of an appropriately normalized F-statistic for the case of large n and fixed T, is also normal. However, we could not find any result regarding the asymptotic distribution of this F-statistic when both n and T are large. This paper suggests the use of the bootstrap F-test, proposed by Mammen (1993b), for testing cross-sectional independence. For this purpose, we adopt the *wild bootstrap* method which is well

[1] These common shocks could be macroeconomic, political, environmental, health, and/or sociological shocks in nature to mention a few, see Andrews (2005).

[2] Note that a large literature on panel cointegration exists with an assumption of cross-section independence (See, e.g., Baltagi and Kao (2000) for a survey, and Baltagi (2008) for a textbook treatment).

developed in the statistical literature. Section 2 introduces the factor model. Section 3 discusses the proposed wild bootstrap F-test. Section 4 presents the Monte Carlo results, while Section 5 concludes. *All the proofs are relegated to the appendix.*

For the asymptotic results in this paper, we use both the sequential limit ($n \to \infty$ following $T \to \infty$, i.e., $(n, T) \overset{\text{seq}}{\to} \infty$) and the joint limit ($n$ and T going to infinity simultaneously, i.e., $(n, T) \to \infty$) depending on the case considered. For details of these methods, see Phillips and Moon (1999). We use $\overset{p}{\to}$ and $\overset{d}{\to}$ to denote convergence in probability and in distribution, respectively. Unless indicated explicitly, we will refer to F_t as the factor (or the global stochastic trend) while F_λ as the F-statistic to avoid any confusion. The bootstrap sample or the bootstrap test statistic will be denoted with superscript star. For example, F_λ^* and P^* indicate the bootstrap F-statistic and the bootstrap probability measure. We also define the matrix that projects onto the orthogonal space of z as $M_z = I_T - z\left(z'z\right)^{-1} z'$. Lastly, $K(\cdot, \cdot)$ denotes the Kolmogorov metric, i.e., $K(P, Q) = \sup_x |P(x) - Q(x)|$ for marginal distributions P and Q.

2 The Model

Consider the panel data factor model

$$y_{it} = x_{it}'\beta + \lambda_i' F_t + u_{it} \text{ for } i = 1, \ldots, n \text{ and } t = 1, \ldots, T \tag{1}$$

where y_{it} is a scalar, x_{it} is a set of k regressors, β is a $k \times 1$ vector of the common slope parameters, λ_i is an $r \times 1$ factor loadings, F_t is an $r \times 1$ vector of latent common factors, and u_{it} is the error. The error terms are assumed to be uncorrelated across cross-section and over time series components. To test the null hypothesis of cross-sectional independence, we set the null

$$H_0 : \lambda_i = 0 \quad \text{for all } i \tag{2}$$

against the alternative that

$$H_a : \lambda_i \neq 0 \quad \text{for some } i.$$

To construct the F-statistic, define $RRSS = \sum_{i=1}^{n} \sum_{t=1}^{T} \widehat{w}_{it}^2$ as the sum of squared residuals from the restricted model:

$$y_{it} = x_{it}'\widetilde{\beta} + \widehat{w}_{it} \tag{3}$$

where $\widetilde{\beta}$ is the least squares estimator of β. Also let $URSS = \sum_{i=1}^{n} \sum_{t=1}^{T} \widehat{u}_{it}^2$ be the sum of squared residuals from the unrestricted model when F_t is not observed:

$$y_{it} = x_{it}'\widehat{\beta} + \widehat{\lambda}_i'\widehat{F}_t + \widehat{u}_{it} \tag{4}$$

where $\widehat{\beta}$, $\widehat{\lambda}_i$, and $\widehat{F}_t$ can be obtained from, e.g., Bai (2009) or Bai, *et al.* (2009). Then, the standard F-statistic is defined as

$$F_\lambda = \frac{nT - k - nr}{nr} \frac{RRSS - URSS}{URSS}. \tag{5}$$

Given this basic setting, the following sections briefly introduce the estimation procedures suggested in the literature under various scenarios depending on whether regressors are included and whether x_{it} and F_t are stationary.

2.1 Case 1: Without regressors

2.1.1 Stationary factors

Let us start from our benchmark case by dropping the regressors in equation (1). That is,

$$y_{it} = \lambda_i' F_t + u_{it} \tag{6}$$

which is *the common factor model.* Rewriting equation (6) in matrix notation, we have

$$y = F\Lambda' + u \tag{7}$$

where y is a $T \times n$ matrix of observed data and u is a $T \times n$ matrix of idiosyncratic errors. The matrices Λ $(n \times r)$ and F $(T \times r)$ are unknown. In fact, Bai (2003) studies the $F_t = I(0)$ case, while Bai (2004) investigates the $F_t = I(1)$ case.[3] The number of factors, r, is assumed to be *known.* If this is not the case, note that r can be consistently estimated as in Bai and Ng (2002).

First, we consider the $F_t = I(0)$ case. It is important to note that F_t $(t = 1, 2, \ldots, T)$ may or may not be observable. If the factors are observable, λ_i can be estimated using least squares. That is,

$$\widetilde{\Lambda} = y' F \left(F' F\right)^{-1}$$

and accordingly we have $RRSS = \sum_{i=1}^{n} \sum_{t=1}^{T} y_{it}^2$ and $URSS = \sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it} - \widetilde{\lambda}_i' F_t\right)^2$. Then, the F-statistic is constructed as follows:

$$F_\lambda = \frac{(nT - n)}{n} \frac{RRSS - URSS}{URSS}.$$

On the other hand, if F_t is not observable, one can estimate F_t using the method

[3]Note that when $F_t = I(1)$, testing $H_0 : \lambda_i = 0$ for all i is not only testing for cross section indepenence, but it is also testing if y_{it} follows an $I(0)$ process.

of PCA subject to the constraint $F^{'}F/T = I_r$. As illustrated in Bai (2003), $\widehat{F}$, the vector of estimated factors, is $\sqrt{T}$ times the eigenvectors corresponding to the r largest eigenvalues of $yy^{'}$. Given $\widehat{F}$, $\widehat{\Lambda}' = \widehat{F}'y/T$ can be obtained as well. Therefore, in this case we have $RRSS = \sum_{i=1}^{n}\sum_{t=1}^{T} y_{it}^2$ and $URSS = \sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it} - \widehat{\lambda}_i^{'}\widehat{F}_t\right)^2$.

2.1.2 Non-stationary factors

Now let us assume that F_t are non-stationary:

$$F_t = F_{t-1} + \eta_t \text{ for } t = 1, 2, \ldots, T \tag{8}$$

where η_t is the idiosyncratic error. If the factors are observable, then we can estimate λ_i using least squares as in the $F_t = I(0)$ case. However, if the factors are unknown, one estimates the factors subject to the constraint $F^{'}F/T^2 = I_r$. As a matter of fact, $\widehat{F}$ is T times the eigenvectors corresponding to the r largest eigenvalues of $yy^{'}$ in this setting. $\widehat{\Lambda}'$ can be also computed by $\widehat{F}'y/T^2$, which is the corresponding matrix of the estimated factor loadings. It is straightforward to construct the F-statistic with estimates of the factors and factor loadings.

2.2 Case 2: With regressors

2.2.1 Stationary regressors and factors

Next we consider *the panel data model with interactive fixed effects,* see Bai (2009), by adding regressors as well as common factors. In matrix notation, we have

$$y = X\beta + F\Lambda^{'} + u.$$

Note that the regressors as well as the interactive fixed effects are assumed to be *stationary*. When F_t are *known*, the estimate of β is easily obtained using least squares as follows:

$$\widetilde{\beta} = \left(\sum_{i=1}^{n} x_i^{'} M_F x_i\right)^{-1} \left(\sum_{i=1}^{n} x_i^{'} M_F y_i\right)$$

where $x_i = (x_{i1}, x_{i2}, \ldots, x_{iT})^{'}$, $y_i = (y_{i1}, y_{i2}, \ldots, y_{iT})^{'}$, and $M_F = I_T - F\left(F^{'}F\right)^{-1} F^{'}$. Given $\widetilde{\beta}$, one can compute $\widetilde{\Lambda}^{'} = \left(F^{'}F\right)^{-1} F^{'}\left(y - X\widetilde{\beta}\right)$. If F_t are not observed, one has the following set of nonlinear equations for estimation subject to the constraint $F^{'}F/T = I_r$.

$$\widehat{\beta} = \left(\sum_{i=1}^{n} x_i^{'} M_{\widehat{F}} x_i\right)^{-1} \left(\sum_{i=1}^{n} x_i^{'} M_{\widehat{F}} y_i\right) \tag{9}$$

and

$$\left[\frac{1}{nT}\sum_{i=1}^{n}\left(y_i - x_i\widehat{\beta}\right)\left(y_i - x_i\widehat{\beta}\right)^{'}\right]\widehat{F} = \widehat{F}V_{nT} \tag{10}$$

where $M_{\widehat{F}} = I_T - \widehat{F}\left(\widehat{F}^{'}\widehat{F}\right)^{-1}\widehat{F}^{'}$ and V_{nT} is a diagonal matrix consisting of the r largest eigenvalues of the matrix in brackets in equation (10). The solution $\left(\widehat{\beta}, \widehat{F}\right)$ for (9) and (10) can be obtained using iteration. From these results, one can compute $\widehat{\Lambda} = \widehat{F}^{'}\left(y - X\widehat{\beta}\right)/T$. For details of this estimation procedure, see Bai (2009). Using these results, one can easily construct the corresponding F-statistic.

2.2.2 Non-stationary regressors and factors

This is the case of *panel cointegration with global stochastic trend* under which both regressors and factors (or global stochastic trends) are assumed to be *non-stationary*. This case is investigated in Bai, *et al.* (2009).[4] More specifically, we have the following equations:

$$\begin{aligned} y_{it} &= x_{it}^{'}\beta + \lambda_i^{'}F_t + u_{it}, \\ x_{it} &= x_{it-1} + \varepsilon_{it}, \end{aligned}$$

[4] For simplicity, the mixed $I(0)/I(1)$ case among x_{it} and F_t will not be considered in this paper although this extension is possible. For details, see Bai, *et al.* (2009).

and

$$F_t = F_{t-1} + \eta_t$$

where x_{it}, F_t, and u_{it} are potentially correlated. The framework here is the *panel cointegration* model so that $u_{it} = y_{it} - x_{it}^{'}\beta - \lambda_i^{'}F_t$ is jointly stationary. Note that a fully-modified (FM) estimator is constructed along the line of Phillips and Hansen (1990) because of possible correlation among x_{it}, F_t, and u_{it}. Let us assume first that F_t are observed. Then, $\widetilde{\beta}_{LSFM}$ can be estimated as follows:

$$\widetilde{\beta}_{LSFM} = \left(\sum_{i=1}^{n} x_i^{'} M_F x_i\right)^{-1} \left(\sum_{i=1}^{n} x_i^{'} M_F \widetilde{y}_i^{+} - T\left(\widetilde{\Delta}_{\varepsilon ui}^{+} - \delta_i^{'}\widetilde{\Delta}_{\eta u}^{+}\right)\right)$$

where $\widetilde{y}^{+}$ and $\widetilde{\Delta}^{+}$ are consistent estimates of y^{+} and Δ^{+} with

$$y_{it}^{+} = y_{it} - \Omega_{ubi}\Omega_{bi}^{-1}\begin{pmatrix} \triangle x_{it} \\ \triangle F_t \end{pmatrix} \text{ and } u_{it}^{+} = u_{it} - \Omega_{ubi}\Omega_{bi}^{-1}\begin{pmatrix} \triangle x_{it} \\ \triangle F_t \end{pmatrix}.$$

Note that Ω_i is the long-run covariance matrix of $w_{it} = \left(u_{it}, \varepsilon_{it}^{'}, \eta_{it}^{'}\right)^{'}$ which is given by,

$$\Omega_i = \sum_{j=-\infty}^{\infty} E\left(w_{i0}w_{ij}^{'}\right) = \begin{bmatrix} \Omega_{ui} & \Omega_{u\varepsilon i} & \Omega_{u\eta i} \\ \Omega_{\varepsilon ui} & \Omega_{\varepsilon i} & \Omega_{\varepsilon\eta i} \\ \Omega_{\eta ui} & \Omega_{\eta\varepsilon i} & \Omega_{\eta} \end{bmatrix}$$

and Δ_i is the one-sided covariance defined as follows:

$$\Delta_i = \sum_{j=0}^{\infty} E\left(w_{i0}w_{ij}^{'}\right) = \begin{bmatrix} \Delta_{ui} & \Delta_{u\varepsilon i} & \Delta_{u\eta i} \\ \Delta_{\varepsilon ui} & \Delta_{\varepsilon i} & \Delta_{\varepsilon\eta i} \\ \Delta_{\eta ui} & \Delta_{\eta\varepsilon i} & \Delta_{\eta} \end{bmatrix}.$$

We also define

$$\Omega_{bi} = \begin{bmatrix} \Omega_{\varepsilon i} & \Omega_{\varepsilon\eta i} \\ \Omega_{\eta\varepsilon i} & \Omega_{\eta} \end{bmatrix}$$

corresponding to ε_{it} and η_t for convenience. We need to estimate the nuisance parameter by using a kernel estimator. Let

$$\widehat{\Omega}_i = \sum_{j=-(T-1)}^{T-1} \omega\left(\frac{j}{K}\right)\widehat{\Gamma}_i(j)$$

and

$$\widehat{\Delta}_i = \sum_{j=0}^{T-1} \omega\left(\frac{j}{K}\right)\widehat{\Gamma}_i(j)$$

where $\widehat{\Gamma}_i(j) = \frac{1}{T}\sum_{t=1}^{T-j} \widehat{w}_{it+j}\widehat{w}_{it}'$ and $\widehat{w}_{it} = \left(\widehat{u}_{it}, \triangle x_{it}', \triangle F_t'\right)'$ with the kernel function $\omega(\cdot)$ and the bandwidth parameter K.[5]

When F_t is not observed, however, one needs to estimate the set of two nonlinear equations:

$$\widehat{\beta} = \left(\sum_{i=1}^{n} x_i' M_{\widehat{F}} x_i\right)^{-1}\left(\sum_{i=1}^{n} x_i' M_{\widehat{F}} y_i\right)$$

and

$$\left[\frac{1}{nT^2}\sum_{i=1}^{n}\left(y_i - x_i\widehat{\beta}\right)\left(y_i - x_i\widehat{\beta}\right)'\right]\widehat{F} = \widehat{F}V_{nT}$$

by iteratively solving for $\widehat{\beta}$ and $\widehat{F}$ subject to the constraint $F'F/T^2 = I_r$. Compared to the *known* F_t case, estimation of the stochastic trends affects the limiting behavior of the estimator, so bias correction becomes essential for estimation. In fact, Bai, *et al.* (2009) propose two FM estimators, i.e., the bias-corrected Cup (continuously updated) estimator, $\widehat{\beta}_{CupBC}$, and the FM Cup estimator, $\widehat{\beta}_{CupFM}$. The details of the estimation procedure can be found in Bai, *et al.* (2009). However, it is worthwhile emphasizing the basic difference between these two estimators. CupBC corrects the

[5] $\widetilde{\beta}_{LSFM}$ can be alternatively written as the bias-corrected estimator, $\widetilde{\beta}_{LSBC}$. For details, see Bai, *et al.* (2009).

bias only in the final stage of iterations while CupFM modifies the data to remove serial correlation and endogeneity in each iteration. CupBC and CupFM have the same asymptotic distribution although constructed in different ways.

3 F-test with Bootstrapped Samples

We discuss the asymptotic behavior of the F-statistic for three cases: (i) fixed n / large T, (ii) large n / fixed T, and (iii) large n / large T. Based on the results, we argue that the F distribution may not be always appropriate to use but the bootstrap F-test can be a good alternative.

3.1 The asymptotics of the F-statistic

To simplify the arguments, we assume that the factors are *known* and stationary. Also, the number of factors is assumed to be one ($r = 1$) unless indicated otherwise. In order to construct the F statistic to test the null $H_0 : \lambda_i = 0$ for all i, we compute $RRSS = \sum_{i=1}^{n} \sum_{t=1}^{T} y_{it}^2$ and $URSS = \sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it} - \widehat{\lambda}_i F_t \right)^2$. Then, we have

$$F_\lambda = \frac{nT - n}{n} \frac{RRSS - URSS}{URSS}.$$

For the case of fixed n / large T, one can rewrite the above formula as follows:

$$F_\lambda = \frac{\frac{\chi_a^2}{a}}{\frac{\chi_b^2}{b}}$$

where $a = n$ and $b = nT - n$. Accordingly, the approximation by a chi-squared distribution is given by,

$$aF_\lambda \xrightarrow{d} \chi_a^2$$

because a is fixed and $b \to \infty$.

We next turn to the case of large n / fixed T. In the statistics literature, Boos and Brownie (1995) and Akritas and Arnold (2000) consider the asymptotic distribution of the ANOVA F-statistic for this case where n and T denote the number of treatments and replications per treatment, respectively. Under their settings, it is shown that

$$\sqrt{n}\,(F-1) \xrightarrow{d} N\left(0, \frac{2T}{T-1}\right)$$

as $n \to \infty$ with fixed T. That is, the F statistic is asymptotically normal with expected value 1. They also show that the asymptotics above hold in a two-way fixed effects model as well. Extending these results to the interaction effects model, Bathke (2004) shows that the limiting normal distribution can be still achievable with the F-statistic centered at 1. Interestingly, in the econometrics literature, Orme and Yamagata (2006) consider a panel data model with one-way fixed effects and derive the same limiting distribution as that of Boos and Brownie (1995) and Akritas and Arnold (2000).

3.1.1 The asymptotics of the F-statistic in a high-dimensional framework

As mentioned earlier, the F-statistic with large n and T have not been explored in the literature. In this section, we sketch the asymptotic properties of the F-statistic under this setting.

Consider the common factor model:

$$y_{it} = \lambda_i F_t + u_{it} \text{ for } i = 1, \ldots, n \text{ and } t = 1, \ldots, T$$

where λ_i and F_t are scalars. Our analysis is based on the following assumptions:

Assumption 1 $u_{it} \overset{i.i.d.}{\sim} (0, \sigma^2)$ *for all i and t with finite fourth order cumulants.*

Assumption 2 *The factor and factor loadings are assumed to be independent of* u_{it} *with* $E(u_{it} \mid F_t) = E(u_{it} \mid \lambda_i) = 0$ *such that:*

1. $0 < \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} F_t^2 = \phi_F < \infty.$

2. $0 < \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \lambda_i^2 = \phi_\lambda < \infty.$

These assumptions are similar to those in Bai (2003).[6] In what follows, we distinguish between cases where the factor F_t is observable or not. If F_t is observable, then one can easily obtain $\widetilde{\lambda}_i$ using least squares. If F_t is not known, one relies on the method of PCA to compute $\widehat{\lambda}_i \widehat{F}_t$. In the lemma below, we consider the limiting distribution of $\widetilde{\lambda}_i$ or $\widehat{\lambda}_i \widehat{F}_t$. Note that the result for $\widehat{\lambda}_i \widehat{F}_t$ is taken from Bai (2003).

Lemma 1 *1. If* F_t *is observable, then for each* i *as* $T \to \infty$

$$\frac{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)}{\left(\sigma^2 \phi_F^{-1}\right)^{1/2}} \xrightarrow{d} N(0,1).$$

2. If F_t *is unobservable, then as* $(n,T) \to \infty$

$$\frac{\delta_{nT}\left(\widehat{\lambda}_i \widehat{F}_t - \lambda_i F_t\right)}{\left(\frac{\delta_{nT}^2}{n} V_{it} + \frac{\delta_{nT}^2}{T} W_{it}\right)^{1/2}} \xrightarrow{d} N(0,1)$$

where $\delta_{nT} = \min\left\{\sqrt{n}, \sqrt{T}\right\}$, $V_{it} = \frac{\lambda_i^2}{\phi_\lambda}\sigma^2$, *and* $W_{it} = \frac{F_t^2}{\phi_F}\sigma^2$.

Lemma 1 shows that we can asymptotically achieve the standard normal distribution whether or not F_t is observable. Note that: (i) We have the limiting distribution of $\widetilde{\lambda}_i$ with *known* F_t. On the other hand, the limiting distribution of $\widehat{\lambda}_i \widehat{F}_t$ is derived for an unknown F_t. (ii) The convergence rate when F_t is unobservable is $\min\left\{\sqrt{n}, \sqrt{T}\right\}$ with no restriction on the relationship between n and T, see Bai (2003). Given the

[6] For simplicity, we assume the i.i.d. error terms, while Bai (2003) allows for time series and cross-section dependence in the error terms.

above results, we derive the asymptotic normality of the F-statistic when the factor is known and unknown, respectively.

Theorem 1 *Assume* $(n,T) \overset{\text{seq}}{\to} \infty$ *and* F_t *is observable. Then*

$$\sqrt{n}\,(F_\lambda - 1) \overset{d}{\to} N(0,2).$$

Theorem 1 shows that the asymptotic distribution of the F-statistic with $(n,T) \overset{\text{seq}}{\to} \infty$ will converge to the normal distribution if F_t is *known*. Note that this result is quite similar to the one reported in previous studies, e.g., the ANOVA literature and Orme and Yamagata (2006) which do not assume the high-dimensional framework, in the sense that the F-statistic gets centered at 1 with the asymptotic normality.

If F_t is not observable, however, one needs to estimate common components $\lambda_i F_t$ using the method of PCA. Next we investigate the limiting distributions of the F-statistic under two specific cases, i.e., $\frac{T}{n} \to 0$ and $\frac{n}{T} \to 0$, following Bai (2003).

Theorem 2 *Assume* $(n,T) \to \infty$ *and* F_t *is not observable.*

1. *If* $\frac{T}{n} \to 0$*, then*

$$\sqrt{nT}\,(F_\lambda - 1) \overset{d}{\to} N(\frac{(F_t^2 - \phi_F)}{\phi_F}, \frac{\psi}{\sigma^4})$$

where $\psi = Var\left(a_{it}^2 - u_{it}^2\right) < \infty$ *and* $a_{it} = F_t\left(\frac{1}{T}\sum_{t=1}^{T} F_t^2\right)^{-1} \frac{1}{\sqrt{T}}\sum_{s=1}^{T} F_s u_{is}$.

2. *If* $\frac{n}{T} \to 0$*, then the asymptotic distribution of* F_λ *is not feasible.*

From Theorem 2, one finds that there will be a shift term in the limiting distribution of the F-statistic. The F-statistic will not be asymptotically centered around 1 any more. Instead, the F-statistic will converge to the sequence, $1 + \frac{F_t^2 - \phi_F}{\phi_F}$ if $\frac{T}{n} \to 0$. For the other case, i.e., $\frac{n}{T} \to 0$, we cannot obtain the asymptotic properties because our current assumption, $\phi_\lambda > 0$, is violated under the null $H_0 : \lambda_i = 0$ for all i.[7]

[7] Note that the asymptotic normality with $\delta_{nT} = \min\left\{\sqrt{n}, \sqrt{T}\right\}$ is also not feasible, under the null, because ϕ_λ cannot be defined.

To conclude this section, we find that using the asymptotic F distribution in a high-dimensional framework may fail especially when the factors are unknown. In the following sections, we discuss the bootstrap procedure as an alternative that avoids all these complexities.

3.2 Bootstrap data generating process

Before we go into the validity of bootstrap F-tests, we briefly discuss a bootstrap data generating process (DGP). Resampling in a regression can be implemented in various ways. One can consider first *the pairs bootstrap*, one of the most general and widely used bootstrap DGP, which is proposed in Freedman (1981). The idea of this method is simply resampling the dependent and independent variables in pairs. However, this method does not condition on the independent variable, X, in a DGP (Instead, each bootstrap sample has a different X^*). As a result, this DGP can be misleading in inference when test statistics depend on X according to MacKinnon (2007). Therefore one may conclude that the pairs bootstrap is not satisfactory for bootstrap inference.

Secondly, *the residual bootstrap* can be considered. Let

$$y_t = x_t\beta + u_t, \qquad u_t \sim IID(0, \sigma^2).$$

The first step of the residual bootstrap is estimating $\widetilde{\beta}$ and the residuals $\widetilde{u}_t$ under the null. After rescaling the residuals, the residual bootstrap DGP can be written as

$$y_t^* = x_t\widetilde{\beta} + u_t^*$$

where u_t^* is obtained from the empirical distribution of rescaled $\widetilde{u}_t$. Note that the validity of this method depends crucially on the assumption $u_t \sim IID(0, \sigma^2)$, i.e., independent and identically distributed error term. Hence, under heteroskedasticity this bootstrap DGP is not recommended.

Finally, with independent but possibly heteroskedastic errors, one can rely on *the wild bootstrap*. First of all, this method is quite simple to implement from its construction. In addition to this, as shown in simulations of Davidson and Flachaire (2008), wild bootstrap tests perform well in practice under heteroskedasticity. In fact, a specific version (using Rademacher distribution) of the wild bootstrap is shown to outperform another version of the wild bootstrap as well as the pairs bootstraps even when the disturbances are homoskedastic.

We adopt the wild bootstrap using Rademacher distribution in our simulations because it is robust to heteroskedasticity. Let

$$y_{it} = x_{it}\beta + u_{it} \text{ where } u_{it} \sim IID(0, \sigma^2),$$

then the corresponding bootstrap DGP is constructed as follows:

$$y_{it}^* = x_{it}\widetilde{\beta} + \widetilde{u}_{it}\varepsilon_{it}^* \tag{11}$$

where y_{it}^* is newly generated data, $\widetilde{u}_{it}$ is the restricted residual, and $\widetilde{\beta}$ is an estimate under the null.[8] ε_{it}^* follows the Rademacher distribution:

$$\varepsilon_{it}^* = \begin{cases} 1 & \text{with probability } 0.5 \\ -1 & \text{with probability } 0.5 \end{cases} \tag{12}$$

which is introduced by Liu (1988) and developed by Davidson and Flachaire (2008).[9]

[8] Note that the model is estimated under the null to obtain restricted estimates $\widetilde{\beta}$. MacKinnon (2006) points out that using the unrestricted residuals is not appropriate because otherwise the bootstrap DGP will not satisfy the null hypothesis.

[9] Alternatively, one may want to use the following bootstrap DGP suggested by Mammen (1993b) especially when the distribution of the error terms is sufficiently asymmetric.

$$\varepsilon_{it}^* = \begin{cases} \frac{-(\sqrt{5}-1)}{2} & \text{with probability } p = \frac{(\sqrt{5}+1)}{2\sqrt{5}} \\ \frac{(\sqrt{5}+1)}{2} & \text{with probability } 1-p \end{cases} .$$

Note that one has $E\left(\varepsilon_{it}^{*}\right)=0$ and $E\left(\varepsilon_{it}^{*2}\right)=1$ with this setting.[10]

Next we describe in some details how to implement the wild bootstrap test for the common factor model.

Step 1: One estimates the common factor model. If F_t are *known*, we simply obtain the OLS residuals. If F_t are not observed, we use the method of PCA. Note that the unrestricted residuals as well as the restricted residuals should be computed in order to calculate the F-statistic. Let this empirical statistic be F_λ.

Step 2: After we obtain the residuals from *step 1*, we re-generate the data using the restricted residuals and an external random variable ε_{it}^{*}. For example, one can generate artificial data for the common factor model by,

$$y_{it}^{*}=u_{it}\varepsilon_{it}^{*}$$

where $i=1,\ldots,n;\ t=1,\ldots,T$. Note that we simply use u_{it} as the restricted residuals which are the same as y_{it} under the null $H_0:\lambda_i=0\quad$ for all i.[11] Now one can compute the bootstrap counterpart of our test statistic, i.e., the bootstrap F statistic. Let us denote this statistic as F_λ^{*}.

Step 3: One repeats *Step 2*, say B times. Then we obtain the distribution of F_λ^{*} and calculate the percentile of F_λ^{*} which are greater than or equal to F_λ. Finally setting this proportion at α^{*}, one can test the null by rejecting $\alpha^{*}<\alpha$, at the 5% significance level.

However, in their simulations Davidson and Flachaire (2008) show that the version we adopt here performs at least as good as this version even when the disturbances are asymmetric.

[10] The further condition $E\left(\varepsilon_{it}^{*3}\right)=1$ is often added for the bootstrap error in the literature.

[11] If we have regressors as well as factors in our equation, then $y_{it}^{*}=x_{it}^{'}\widetilde{\beta}+\widehat{w}_{it}\varepsilon_{it}^{*}$ where $\widehat{w}_{it}=y_{it}-x_{it}^{'}\widetilde{\beta}$ are the restricted residuals under the null $H_0:\lambda_i=0$ for all i.

3.3 The validity of the bootstrap F-test

Mammen (1993b) seems to be the first to show that under some regularity conditions the asymptotic distribution of the F statistic is equivalent to that of the wild bootstrap counterpart in a high-dimensional framework. Using simulation results, Flachaire (2005) shows that the wild bootstrap F-test performs well compared to other bootstrap methods such as the pairs bootstrap. We sketch the validity of the bootstrap F-test for cross-sectional dependence relying on the results of Mammen (1993b).

Consider first a simple regression model

$$y_t = x_t^{'}\beta + \varepsilon_t \text{ for } t = 1, \ldots, n$$

where β is a k-dimensional parameter and ε_t is the error. Mammen (1993b) studies the case in which k may also increase as n increases. For the testing problem $\beta \in H_0$ versus $\beta \in H_1$, the F-statistic can be constructed by,

$$F = \frac{\left[\sum_{t=1}^{n}\left(y_t - x_{1,t}^{'}\widehat{\beta}_1\right)^2 - \sum_{t=1}^{n}\left(y_t - x_{0,t}^{'}\widehat{\beta}_0\right)^2\right]/(k_1 - k_0)}{\sum_{t=1}^{n}\left(y_t - x_{1,t}^{'}\widehat{\beta}_1\right)^2/(n-k_1)} \tag{13}$$

where each squared sum indicates the square of the projection of y onto H_0 and H_1. Under the hypothesis H_i, $\widehat{\beta}_i$ is the least squares estimator, k_i is the dimension of the parameters, and $x_{i,t}$ is a set of k_i regressors. It is important to note that the degrees of freedom of both the numerator and denominator cannot be assumed to be fixed and that simply applying the F distribution in testing may fail. Mammen (1993b) shows that the asymptotic distribution of the bootstrap F-statistic is consistent for that of the F-statistic.[12]

[12] To compute the bootstrap F statistic, $F^* = \frac{\left[\sum_{t=1}^{n}\left(y_t^* - x_{1,t}^{'}\widehat{\beta}_1^*\right)^2 - \sum_{t=1}^{n}\left(y_t^* - x_{0,t}^{'}\widehat{\beta}_0^*\right)^2\right]/(k_1-k_0)}{\sum_{t=1}^{n}\left(y_t^* - x_{1,t}^{'}\widehat{\beta}_1^*\right)^2/(n-k_1)}$ where

One may observe that Mammen's results can be readily extended to our case if the factors are observable. For the common factor model with a single factor,[13] the F-statistic is defined as:

$$F_{\lambda} = \frac{\left[\sum_{i=1}^{n}\sum_{t=1}^{T} y_{it}^{2} - \sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it} - \widetilde{\lambda}_i F_t\right)^2\right]/n}{\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it} - \widetilde{\lambda}_i F_t\right)^2 / (nT-n)} \tag{14}$$

where the F_t is *known*. The bootstrap F-statistic can be constructed by,

$$F_{\lambda}^{*} = \frac{\left[\sum_{i=1}^{n}\sum_{t=1}^{T} y_{it}^{*2} - \sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it}^{*} - \widetilde{\lambda}_i^{*} F_t\right)^2\right]/n}{\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it}^{*} - \widetilde{\lambda}_i^{*} F_t\right)^2 / (nT-n)}$$

where $\widetilde{\lambda}_i^{*}$ denotes the bootstrap estimate which is the least squares estimator for $\widetilde{\lambda}_i$ from $y_{it}^{*} = \widetilde{\lambda}_i F_t + u_{it}\varepsilon_{it}^{*}$. In fact, equation (14) is a set up similar to (13): (i) Degrees of freedom of both the numerator and denominator are not bounded. For example, n, the number of factor loadings, corresponds to k_1 in equation (13) with $k_0 = 0$. Also nT, the number of total observations, is the counterpart of n in equation (13) as well. (ii) Both (13) and (14) are obtained from least squares estimation. (iii) Note that one of the key conditions in Mammen (1993b) to identify the parameters under a high-dimensional framework, i.e., $\frac{k_1}{n} \to 0$, is automatically satisfied in our panel data model, because $\frac{n}{nT} = \frac{1}{T} \to 0$.[14]

Proposition 3 *Assume* $(n, T) \overset{\text{seq}}{\to} \infty$. *If Assumptions 1-2 hold and* F_t *is observable,*

$\widehat{\beta}_1^{*}$ and $\widehat{\beta}_0^{*}$ denote the least squares estimators from newly generated bootstrap data under the null and alternative, respectively.

[13] The dimension k of β is not a concern in this paper and is assumed to be fixed. Therefore, we only consider the common factor model dropping the regressors without loss of generality.

[14] Our model is similar to that of Mammen in that we have an infinite number of parameters to estimate as the sample size tends to infinity. However, our model is also different from that of Mammen, because the number of F_t is assumed to be fixed. Hence, λ_i for all i can be estimated with large T. Therefore, we do not need the corresponding condition ($\frac{k_1}{n} \to 0$) in Mammen (1993b).

then

$$K\left(\mathcal{L}\left(F_{\lambda}\right), \mathcal{L}^{*}\left(F_{\lambda}^{*}\right)\right) \xrightarrow{p} 0$$

where $\mathcal{L}\left(F_{\lambda}\right)=P\left(\sqrt{n}\left(F_{\lambda}-1\right) \leq x\right)$ *and* $\mathcal{L}^{*}\left(F_{\lambda}^{*}\right)=P^{*}\left(\sqrt{n}\left(F_{\lambda}^{*}-1\right) \leq x\right)$.

Proposition 3 provides the consistency of the bootstrap distribution of the F-statistic. Hence, one can infer that the bootstrap method can be justified in testing cross-sectional dependence when the factors are *known*. We also notice that one does not necessarily have to theoretically derive the limiting distribution of the F-statistic now that the distribution of the bootstrap statistic can mimic it asymptotically. In fact, the asymptotic distribution of the F-statistic can remain unknown, while one can still properly test the null using the bootstrap F-statistic.

3.3.1 Bootstrapping PCA

When the factors are not observed, another important issue needs to be taken into consideration. With unknown factors one computes $URSS = \sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it}-\widehat{\lambda}_i\widehat{F}_t\right)^2$ by the method of PCA instead of using least squares. Namely, we need to bootstrap the PCA estimators. Diaconis and Efron (1983) introduce an application of bootstrap to principal component analysis and illustrate how to bootstrap the eigenvalue and eigenvector components. However, this is done without theoretical justification.

Recently, Gonçalves and Perron (2010) establish the asymptotic validity of the bootstrap for factor-augmented regressions under a high-dimensional framework. They provide an appropriate set of assumptions under which the wild bootstrap procedure can be used to estimate the bootstrap factors by principal components.[15] Note also that Mammen (1993a) shows that the wild bootstrap in a high-dimensional model is valid as long as the asymptotic normality holds. To carry this point, we recall that the

[15] Note that Gonçalves and Perron (2010) focus on the factors which cannot be identified separately with the factor loadings. However, identification problem is not the concern of this paper. In fact, in order to construct the F-statistic, we only need to estimate the common components ($\lambda_i F_t$, not F_t) which are identifiable.

asymptotic normality of the estimated common components for $\lambda_i F_t$ as $(n,T) \to \infty$ can be achieved, as shown in Lemma 1.

Proposition 4 *Assume* $(n,T) \to \infty$ *and* $\frac{T}{n} \to 0$*. If Assumptions 1-2 hold and* F_t *is unobservable, then*

$$K\left(H_{nT}, H_{Boot}\right) \xrightarrow{p} 0$$

where $H_{nT} = P\left(\tau \leq x\right)$ *is the c.d.f. with a functional* $\tau = \sqrt{T}\left(\widehat{\lambda}_i \widehat{F}_t - \lambda_i F_t\right)$, $H_{Boot} = P^*\left(\tau^* \leq x\right)$ *is the empirical c.d.f. with* $\tau^* = \sqrt{T}\left(\widehat{\lambda}_i^* \widehat{F}_t^* - \widehat{\lambda}_i \widehat{F}_t\right)$.

Proposition 4 indicates the consistency of bootstrapping PCA. Hence, this implies that the bootstrap F-statistic can be used for the unobservable F_t case. Note that the condition $\frac{T}{n} \to 0$ is required to achieve the asymptotic normality of the estimated common components under a high-dimensional framework. Based on this, we also check the consistency of the distribution of the bootstrap F-test using PCA.

Proposition 5 *Under the assumptions of Proposition 4,*

$$K\left(\mathcal{L}\left(F_\lambda\right), \mathcal{L}^*\left(F_\lambda^*\right)\right) \xrightarrow{p} 0$$

where $\mathcal{L}\left(F_\lambda\right) = P\left(\sqrt{n}\left(F_\lambda - 1\right) \leq x\right)$ *and* $\mathcal{L}^*\left(F_\lambda^*\right) = P^*\left(\sqrt{n}\left(F_\lambda^* - 1\right) \leq x\right)$.

According to Proposition 5, the distribution of the bootstrap F-statistic will uniformly converge to the distribution of the empirical F-statistic. Hence, combining this result with that of Proposition 3, one can conclude that the bootstrap F-statistic can be used in testing cross-sectional dependence whether the factors are known or not. The following section presents the various simulation results in support of this conclusion.

4 Monte Carlo Results

4.1 Experiment design

We consider the following equation:

$$y_{it} = x_{it}\beta + \lambda_i F_t + u_{it} \text{ for } i = 1, \ldots, n \text{ and } t = 1, \ldots, T$$

where $\beta = 2$ and $\lambda_i = 0$ for all i. x_{it} and F_t follow either $I(0)$ or $I(1)$ processes. For simplicity, we assume that both λ_i and β are scalars. u_{it} is generated by $IIDN(0,1)$ for our benchmark case. In the common factor model, the regressor, x_{it}, is simply dropped. We study the finite sample properties of the F-statistic for $H_0 : \lambda_i = 0$ for all i; based on various estimators discussed in Section 2. We denote the empirical F statistic and the bootstrap F statistic as EF and BF, respectively. The sample sizes n and T are varied over the range $\{10, 50, 100\}$ for the model without the regressor, and $\{10, 20, 50\}$ for the model with the regressor.

For each experiment, we perform $1,000$ replications and 200 bootstrap iterations. GAUSS 7.0.6 is used to perform the simulations. Random numbers for u_{it}, F_t, and x_{it} are generated by the GAUSS procedure RNDNS. We generate $n(T + 1000)$ random numbers and then split them into n series so that each series has the same mean and variance. The first $1,000$ observations are discarded for each series.

4.2 Case 1: Without the regressor

This section runs Monte Carlo experiments for the common factor model:

$$y_{it} = \lambda_i F_t + u_{it} \text{ for } i = 1, \ldots, n \text{ and } t = 1, \ldots, T.$$

Note that in this case we generate the bootstrap data from $y_{it}^* = \widetilde{u}_{it}\varepsilon_{it}^*$ where $\widetilde{u}_{it}$ is simply y_{it} itself (or u_{it}) under the null. We discuss the case of stationary and non-stationary factors subsequently.

4.2.1 Stationary factors

Let us first consider the benchmark case under which both F_t and u_{it} are generated from $IIDN(0,1)$.[16] Table 1 shows the empirical size of EF and BF when $F_t = I(0)$ with true size 5%. Given this setting, we find the following: (i) If F_t is known, both EF and BF are quite close to their true size. (ii) In contrast, when F_t is unknown, EF gets extremely shifted to the right so that its size becomes almost 100%, which implies rejection for almost all cases. BF, however, mimics the empirical F distribution quite well so that its size stays very close to 5%. For example, with $(n,T) = (50,100)$ the size of EF is 99.9% while that of BF is 4.9% when the factors are not observed.

Next, in order to examine the power of the F-test under some alternative hypotheses, we divide our cases into strong and weak cross section dependence. Weak dependence is set at $\lambda_i \sim IIDU\,(0.01, 0.2)$ while strong dependence at $\lambda_i \sim IIDU\,(0.2, 0.5)$. All the results are reported in Table 2. Overall, the power of the F test seems satisfactory: (i) The power increases as λ_i increases as expected. (ii) Also, the power increases as n or T increases. (iii) With weak dependence, both EF and BF have no power or very low power if any, when F_t is unknown. In fact, even in the largest sample size of our experiments, $(n,T) = (100,100)$, the power of EF and BF is no more than 46%.

We also check robustness of our benchmark results to heteroskedasticity and serial correlation in the error terms. We first introduce heteroskedasticity into the error as follows:

$$u_{it} = \sigma_i v_{it}$$

where v_{it} is generated from $N(0,1)$ and σ_i is set as either standard normal or simply

[16] We also run experiments with AR(1) factors and linear trended factors. All the results are similar to those when the factors follow $IIDN(0,1)$.

10. That is,

$$\sigma_i \begin{cases} \sim N(0,1) \text{ for } i = 1, \ldots, \frac{4n}{5} \\ = 10 \text{ for } i = \frac{4n}{5} + 1, \ldots, n \end{cases}.$$

Notice that we do not correct for heteroskedasticity to compute the residuals.[17] All the results are reported in Table 3. We find that BF stays robust despite huge heteroskedasticity. More specifically, the following can be observed: (i) With heteroskedasticity, EF gets over-sized although F_t is *known*. In fact, the empirical size of EF varies from 13 to 20%. This is different from our benchmark case where the size of EF stays close to 5% when F_t is *known*. (ii) When F_t is unknown, as expected, EF shows extreme over-rejection like in the benchmark case. However, BF behaves well whether or not the factors are observable. In fact, the empirical size of BF consistently stays robust varying from 4-6% for all experiments. Therefore, we conclude that bootstrap F-test in the common factor model can be used under heteroskedasticity.

For serial correlation, the error terms are set as follows:

$$u_{it} = \rho u_{it-1} + \nu_{it}$$

where $\rho = 0.4$ and $\nu_{it} \sim N(0,1)$. Again we do not correct for serial correlation. In Table 4, one can observe that: (i) Overall, it appears that both EF and BF are not appropriate to use because of considerable over-rejections. In fact, they get more over-sized as n increases.[18] (ii) More specifically, we have the empirical size of EF and BF varying between 5 to 16% even when the factors are *known*. (iii) This is an expected result in the sense that the wild bootstrap method used in this paper is not

[17] Since our concern in this paper is consistency, we do not go into details into the efficiency problem. Note that Choi (2008) proposes efficient estimation of factor models (when the factors are unknown), the so-called generalized principal component estimators (GPCEs). In fact, he uses maximum likelihood estimation of the factors and factor loadings under the assumption of normal error terms.

[18] We run also $\rho = \{0.2, 0.8, 0.99\}$ and find that EF and BF get more over-sized as ρ increases.

designed for the serially correlated case. Note that Gonçalves and Perron (2010) also obtain some noticeable size distortions for the serially correlated error terms. Hence, one needs to explore alternative bootstrap methods (such as the block bootstrap) rather than the wild bootstrap for this case.

4.2.2 Non-stationary factors

This section considers non-stationary factors, i.e., $F_t = F_{t-1} + \eta_t$ where η_t is generated by $IIDN(0,1)$. From Table 5, one can observe the following: (i) Basically, the results are similar to the case of stationary factors. With observable F_t both EF and BF are quite close to their true size. However, with unobservable F_t, EF gets extremely over-sized while the size of BF stays close to 5% varying from 4 to 6%. (ii) In addition to this, note that one obtains exactly the same size of EF and BF whether $F_t = I(0)$ or $F_t = I(1)$ if F_t is unknown. This is because the restricted bootstrap residuals are used to compute the F-statistic. That is, estimates of the factor and factor loadings are calculated based on y_{it} which is the same under the null whether F_t is $I(0)$ or $I(1)$.

To compute the size-adjusted power, we again divide our cases into strong and weak cross section dependence as in the previous section. All the results are reported in Table 6 and we find the following: (i) The power increases as λ_i increases, as in the $F_t = I(0)$ case. (ii) The overall power is higher with the non-stationary factors as compared with the stationary factors for each sample size. This may be due to the fact that the explanatory power of the estimated model increases because the signal with an $I(1)$ process is stronger than the one with an $I(0)$ process.[19]

We also check robustness to heteroskedasticity and serial correlation. Table 7 reports the results for heteroskedasticity. We again find that BF stays robust despite huge heteroskedasticity, while EF shows over-sized results even when the factors are

[19] Note that with the non-stationary regressor it is easier to identify coefficient estimates because of the stronger signal.

observable. In fact, the size of BF varies from 4 to 6% while that of EF from 14 to 21%. Combining this with Table 3, one concludes that BF in the common factor model can be used with heteroskedasticity *regardless of whether or not the factors are stationary and whether or not the factors are known.*

Table 8 re-confirms that some alternative bootstrap methods should be investigated for the serially correlated case. Figures 1 to 6 overlap the F distribution (Theoretical F), the empirical F distribution, and the bootstrap F distribution for the benchmark case depending on observability and stationarity of F_t. From the graphical illustrations, it can be seen that we have the consistent results with the previous literature, e.g., ANOVA literature and Orme and Yamagata (2006). In fact, when we vary n from 5 to 100, the distribution of EF converges to the normal shaped curve centered at 1 if F_t is known.

4.3 Case 2: With the regressor

4.3.1 Stationary regressor and factors

In this case, we add the regressor, x_{it}, as well as F_t:

$$y_{it} = x_{it}\beta + \lambda_i F_t + u_{it} \text{ for } i = 1, \ldots, n \text{ and } t = 1, \ldots, T$$

where $\beta = 2$ and $\lambda_i = 0$ for all i. Both x_{it} and F_t follow $I(0)$ processes and are generated from $N(0,1)$. For the benchmark case, we first generate u_{it} from $IIDN(0,1)$. The maximum number of iterations (when F_t is unobserved, for the interactive fixed effects estimator for β) is set at 5. Table 9 reports the empirical size of EF and BF. We basically observe similar results as in Case 1: (i) If F_t is known, the size of EF and BF are quite close to the true size (5%). (ii) If F_t is unknown, EF gets extremely over-sized while BF mimics the distribution of EF pretty well with huge improvements in size. Again, Figures 7 to 10 overlap the F distribution, EF, and BF for varying $n = \{5, 50\}$ and $T = \{10, 20, 40, 50\}$. One can easily check that we have a similar pattern with that in Section 4.2.1.

Table 10 indicates the size-adjusted power under strong and weak dependence. Again, the power seems good under strong dependence especially when F_t is known. Under weak dependence, however, both EF and BF have much less power. In fact, if F_t is unobserved, then the size-adjusted power of EF and BF ranges between 4 and 9%. Also note that the power increases as λ_i or the sample size increases. Heteroskedasticity and serial correlation are again introduced into the error terms (Table 11 and 12) and we have the similar findings as in Case 1.

4.3.2 Non-stationary regressor and factors

In this section, x_{it} and F_t are assumed to be non-stationary. The data are generated as follows: For $i = 1, \ldots, n$ and $t = 1, \ldots, T$,

$$\begin{aligned} y_{it} &= 2x_{it} + \lambda_i F_t + u_{it}, \\ x_{it} &= x_{it-1} + \varepsilon_{it}, \end{aligned}$$

and

$$F_t = F_{t-1} + \eta_t$$

where $\begin{pmatrix} u_{it} \\ \varepsilon_{it} \\ \eta_t \end{pmatrix} \overset{iid}{\sim} N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & 1 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & 1 \end{bmatrix} \right).$

We follow most of the settings in Bai, *et al.* (2009) for the simulation. In particular, we set σ_{32} at 0.4 while varying σ_{21} and σ_{31} over $\{0, 0.2, 0.8\}$. The long-run covariance matrix is estimated using the KERNEL procedure in COINT 2.0. We use the Bartlett window with the truncation set at 5. The maximum number of iterations to estimate β (when F_t is unknown) is also set at 5. The empirical size for each case is reported Tables 13 to 15 depending on the combination of σ_{21} and σ_{31}. We find the following: (i) Suppose first that F_t is observable. In this case, if σ_{21} is low ($\sigma_{21} = 0$ or $\sigma_{21} = 0.2$), each of EF and BF shows the correct size (Table 13 and 14). However, for $\sigma_{21} = 0.8$, both EF and BF get over-sized in relatively small sample sizes although this distortion seems to quickly get better as T increases (Table 15). In fact, for $(n, T) = (50, 50)$ the size of EF and BF varies between 8 and 9%. This can be explained by the fact that one needs to have enough samples to estimate the long-run covariance matrix. One can also observe that σ_{21} rather than σ_{31} affects the performance of EF and BF. This phenomenon stems from the fact that σ_{31} does not matter much under the null. (ii) If F_t is unobserved, EF almost always rejects the null like in the previous cases. BF shows the correct size for $\sigma_{21} = 0$ or $\sigma_{21} = 0.2$. Interestingly, if $\sigma_{21} = 0.8$, the performance of BF using CupFM is quite different from that using CupBC although the size of both improves as T increases. In fact, CupFM leads to the reasonable size varying between 3-8% while CupBC causes considerable over-sizing.[20] Note that the distortion using CupBC gets worse with larger n and smaller T. This implies that correcting for endogeneity and serial correlation at every iteration (CupFM), not only at the final stage (CupBC), is helpful in improving the goodness of the long-run covariance matrix estimation. (iii) Overall, with low σ_{21}, similar conclusions with the previous sections continue to hold. Both EF and BF (with LSFM) can be used if the factors are *known*, while only BF should be used if the factors are unknown. However, with high σ_{21}, using CupFM instead of CupBC seems to be more appropriate. (iv) Lastly, note that the results when $\sigma_{21} = 0$ and $\sigma_{31} = 0$ are graphically displayed in Figures 11 to 16.

Tables 16 to 24 present the size-adjusted power for each case. The results seem

[20]We compute the signal-to-noise ratio $= \frac{2+2\sigma_{23}}{5+4\sigma_{21}}$ and observe that we have the lower signal-to-noise ratio as σ_{21} increases (so more size distortion is expected). In fact, we vary σ_{21} over $\{0, 0.2, 0.4, 0.6, 0.8\}$ although not reported here. The size of CupBC clearly gets worse as σ_{21} increases but seems to be relatively robust until $\sigma_{21} = 0.4$. In contrast, increasing σ_{23} (the stronger signal) leads to slight improvement in size.

satisfactory and one can basically draw the same conclusion as in the previous sections. We also check robustness to heteroskedasticity and serial correlation. The empirical size under heteroskedasticity is reported in Tables 25 to 27. One observes the following: (i) When F_t is observed, EF becomes over-sized and this gets worse with higher σ_{21}. However, even with high σ_{21}, BF is much less over-sized than EF. In fact, with $\sigma_{21} = 0.8$ the size of BF gets quickly closer to true size 5% as T increases (Table 27). (ii) When F_t is not observed, EF gets extremely over-sized again. The size of BF using CupFM, however, stays relatively robust varying from 3 to 9% and clearly improves as the sample size increases. Hence, the size of BF using CupFM under heteroskedasticity seems to perform well whether the regressor is included or not and whether x_{it} and F_t follow $I(0)$ or $I(1)$. In contrast, BF is consistently over-sized for all the experiments and gets worse as the sample size increases when serial correlation is present, see Tables 28 to 30.

5 Conclusion

High-dimensional data analysis for large n / large T has become an integral part of the macro panel data literature. This paper suggests using the bootstrap F-test to test for cross-sectional independence. This circumvents the difficulty of deriving the asymptotic distribution of this statistic with large n / large T . The simulation results show that the bootstrap F-test performs well in testing cross-sectional independence and is recommended in practice. This F-test has the added advantage of being feasible even when we do not observe the factors. Extensive simulations show that the wild bootstrap F-test is robust to heteroskedasticity but sensitive to serial correlation.

References

[1] Andrews, D. W. K. (2005), "Cross-section Regression with Common Shocks," *Econometrica*, 73, 1551-1585.

[2] Akritas, M., and Arnold, S. (2000), "Asymptotics for Analysis of Variance When the Number of Levels is Large," *Journal of the American Statistical Association*, 449, 212-226.

[3] Bai, J. (2003), "Inferential Theory for Factor Models of Large Dimensions," *Econometrica*, 71, 135-171.

[4] Bai, J. (2004), "Estimating Cross-section Common Stochastic Trends in Nonstationary Panel Data," *Journal of Econometrics*, 122, 137-183.

[5] Bai, J. (2009), "Panel Data Models with Interactive Fixed Effects," *Econometrica*, 77, 1229-1279.

[6] Bai, J., Kao, C., and Ng, S. (2009), "Panel Cointegration with Global Stochastic Trends," *Journal of Econometrics*, 149, 82-99.

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

[8] Baltagi, B. H. (2008), *Econometric Analysis of Panel Data*, New York, Wiley.

[9] Baltagi, B. H., and Kao, C. (2000), "Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey," *Advances in Econometrics*, 15, 7-51.

[10] Bathke, A. (2004), "The ANOVA F Test Can Still Be Used in Some Balanced Designs with Unequal Variances and Nonnormal Data," *Journal of Statistical Planning and Inference*, 126, 413-422.

[11] Boos, D. D., and Brownie, C. (1995), "ANOVA and Rank Tests When the Number of Treatments is Large," *Statistics & Probability Letters*, 23, 183-191.

[12] Choi, I. (2008), "Efficient Estimation of Factor Model," *Working Paper*, Hong Kong University of Science and Technology.

[13] Davidson, R., and Flachaire, E. (2008), "Wild Bootstrap, Tamed at Last," *Journal of Econometrics*, 146, 162-169.

[14] Diaconis, P., and Efron, B. (1983), "Computer-intensive Methods in Statistics," *Scientific American*, 248, 116-130.

[15] Flachaire, E. (2005), "Bootstrapping Heteroskedastic Regression Models: Wild Bootstrap vs. Pairs Bootstrap," *Computational Statistics & Data Analysis*, 49, 361-376.

[16] Freedman, D. A. (1981), "Bootstrapping Regression Models," *Annals of Statistics*, 9, 1218-1228.

[17] Gonçalves, S., and Perron, B. (2010), "Bootstrapping Factor-augmented Regression Models," *Working Paper*, Université de Montréal.

[18] Lehmann, E. L. (1999), *Elements of Large-Sample Theory*, New York, Springer.

[19] Liu, R. Y. (1988), "Bootstrap Procedures under Some Non-I.I.D. Models," *Annals of Statistics*, 16, 1696-1708.

[20] MacKinnon, J. G. (2006), "Bootstrap Methods in Econometrics," *Econometric Record*, 82, S2-S18

[21] MacKinnon, J. G. (2007), "Bootstrap Hypothesis Testing," *Working Paper*, Queen's University.

[22] Mammen, E. (1993a), *When Does Bootstrap Work? Asymptotic Results and Simulations*, New York, Springer.

[23] Mammen, E. (1993b), "Bootstrap and Wild Bootstrap for High Dimensional Linear Models," *Annals of Statistics*, 21, 255-285.

[24] Orme, C. D., and Yamagata, T. (2006), "The asymptotic distribution of the F-test statistic for individual effects," *Econometrics Journal*, 9, 404-422.

[25] Phillips, P. C. B., and Hansen, B. E. (1990), "Statistical Inference in Instrumental Variables Regression with I(1) Processes," *Review of Economic Studies*, 57, 99-125.

[26] Phillips, P. C. B., and Moon, H. R. (1999), "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica*, 67, 1057-1111.

Figure 1: Case 1, The Histogram of Bootstrap F When F_t Is $I(0)$ and Known ($n = 5$)

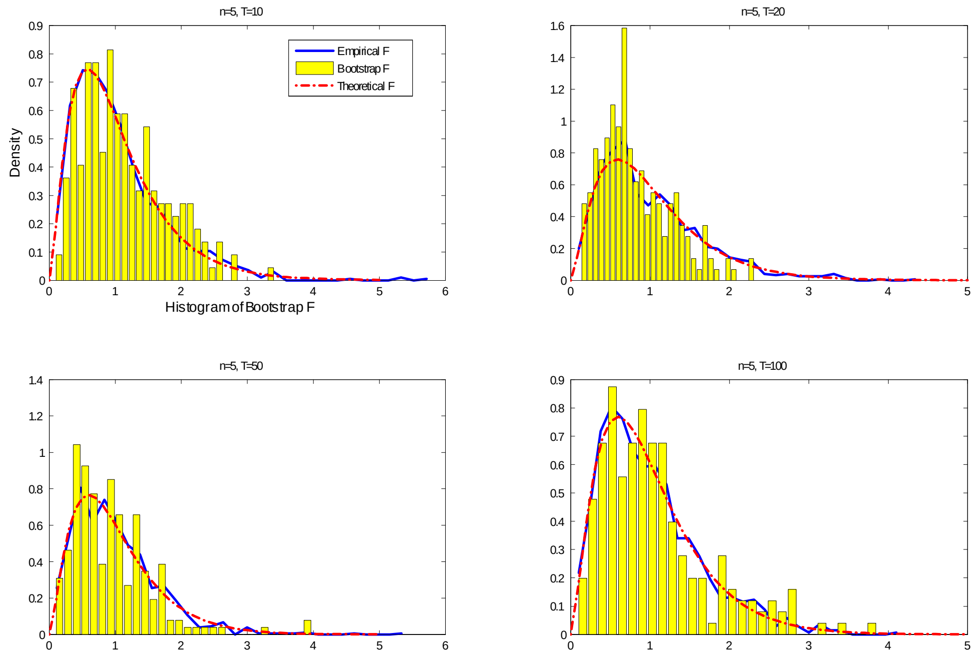

Figure 2: Case 1, The Histogram of Bootstrap F When F_t Is $I(0)$ and Known ($n = 100$)

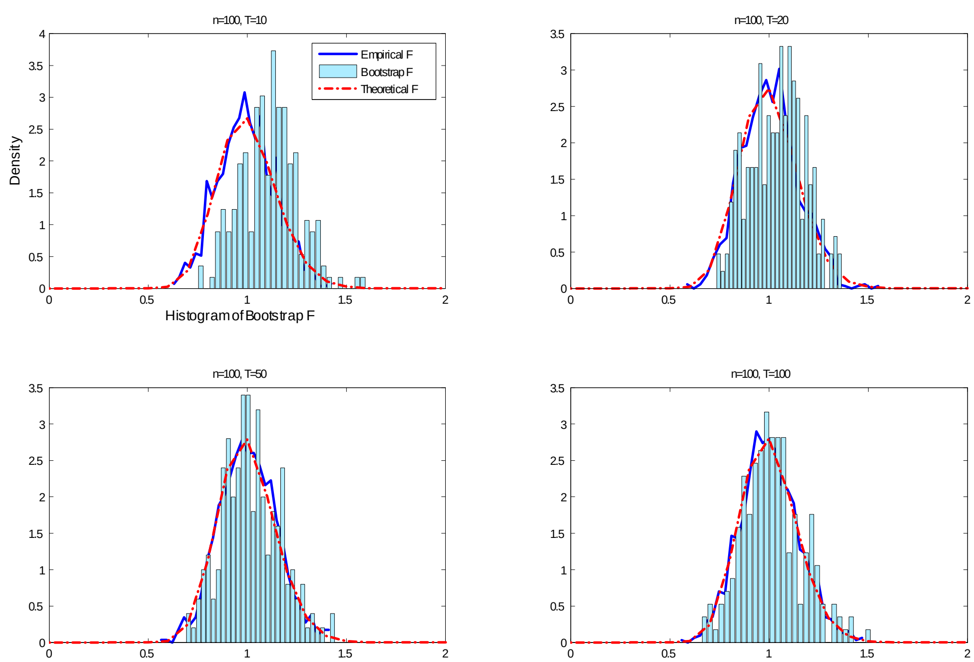

Figure 3: Case 1, The Histogram of Bootstrap F When F_t Is $I(1)$ and Known ($n = 5$)

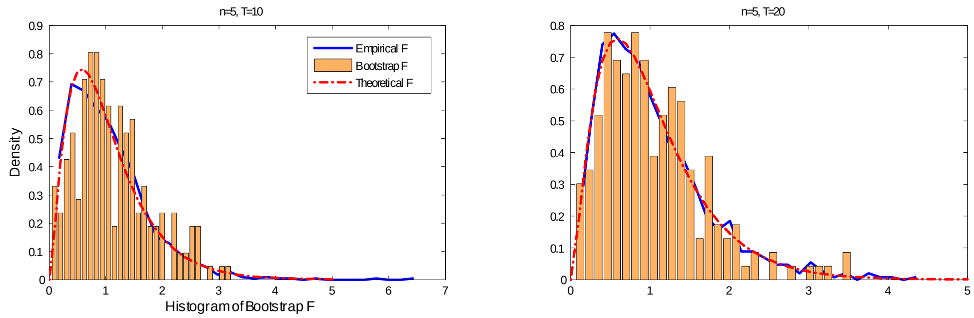

Figure 4: Case 1, The Histogram of Bootstrap F When F_t Is $I(1)$ and Known ($n = 100$)

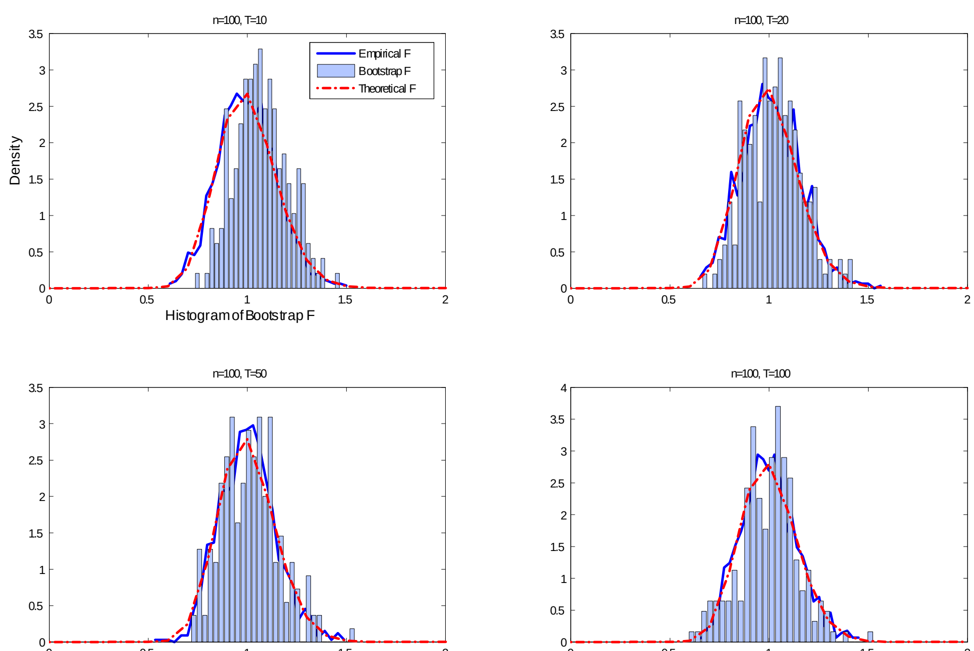

Figure 5: Case 1, The Histogram of Bootstrap F When F_t Is Unknown ($n = 5$)

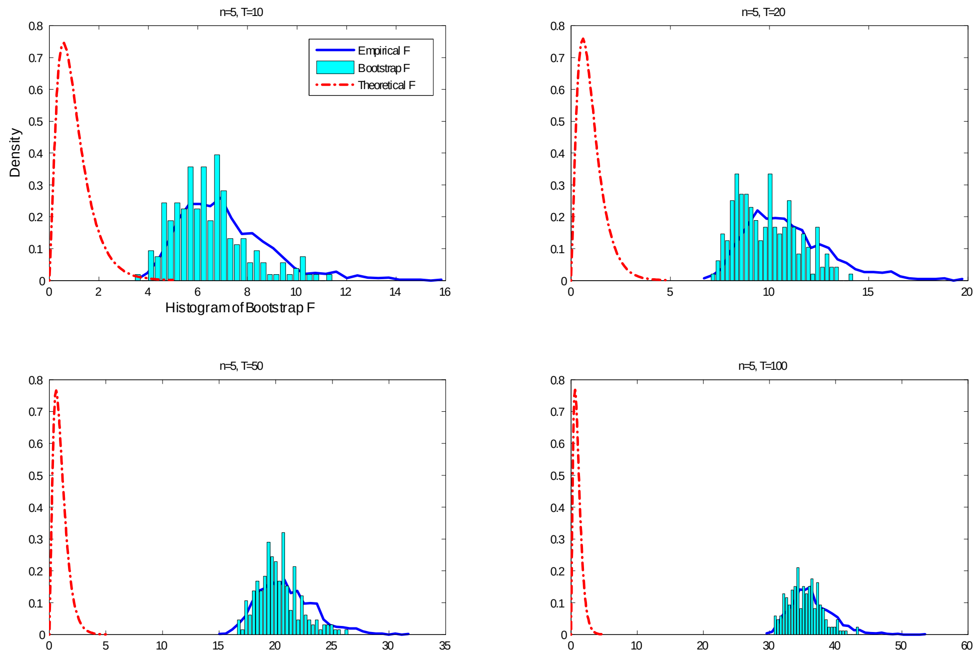

Figure 6: Case 1, The Histogram of Bootstrap F When F_t Is Unknown ($n = 100$)

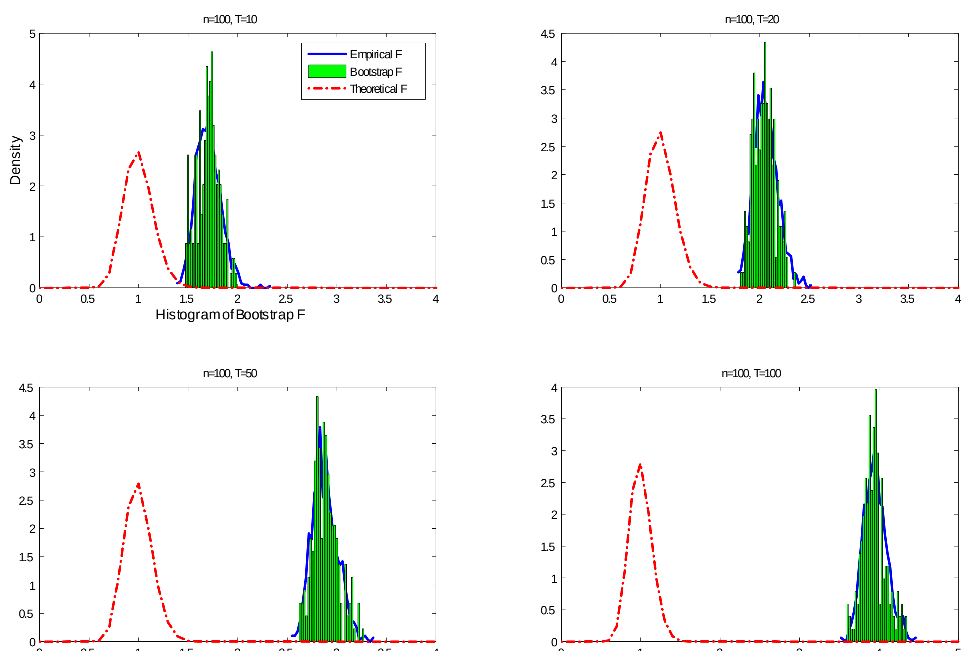

Figure 7: Case 2, The Histogram of Bootstrap F When F_t Is $I\left(0\right)$ and Known $\left(n=5\right)$

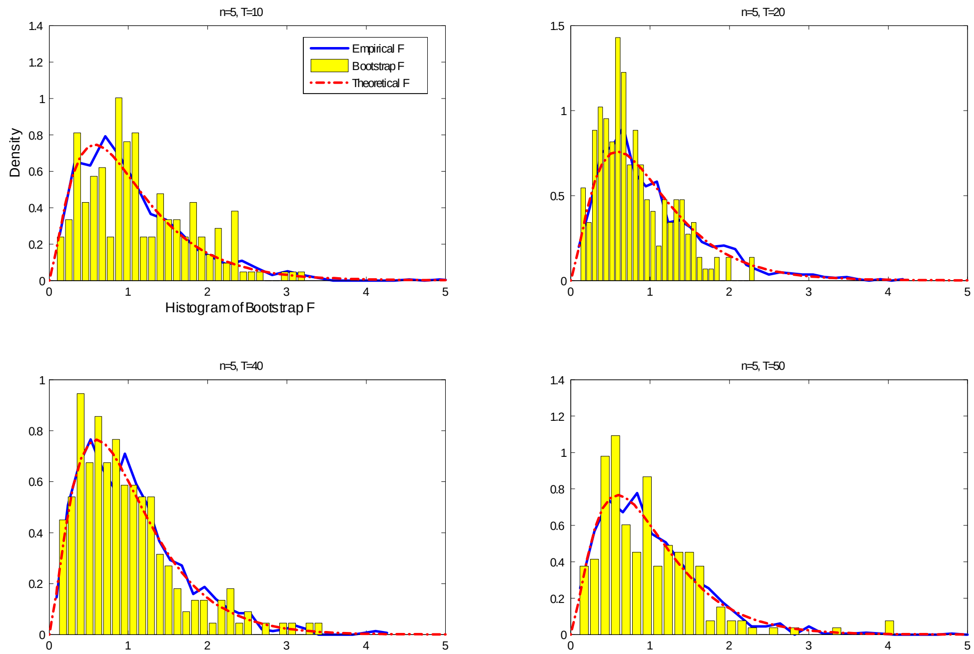

Figure 8: Case 2, The Histogram of Bootstrap F When F_t Is $I\left(0\right)$ and Known $\left(n=50\right)$

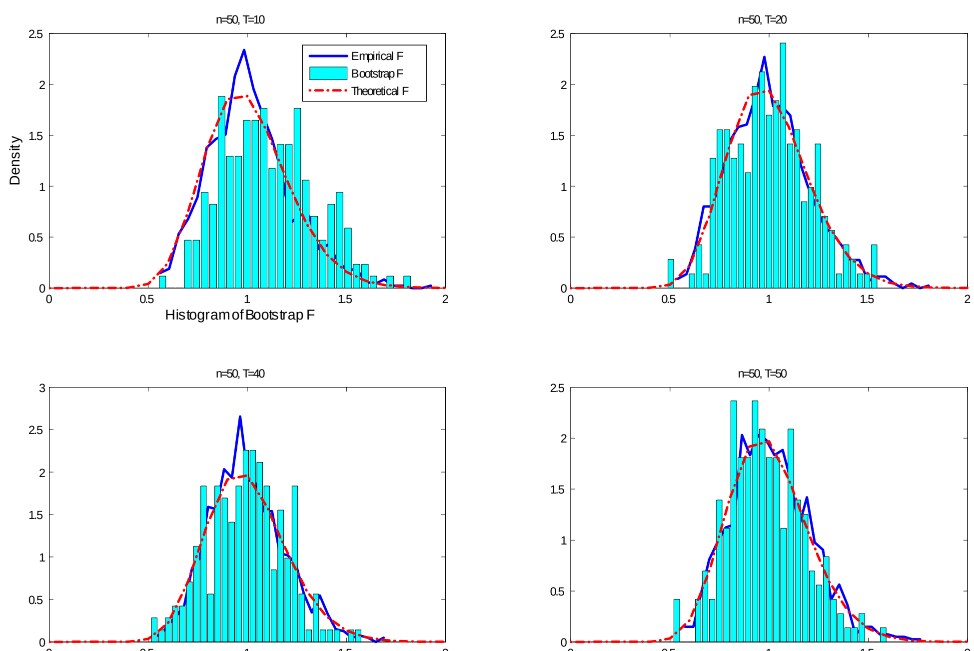

Figure 9: Case 2, The Histogram of Bootstrap F When F_t Is $I(0)$ and Unknown $(n = 5)$

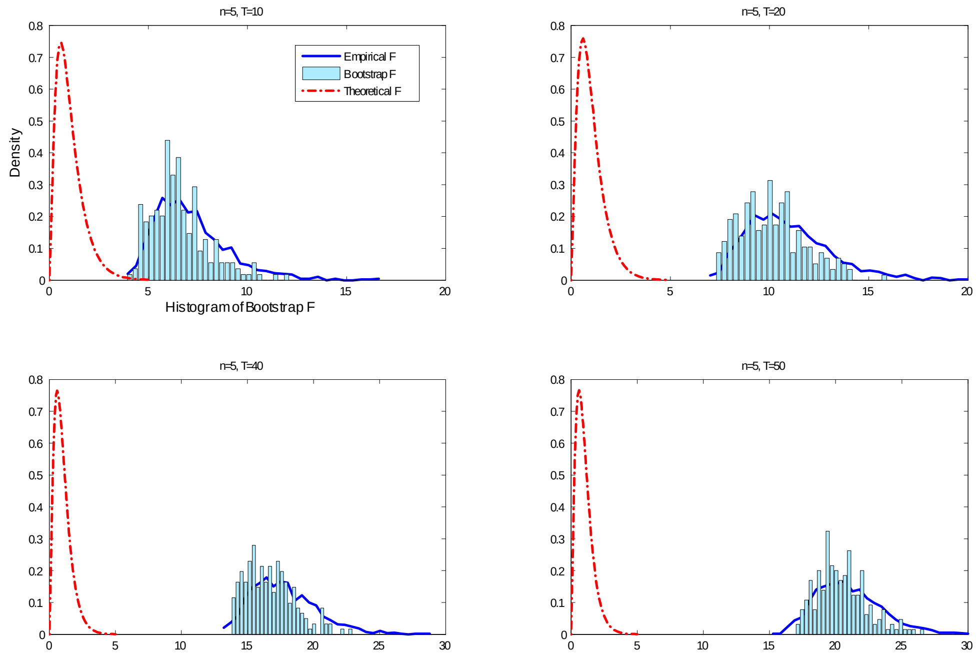

Figure 10: Case 2, The Histogram of Bootstrap F When F_t Is $I(0)$ and Unknown $(n = 50)$

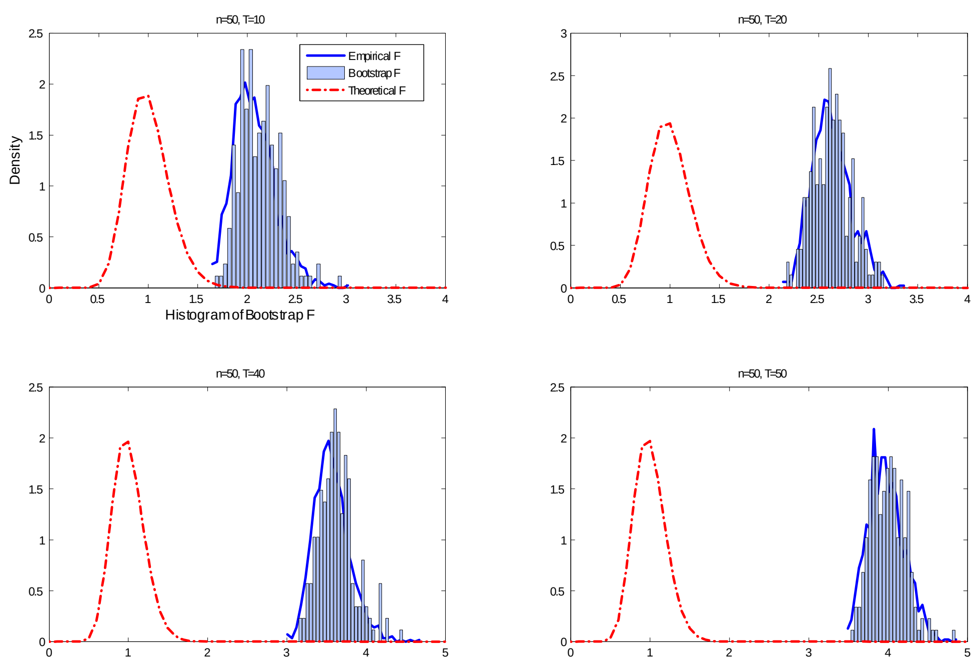

Figure 11: Case 2, The Histogram of Bootstrap F When F_t Is $I(1)$ and Known $(n = 5)$

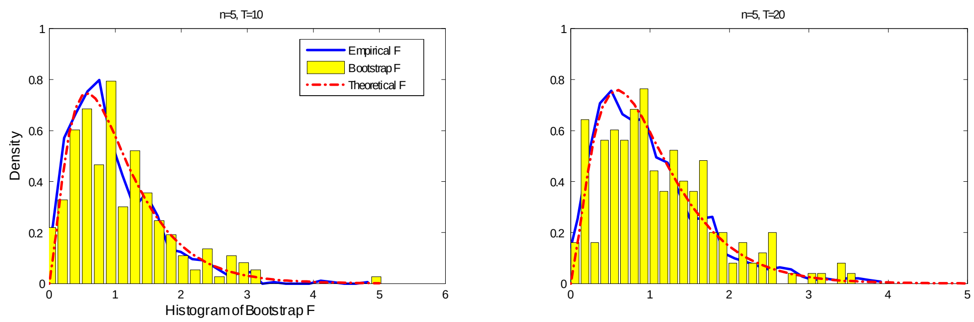

Figure 12: Case 2, The Histogram of Bootstrap F When F_t Is $I(1)$ and Known $(n = 50)$

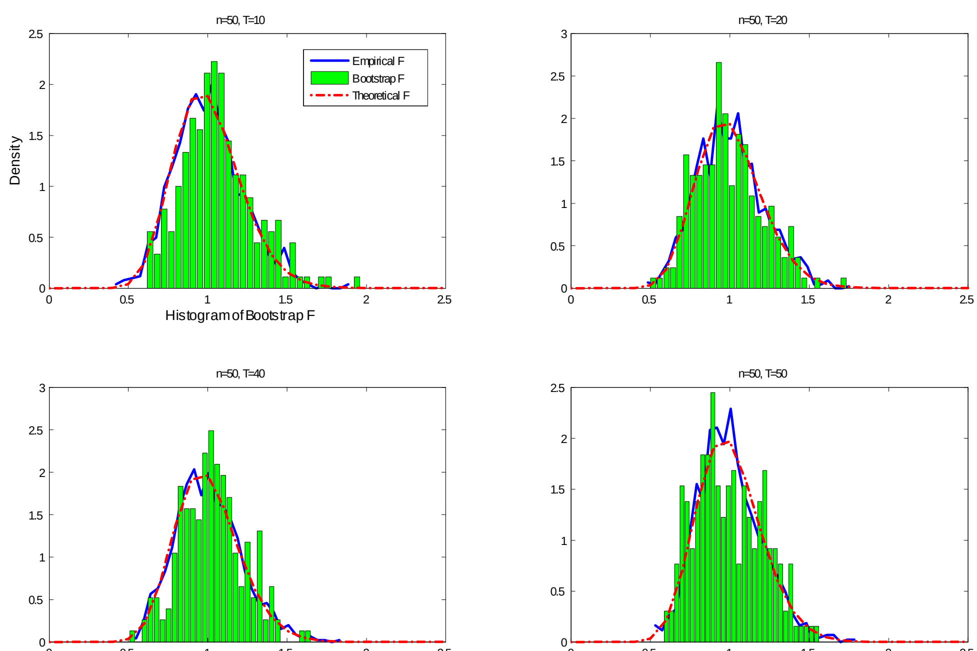

Figure 13: Case 2, The Histogram of Bootstrap F When F_t Is $I(1)$ and Unknown (CupBC, $n = 5$)

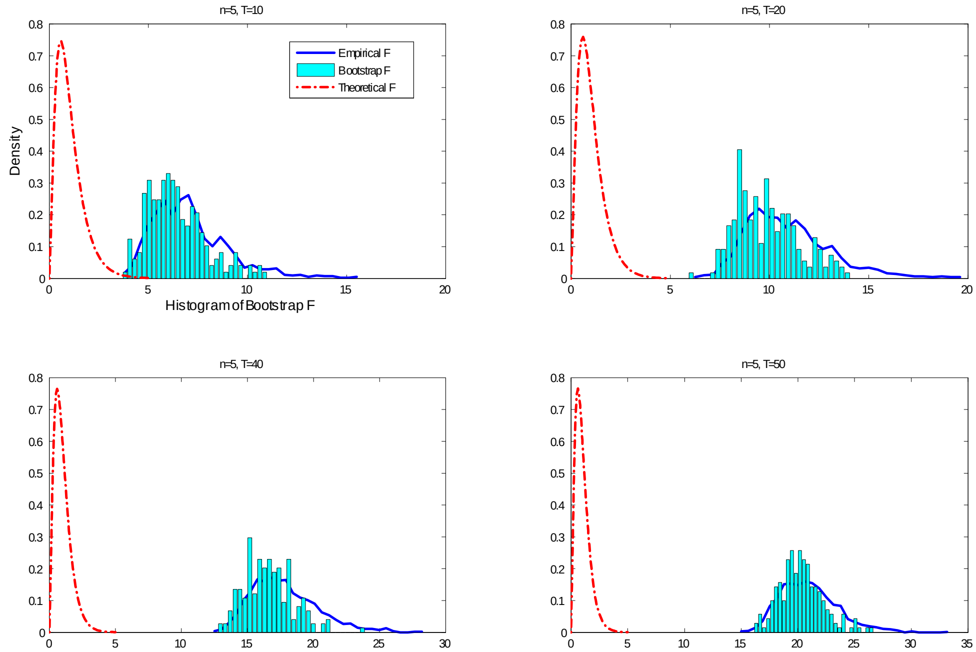

Figure 14: Case 2, The Histogram of Bootstrap F When F_t Is $I(1)$ and Unknown (CupBC, $n = 50$)

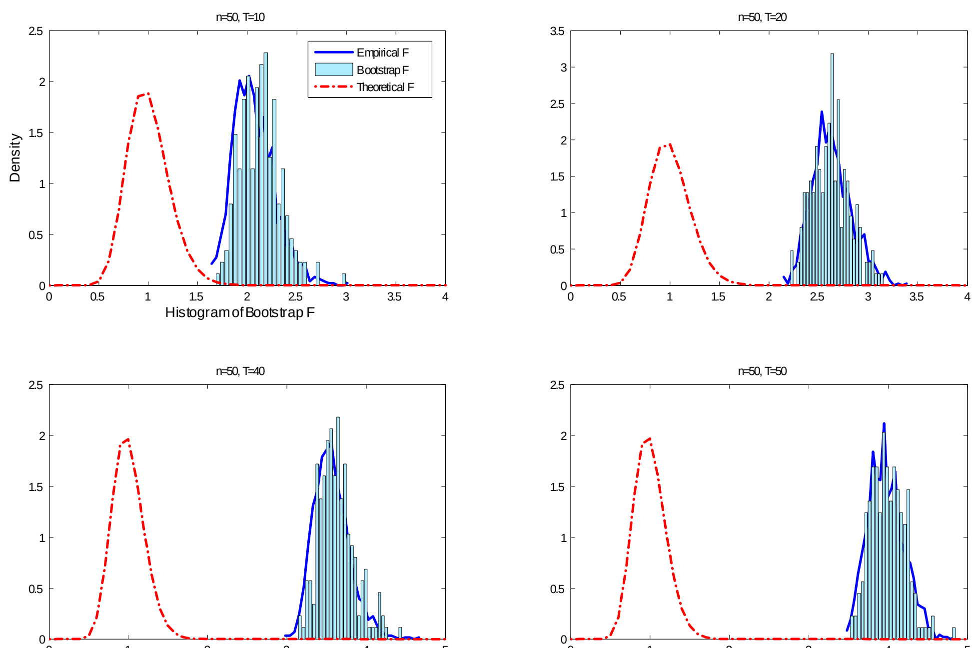

Figure 15: Case 2, The Histogram of Bootstrap F When F_t Is $I(1)$ and Unknown (CupFM, $n = 5$)

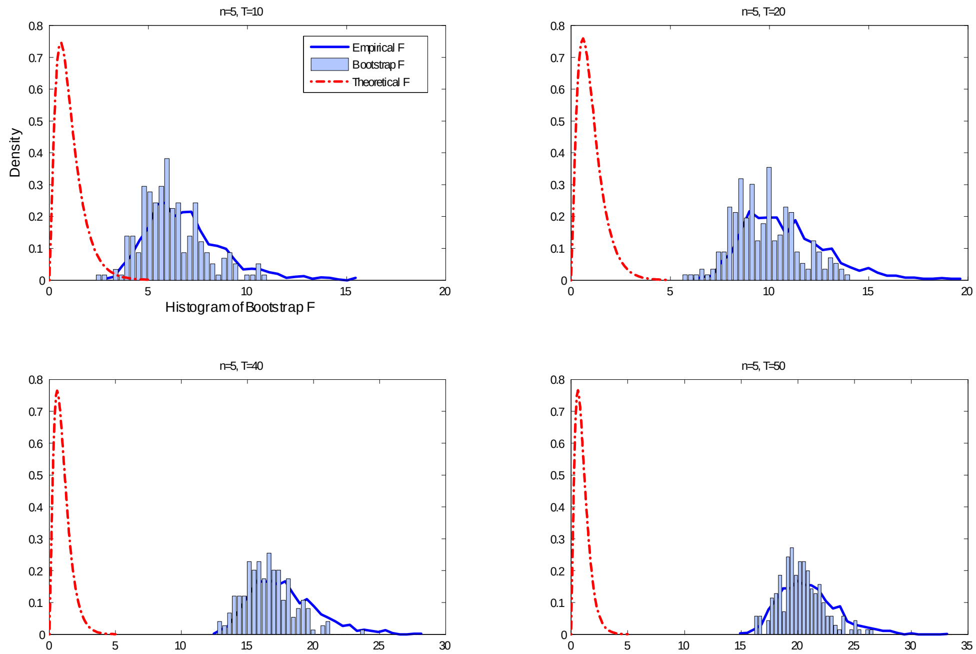

Figure 16: Case 2, The Histogram of Bootstrap F When F_t Is $I(1)$ and Unknown (CupFM, $n = 50$)

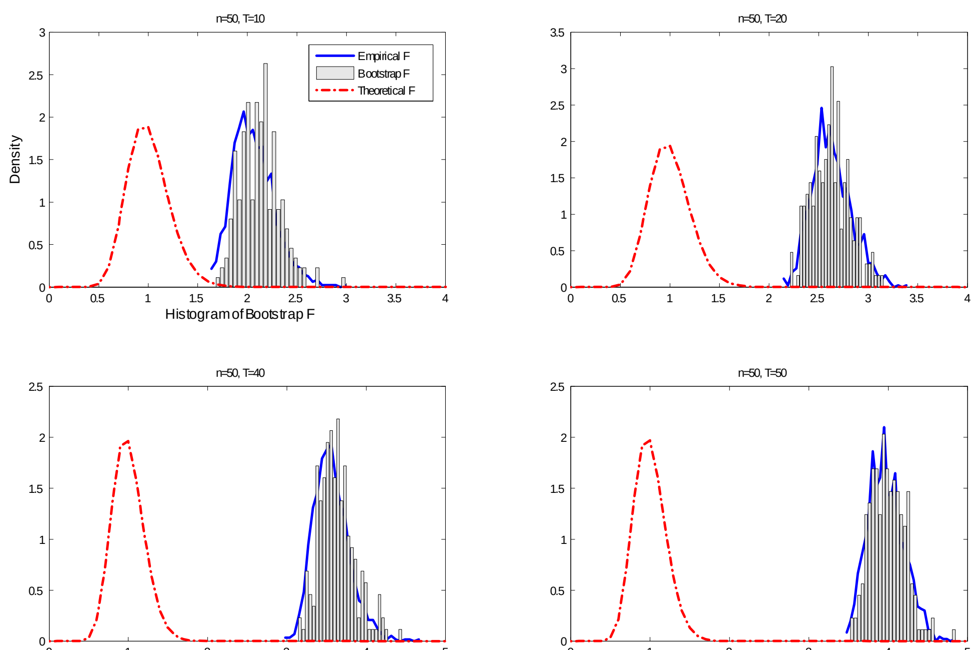

Table 1: Case 1, The Size (%) of F-test When F_t Is $I(0)$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 90$	$10, 490$	$10, 990$	$50, 450$	$50, 2450$	$50, 4950$	$100, 900$	$100, 4900$	$100, 9900$
(n, T)		$(10, 10)$	$(10, 50)$	$(10, 100)$	$(50, 10)$	$(50, 50)$	$(50, 100)$	$(100, 10)$	$(100, 50)$	$(100, 100)$
Known F_t	EF	4.1	4.7	5.3	6.3	5.1	4.3	3.6	5.0	5.6
	BF	5.1	5.8	5.2	6.2	5.6	4.6	5.1	5.3	5.6
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
	BF	4.5	4.8	4.5	6.2	5.8	4.9	5.4	5.9	5.6

Note: True size is 5%

Table 2: Case 1, The Size-adjusted Power (%) of F-test When F_t Is $I(0)$ ($H_a : \lambda_i \neq 0$ for all i)

Strong dependence: $\lambda_i \sim IIDU(0.2, 0.5)$										
(n, T)		$(10, 10)$	$(10, 50)$	$(10, 100)$	$(50, 10)$	$(50, 50)$	$(50, 100)$	$(100, 10)$	$(100, 50)$	$(100, 100)$
Known F_t	EF	59.5	99.8	100.0	92.6	100.0	100.0	98.6	100.00	100.00
	BF	64.1	99.9	100.0	95.6	100.0	100.0	98.6	100.00	100.00
Unknown F_t	EF	15.5	73.8	95.8	62.7	100.0	100.0	86.4	100.00	100.00
	BF	20.4	73.8	95.7	66.9	100.0	100.0	86.1	100.00	100.00
Weak dependence: $\lambda_i \sim IIDU(0.01, 0.2)$										
Known F_t	EF	8.7	39.4	72.3	14.4	85.4	99.4	28.2	97.7	100.0
	BF	10.1	40.9	72.5	20.9	86.8	99.3	27.4	98.3	100.0
Unknown F_t	EF	5.1	6.0	6.7	6.0	9.0	17.0	6.7	16.7	45.3
	BF	5.4	5.5	6.2	6.7	9.6	17.2	7.3	19.1	46.1

Table 3: Case 1, The Size (%) of F-test under Heteroskedasticity When F_t Is $I(0)$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 90$	$10, 490$	$10, 990$	$50, 450$	$50, 2450$	$50, 4950$	$100, 900$	$100, 4900$	$100, 9900$
(n, T)		$(10, 10)$	$(10, 50)$	$(10, 100)$	$(50, 10)$	$(50, 50)$	$(50, 100)$	$(100, 10)$	$(100, 50)$	$(100, 100)$
Known F_t	EF	15.8	16.8	13.9	19.7	19.6	19.6	19.0	18.1	20.9
	BF	4.5	6.2	5.7	4.8	4.3	5.3	4.7	4.8	5.0
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
	BF	5.2	5.2	4.3	5.8	5.9	5.0	5.6	5.0	5.1

Table 4: Case 1, The Size (%) of F-test under Serial Correlation When F_t Is $I(0)$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 90$	$10, 490$	$10, 990$	$50, 450$	$50, 2450$	$50, 4950$	$100, 900$	$100, 4900$	$100, 9900$
(n, T)		$(10, 10)$	$(10, 50)$	$(10, 100)$	$(50, 10)$	$(50, 50)$	$(50, 100)$	$(100, 10)$	$(100, 50)$	$(100, 100)$
Known F_t	EF	6.1	5.8	6.3	14.1	9.0	6.8	16.3	10.5	8.2
	BF	7.4	5.7	6.5	15.1	9.4	7.2	16.7	10.4	8.8
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
	BF	31.4	29.6	28.7	99.4	98.3	98.6	100.0	100.0	100.0

Table 5: Case 1, The Size (%) of F-test When F_t Is $I(1)$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 90$	$10, 490$	$10, 990$	$50, 450$	$50, 2450$	$50, 4950$	$100, 900$	$100, 4900$	$100, 9900$
(n, T)		$(10, 10)$	$(10, 50)$	$(10, 100)$	$(50, 10)$	$(50, 50)$	$(50, 100)$	$(100, 10)$	$(100, 50)$	$(100, 100)$
Known F_t	EF	4.5	4.3	5.3	6.5	5.1	4.7	4.0	4.8	4.9
	BF	6.3	5.7	5.9	6.0	5.4	5.6	4.5	5.2	5.3
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
	BF	4.5	4.8	4.5	6.2	5.8	4.9	5.4	5.9	5.6

Note: True size is 5%

Table 6: Case 1, The Size-adjusted Power (%) of F-test When F_t Is $I(1)$ ($H_a : \lambda_i \neq 0$ for all i)

Strong dependence: $\lambda_i \sim IIDU(0.2, 0.5)$										
(n, T)		$(10, 10)$	$(10, 50)$	$(10, 100)$	$(50, 10)$	$(50, 50)$	$(50, 100)$	$(100, 10)$	$(100, 50)$	$(100, 100)$
Known F_t	EF	88.6	100.0	100.0	97.5	100.0	100.0	99.0	100.0	100.0
	BF	89.3	100.0	100.0	98.6	100.0	100.0	99.1	100.0	100.0
Unknown F_t	EF	64.0	100.0	100.0	89.1	100.0	100.0	96.1	100.0	100.0
	BF	66.0	100.0	100.0	90.5	100.0	100.0	95.9	100.0	100.0
Weak dependence: $\lambda_i \sim IIDU(0.01, 0.2)$										
Known F_t	EF	33.9	97.9	100.0	55.7	100.0	100.0	69.0	100.0	100.0
	BF	35.2	97.7	100.0	60.6	100.0	100.0	70.4	100.0	100.0
Unknown F_t	EF	12.8	75.8	94.9	29.0	96.5	100.0	42.2	98.9	100.0
	BF	15.0	75.0	95.1	33.1	96.5	100.0	43.8	99.1	100.0

Table 7: Case 1, The Size (%) of F-test under Heteroskedasticity When F_t Is $I(1)$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 90$	$10, 490$	$10, 990$	$50, 450$	$50, 2450$	$50, 4950$	$100, 900$	$100, 4900$	$100, 9900$
(n, T)		$(10, 10)$	$(10, 50)$	$(10, 100)$	$(50, 10)$	$(50, 50)$	$(50, 100)$	$(100, 10)$	$(100, 50)$	$(100, 100)$
Known F_t	EF	16.9	16.7	14.9	17.6	17.4	18.8	20.2	21.2	20.0
	BF	5.3	5.6	5.5	4.0	4.6	5.6	5.4	5.3	6.1
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
	BF	5.2	5.2	4.3	5.8	5.9	5.0	5.6	5.0	5.1

Table 8: Case 1, The Size (%) of F-test under Serial Correlation When F_t Is $I(1)$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 90$	$10, 490$	$10, 990$	$50, 450$	$50, 2450$	$50, 4950$	$100, 900$	$100, 4900$	$100, 9900$
(n, T)		$(10, 10)$	$(10, 50)$	$(10, 100)$	$(50, 10)$	$(50, 50)$	$(50, 100)$	$(100, 10)$	$(100, 50)$	$(100, 100)$
Known F_t	EF	48.2	59.1	63.1	84.8	98.9	98.8	91.6	99.9	99.9
	BF	49.6	60.5	64.4	87.8	98.9	98.8	92.6	99.9	100.0
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
	BF	31.4	29.6	28.7	97.4	98.3	98.6	100.0	100.0	100.0

Table 9: Case 2, The Size (%) of F-test When F_t Is $I(0)$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 89$	$10, 189$	$10, 489$	$20, 179$	$20, 379$	$20, 979$	$50, 449$	$50, 949$	$50, 2449$
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	4.2	5.8	4.4	4.2	6.2	5.6	6.4	5.5	5.4
	BF	3.8	6.0	5.8	5.6	6.5	5.8	6.0	5.8	5.4
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
	BF	5.1	5.4	4.8	5.1	5.5	4.1	5.9	5.3	5.6

Note: True size is 5%

Table 10: Case 2, The Size-adjusted Power (%) of F-test When F_t Is $I(0)$ ($H_a : \lambda_i \neq 0$ for all i)

Strong dependence: $\lambda_i \sim IIDU(0.2, 0.5)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	59.4	88.9	99.9	77.6	97.4	100.0	92.0	100.0	100.0
	BF	63.4	90.3	99.9	81.7	98.0	100.0	95.7	100.0	100.0
Unknown F_t	EF	15.8	29.1	73.2	31.0	60.0	97.1	62.5	93.3	100.0
	BF	20.4	34.1	73.8	33.6	63.5	97.0	67.1	93.8	100.0
Weak dependence: $\lambda_i \sim IIDU(0.01, 0.2)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	8.2	15.9	39.4	12.6	20.7	54.3	14.4	33.6	85.1
	BF	9.5	16.6	40.2	15.4	24.6	59.0	21.1	39.9	86.9
Unknown F_t	EF	4.9	5.1	6.2	5.3	5.5	6.9	6.1	6.0	9.1
	BF	5.4	5.5	5.7	5.5	6.0	6.0	6.7	6.2	9.7

Table 11: Case 2, The Size (%) of F-test under Heteroskedasticity When F_t Is $I(0)$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 89$	$10, 189$	$10, 489$	$20, 179$	$20, 379$	$20, 979$	$50, 449$	$50, 949$	$50, 2449$
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	15.6	17.1	16.7	17.4	17.7	18.9	19.7	18.3	19.7
	BF	4.6	5.6	6.2	4.9	5.5	6.5	4.7	5.6	4.5
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
	BF	6.2	6.8	5.8	5.1	6.7	5.9	6.0	4.8	5.9

Table 12: Case 2, The Size (%) of F-test under Serial Correlation When F_t Is $I(0)$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 89$	$10, 189$	$10, 489$	$20, 179$	$20, 379$	$20, 979$	$50, 449$	$50, 949$	$50, 2449$
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	6.2	6.7	5.7	10.2	7.9	5.4	14.1	10.5	9.0
	BF	7.1	6.5	6.1	11.4	8.6	5.2	15.1	11.5	9.4
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
	BF	31.1	29.8	29.7	57.9	61.1	61.8	97.2	98.1	98.3

Table 13: Case 2, The Size (%) of F-test When F_t Is $I(1)$ Where $\sigma_{21}=0$ ($H_0:\lambda_i=0$ for all i)

$D.F.(num, den)$		$10,89$	$10,189$	$10,489$	$20,179$	$20,379$	$20,979$	$50,449$	$50,949$	$50,2449$
(n,T)		$(10,10)$	$(10,20)$	$(10,50)$	$(20,10)$	$(20,20)$	$(20,50)$	$(50,10)$	$(50,20)$	$(50,50)$
$\sigma_{31}=0$										
Known F_t	EF	4.7	5.6	5.2	4.6	4.5	6.7	5.7	5.7	5.0
(LSFM)	BF	6.2	6.1	4.3	4.5	5.2	6.8	5.6	5.9	5.4
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	4.6	5.3	4.8	5.3	5.5	4.3	6.0	5.5	5.7
Unknown F_t	EF	99.7	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	4.6	5.3	4.9	5.2	5.5	4.3	6.1	5.5	5.7
$\sigma_{31}=0.2$										
Known F_t	EF	5.0	5.5	5.0	4.0	4.5	6.7	5.9	6.2	4.9
(LSFM)	BF	5.2	6.6	4.8	4.7	5.7	6.9	5.8	6.9	5.0
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	4.6	5.3	4.8	5.3	5.5	4.3	6.0	5.5	5.7
Unknown F_t	EF	99.7	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	4.6	5.3	4.9	5.2	5.5	4.3	6.1	5.5	5.7
$\sigma_{31}=0.8$										
Known F_t	EF	5.6	4.3	5.5	4.5	5.6	5.3	4.8	5.2	4.0
(LSFM)	BF	6.3	5.7	6.5	4.9	6.0	5.5	5.0	6.4	4.9
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	4.6	5.3	4.7	5.3	5.5	4.3	6.0	5.5	5.7
Unknown F_t	EF	99.7	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	4.5	5.3	4.9	5.2	5.6	4.3	6.0	5.5	5.7

Note: True size is 5%

Table 14: *Case 2, The Size (%) of F-test When F_t Is $I(1)$ Where $\sigma_{21} = 0.2$ ($H_0 : \lambda_i = 0$ for all i)*

$D.F.(num, den)$		10, 89	10, 189	10, 489	20, 179	20, 379	20, 979	50, 449	50, 949	50, 2449
(n, T)		(10, 10)	(10, 20)	(10, 50)	(20, 10)	(20, 20)	(20, 50)	(50, 10)	(50, 20)	(50, 50)
$\sigma_{31} = 0$										
Known F_t	EF	5.4	6.2	5.4	5.0	4.0	6.1	5.7	6.1	5.4
(LSFM)	BF	6.1	6.5	5.5	5.4	4.9	6.4	6.1	6.2	6.0
Unknown F_t	EF	99.8	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	4.6	5.2	5.1	6.5	6.3	5.0	6.5	5.3	6.4
Unknown F_t	EF	99.8	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	4.6	5.1	5.1	6.0	5.9	4.6	5.5	4.8	5.7
$\sigma_{31} = 0.2$										
Known F_t	EF	5.8	6.3	5.0	4.4	4.0	6.2	5.4	6.3	5.2
(LSFM)	BF	6.3	7.2	5.3	5.0	5.2	6.6	6.1	6.4	5.4
Unknown F_t	EF	99.8	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	4.6	5.2	5.2	6.5	6.3	5.0	6.5	5.3	6.4
Unknown F_t	EF	99.8	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	4.6	5.1	5.0	6.1	5.9	4.6	5.4	4.8	5.7
$\sigma_{31} = 0.8$										
Known F_t	EF	5.1	4.4	5.6	4.5	5.7	4.8	4.5	5.1	4.6
(LSFM)	BF	6.5	5.2	6.0	5.3	5.4	5.7	4.7	5.9	5.1
Unknown F_t	EF	99.8	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	4.6	5.1	5.2	6.6	6.3	5.0	6.5	5.3	6.4
Unknown F_t	EF	99.7	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	4.6	5.1	5.0	6.2	5.9	4.6	5.4	4.8	5.7

Note: True size is 5%

Table 15: Case 2, The Size (%) of F-test When F_t Is $I(1)$ Where $\sigma_{21} = 0.8$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 89$	$10, 189$	$10, 489$	$20, 179$	$20, 379$	$20, 979$	$50, 449$	$50, 949$	$50, 2449$
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
$\sigma_{31} = 0$										
Known F_t	EF	22.7	20.6	11.7	19.8	15.1	10.3	22.6	18.4	8.4
(LSFM)	BF	11.3	11.2	7.3	12.0	11.5	8.7	19.0	16.0	8.8
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.8	99.9	99.9
(CupBC)	BF	30.4	22.9	14.5	57.0	49.8	22.6	85.8	82.9	47.7
Unknown F_t	EF	98.5	99.9	99.9	96.4	99.9	99.9	89.2	99.9	99.9
(CupFM)	BF	8.4	5.7	5.1	5.9	5.4	4.5	4.9	3.5	5.3
$\sigma_{31} = 0.2$										
Known F_t	EF	21.9	19.7	12.5	20.9	15.9	9.2	22.8	19.6	8.7
(LSFM)	BF	11.3	11.0	6.9	13.3	9.9	9.2	19.6	15.8	9.8
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.8	99.9	99.9
(CupBC)	BF	30.4	23.0	14.5	57.0	49.8	22.6	85.9	82.9	47.7
Unknown F_t	EF	98.5	99.9	99.9	96.4	99.9	99.9	89.2	99.9	99.5
(CupFM)	BF	8.4	5.9	5.1	5.9	5.4	4.5	4.9	3.5	5.2
$\sigma_{31} = 0.8$										
Known F_t	EF	23.2	18.7	11.9	21.6	16.3	9.1	23.8	19.0	9.4
(LSFM)	BF	13.2	9.7	7.5	14.2	10.8	8.0	20.1	14.7	9.8
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.8	99.9	99.9
(CupBC)	BF	31.0	23.3	14.6	56.9	49.6	23.0	85.8	82.9	47.6
Unknown F_t	EF	98.4	99.9	99.9	96.4	99.9	99.9	89.3	99.9	99.9
(CupFM)	BF	8.3	5.9	5.1	6.0	5.5	4.4	4.9	3.5	5.2

Note: True size is 5%

Table 16: Case 2, The Size-adjusted Power (%) of F-test Where $\sigma_{21} = 0$ and $\sigma_{31} = 0$ ($H_a : \lambda_i \neq 0$ for all i)

Strong dependence: $\lambda_i \sim IIDU\,(0.2, 0.5)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	86.9	99.6	100.0	94.3	99.8	100.0	97.6	100.0	100.0
(LSFM)	BF	89.0	99.6	100.0	94.4	99.8	100.0	98.6	100.0	100.0
Unknown F_t	EF	64.0	92.0	100.0	77.1	97.0	100.0	89.1	99.4	100.0
(CupBC)	BF	65.9	92.9	100.0	79.6	96.8	100.0	90.1	99.4	100.0
Unknown F_t	EF	64.0	91.9	100.0	77.2	97.0	100.0	88.8	99.4	100.0
(CupFM)	BF	65.3	92.9	100.0	79.7	96.8	100.0	90.2	99.4	100.0
Weak dependence: $\lambda_i \sim IIDU\,(0.01, 0.2)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	33.0	66.7	98.1	43.4	81.2	99.3	56.6	89.3	100.0
(LSFM)	BF	35.0	69.9	97.8	45.2	81.8	99.3	60.7	89.0	100.0
Unknown F_t	EF	13.1	31.0	75.1	15.9	46.4	84.9	29.3	62.4	96.4
(CupBC)	BF	13.3	34.6	75.5	18.6	49.4	85.6	33.6	62.6	96.3
Unknown F_t	EF	13.9	30.9	75.2	15.9	46.3	84.9	29.4	62.3	96.4
(CupFM)	BF	13.8	33.9	75.5	18.6	49.2	85.5	33.3	62.5	96.3

Note: F_t follows an $I\,(1)$ process hereafter

Table 17: Case 2, The Size-adjusted Power (%) of F-test Where $\sigma_{21} = 0$ and $\sigma_{31} = 0.2$ ($H_a : \lambda_i \neq 0$ for all i)

Strong dependence: $\lambda_i \sim IIDU\,(0.2, 0.5)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	86.1	99.3	100.0	95.0	99.8	100.0	97.2	100.0	100.0
(LSFM)	BF	88.6	99.6	100.0	95.2	99.8	100.0	98.8	100.0	100.0
Unknown F_t	EF	64.3	91.3	100.0	78.0	96.4	100.0	89.2	99.4	100.0
(CupBC)	BF	65.6	92.7	100.0	79.8	96.6	100.0	90.7	99.5	100.0
Unknown F_t	EF	63.8	91.4	100.0	78.0	96.4	100.0	89.2	99.4	100.0
(CupFM)	BF	65.3	92.5	100.0	79.7	96.6	100.0	90.8	99.5	100.0
Weak dependence: $\lambda_i \sim IIDU\,(0.01, 0.2)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	32.8	67.6	98.0	43.3	80.8	99.3	55.8	90.3	100.0
(LSFM)	BF	35.6	70.4	98.0	45.0	82.4	99.4	61.5	90.9	100.0
Unknown F_t	EF	13.4	30.6	75.7	15.5	45.9	85.6	32.0	63.6	96.3
(CupBC)	BF	14.3	33.1	75.6	19.0	48.5	85.6	34.3	64.1	96.2
Unknown F_t	EF	14.0	30.4	75.7	15.3	45.2	85.5	31.7	63.8	96.3
(CupFM)	BF	14.7	32.5	75.6	18.7	48.6	85.5	34.2	63.7	96.3

Table 18: Case 2, The Size-adjusted Power (%) of F-test Where $\sigma_{21} = 0$ and $\sigma_{31} = 0.8$ ($H_a : \lambda_i \neq 0$ for all i)

Strong dependence: $\lambda_i \sim IIDU\,(0.2, 0.5)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	85.4	99.3	100.0	92.9	99.9	100.0	97.9	100.0	100.0
(LSFM)	BF	88.3	99.4	100.0	93.5	100.0	100.0	98.9	100.0	100.0
Unknown F_t	EF	63.9	92.7	100.0	77.2	96.6	100.0	89.7	99.5	100.0
(CupBC)	BF	65.0	93.4	100.0	78.9	96.7	100.0	90.8	99.6	100.0
Unknown F_t	EF	63.9	92.7	100.0	77.3	96.5	100.0	89.6	99.5	100.0
(CupFM)	BF	64.5	93.4	100.0	78.8	96.7	100.0	90.7	99.6	100.0
Weak dependence: $\lambda_i \sim IIDU\,(0.01, 0.2)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	30.1	71.1	97.3	43.8	77.6	99.2	59.8	91.4	99.9
(LSFM)	BF	34.4	71.5	97.6	44.5	80.2	99.2	63.8	91.4	99.9
Unknown F_t	EF	13.2	29.6	75.9	15.0	46.2	87.8	32.7	65.3	95.8
(CupBC)	BF	13.7	33.4	75.8	17.7	46.2	88.1	35.0	65.5	96.0
Unknown F_t	EF	13.1	28.8	75.8	14.9	45.9	87.8	32.4	65.2	95.8
(CupFM)	BF	13.5	32.5	75.6	17.6	46.1	88.1	34.6	65.3	95.9

Table 19: Case 2, The Size-adjusted Power (%) of F-test Where $\sigma_{21} = 0.2$ and $\sigma_{31} = 0$ ($H_a : \lambda_i \neq 0$ for all i)

Strong dependence: $\lambda_i \sim IIDU\,(0.2, 0.5)$										
(n, T)		$(10,10)$	$(10,20)$	$(10,50)$	$(20,10)$	$(20,20)$	$(20,50)$	$(50,10)$	$(50,20)$	$(50,50)$
Known F_t	EF	85.5	99.6	100.0	94.2	99.8	100.0	97.7	100.0	100.0
(LSFM)	BF	89.1	99.7	100.0	94.9	99.8	100.0	98.2	100.0	100.0
Unknown F_t	EF	63.6	92.1	100.0	77.4	96.9	100.0	89.0	99.4	100.0
(CupBC)	BF	66.4	93.4	100.0	80.0	96.7	100.0	90.2	99.4	100.0
Unknown F_t	EF	63.2	92.2	100.0	76.8	96.8	100.0	88.9	99.4	100.0
(CupFM)	BF	65.6	93.1	100.0	79.4	96.6	100.0	90.0	99.4	100.0
Weak dependence: $\lambda_i \sim IIDU\,(0.01, 0.2)$										
(n, T)		$(10,10)$	$(10,20)$	$(10,50)$	$(20,10)$	$(20,20)$	$(20,50)$	$(50,10)$	$(50,20)$	$(50,50)$
Known F_t	EF	30.6	64.5	97.8	43.2	81.5	99.2	56.5	89.1	100.0
(LSFM)	BF	34.4	70.2	97.9	45.9	82.3	99.3	59.4	89.7	100.0
Unknown F_t	EF	13.1	32.2	75.3	16.0	47.2	85.4	30.6	62.2	96.6
(CupBC)	BF	14.3	34.2	75.6	19.8	50.5	85.9	35.4	64.1	96.7
Unknown F_t	EF	12.7	31.2	75.0	15.3	47.0	85.5	30.0	61.8	96.4
(CupFM)	BF	13.9	32.7	75.3	18.4	48.9	85.3	33.1	63.0	96.5

Table 20: Case 2, The Size-adjusted Power (%) of F-test Where $\sigma_{21} = 0.2$ and $\sigma_{31} = 0.2$ ($H_a : \lambda_i \neq 0$ for all i)

Strong dependence: $\lambda_i \sim IIDU\,(0.2, 0.5)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	84.5	99.3	100.0	94.5	99.8	100.0	97.3	100.0	100.0
(LSFM)	BF	87.8	99.6	100.0	95.0	99.7	100.0	98.4	100.0	100.0
Unknown F_t	EF	64.4	91.6	99.9	77.8	96.5	100.0	89.2	99.4	100.0
(CupBC)	BF	65.5	93.0	100.0	79.9	96.7	100.0	90.6	99.4	100.0
Unknown F_t	EF	63.4	91.6	99.9	77.5	96.5	100.0	89.1	99.3	100.0
(CupFM)	BF	64.8	92.8	100.0	79.3	96.4	100.0	90.2	99.4	100.0
Weak dependence: $\lambda_i \sim IIDU\,(0.01, 0.2)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	31.0	67.3	98.0	43.9	81.0	99.4	55.9	89.8	100.0
(LSFM)	BF	35.7	70.8	98.1	45.6	82.1	99.5	60.5	90.8	100.0
Unknown F_t	EF	12.9	31.2	75.5	16.4	44.8	85.5	32.8	63.8	96.0
(CupBC)	BF	14.2	33.6	75.9	19.8	48.9	86.0	36.2	64.5	96.3
Unknown F_t	EF	12.6	30.7	75.1	15.6	45.2	85.7	32.2	63.4	96.1
(CupFM)	BF	14.3	32.5	75.6	18.4	48.0	85.5	33.5	63.7	96.0

Table 21: Case 2, The Size-adjusted Power (%) of F-test Where $\sigma_{21} = 0.2$ and $\sigma_{31} = 0.8$ ($H_a : \lambda_i \neq 0$ for all i)

Strong dependence: $\lambda_i \sim IIDU\,(0.2, 0.5)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	85.7	99.3	100.0	92.9	99.7	100.0	98.9	100.0	100.0
(LSFM)	BF	88.5	99.7	100.0	92.9	99.9	100.0	98.9	100.0	100.0
Unknown F_t	EF	62.4	91.7	100.0	75.7	96.6	100.0	90.8	99.3	100.0
(CupBC)	BF	65.1	92.7	100.0	78.3	97.0	100.0	92.2	99.3	100.0
Unknown F_t	EF	61.4	91.5	100.0	75.3	96.4	100.0	90.7	99.3	100.0
(CupFM)	BF	64.7	92.5	100.0	78.0	96.8	100.0	91.5	99.3	100.0
Weak dependence: $\lambda_i \sim IIDU\,(0.01, 0.2)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	31.8	70.1	97.5	42.8	77.4	99.2	60.4	90.1	100.0
(LSFM)	BF	34.7	70.7	97.8	44.5	80.2	99.2	62.2	90.6	100.0
Unknown F_t	EF	12.7	31.5	76.4	15.9	46.0	87.1	33.2	64.0	95.8
(CupBC)	BF	14.6	34.4	76.5	18.4	48.4	87.4	36.4	65.4	96.4
Unknown F_t	EF	12.2	30.6	76.4	15.3	46.0	87.0	33.3	63.0	95.8
(CupFM)	BF	14.1	33.8	75.9	17.3	47.0	86.8	35.1	63.9	96.3

Table 22: Case 2, The Size-adjusted Power (%) of F-test Where $\sigma_{21} = 0.8$ and $\sigma_{31} = 0$ ($H_a : \lambda_i \neq 0$ for all i)

Strong dependence: $\lambda_i \sim IIDU\,(0.2, 0.5)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	82.4	99.2	100.0	91.1	99.9	100.0	94.4	99.9	100.0
(LSFM)	BF	88.0	99.6	100.0	94.9	99.9	100.0	98.5	99.9	100.0
Unknown F_t	EF	59.3	90.5	99.9	68.3	94.9	100.0	78.0	97.8	100.0
(CupBC)	BF	77.3	94.5	100.0	89.4	99.2	100.0	98.3	100.0	100.0
Unknown F_t	EF	61.2	91.2	99.9	76.3	97.1	100.0	88.1	99.4	100.0
(CupFM)	BF	67.5	92.1	99.9	78.0	97.4	100.0	87.5	99.2	100.0
Weak dependence: $\lambda_i \sim IIDU\,(0.01, 0.2)$										
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
Known F_t	EF	21.7	59.5	97.9	26.0	76.0	99.4	31.5	85.1	99.9
(LSFM)	BF	32.8	71.2	97.8	39.2	82.6	99.6	53.6	92.7	100.0
Unknown F_t	EF	11.4	32.6	74.2	14.7	45.2	85.5	16.1	51.2	92.5
(CupBC)	BF	41.6	52.4	81.9	64.4	78.0	91.4	86.0	94.0	99.0
Unknown F_t	EF	12.5	29.4	72.9	19.5	46.4	83.1	33.2	64.8	96.1
(CupFM)	BF	17.4	33.1	74.5	22.3	47.1	83.9	32.2	62.8	96.2

Table 23: Case 2, The Size-adjusted Power (%) of F-test Where $\sigma_{21} = 0.8$ and $\sigma_{31} = 0.2$ ($H_a : \lambda_i \neq 0$ for all i)

Strong dependence: $\lambda_i \sim IIDU\,(0.2, 0.5)$										
(n, T)		$(10,10)$	$(10,20)$	$(10,50)$	$(20,10)$	$(20,20)$	$(20,50)$	$(50,10)$	$(50,20)$	$(50,50)$
Known F_t	EF	80.7	98.8	100.0	90.9	99.7	100.0	93.5	99.9	100.0
(LSFM)	BF	88.8	99.3	100.0	95.5	99.7	100.0	99.0	100.0	100.0
Unknown F_t	EF	58.0	90.8	99.8	67.7	94.4	100.0	77.0	97.9	100.0
(CupBC)	BF	76.3	95.6	100.0	89.8	98.3	100.0	97.9	100.0	100.0
Unknown F_t	EF	60.4	91.0	99.9	75.8	96.3	100.0	87.7	99.1	100.0
(CupFM)	BF	65.4	92.1	99.9	77.6	96.4	100.0	87.4	99.1	100.0
Weak dependence: $\lambda_i \sim IIDU\,(0.01, 0.2)$										
(n, T)		$(10,10)$	$(10,20)$	$(10,50)$	$(20,10)$	$(20,20)$	$(20,50)$	$(50,10)$	$(50,20)$	$(50,50)$
Known F_t	EF	20.6	57.8	97.6	27.3	76.0	99.5	33.4	85.1	100.0
(LSFM)	BF	32.7	70.7	98.0	41.4	85.3	99.6	55.7	93.3	100.0
Unknown F_t	EF	11.9	33.3	75.2	14.4	44.9	86.5	16.3	51.1	91.9
(CupBC)	BF	42.8	53.5	82.4	64.0	78.3	92.1	86.2	94.6	98.7
Unknown F_t	EF	12.6	29.7	72.8	19.4	47.1	84.9	32.7	62.8	95.6
(CupFM)	BF	18.2	33.2	74.2	21.9	47.4	84.9	32.0	60.9	95.6

Table 24: Case 2, The Size-adjusted Power (%) of F-test Where $\sigma_{21} = 0.8$ and $\sigma_{31} = 0.8$ ($H_a : \lambda_i \neq 0$ for all i)

Strong dependence: $\lambda_i \sim IIDU\,(0.2, 0.5)$										
(n, T)		$(10,10)$	$(10,20)$	$(10,50)$	$(20,10)$	$(20,20)$	$(20,50)$	$(50,10)$	$(50,20)$	$(50,50)$
Known F_t	EF	79.0	99.1	100.0	88.6	99.5	100.0	93.6	100.0	100.0
(LSFM)	BF	87.9	99.8	100.0	94.2	99.9	100.0	98.5	100.0	100.0
Unknown F_t	EF	55.3	89.3	99.9	66.7	94.2	100.0	76.0	97.7	100.0
(CupBC)	BF	76.2	94.8	99.9	89.1	98.5	100.0	98.9	99.9	100.0
Unknown F_t	EF	59.0	90.6	99.9	75.5	96.2	100.0	88.2	99.7	100.0
(CupFM)	BF	65.8	91.4	99.9	77.4	96.3	100.0	87.9	99.6	100.0
Weak dependence: $\lambda_i \sim IIDU\,(0.01, 0.2)$										
(n, T)		$(10,10)$	$(10,20)$	$(10,50)$	$(20,10)$	$(20,20)$	$(20,50)$	$(50,10)$	$(50,20)$	$(50,50)$
Known F_t	EF	19.7	60.9	97.7	26.2	70.7	98.9	28.4	82.5	100.0
(LSFM)	BF	33.9	70.1	98.0	41.9	80.8	99.1	53.1	92.6	100.0
Unknown F_t	EF	10.9	34.9	74.6	14.1	43.8	87.4	17.4	51.1	91.6
(CupBC)	BF	40.3	54.8	81.1	62.9	76.2	93.1	86.5	95.6	98.4
Unknown F_t	EF	12.5	29.8	73.0	18.4	44.2	86.8	32.5	62.6	95.5
(CupFM)	BF	18.6	33.4	73.7	21.0	44.4	86.1	31.1	61.4	95.3

Table 25: *Case 2, The Size (%) of F-test under Heteroskedasticity Where $\sigma_{21} = 0$ ($H_0 : \lambda_i = 0$ for all i)*

$D.F.(num, den)$		$10, 89$	$10, 189$	$10, 489$	$20, 179$	$20, 379$	$20, 979$	$50, 449$	$50, 949$	$50, 2449$
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
$\sigma_{31} = 0$										
Known F_t	EF	18.3	14.8	14.2	17.6	17.2	18.9	17.2	17.9	17.7
(LSFM)	BF	5.5	5.7	4.6	7.0	5.4	5.4	4.8	4.1	4.5
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	7.0	6.4	5.8	4.8	6.9	5.7	5.9	5.0	5.9
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	6.8	6.4	5.8	4.6	6.9	5.7	5.9	5.0	5.9
$\sigma_{31} = 0.2$										
Known F_t	EF	17.0	16.4	15.8	17.6	17.0	18.4	18.3	18.7	18.8
(LSFM)	BF	6.1	5.5	5.9	6.4	6.2	6.8	4.6	4.8	5.0
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	7.0	6.4	5.8	4.8	6.9	5.7	5.9	5.0	5.9
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	6.8	6.4	5.8	4.6	6.9	5.7	5.9	5.0	5.9
$\sigma_{31} = 0.8$										
Known F_t	EF	14.9	16.1	15.7	16.5	16.6	17.6	18.5	20.8	19.0
(LSFM)	BF	6.0	5.9	5.7	6.4	4.6	5.8	5.4	5.4	4.9
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	7.0	6.4	5.8	4.8	6.9	5.7	5.9	5.0	5.9
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	6.8	6.4	5.8	4.6	6.9	5.7	5.9	5.0	5.9

Table 26: *Case 2, The Size (%) of F-test under Heteroskedasticity Where $\sigma_{21} = 0.2$ ($H_0 : \lambda_i = 0$ for all i)*

$D.F.(num, den)$		$10, 89$	$10, 189$	$10, 489$	$20, 179$	$20, 379$	$20, 979$	$50, 449$	$50, 949$	$50, 2449$
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
$\sigma_{31} = 0$										
Known F_t	EF	28.8	28.6	23.8	24.8	25.1	20.3	21.6	22.8	16.3
(LSFM)	BF	5.1	6.1	5.4	6.5	5.2	5.0	6.3	5.8	4.5
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	7.2	7.4	6.4	6.9	8.7	7.2	13.2	9.3	8.7
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	6.4	6.3	5.3	4.0	6.8	6.0	4.8	4.5	6.1
$\sigma_{31} = 0.2$										
Known F_t	EF	27.0	26.7	25.3	25.5	23.4	22.1	20.6	22.2	17.3
(LSFM)	BF	6.0	6.6	6.4	6.4	6.3	6.2	5.8	5.7	4.7
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	7.2	7.4	6.4	6.9	8.8	7.2	13.2	9.3	8.7
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	6.4	6.3	5.5	4.0	6.7	6.0	4.8	4.5	6.1
$\sigma_{31} = 0.8$										
Known F_t	EF	25.0	28.6	24.2	26.7	22.8	23.6	21.8	19.5	21.6
(LSFM)	BF	11.1	7.4	6.3	6.5	8.4	6.9	6.1	5.4	7.6
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	7.3	7.3	6.4	7.0	8.8	7.2	13.3	9.3	8.6
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	6.5	6.3	5.4	3.9	6.7	6.0	4.8	4.5	6.1

Table 27: Case 2, The Size (%) of F-test under Heteroskedasticity Where $\sigma_{21} = 0.8$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 89$	$10, 189$	$10, 489$	$20, 179$	$20, 379$	$20, 979$	$50, 449$	$50, 949$	$50, 2449$
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
$\sigma_{31} = 0$										
Known F_t	EF	47.2	42.2	31.5	44.4	41.1	27.4	43.4	37.9	24.0
(LSFM)	BF	7.8	7.5	6.3	8.8	9.8	4.1	12.2	10.2	6.2
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	14.1	12.4	8.8	27.0	24.2	14.1	69.7	54.1	26.5
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	98.9	99.9	99.9
(CupFM)	BF	9.7	7.2	5.6	7.0	6.7	5.5	4.7	3.4	5.9
$\sigma_{31} = 0.2$										
Known F_t	EF	46.6	42.3	34.1	47.5	36.4	29.2	42.5	40.0	24.7
(LSFM)	BF	9.4	9.6	6.3	11.1	8.3	7.3	9.5	10.7	7.2
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	13.9	12.4	8.7	26.9	24.1	14.1	69.7	54.1	26.5
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	98.9	99.9	99.9
(CupFM)	BF	9.8	7.2	5.7	7.0	6.7	5.5	4.7	3.4	5.9
$\sigma_{31} = 0.8$										
Known F_t	EF	41.7	41.9	32.5	45.6	31.6	28.9	46.2	39.9	25.9
(LSFM)	BF	17.2	10.6	5.9	11.2	14.9	8.0	13.9	12.2	10.4
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	13.9	12.6	8.7	26.9	23.9	14.1	69.7	54.0	26.6
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	98.8	99.9	99.9
(CupFM)	BF	9.9	6.9	5.7	6.9	6.7	5.5	4.6	3.4	5.9

Table 28: Case 2, The Size (%) of F-test under Serial Correlation Where $\sigma_{21} = 0$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 89$	$10, 189$	$10, 489$	$20, 179$	$20, 379$	$20, 979$	$50, 449$	$50, 949$	$50, 2449$
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
$\sigma_{31} = 0$										
Known F_t	EF	44.6	53.3	58.7	64.3	72.7	81.9	84.8	93.9	98.5
(LSFM)	BF	45.6	54.3	60.6	67.9	76.0	83.6	87.1	94.7	98.8
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	32.4	31.7	32.1	59.5	61.1	62.1	97.5	97.9	98.6
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	31.4	30.4	32.0	58.7	60.8	61.7	97.0	97.8	98.6
$\sigma_{31} = 0.2$										
Known F_t	EF	45.7	54.3	59.2	63.1	73.7	83.0	84.5	94.4	98.2
(LSFM)	BF	46.8	56.4	60.9	67.3	76.4	84.5	87.4	94.9	98.4
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	32.4	31.6	32.1	59.4	61.1	62.1	97.5	97.9	98.6
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	31.3	30.6	32.0	58.7	60.8	61.7	97.0	97.8	98.6
$\sigma_{31} = 0.8$										
Known F_t	EF	44.1	52.4	57.4	62.6	74.5	81.8	86.6	94.4	98.2
(LSFM)	BF	44.8	55.4	59.5	65.9	77.8	83.2	88.0	95.7	98.6
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	32.3	31.3	32.1	59.6	61.1	62.2	97.5	97.9	98.6
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	31.4	30.2	31.8	58.6	60.8	61.9	97.0	97.8	98.6

Table 29: Case 2, The Size (%) of F-test under Serial Correlation Where $\sigma_{21} = 0.2$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 89$	$10, 189$	$10, 489$	$20, 179$	$20, 379$	$20, 979$	$50, 449$	$50, 949$	$50, 2449$
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
$\sigma_{31} = 0$										
Known F_t	EF	44.5	55.6	62.2	64.9	76.0	83.9	85.0	95.2	98.9
(LSFM)	BF	45.7	56.0	63.8	68.8	78.7	84.9	88.2	95.5	98.8
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	33.3	32.9	32.9	61.3	64.3	64.3	97.3	99.2	98.5
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	32.5	31.2	31.6	60.2	61.6	62.3	96.6	98.6	98.2
$\sigma_{31} = 0.2$										
Known F_t	EF	44.1	55.8	63.0	65.1	75.7	84.7	84.9	94.3	98.6
(LSFM)	BF	46.1	57.5	64.9	68.0	78.7	85.4	88.3	95.2	98.6
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	33.3	32.9	32.9	61.3	64.3	64.3	97.3	99.2	98.5
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	32.4	31.1	31.6	60.1	61.7	62.3	96.5	98.6	98.2
$\sigma_{31} = 0.8$										
Known F_t	EF	42.5	53.6	61.7	63.2	75.7	83.3	87.5	95.0	98.8
(LSFM)	BF	44.6	55.6	63.4	65.7	77.4	85.6	89.5	96.7	98.8
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	33.3	32.9	32.8	61.4	64.4	64.2	97.3	99.2	98.5
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	32.0	31.4	31.7	60.0	61.7	62.1	96.6	98.6	98.2

Table 30: Case 2, The Size (%) of F-test under Serial Correlation Where $\sigma_{21} = 0.8$ ($H_0 : \lambda_i = 0$ for all i)

$D.F.(num, den)$		$10, 89$	$10, 189$	$10, 489$	$20, 179$	$20, 379$	$20, 979$	$50, 449$	$50, 949$	$50, 2449$
(n, T)		$(10, 10)$	$(10, 20)$	$(10, 50)$	$(20, 10)$	$(20, 20)$	$(20, 50)$	$(50, 10)$	$(50, 20)$	$(50, 50)$
$\sigma_{31} = 0$										
Known F_t	EF	51.0	68.1	79.8	64.0	85.3	94.0	82.9	97.2	99.7
(LSFM)	BF	51.4	71.3	82.0	65.4	85.9	94.7	84.0	97.4	99.8
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	77.2	80.0	66.4	97.0	98.8	95.8	100.0	100.0	100.0
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	54.2	51.9	41.7	78.8	80.9	72.1	99.1	99.2	99.4
$\sigma_{31} = 0.2$										
Known F_t	EF	50.5	70.2	80.2	63.7	84.1	94.8	81.8	97.4	99.7
(LSFM)	BF	51.0	71.1	80.1	64.7	86.0	94.5	82.4	97.3	99.8
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	77.0	80.3	66.3	97.0	98.8	95.9	100.0	100.0	100.0
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	54.5	52.0	41.6	78.8	80.8	72.3	99.1	99.2	99.4
$\sigma_{31} = 0.8$										
Known F_t	EF	49.7	67.0	79.8	64.7	84.6	93.9	79.0	96.6	99.4
(LSFM)	BF	50.1	69.2	81.2	65.6	85.9	94.7	81.2	97.5	99.5
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupBC)	BF	76.7	80.4	66.4	97.1	98.8	96.0	100.0	100.0	100.0
Unknown F_t	EF	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
(CupFM)	BF	54.7	51.7	41.2	79.0	80.9	72.0	99.1	99.2	99.4

Appendix: Proofs of Lemmas, Theorems, and Propositions

This appendix includes proofs for the main results in the text.

A Proof of Lemma 1

Proof. It is straightforward to prove part 1 with the given assumptions, so omitted here. For part 2, one can find the complete proof in Bai (2003). ■

B Proof of Theorem 1

We start from the lemma below. In this lemma, we check the consistency of the F-statistic when F_t is observable. First, note that given our assumptions, we have the following results using central limit theorem (CLT).

For each t, as $n \to \infty$,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \lambda_i u_{it} \xrightarrow{d} N(0, \sigma^2 \phi_\lambda).$$

For each i, as $T \to \infty$,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_t u_{it} \xrightarrow{d} N(0, \sigma^2 \phi_F).$$

Lemma 1 (B) *Assume* $(n, T) \overset{\text{seq}}{\to} \infty$*. If* F_t *is observable, then*

$$F_\lambda = \frac{RRSS - URSS}{URSS} \frac{(nT - n)}{n} \xrightarrow{p} 1.$$

Proof. Because λ_i can be estimated using least squares, we have

$$\begin{aligned}
\widetilde{\lambda}_i &= \frac{\sum_{t=1}^T F_t u_{it}}{\sum_{t=1}^T F_t^2}, \\
RRSS &= \sum_{i=1}^n \sum_{t=1}^T y_{it}^2,
\end{aligned}$$

and

$$URSS = \sum_{i=1}^n \sum_{t=1}^T \left(y_{it} - \widetilde{\lambda}_i F_t \right)^2.$$

Then,

$$F_\lambda = \frac{(RRSS - URSS)/n}{URSS/(nT-n)}$$

can be readily obtained.

1. First, we consider the denominator.

$$\begin{aligned}
\frac{URSS}{(nT-n)} &= \frac{1}{n(T-1)} \sum_{i=1}^n \sum_{t=1}^T \left(y_{it} - \widetilde{\lambda}_i F_t \right)^2 \\
&= \frac{1}{n(T-1)} \sum_{i=1}^n \sum_{t=1}^T \left[u_{it} - \left(\widetilde{\lambda}_i - \lambda_i \right) F_t \right]^2 \\
&= \frac{1}{n(T-1)} \sum_{i=1}^n \sum_{t=1}^T \left[u_{it}^2 + \left(\widetilde{\lambda}_i - \lambda_i \right)^2 F_t^2 - 2 \left(\widetilde{\lambda}_i - \lambda_i \right) u_{it} F_t \right] \\
&= \frac{1}{n(T-1)} \sum_{i=1}^n \sum_{t=1}^T u_{it}^2 + \frac{1}{n(T-1)} \sum_{i=1}^n \left(\widetilde{\lambda}_i - \lambda_i \right)^2 \sum_{t=1}^T F_t^2 \\
&\quad - \frac{2}{n(T-1)} \sum_{i=1}^n \left(\widetilde{\lambda}_i - \lambda_i \right) \sum_{t=1}^T u_{it} F_t \\
&= I + II + III
\end{aligned}$$

Note that $\frac{1}{\sqrt{T}} \sum_{t=1}^T u_{it} F_t = O_p(1)$ by a CLT since there is no correlation between u_{it} and F_t.

Consider I. It is easy to see that

$$\begin{aligned} I &= \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} u_{it}^2 \approx \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} u_{it}^2 \\ &\xrightarrow{p} \sigma^2 \end{aligned}$$

as $(n, T) \overset{\text{seq}}{\rightarrow} \infty$.

For II and III, one can show that

$$\begin{aligned} II &= \frac{1}{n(T-1)} \sum_{i=1}^{n} \left(\widetilde{\lambda}_i - \lambda_i\right)^2 \sum_{t=1}^{T} F_t^2 \\ &\approx \frac{1}{T} \frac{1}{n} \sum_{i=1}^{n} \left\{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right\}^2 \frac{1}{T} \sum_{t=1}^{T} F_t^2 = O_p\left(\frac{1}{T}\right) = o_p(1) \end{aligned}$$

and

$$\begin{aligned} III &= \frac{2}{n(T-1)} \sum_{i=1}^{n} \left(\widetilde{\lambda}_i - \lambda_i\right) \sum_{t=1}^{T} u_{it} F_t \\ &\approx \frac{1}{T} \frac{2}{n} \sum_{i=1}^{n} \left\{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right\} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_{it} F_t = O_p\left(\frac{1}{T}\right) = o_p(1) \end{aligned}$$

using

$$\frac{1}{n} \sum_{i=1}^{n} \left\{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right\}^2 = O_p(1),$$

$$\frac{1}{T} \sum_{t=1}^{T} F_t^2 = O_p(1),$$

and

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_{it} F_t = O_p(1)$$

provided u_{it} and F_t are uncorrelated. Hence, for the denominator we conclude

$$\frac{URSS}{(nT-n)} \xrightarrow{p} \sigma^2$$

as $(n,T) \overset{\text{seq}}{\to} \infty$.

2. Next, we turn to the numerator.

$$\begin{aligned}
\frac{(RRSS - URSS)}{n} &= \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T} y_{it}^2 - \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it} - \widetilde{\lambda}_i F_t\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[u_{it}^2 - u_{it}^2 - \left(\widetilde{\lambda}_i - \lambda_i\right)^2 F_t^2 + 2\left(\widetilde{\lambda}_i - \lambda_i\right) u_{it}F_t\right] \\
&= -\frac{1}{n}\sum_{i=1}^{n}\left\{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right\}^2 \frac{1}{T}\sum_{t=1}^{T} F_t^2 \\
&\quad + \frac{2}{n}\sum_{i=1}^{n}\left\{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right\} \frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_{it}F_t \\
&= I + II.
\end{aligned}$$

Note that in the constrained regression, we have $y_{it} = \lambda_i F_t + u_{it} = u_{it}$ with the restriction $\lambda_i = 0$ for all i.

Consider I first. For a fixed n, we have

$$\begin{aligned}
I &= -\frac{1}{n}\sum_{i=1}^{n}\left\{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right\}^2 \frac{1}{T}\sum_{t=1}^{T} F_t^2 \\
&\overset{p}{\to} -\frac{1}{n}\sum_{i=1}^{n} Z_i^2 \phi_F
\end{aligned}$$

where $\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right) \overset{d}{\to} Z_i \sim N(0, \sigma^2\phi_F^{-1})$ and $\frac{1}{T}\sum_{t=1}^{T} F_t^2 \overset{p}{\to} \phi_F$ as $T \to \infty$. As a result, we have $\frac{1}{n}\sum_{i=1}^{n} Z_i^2\phi_F \to \phi_F E(Z_i^2) = \phi_F \sigma^2 \phi_F^{-1} = \sigma^2$ as $n \to \infty$.

Hence, one concludes that

$$I \overset{p}{\to} -\sigma^2$$

as $(n,T) \overset{\text{seq}}{\to} \infty$.

For II, it can be shown that

$$\begin{aligned} II &= \frac{2}{n}\sum_{i=1}^{n}\left\{\sqrt{T}\left(\widetilde{\lambda}_i-\lambda_i\right)\right\}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_{it}F_t \\ &\overset{p}{\to} 2\sigma^2 \end{aligned}$$

as $(n,T)\overset{\text{seq}}{\to}\infty$ using $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_{it}F_t \overset{d}{\to} N\left(0,\sigma^2\phi_F\right)$ as $T\to\infty$.

Combining the results, we obtain that

$$\frac{(RRSS-URSS)}{n}\overset{p}{\to}\sigma^2$$

as $(n,T)\overset{\text{seq}}{\to}\infty$.

3. Finally, we conclude that

$$F_\lambda = \frac{RRSS-URSS}{URSS}\frac{(nT-n)}{n}\overset{p}{\to}1$$

which implies that the F-statistic gets centered at 1 as $(n,T)\overset{\text{seq}}{\to}\infty$.

■

Given the above results, for the proof of Theorem 1 we write

$$F_\lambda = \frac{R_\lambda}{\widehat{\sigma}^2}$$

where $R_\lambda = \frac{(RRSS-URSS)}{n}$ and $\widehat{\sigma}^2 = \frac{URSS}{(nT-n)}$ using a set up which is similar to Orme and Yamagata (2006). Rearranging the terms, we have

$$\sqrt{n}\left(F_\lambda-1\right) = \frac{1}{\widehat{\sigma}^2}\sqrt{n}\left(R_\lambda-\widehat{\sigma}^2\right).$$

Proof. Expanding the equations, we have

$$\begin{aligned}
R_\lambda - \widehat{\sigma}^2 &= \frac{(RRSS - URSS)}{n} - \frac{URSS}{(nT-n)} \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[-\left(\widetilde{\lambda}_i - \lambda_i\right)^2 F_t^2 + 2\left(\widetilde{\lambda}_i - \lambda_i\right) u_{it}F_t\right] \\
&\quad - \frac{1}{n\left(T-1\right)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[u_{it}^2 + \left(\widetilde{\lambda}_i - \lambda_i\right)^2 F_t^2 - 2\left(\widetilde{\lambda}_i - \lambda_i\right) u_{it}F_t\right] \\
&= -\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right\}^2 F_t^2 \\
&\quad + \frac{2}{n\sqrt{T}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right\} u_{it}F_t \\
&\quad - \frac{1}{n\left(T-1\right)}\sum_{i=1}^{n}\sum_{t=1}^{T} u_{it}^2 - \frac{1}{nT(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right\}^2 F_t^2 \\
&\quad + \frac{2}{n\sqrt{T}\left(T-1\right)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right\} u_{it}F_t \\
&= I + II + III + IV + V.
\end{aligned}$$

Consider I. It can be shown that

$$\begin{aligned}
I &= -\frac{1}{n}\sum_{i=1}^{n}\left[\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right]^2\left[\frac{1}{T}\sum_{t=1}^{T}F_t^2\right] \\
&= -\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_{it}F_t\right]^2\left[\frac{1}{T}\sum_{t=1}^{T}F_t^2\right]^{-1} = O_p\left(1\right).
\end{aligned}$$

For II,

$$\begin{aligned}
II &= \frac{2}{n\sqrt{T}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right\} u_{it}F_t \\
&= \frac{2}{n}\sum_{i=1}^{n}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_{it}F_t\right]^2\left[\frac{1}{T}\sum_{t=1}^{T}F_t^2\right]^{-1} = O_p\left(1\right).
\end{aligned}$$

For III,

$$III = -\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T-1}\sum_{t=1}^{T}u_{it}^{2} = O_p\left(1\right).$$

For IV and V, as already shown above,

$$IV \approx -\frac{1}{n}\sum_{i=1}^{n}\left\{\sqrt{T}\left(\widetilde{\lambda}_i-\lambda_i\right)\right\}^2\frac{1}{T^2}\sum_{t=1}^{T}F_t^2 = O_p\left(\frac{1}{T}\right) = o_p\left(1\right)$$

and

$$V \approx \frac{2}{n}\sum_{i=1}^{n}\left\{\sqrt{T}\left(\widetilde{\lambda}_i-\lambda_i\right)\right\}\frac{1}{T^{3/2}}\sum_{t=1}^{T}u_{it}F_t = O_p\left(\frac{1}{T}\right) = o_p\left(1\right).$$

After rearranging all the terms, one has

$$\begin{aligned} R_\lambda - \widehat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_{it}F_t\right]^2\left[\frac{1}{T}\sum_{t=1}^{T}F_t^2\right]^{-1} \\ &\quad -\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T-1}\sum_{t=1}^{T}u_{it}^2 + o_p\left(1\right) \end{aligned}$$

and accordingly

$$\begin{aligned} \sqrt{n}\left(R_\lambda - \widehat{\sigma}^2\right) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_{it}F_t\right]^2\left[\frac{1}{T}\sum_{t=1}^{T}F_t^2\right]^{-1} \\ &\quad -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{T-1}\sum_{t=1}^{T}u_{it}^2 + o_p\left(1\right). \end{aligned}$$

To apply the CLT, we recall that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_{it}F_t \xrightarrow{d} W_i \sim N(0,\sigma^2\phi_F)$$

for all i as $T\rightarrow\infty$ and establish the standard normal random variables such that

$$\frac{W_i}{\sqrt{\sigma^2\phi_F}} \sim N(0,1).$$

Furthermore, one can construct the random variables, $\frac{W_i^2}{\sigma^2 \phi_F}$, such that

$$E\left(\frac{W_i^2}{\sigma^2 \phi_F}\right) = 1$$

and

$$Var\left(\frac{W_i^2}{\sigma^2 \phi_F}\right) = 2$$

because $\frac{W_i^2}{\sigma^2 \phi_F}$ follows a chi-squared distribution. Rewriting above,

$$E\left(W_i^2\right) = \sigma^2 \phi_F$$

and

$$Var\left(W_i^2\right) = 2\sigma^4 \phi_F^2.$$

Hence, we have

$$\begin{aligned}\sqrt{n}\left(R_\lambda - \widehat{\sigma}^2\right) &= \frac{1}{\sqrt{n}}\left[\sum_{i=1}^{n} W_i^2 \phi_F^{-1} - \sum_{i=1}^{n} \sigma^2\right] + o_p(1) \\ &\xrightarrow{d} N(0, 2\sigma^4)\end{aligned}$$

as $n \to \infty$. This is because $E\left(W_i^2 \phi_F^{-1} - \sigma^2\right) = \sigma^2 - \sigma^2 = 0$ and $Var\left(W_i^2 \phi_F^{-1} - \sigma^2\right) = 2\sigma^4$.

Finally, we obtain

$$\sqrt{n}\left(F_\lambda - 1\right) = \frac{1}{\widehat{\sigma}^2}\sqrt{n}\left(R_\lambda - \widehat{\sigma}^2\right) \xrightarrow{d} N(0, 2)$$

as $(n, T) \xrightarrow{\text{seq}} \infty$ using $\widehat{\sigma}^2 \xrightarrow{p} \sigma^2$. ∎

C Proof of Theorem 2

Next let us assume that F_t is unknown. Again, we first look at the consistency of the F-statistic in the lemma below. Note that in this case we cannot simply use least squares estimation and need to use the method of PCA.

Lemma 2 (B) *Assume $(n, T) \to \infty$ and F_t is not observable.*

1. *If $\frac{T}{n} \to 0$, then*

$$F_\lambda = \frac{RRSS - URSS}{URSS} \frac{(nT - n)}{n} \xrightarrow{p} 1 + \frac{F_t^2 - \phi_F}{\phi_F}.$$

2. *If $\frac{n}{T} \to 0$,*

Not feasible.

Proof. We check two specific cases separately, i.e., $\frac{T}{n} \to 0$ and $\frac{n}{T} \to 0$, following Bai (2003).

1. Assume $\frac{T}{n} \to 0$.

   Consider the denominator. We have

$$\begin{aligned}
\frac{URSS}{(nT - n)} &= \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it} - \widehat{\lambda}_i \widehat{F}_t\right)^2 \\
&= \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[u_{it} - \left(\widehat{\lambda}_i \widehat{F}_t - \lambda_i F_t\right)\right]^2 \\
&= \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[u_{it}^2 + \left(\widehat{\lambda}_i \widehat{F}_t - \lambda_i F_t\right)^2 - 2\left(\widehat{\lambda}_i \widehat{F}_t - \lambda_i F_t\right) u_{it}\right] \\
&= \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} u_{it}^2 + \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(\widehat{\lambda}_i \widehat{F}_t - \lambda_i F_t\right)^2 \\
&\quad - \frac{2}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(\widehat{\lambda}_i \widehat{F}_t - \lambda_i F_t\right) u_{it} \\
&= I + II + III.
\end{aligned}$$

Firstly, note that if $\frac{T}{n} \to 0$, then $\sqrt{T}\left(\widehat{\lambda}_i\widehat{F}_t - \lambda_i F_t\right) \xrightarrow{d} N\left(0, W_{it}\right)$ as $(n, T) \to \infty$ where $W_{it} = \frac{F_t^2}{\phi_F^2}\sigma^2\phi_F = \frac{F_t^2}{\phi_F}\sigma^2$ by, e.g., Bai (2003).

Consider I. One can easily verify that

$$I = \frac{1}{n\left(T-1\right)}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it}^2 \approx \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it}^2 \xrightarrow{p} \sigma^2$$

as $(n, T) \to \infty$.

For II and III,

$$\begin{aligned} II &= \frac{1}{n\left(T-1\right)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\widehat{\lambda}_i\widehat{F}_t - \lambda_i F_t\right)^2 \\ &\approx \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i\widehat{F}_t - \lambda_i F_t\right)\right\}^2 = O_p\left(\frac{1}{T}\right) \end{aligned}$$

and

$$\begin{aligned} III &= \frac{2}{n\left(T-1\right)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\widehat{\lambda}_i\widehat{F}_t - \lambda_i F_t\right)u_{it} \\ &\approx \frac{2}{nT^{3/2}}\sum_{i=1}^{n}\sum_{t=1}^{T}\sqrt{T}\left(\widehat{\lambda}_i\widehat{F}_t - \lambda_i F_t\right)u_{it} \\ &= \frac{2}{nT^{3/2}}\sum_{i=1}^{n}\sum_{t=1}^{T}Q_{it}u_{it} = O_p\left(\frac{1}{\sqrt{T}}\right) \end{aligned}$$

where $Q_{it} = \sqrt{T}\left(\widehat{\lambda}_i\widehat{F}_t - \lambda_i F_t\right)$. Note that

$$\begin{aligned} Q_{it} &= \sqrt{T}\left(\widehat{\lambda}_i\widehat{F}_t - \lambda_i F_t\right) \\ &= F_t\left(\frac{1}{T}\sum_{t=1}^{T}F_t^2\right)^{-1}\frac{1}{\sqrt{T}}\sum_{s=1}^{T}F_s u_{is} + o_p\left(1\right) \\ &\xrightarrow{d} N\left(0, \frac{F_t^2}{\phi_F}\sigma^2\right) \end{aligned}$$

as $(n, T) \to \infty$.

Accordingly, one can obtain that

$$\begin{aligned}
\frac{URSS}{(nT-n)} &= I + O_p\left(\frac{1}{T}\right) + O_p\left(\frac{1}{\sqrt{T}}\right) \\
&= I + O_p\left(\frac{1}{\sqrt{T}}\right) \xrightarrow{p} \sigma^2
\end{aligned}$$

as $(n, T) \to \infty$.

Next, consider the numerator.

$$\begin{aligned}
\frac{(RRSS - URSS)}{n} &= \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T} y_{it}^2 - \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it} - \widehat{\lambda}_i \widehat{F}_t\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[u_{it}^2 - u_{it}^2 - \left(\widehat{\lambda}_i\widehat{F}_t - \lambda_i F_t\right)^2 + 2\left(\widehat{\lambda}_i\widehat{F}_t - \lambda_i F_t\right)u_{it}\right] \\
&= -\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i\widehat{F}_t - \lambda_i F_t\right)\right\}^2 \\
&\quad + \frac{2}{n\sqrt{T}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i\widehat{F}_t - \lambda_i F_t\right)\right\}u_{it} \\
&= I + II.
\end{aligned}$$

Consider I first.

$$I = -\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} Q_{it}^2 \to -E\left(Q_{it}^2\right) = -\frac{F_t^2}{\phi_F}\sigma^2$$

as $(n, T) \to \infty$ where $Q_{it} \xrightarrow{d} N\left(0, \frac{F_t^2}{\phi_F}\sigma^2\right)$.

For II,

$$\begin{aligned}
II &= \frac{2}{n\sqrt{T}}\sum_{i=1}^{n}\sum_{t=1}^{T} Q_{it}u_{it} \\
&= \frac{2}{n\sqrt{T}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[F_t\left(\frac{1}{T}\sum_{t=1}^{T}F_t^2\right)^{-1}\frac{1}{\sqrt{T}}\sum_{s=1}^{T}F_s u_{is}\right]u_{it} \\
&= \frac{2}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}F_t u_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}F_s u_{is}\right)\left(\frac{1}{T}\sum_{t=1}^{T}F_t^2\right)^{-1} \\
&= \frac{2}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}F_t u_{it}\right)^2\left(\frac{1}{T}\sum_{t=1}^{T}F_t^2\right)^{-1} \xrightarrow{p} \frac{2F_t^2}{\phi_F}\sigma^2
\end{aligned}$$

as $(n,T)\to\infty$. To see why, it can be shown that

$$\begin{aligned}
&\frac{2}{n}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T}F_t^2\right)^{-1}\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}F_s u_{is}\right)^2 \\
&= \frac{2}{n}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T}F_t^2\right)\left(\frac{1}{T}\sum_{t=1}^{T}F_t^2\right)^{-2}\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}F_s u_{is}\right)^2 \\
&= \frac{2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}F_t^2\left(\frac{1}{T}\sum_{t=1}^{T}F_t^2\right)^{-2}\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}F_s u_{is}\right)^2 \\
&= \frac{2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[F_t\left(\frac{1}{T}\sum_{t=1}^{T}F_t^2\right)^{-1}\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}F_s u_{is}\right)\right]^2 \\
&= \frac{2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}Q_{it}^2 \ (= 2 \text{ times of term I}).
\end{aligned}$$

Combining the results, we obtain that

$$\frac{(RRSS-URSS)}{n} \xrightarrow{p} \frac{F_t^2}{\phi_F}\sigma^2.$$

Hence, one concludes that

$$F_\lambda = \frac{RRSS-URSS}{URSS}\frac{(nT-n)}{n} \xrightarrow{p} \frac{F_t^2}{\phi_F} = 1 + \frac{F_t^2-\phi_F}{\phi_F}$$

as $(n,T)\to\infty$. Clearly, now we have the shift term which cannot be specified.

2. Assume $\frac{n}{T}\to 0$. Note that if this is the case, then

$$\sqrt{n}\left(\widehat{\lambda}_i\widehat{F}_t-\lambda_iF_t\right)\xrightarrow{d}N\left(0,V_{it}\right)$$

where $V_{it}=\frac{\lambda_i^2}{\phi_\lambda^2}\sigma^2\phi_\lambda=\frac{\lambda_i^2}{\phi_\lambda}\sigma^2$ as in Bai (2003). However, we cannot obtain V_{it} since $\lambda_i=0$ for all $i=1,\ldots,n$ under the null and ϕ_λ cannot be defined.

■

Next, we check the asymptotic normality of the F-statistic by proving Theorem 2 with an assumption $\frac{T}{n}\to 0$.

Proof. For the sketch of proof, we write

$$\begin{aligned}
R_\lambda-\widehat{\sigma}^2 &= \frac{(RRSS-URSS)}{n}-\frac{URSS}{(nT-n)}\\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[-\left(\widehat{\lambda}_i\widehat{F}_t-\lambda_iF_t\right)^2+2\left(\widehat{\lambda}_i\widehat{F}_t-\lambda_iF_t\right)u_{it}\right]\\
&\quad -\frac{1}{n\left(T-1\right)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[u_{it}^2+\left(\widehat{\lambda}_i\widehat{F}_t-\lambda_iF_t\right)^2-2\left(\widehat{\lambda}_i\widehat{F}_t-\lambda_iF_t\right)u_{it}\right]\\
&= -\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i\widehat{F}_t-\lambda_iF_t\right)\right\}^2\\
&\quad +\frac{2}{n\sqrt{T}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i\widehat{F}_t-\lambda_iF_t\right)u_{it}\right\}\\
&\quad -\frac{1}{n\left(T-1\right)}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it}^2\\
&\quad -\frac{1}{nT(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i\widehat{F}_t-\lambda_iF_t\right)\right\}^2\\
&\quad +\frac{2}{n\sqrt{T}\left(T-1\right)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i\widehat{F}_t-\lambda_iF_t\right)u_{it}\right\}\\
&= I+II+III+IV+V.
\end{aligned}$$

Note that $IV = O_p\left(\frac{1}{T}\right)$ and $V = O_p\left(\frac{1}{\sqrt{T}}\right)$, as shown above. Next, to apply the CLT,

$$\sqrt{nT}\left(R_\lambda - \widehat{\sigma}^2\right) \approx \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T} Q_{it}^2 - \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T} u_{it}^2 + o_p(1).$$

Again, the standard normal random variables can be defined as follows:

$$\frac{Q_{it}}{\sqrt{\frac{F_t^2}{\phi_F}\sigma^2}} \xrightarrow{d} N(0,1)$$

as $(n,T) \to \infty$. It can be also shown that

$$E\left(\frac{Q_{it}^2}{\frac{F_t^2}{\phi_F}\sigma^2}\right) = 1$$

and

$$Var\left(\frac{Q_{it}^2}{\frac{F_t^2}{\phi_F}\sigma^2}\right) = 2$$

since $\frac{Q_{it}^2}{\sigma^2\phi_F}$ follows a chi-squared distribution. Then above can be rewritten by,

$$E\left(Q_{it}^2\right) = \frac{F_t^2}{\phi_F}\sigma^2$$

and

$$Var\left(Q_{it}^2\right) = 2\frac{F_t^4}{\phi_F^2}\sigma^4.$$

Hence, one can see that

$$\begin{aligned}\sqrt{nT}\left(R_\lambda - \widehat{\sigma}^2\right) &= \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[Q_{it}^2 - u_{it}^2\right] + o_p(1) \\ &\xrightarrow{d} N\left(\frac{(F_t^2 - \phi_F)}{\phi_F}\sigma^2, \psi\right)\end{aligned}$$

if $E\left(Q_{it}^2 - u_{it}^2\right) = \frac{F_t^2}{\phi_F}\sigma^2 - \sigma^2 = \frac{\left(F_t^2 - \phi_F\right)}{\phi_F}\sigma^2$ and $\psi = Var\left(Q_{it}^2 - u_{it}^2\right) < \infty$ as $(n,T) \to \infty$. Finally, we obtain

$$\sqrt{nT}\left(F_\lambda - 1\right) = \frac{1}{\widehat{\sigma}^2}\sqrt{nT}\left(R_\lambda - \widehat{\sigma}^2\right) \xrightarrow{d} N\left(\frac{\left(F_t^2 - \phi_F\right)}{\phi_F}, \frac{\psi}{\sigma^4}\right)$$

as $(n,T) \to \infty$ using $\widehat{\sigma}^2 \xrightarrow{p} \sigma^2$. ■

D Proof of Proposition 3

We next consider the limiting distribution of the bootstrap F-statistic when F_t is known. With the assumption $(n,T) \overset{\text{seq}}{\to} \infty$, consider the bootstrap DGP like the following:

$$y_{it}^* = \widetilde{\lambda}_i F_t + u_{it}\varepsilon_{it}^*$$

where $\widetilde{\lambda}_i^*$ denotes the bootstrap least squares estimator.

We first write

$$F_\lambda^* = \frac{R_\lambda^*}{\widehat{\sigma}^{*2}}$$

where $R_\lambda = \frac{(RRSS^* - URSS^*)}{n}$ and $\widehat{\sigma}^{*2} = \frac{URSS^*}{(nT-n)}$ with

$$RRSS^* = \sum_{i=1}^{n}\sum_{t=1}^{T} y_{it}^{*2}$$

and

$$URSS^* = \sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it}^* - \widetilde{\lambda}_i^* F_t\right)^2.$$

Rearranging terms, we have

$$\sqrt{n}\left(F_\lambda^* - 1\right) = \frac{\sqrt{n}\left(R_\lambda^* - \widehat{\sigma}^{*2}\right)}{\widehat{\sigma}^{*2}}.$$

Before we sketch the proof of the consistency of the bootstrap F-statistic, we first derive the asymptotic distribution of it in the following lemma.

Lemma 3 (B) *Assume* $(n,T) \overset{\text{seq}}{\to} \infty$ *and* F_t *is observable. Then*

$$\sqrt{n}\left(F_\lambda^* - 1\right) \overset{d}{\to} N(0,2).$$

Proof. Expanding the terms, one has

$$\begin{aligned}
R_\lambda^* - \widehat{\sigma}^{*2} &= \frac{(RRSS^* - URSS^*)}{n} - \frac{URSS^*}{(nT-n)} \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[u_{it}^2\varepsilon_{it}^{*2} - u_{it}^2\varepsilon_{it}^{*2} - \left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right)^2 F_t^2 + 2\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right)u_{it}F_t\varepsilon_{it}^*\right] \\
&\quad - \frac{1}{n(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[u_{it}\varepsilon_{it}^* - \left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right)F_t\right]^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[-\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right)^2 F_t^2 + 2\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right)u_{it}F_t\varepsilon_{it}^*\right] \\
&\quad - \frac{1}{n(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[u_{it}^2\varepsilon_{it}^{*2} + \left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right)^2 F_t^2 - 2\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right)u_{it}F_t\varepsilon_{it}^*\right] \\
&= -\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right)\right\}^2 F_t^2 + \frac{2}{n\sqrt{T}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right)\right\}u_{it}F_t\varepsilon_{it}^* \\
&\quad - \frac{1}{n(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it}^2\varepsilon_{it}^{*2} - \frac{1}{nT(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right)\right\}^2 F_t^2 \\
&\quad + \frac{2}{n\sqrt{T}(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widetilde{\lambda}_i - \lambda_i\right)\right\}u_{it}F_t\varepsilon_{it}^* \\
&= I + II + III + IV + V.
\end{aligned}$$

Consider I.

$$\begin{aligned}
I &= -\frac{1}{n}\sum_{i=1}^{n}\left[\sqrt{T}\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right)\right]^2\left[\frac{1}{T}\sum_{t=1}^{T}F_t^2\right] \\
&= -\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}F_t u_{it}\varepsilon_{it}^*\right]^2\left[\frac{1}{T}\sum_{t=1}^{T}F_t^2\right]^{-1} = O_p(1)
\end{aligned}$$

since we know $\sqrt{T}\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right) = O_p(1)$.

For II,

$$\begin{aligned} II &= \frac{2}{n\sqrt{T}} \sum_{i=1}^{n} \sum_{t=1}^{T} \left\{ \sqrt{T}\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right) \right\} u_{it} F_t \varepsilon_{it}^* \\ &= \frac{2}{n} \sum_{i=1}^{n} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_t u_{it} \varepsilon_{it}^* \right]^2 \left[\frac{1}{T} \sum_{t=1}^{T} F_t^2 \right]^{-1} = O_p(1). \end{aligned}$$

For III,

$$III = -\frac{1}{n} \sum_{i=1}^{n} \frac{1}{T-1} \sum_{t=1}^{T} u_{it}^2 \varepsilon_{it}^{*2} = O_p(1).$$

For IV and V, note that

$$IV \approx -\frac{1}{n} \sum_{i=1}^{n} \left\{ \sqrt{T}\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right) \right\}^2 \frac{1}{T^2} \sum_{t=1}^{T} F_t^2 = O_p\left(\frac{1}{T}\right) = o_p(1)$$

and

$$V \approx \frac{2}{n} \sum_{i=1}^{n} \left\{ \sqrt{T}\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right) \right\} \frac{1}{T^{3/2}} \sum_{t=1}^{T} F_t u_{it} \varepsilon_{it}^* = O_p\left(\frac{1}{T}\right) = o_p(1).$$

After rearranging all the terms, we have

$$R_\lambda^* - \widehat{\sigma}^{*2} = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_t u_{it} \varepsilon_{it}^* \right]^2 \left[\frac{1}{T} \sum_{t=1}^{T} F_t^2 \right]^{-1} - \frac{1}{n} \sum_{i=1}^{n} \frac{1}{T-1} \sum_{t=1}^{T} u_{it}^2 \varepsilon_{it}^{*2} + o_p(1)$$

and

$$\begin{aligned} \sqrt{n}\left(R_\lambda^* - \widehat{\sigma}^{*2}\right) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_t u_{it} \varepsilon_{it}^* \right]^2 \left[\frac{1}{T} \sum_{t=1}^{T} F_t^2 \right]^{-1} \\ &\quad - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{T-1} \sum_{t=1}^{T} u_{it}^2 \varepsilon_{it}^{*2} + o_p(1). \end{aligned}$$

To apply the CLT, note that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_t u_{it} \varepsilon_{it}^* \xrightarrow{d} W_i^* \sim N(0, \sigma^2 \phi_F)$$

as $T \to \infty$ using the fact that ε_{it}^{*} is an external random variable with $E\left(\varepsilon_{it}^{*}\right) = 0$ and $E\left(\varepsilon_{it}^{*2}\right) = 1$.

Hence, the standard normal random variables can be defined as

$$\frac{W_i^*}{\sqrt{\sigma^2 \phi_F}} \sim N(0,1)$$

and we can construct the random variable, $\frac{W_i^{*2}}{\sigma^2 \phi_F}$, satisfying

$$E\left(\frac{W_i^{*2}}{\sigma^2 \phi_F}\right) = 1$$

and

$$Var\left(\frac{W_i^{*2}}{\sigma^2 \phi_F}\right) = 2$$

because $\frac{W_i^{*2}}{\sigma^2 \phi_F}$ follows a chi-squared distribution.

Rewriting above, it can be shown that

$$E\left(W_i^{*2}\right) = \sigma^2 \phi_F$$

and

$$Var\left(W_i^{*2}\right) = 2\sigma^4 \phi_F^2.$$

Next, we consider

$$\begin{aligned}
\widehat{\sigma}^{*2} &= \frac{URSS^*}{(nT-n)} = \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(u_{it}^{*} - \left(\widetilde{\lambda}_i^{*} - \widetilde{\lambda}_i\right) F_t\right)^2 \\
&= \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} u_{it}^2 \varepsilon_{it}^{*2} + \frac{1}{n(T-1)} \sum_{i=1}^{n} \left(\widetilde{\lambda}_i^{*} - \widetilde{\lambda}_i\right)^2 \sum_{t=1}^{T} F_t^2 \\
&\quad - \frac{2}{n(T-1)} \sum_{i=1}^{n} \left(\widetilde{\lambda}_i - \lambda_i\right) \sum_{t=1}^{T} u_{it} F_t \varepsilon_{it}^{*} \\
&= I + II + III.
\end{aligned}$$

For I, it is easy to see that

$$\frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} u_{it}^2 \varepsilon_{it}^{*2} \approx \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} u_{it}^2 \varepsilon_{it}^{*2} \xrightarrow{p} \sigma^2.$$

For II and III, it is straightforward to show $II = O_p\left(\frac{1}{T}\right) = o_p(1)$ and $III = O_p\left(\frac{1}{T}\right) = o_p(1)$ using $\sqrt{T}\left(\widetilde{\lambda}_i^* - \widetilde{\lambda}_i\right) = O_p(1)$.

Combining the results above, we have

$$\begin{aligned} \sqrt{n}\left(R_\lambda^* - \widehat{\sigma}^{*2}\right) &= \frac{1}{\sqrt{n}}\left[\sum_{i=1}^{n}\left(W_i^{*2}\phi_F^{-1} - \sigma^2\right)\right] + o_p(1) \\ &\xrightarrow{d} N(0, 2\sigma^4) \end{aligned}$$

as $n \to \infty$. This is because $E\left(W_i^{*2}\phi_F^{-1} - \sigma^2\right) = \sigma^2 - \sigma^2 = 0$ and $Var\left(W_i^{*2}\phi_F^{-1} - \sigma^2\right) = 2\sigma^4$. Finally, we obtain

$$\sqrt{n}\left(F_\lambda^* - 1\right) = \frac{1}{\widehat{\sigma}^{*2}}\sqrt{n}\left(R_\lambda^* - \widehat{\sigma}^{*2}\right) \xrightarrow{d} N(0, 2)$$

as $(n, T) \overset{seq}{\to} \infty$ using $\widehat{\sigma}^{*2} \xrightarrow{p} \sigma^2$. ■

From above, we can see that the asymptotic distribution of the bootstrap F-statistic coincides with the empirical one: convergence to the normal distribution. Based on this, we next sketch the proof for the validity of the bootstrap F-statistic. In this proof, we use Kolmogorov metric which is defined as $K(F, G) = \sup_x |F(x) - G(x)|$.

With an assumption $(n, T) \overset{seq}{\to} \infty$, the F-statistc and the bootstrap counterpart can be defined as follows:

$$F_\lambda = \frac{(RRSS - URSS)/n}{URSS/(nT - n)} = \frac{R_\lambda}{\widehat{\sigma}^2}$$

and

$$F_\lambda^* = \frac{(RRSS^* - URSS^*)/n}{URSS^*/(nT-n)} = \frac{R_\lambda^*}{\widehat{\sigma}^{*2}}.$$

Proof. We consider the denominator and the numerator subsequently.

1. We first treat the denominators of F_λ and F_λ^*.

   For the denominators of F_λ and F_λ^*, it is already shown that

   $$\frac{URSS}{(nT-n)} = \widehat{\sigma}^2 = \sigma^2 + O_p\left(\frac{1}{T}\right)$$

   and

   $$\frac{URSS^*}{(nT-n)} = \widehat{\sigma}^{*2} = \sigma^2 + O_p\left(\frac{1}{T}\right)$$

   as $(n,T) \overset{\text{seq}}{\to} \infty$.

2. Now we treat the numerators of F_λ and F_λ^*. Notice that $\sqrt{n}$ is the right norming factor in this case.

   Recall that we have

   $$\sqrt{n}\left(R_\lambda - \widehat{\sigma}^2\right) \overset{d}{\to} N(0, 2\sigma^4)$$

   and

   $$\frac{\sqrt{n}\left(R_\lambda - \widehat{\sigma}^2\right)}{(2\sigma^4)^{1/2}} \overset{d}{\to} N(0,1)$$

   as $(n,T) \overset{\text{seq}}{\to} \infty$.

   We define $\mathcal{L}(R_\lambda) = P\left(\sqrt{n}\left(R_\lambda - \widehat{\sigma}^2\right) \le x\right)$ and $\mathcal{L}^*(R_\lambda^*) = P^*\left(\sqrt{n}\left(R_\lambda^* - \widehat{\sigma}^{*2}\right) \le x\right)$, respectively. Note that P^* denotes an empirical (or bootstrap) distribution.

Then we write

$$\begin{aligned}
K\left(\mathcal{L}\left(R_\lambda\right), \mathcal{L}^*\left(R_\lambda^*\right)\right) &= \sup_x \left| \begin{array}{c} P\left(\frac{\sqrt{n}\left(R_\lambda - \widehat{\sigma}^2\right)}{\left(2\sigma^4\right)^{1/2}} \leq \frac{x}{\left(2\sigma^4\right)^{1/2}}\right) \\ -P^*\left(\frac{\sqrt{n}\left(R_\lambda^* - \widehat{\sigma}^2\right)}{\left(s^2\right)^{1/2}} \leq \frac{x}{\left(s^2\right)^{1/2}}\right) \end{array} \right| \\
&= \sup_x \left| \begin{array}{c} \left[P\left(\frac{\sqrt{n}\left(R_\lambda - \widehat{\sigma}^2\right)}{\left(2\sigma^4\right)^{1/2}} \leq \frac{x}{\left(2\sigma^4\right)^{1/2}}\right) - \Phi\left(\frac{x}{\left(2\sigma^4\right)^{1/2}}\right)\right] \\ +\left[\Phi\left(\frac{x}{\left(2\sigma^4\right)^{1/2}}\right) - \Phi\left(\frac{x}{\left(s^2\right)^{1/2}}\right)\right] \\ +\left[\Phi\left(\frac{x}{\left(s^2\right)^{1/2}}\right) - P^*\left(\frac{\sqrt{n}\left(R_\lambda^* - \widehat{\sigma}^{*2}\right)}{\left(s^2\right)^{1/2}} \leq \frac{x}{\left(s^2\right)^{1/2}}\right)\right] \end{array} \right| \\
&\leq \sup_x \left| P\left(\frac{\sqrt{n}\left(R_\lambda - \widehat{\sigma}^2\right)}{\left(2\sigma^4\right)^{1/2}} \leq \frac{x}{\left(2\sigma^4\right)^{1/2}}\right) - \Phi\left(\frac{x}{\left(2\sigma^4\right)^{1/2}}\right) \right| \\
&\quad + \sup_x \left| \Phi\left(\frac{x}{\left(2\sigma^4\right)^{1/2}}\right) - \Phi\left(\frac{x}{\left(s^2\right)^{1/2}}\right) \right| \\
&\quad + \sup_x \left| \Phi\left(\frac{x}{\left(s^2\right)^{1/2}}\right) - P^*\left(\frac{\sqrt{n}\left(R_\lambda^* - \widehat{\sigma}^{*2}\right)}{\left(s^2\right)^{1/2}} \leq \frac{x}{\left(s^2\right)^{1/2}}\right) \right| \\
&= I + II + III
\end{aligned}$$

using triangle inequality where $\Phi(\cdot)$ indicates the c.d.f. of a standard normal distribution and $s^2 = Var^*\left(\sqrt{n}\left(R_\lambda^* - \widehat{\sigma}^{*2}\right)\right)$.

Consider I. It can be shown that

$$I = \sup_x \left| P\left(\frac{\sqrt{n}\left(R_\lambda - \widehat{\sigma}^2\right)}{\left(2\sigma^4\right)^{1/2}} \leq \frac{x}{\left(2\sigma^4\right)^{1/2}}\right) - \Phi\left(\frac{x}{\left(2\sigma^4\right)^{1/2}}\right) \right| \to 0$$

by Polya's theorem since $P\left(\frac{\sqrt{n}\left(R_\lambda - \widehat{\sigma}^2\right)}{\left(2\sigma^4\right)^{1/2}} \leq \frac{x}{\left(2\sigma^4\right)^{1/2}}\right) \xrightarrow{d} \Phi\left(\frac{x}{\left(2\sigma^4\right)^{1/2}}\right)$ and $\Phi(\cdot)$ is a continuous cdf. For details of Polya's theorem, see, e.g., Lehmann (1999).

For II, obviously

$$II = \sup_x \left| \Phi\left(\frac{x}{\left(2\sigma^4\right)^{1/2}}\right) - \Phi\left(\frac{x}{\left(s^2\right)^{1/2}}\right) \right| \to 0$$

using continuous mapping theorem (CMT) because $s^2 = Var^*\left(\sqrt{n}\left(R_\lambda^* - \widehat{\sigma}^{*2}\right)\right) \xrightarrow{p}$

$2\sigma^4$.

Lastly, one can apply Berry-Esseen theorem (see, e.g., Lehmann (1999)) to show $III \to 0$. That is, there exists a positive constant C such that,

$$III \leq \frac{C}{\sqrt{n}} \frac{\Gamma_{nT}^3}{(s)^3} \to 0$$

where $\Gamma_{nT}^3 = E^* \left| R_\lambda^* - \widehat{\sigma}^{*2} \right|^3$.

Combining above results, we have

$$I + II + III \xrightarrow{p} 0$$

and hence,

$$K \left(\mathcal{L} \left(R_\lambda \right), \mathcal{L}^* \left(R_\lambda^* \right) \right) \xrightarrow{p} 0$$

as $(n, T) \overset{\text{seq}}{\to} \infty$.

■

E Proof of Proposition 4

Proof. We sketch the proof of the consistency of bootstrapping PCA with an assumption $\frac{T}{n} \to 0$.

Let us define first

$$H_{nT} = P \left(\tau \leq x \right)$$

where a functional $\tau = \sqrt{T} \left(\widehat{\lambda}_i \widehat{F}_t - \lambda_i F_t \right)$.

Accordingly, the bootstrap counterpart can be defined as

$$H_{Boot} = P^* \left(\tau^* \leq x \right)$$

where $\tau^* = \sqrt{T}\left(\widehat{\lambda}_i^* \widehat{F}_t^* - \widehat{\lambda}_i \widehat{F}_t\right)$. Note that P^* denotes the empirical (or bootstrap) distribution.

Recall from Bai (2003) that

$$\frac{\sqrt{T}\left(\widehat{\lambda}_i \widehat{F}_t - \lambda_i F_t\right)}{\left(W_{it}\right)^{1/2}} \xrightarrow{d} N(0,1)$$

where $W_{it} = \frac{F_t^2}{\phi_F}\sigma^2$. Then we can write the following as in Proposition 3:

$$\begin{aligned}
K\left(H_{nT}, H_{Boot}\right) &= \sup_x \left| P\left(\frac{\tau}{\left(W_{it}\right)^{1/2}} \leq \frac{x}{\left(W_{it}\right)^{1/2}}\right) - P^*\left(\frac{\tau^*}{\left(\widehat{W}_{it}\right)^{1/2}} \leq \frac{x}{\left(\widehat{W}_{it}\right)^{1/2}}\right)\right| \\
&= \sup_x \left| \begin{array}{c} \left[P\left(\frac{\tau}{(W_{it})^{1/2}} \leq \frac{x}{(W_{it})^{1/2}}\right) - \Phi\left(\frac{x}{(W_{it})^{1/2}}\right)\right] \\ + \left[\Phi\left(\frac{x}{(W_{it})^{1/2}}\right) - \Phi\left(\frac{x}{\left(\widehat{W}_{it}\right)^{1/2}}\right)\right] \\ + \left[\Phi\left(\frac{x}{\left(\widehat{W}_{it}\right)^{1/2}}\right) - P^*\left(\frac{\tau^*}{\left(\widehat{W}_{it}\right)^{1/2}} \leq \frac{x}{\left(\widehat{W}_{it}\right)^{1/2}}\right)\right] \end{array} \right| \\
&\leq \sup_x \left| P\left(\frac{\tau}{\left(W_{it}\right)^{1/2}} \leq \frac{x}{\left(W_{it}\right)^{1/2}}\right) - \Phi\left(\frac{x}{\left(W_{it}\right)^{1/2}}\right)\right| \\
&\quad + \sup_x \left| \Phi\left(\frac{x}{\left(W_{it}\right)^{1/2}}\right) - \Phi\left(\frac{x}{\left(\widehat{W}_{it}\right)^{1/2}}\right)\right| \\
&\quad + \sup_x \left| \Phi\left(\frac{x}{\left(\widehat{W}_{it}\right)^{1/2}}\right) - P^*\left(\frac{\tau^*}{\left(\widehat{W}_{it}\right)^{1/2}} \leq \frac{x}{\left(\widehat{W}_{it}\right)^{1/2}}\right)\right| \\
&= I + II + III.
\end{aligned}$$

Consider I. It can be shown that

$$I = \sup_x \left| P\left(\frac{T}{\left(W_{it}\right)^{1/2}} \leq \frac{x}{\left(W_{it}\right)^{1/2}}\right) - \Phi\left(\frac{x}{\left(W_{it}\right)^{1/2}}\right)\right| \to 0$$

since $P\left(\frac{T}{(W_{it})^{1/2}} \leq \frac{x}{(W_{it})^{1/2}}\right) \xrightarrow{d} \Phi\left(\frac{x}{(W_{it})^{1/2}}\right)$ and $\Phi\left(\cdot\right)$ is a continuous cdf.

We turn to II. Obviously,

$$II = \sup_x \left| \Phi\left(\frac{x}{(W_{it})^{1/2}}\right) - \Phi\left(\frac{x}{\left(\widehat{W}_{it}\right)^{1/2}}\right) \right| \to 0$$

using CMT because $\widehat{W}_{it} \xrightarrow{p} W_{it}$. For details of consistency of $\widehat{W}_{it}$ for W_{it}, see Bai (2003).

Lastly, we can apply Berry-Esseen theorem to show $III \to 0$. That is, there exists a positive constant C such that,

$$III \leq \frac{C}{\sqrt{T}} \frac{\Gamma_{nT}^3}{\left(\widehat{W}_{it}\right)^{3/2}} \to 0$$

if $\Gamma_{nT}^3 = E^* \left| \widehat{\lambda}_i^* \widehat{F}_t^* - \widehat{\lambda}_i \widehat{F}_t \right|^3$ and $var^* \left(\sqrt{T} \left(\widehat{\lambda}_i^* \widehat{F}_t^* - \widehat{\lambda}_i \widehat{F}_t \right) \right) = \widehat{W}_{it}$.

Therefore, we obtain

$$I + II + III \xrightarrow{p} 0.$$

and conclude that

$$K\left(H_{nT}, H_{Boot}\right) \xrightarrow{p} 0$$

as $(n, T) \to \infty$. ■

F Proof of Proposition 5

This section considers the validity of the bootstrap F-statistic when F_t is unknown. With an assumption $(n, T) \to \infty$, consider the bootstrap DGP as follows:

$$y_{it}^* = \widehat{\lambda}_i \widehat{F}_t + u_{it} \varepsilon_{it}^*$$

where $\widehat{\lambda}_i^* \widehat{F}_t^*$ denotes the bootstrap principal component estimates of above equation.

We again write

$$F_\lambda^* = \frac{R_\lambda^*}{\widehat{\sigma}^{*2}}$$

where $R_\lambda = \frac{(RRSS^* - URSS^*)}{n}$ and $\widehat{\sigma}^{*2} = \frac{URSS^*}{(nT-n)}$ with

$$RRSS^* = \sum_{i=1}^{n} \sum_{t=1}^{T} y_{it}^{*2}$$

and

$$URSS^* = \sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it}^* - \widehat{\lambda}_i^* \widehat{F}_t^* \right)^2 .$$

We first derive the asymptotic distribution of the bootstrap F-statistic.

Lemma 4 (B) *Assume* $(n, T) \to \infty$ *and* $\frac{T}{n} \to 0$ *with unobservable* F_t*. Then,*

$$\sqrt{nT} \left(F_\lambda^* - 1 \right) \xrightarrow{d} N\left(\frac{(F_t^2 - \phi_F)}{\phi_F}, \frac{\psi}{\sigma^4}\right)$$

where $\psi = Var \left(Q_{it}^{*2} - u_{it}^2 \varepsilon_{it}^{*2} \right) < \infty$ *and*

$$\begin{aligned} Q_{it}^* &= \sqrt{T} \left(\widehat{\lambda}_i^* \widehat{F}_t^* - \widehat{\lambda}_i \widehat{F}_t \right) \\ &= \sum_{t=1}^{T} \widehat{F}_t \left(\frac{1}{T} \sum_{t=1}^{T} \widehat{F}_t^2 \right)^{-1} \frac{1}{\sqrt{T}} \sum_{s=1}^{T} \widehat{F}_s u_{is} \varepsilon_{is}^* + o_p(1) . \end{aligned}$$

Proof. For the limiting distribution of the bootstrap F-statistic, consider

$$
\begin{aligned}
R_\lambda^* - \widehat{\sigma}^{*2} &= \frac{(RRSS^* - URSS^*)}{n} - \frac{URSS^*}{(nT-n)} \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[-\left(\widehat{\lambda}_i^*\widehat{F}_t^* - \widehat{\lambda}_i\widehat{F}_t\right)^2 + 2\left(\widehat{\lambda}_i^*\widehat{F}_t^* - \widehat{\lambda}_i\widehat{F}_t\right)u_{it}\varepsilon_{it}^*\right] \\
&\quad -\frac{1}{n\,(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[u_{it}^2\varepsilon_{it}^{*2} + \left(\widehat{\lambda}_i^*\widehat{F}_t^* - \widehat{\lambda}_i\widehat{F}_t\right)^2 - 2\left(\widehat{\lambda}_i^*\widehat{F}_t^* - \widehat{\lambda}_i\widehat{F}_t\right)u_{it}\varepsilon_{it}^*\right] \\
&= -\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i^*\widehat{F}_t^* - \widehat{\lambda}_i\widehat{F}_t\right)\right\}^2 \\
&\quad +\frac{2}{n\sqrt{T}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i^*\widehat{F}_t^* - \widehat{\lambda}_i\widehat{F}_t\right)u_{it}\varepsilon_{it}^*\right\} \\
&\quad -\frac{1}{n\,(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it}^2\varepsilon_{it}^{*2} - \frac{1}{nT(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i^*\widehat{F}_t^* - \widehat{\lambda}_i\widehat{F}_t\right)\right\}^2 \\
&\quad +\frac{2}{n\sqrt{T}\,(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i^*\widehat{F}_t^* - \widehat{\lambda}_i\widehat{F}_t\right)u_{it}\varepsilon_{it}^*\right\} \\
&= I + II + III + IV + V.
\end{aligned}
$$

For I,

$$
\begin{aligned}
I &= -\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i^*\widehat{F}_t^* - \widehat{\lambda}_i\widehat{F}_t\right)\right\}^2 \\
&= -\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}Q_{it}^{*2} \to -E\left(Q_{it}^{*2}\right) = -\frac{F_t^2}{\phi_F}\sigma^2
\end{aligned}
$$

as $(n,T)\to\infty$ using the consistency of bootstrapping PCA.

For II,

$$\begin{aligned}
II &= \frac{2}{n\sqrt{T}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\sqrt{T}\left(\widehat{\lambda}_i^*\widehat{F}_t^* - \widehat{\lambda}_i\widehat{F}_t\right)u_{it}\varepsilon_{it}^*\right\} \\
&= \frac{2}{n\sqrt{T}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\widehat{F}_t\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_t^2\right)^{-1}\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\widehat{F}_s u_{is}\varepsilon_{is}^*\right]u_{it}\varepsilon_{it}^* \\
&= \frac{2}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\widehat{F}_t u_{it}\varepsilon_{it}^*\right)\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\widehat{F}_s u_{is}\varepsilon_{is}^*\right)\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_t^2\right)^{-1} \\
&= \frac{2}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\widehat{F}_t u_{it}\varepsilon_{it}^*\right)^2\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_t^2\right)^{-1} \\
&= \frac{2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}Q_{it}^{*2}.
\end{aligned}$$

Note that the last equality can be shown by

$$\begin{aligned}
&\frac{2}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\widehat{F}_t u_{it}\varepsilon_{it}^*\right)^2\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_t^2\right)^{-1} \\
= &\frac{2}{n}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_t^2\right)\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_t^2\right)^{-2}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\widehat{F}_t u_{it}\varepsilon_{it}^*\right)^2 \\
= &\frac{2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{F}_t^2\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_t^2\right)^{-2}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\widehat{F}_t u_{it}\varepsilon_{it}^*\right)^2 \\
= &\frac{2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\widehat{F}_t\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_t^2\right)^{-1}\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\widehat{F}_s u_{is}\varepsilon_{is}^*\right]^2 \\
= &\frac{2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}Q_{it}^{*2}
\end{aligned}$$

using $Q_{it}^* = \sqrt{T}\left(\widehat{\lambda}_i^*\widehat{F}_t^* - \widehat{\lambda}_i\widehat{F}_t\right) = \sum_{t=1}^{T}\widehat{F}_t\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_t^2\right)^{-1}\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\widehat{F}_s u_{is}\varepsilon_{is}^* + o_p(1)$ as $(n,T)\to\infty$.

For III, one can easily find that

$$III = -\frac{1}{n(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it}^2\varepsilon_{it}^{*2} \xrightarrow{p} -\sigma^2$$

using an external random variable with $E\left(\varepsilon_{it}^{*2}\right)=1$.

For IV and V, It can be shown that $IV=O_p\left(\frac{1}{T}\right)$ and $V=O_p\left(\frac{1}{\sqrt{T}}\right)$ using $\sqrt{T}\left(\widehat{\lambda}_i^*\widehat{F}_t^*-\widehat{\lambda}_i\widehat{F}_t\right)=O_p\left(1\right)$.

We finally conclude that

$$\begin{aligned}\sqrt{nT}\left(R_\lambda^*-\widehat{\sigma}^{*2}\right) &\approx \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}Q_{it}^{*2}-\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it}^2\varepsilon_{it}^{*2}+o_p\left(1\right)\\ &= \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[Q_{it}^{*2}-u_{it}^2\varepsilon_{it}^{*2}\right]+o_p\left(1\right).\end{aligned}$$

Following a similar process to that in Theorem 2 using the consistency of bootstrapping PCA (i.e., Q_{it}^* for Q_{it}), it can be shown that

$$\sqrt{nT}\left(R_\lambda^*-\widehat{\sigma}^{*2}\right)\xrightarrow{d}N(\frac{(F_t^2-\phi_F)}{\phi_F}\sigma^2,\psi^*)$$

where $E\left(Q_{it}^{*2}-u_{it}^2\varepsilon_{it}^{*2}\right)=\frac{F_t^2}{\phi_F}\sigma^2-\sigma^2=\frac{\left(F_t^2-\phi_F\right)}{\phi_F}\sigma^2$ and $\psi^*=Var\left(Q_{it}^{*2}-u_{it}^2\varepsilon_{it}^{*2}\right)$ as $(n,T)\to\infty$.

Hence, we obtain

$$\sqrt{nT}\left(F_\lambda^*-1\right)=\frac{1}{\widehat{\sigma}^{*2}}\sqrt{nT}\left(R_\lambda^*-\widehat{\sigma}^{*2}\right)\xrightarrow{d}N(\frac{(F_t^2-\phi_F)}{\phi_F},\frac{\psi}{\sigma^4})$$

as $(n,T)\to\infty$ by showing $\widehat{\sigma}^{*2}\xrightarrow{p}\sigma^2$ and $\psi^*\xrightarrow{p}\psi=Var\left(Q_{it}^2-u_{it}^2\right)$. ■

Now we check the validity of the bootstrap F-statistic when F_t is unknown. With $(n,T)\to\infty$ and $\frac{T}{n}\to 0$, consider again

$$F_\lambda=\frac{\left(RRSS-URSS\right)/n}{URSS/\left(nT-n\right)}$$

and

$$F_\lambda^*=\frac{\left(RRSS^*-URSS^*\right)/n}{URSS^*/\left(nT-n\right)}.$$

Proof. We treat the denominator first and then the numerator in a similar fashion with the case of known factors.

1. Consider the denominators of F_λ and F_λ^*.

   For the denominators of F_λ and F_λ^*, it is already shown that

$$\frac{URSS}{(nT-n)} = \widehat{\sigma}^2 = \sigma^2 + O_p\left(\frac{1}{\sqrt{T}}\right)$$

   and

$$\frac{URSS^*}{(nT-n)} = \widehat{\sigma}^{*2} = \sigma^2 + O_p\left(\frac{1}{\sqrt{T}}\right)$$

   as $(n,T) \to \infty$.

2. Now we treat the numerator. We normalize it so that we have the normal distribution with zero mean, which is given by,

$$\sqrt{nT}\left(R_\lambda - \widehat{\sigma}^2 - \frac{(F_t^2 - \phi_F)}{\phi_F}\sigma^2\right) \xrightarrow{d} N(0, 2\frac{F_t^4}{\phi_F^2}\sigma^4)$$

   and hence,

$$\frac{\sqrt{nT}\left(R_\lambda - \widehat{\sigma}^2 - \frac{(F_t^2-\phi_F)}{\phi_F}\sigma^2\right)}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}} \xrightarrow{d} N(0,1)$$

   as $(n,T) \to \infty$. Notice that $\sqrt{nT}$ is the right norming factor in this case.

   We again define

$$\mathcal{L}\left(R_\lambda\right) = P\left(\sqrt{nT}\left(R_\lambda - \widehat{\sigma}^2 - \frac{(F_t^2 - \phi_F)}{\phi_F}\sigma^2\right) \le x\right)$$

   and

$$\mathcal{L}^*\left(R_\lambda^*\right) = P^*\left(\sqrt{nT}\left(R_\lambda^* - \widehat{\sigma}^{*2} - \frac{(F_t^2 - \phi_F)}{\phi_F}\sigma^2\right) \le x\right).$$

Using Kolmogorov metric, one writes

$$
\begin{aligned}
& K\left(\mathcal{L}\left(R_\lambda\right), \mathcal{L}^*\left(R_\lambda^*\right)\right) \\
= & \sup_x \left| \begin{array}{l} P\left(\frac{\sqrt{nT}\left(R_\lambda-\widehat{\sigma}^2-\frac{\left(F_t^2-\phi_F\right)}{\phi_F}\sigma^2\right)}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}} \leq \frac{x}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\right) \\ -P^*\left(\frac{\sqrt{nT}\left(R_\lambda^*-\widehat{\sigma}^{*2}-\frac{\left(F_t^2-\phi_F\right)}{\phi_F}\sigma^2\right)}{\left(s'^2\right)^{1/2}} \leq \frac{x}{\left(s'^2\right)^{1/2}}\right) \end{array} \right| \\
= & \sup_x \left| \begin{array}{c} \left[P\left(\frac{\sqrt{nT}\left(R_\lambda-\widehat{\sigma}^2-\frac{\left(F_t^2-\phi_F\right)}{\phi_F}\sigma^2\right)}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}} \leq \frac{x}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\right)-\Phi\left(\frac{x}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\right)\right] \\ +\left[\Phi\left(\frac{x}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\right)-\Phi\left(\frac{x}{\left(s'^2\right)^{1/2}}\right)\right] \\ +\left[\Phi\left(\frac{x}{\left(s'^2\right)^{1/2}}\right)-P^*\left(\frac{\sqrt{nT}\left(R_\lambda^*-\widehat{\sigma}^{*2}-\frac{\left(F_t^2-\phi_F\right)}{\phi_F}\sigma^2\right)}{\left(s'^2\right)^{1/2}} \leq \frac{x}{\left(s'^2\right)^{1/2}}\right)\right] \end{array} \right| \\
\leq & \sup_x \left| P\left(\frac{\sqrt{nT}\left(R_\lambda-\widehat{\sigma}^2-\frac{\left(F_t^2-\phi_F\right)}{\phi_F}\sigma^2\right)}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}} \leq \frac{x}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\right)-\Phi\left(\frac{x}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\right)\right| \\
& +\sup_x \left|\Phi\left(\frac{x}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\right)-\Phi\left(\frac{x}{\left(s'^2\right)^{1/2}}\right)\right| \\
& +\sup_x \left|\Phi\left(\frac{x}{\left(s'^2\right)^{1/2}}\right)-P^*\left(\frac{\sqrt{nT}\left(R_\lambda^*-\widehat{\sigma}^{*2}-\frac{\left(F_t^2-\phi_F\right)}{\phi_F}\sigma^2\right)}{\left(s'^2\right)^{1/2}} \leq \frac{x}{\left(s'^2\right)^{1/2}}\right)\right| \\
= & I+II+III
\end{aligned}
$$

where $s'^2 = Var^*\left(\sqrt{nT}\left(R_\lambda^* - \widehat{\sigma}^{*2} - \frac{\left(F_t^2-\phi_F\right)}{\phi_F}\sigma^2\right)\right) \xrightarrow{p} 2\frac{F_t^4}{\phi_F^2}\sigma^4$.

Consider I. It can be shown that

$$I=\sup_x\left|\begin{array}{c} P\left(\frac{\sqrt{nT}\left(R_\lambda-\widehat{\sigma}^2-\frac{\left(F_t^2-\phi_F\right)}{\phi_F}\sigma^2\right)}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\leq\frac{x}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\right)\\ -\Phi\left(\frac{x}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\right)\end{array}\right|\to 0$$

since $P\left(\frac{\sqrt{nT}\left(R_\lambda-\widehat{\sigma}^2-\frac{\left(F_t^2-\phi_F\right)}{\phi_F}\sigma^2\right)}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\leq\frac{x}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\right)\xrightarrow{d}\Phi\left(\frac{x}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\right)$.

For II, obviously

$$II=\sup_x\left|\Phi\left(\frac{x}{\left(2\frac{F_t^4}{\phi_F^2}\sigma^4\right)^{1/2}}\right)-\Phi\left(\frac{x}{\left(s'^2\right)^{1/2}}\right)\right|\to 0$$

using $s'^2\xrightarrow{p}2\frac{F_t^4}{\phi_F^2}\sigma^4$.

For III, by Berry-Esseen theorem it can be shown that there exists a positive constant C such that,

$$III\leq\frac{C}{\sqrt{nT}}\frac{\Gamma_{nT}^3}{\left(s'\right)^3}\to 0$$

where $\Gamma_{nT}^3=E^*\left|R_\lambda^*-\widehat{\sigma}^{*2}-\frac{\left(F_t^2-\phi_F\right)}{\phi_F}\sigma^2\right|^3$. Hence, we obtain

$$I+II+III\xrightarrow{p}0,$$

and conclude that

$$K\left(\mathcal{L}\left(R_\lambda\right),\mathcal{L}^*\left(R_\lambda^*\right)\right)\xrightarrow{p}0$$

as $(n,T)\to\infty$.

■

VITA

Name of Author: Sang Gon Na

Place of Birth: Seoul, Republic of Korea

Date of Birth: Feburary 26, 1975

Graduate and Undergraduate Schools Attended:

Syracuse University, Syracuse, New York

Yonsei University, Seoul, Republic of Korea

Degrees Awarded:

Master of Arts in Economics, 2010, Syracuse University

Executive Master of Public Administration, 2009, Syracuse University

Bachelor of Arts in Applied Statistics, 1999, Yonsei University

Awards and Honors:

Excellent Official Award from Prime Minister in Recognition of Outstanding Performance, 2006

Excellent Official Award from Deputy Prime Minister, 2002

High Honors Student Award from Yonsei University, 1996

Professional Experience:

Teaching Assistant, Economics Department, Syracuse University, 2009-2011

Deputy Director, Ministry of Strategy and Finance, Korea, 2008-present

Deputy Director, Ministry of Finance and Economy, Korea, 2000-2008