2006

The Asymptotics for Panel Models with Common Shocks

Chihwa Kao
Syracuse University. Center for Policy Research, cdkao@maxwell.syr.edu

Lorenzo Trapani
City University, Cass Business School (London, England) ; Bergamo University (Italy), l.trapani@city.ac.uk

Giovanni Urga
City University, Cass Business School (London, England), g.urga@city.ac.uk

ISSN: 1525-3066

Center for Policy Research
Working Paper No. 77

THE ASYMPTOTICS FOR PANEL MODELS WITH COMMON SHOCKS∗

Chihwa Kao
Lorenzo Trapani
Giovanni Urga

Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
426 Eggers Hall
Syracuse, New York 13244-1020
(315) 443-3114 | Fax (315) 443-1081
e-mail: ctrpol@syr.edu

February 2006

$5.00

CENTER FOR POLICY RESEARCH – Spring 2006

Timothy Smeeding, Director
Professor of Economics & Public Administration

Associate Directors

Margaret Austin
Associate Director
Budget and Administration

Douglas Wolf
Professor of Public Administration
Associate Director, Aging Studies Program

John Yinger
Professor of Economics and Public Administration
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Badi Baltagi .... Economics
Pablo Beramendi .... Political Science
Dan Black .... Economics
Lloyd Blanchard .... Public Administration
William Duncombe .... Public Administration
Gary Engelhardt .... Economics
Deborah Freund .... Public Administration
Madonna Harrington Meyer .... Sociology
Christine Himes .... Sociology
William C. Horrace .... Economics
Bernard Jump .... Public Administration
Duke Kao .... Economics
Eric Kingson .... Social Work
Thomas Kniesner .... Economics
Tom Krebs .... Economics
Jeffrey Kubik .... Economics
Andrew London .... Sociology
Len Lopoo .... Public Administration
Jerry Miner .... Economics
Jan Ondrich .... Economics
John Palmer .... Public Administration
Lori Ploutz-Snyder .... Exercise Science
David Popp .... Public Administration
Grant Reeher .... Political Science
Stuart Rosenthal .... Economics
Ross Rubenstein .... Public Administration
Margaret Usdansky .... Sociology
Michael Wasylenko .... Economics
Janet Wilmoth .... Sociology

GRADUATE ASSOCIATES

Javier Baez .... Economics
Sonali Ballal .... Public Administration
Dana Balter .... Public Administration
Jason Bladen .... Economics
Jesse Bricker .... Economics
Maria Brown .... Social Science
Yong Chen .... Economics
Ginnie Cronin .... Social Science
Ana Dammert .... Economics
Mike Eriksen .... Economics
Katie Fitzpatrick .... Economics
Alexandre Genest .... Public Administration
Julie Anna Golebiewski .... Economics
Nadia Greenhalgh-Stanley .... Economics
Yue Hu .... Economics
Becky Lafrancois .... Economics
Joseph Marchand .... Economics
Neelakshi Medhi .... Social Science
Larry Miller .... Public Administration
Wendy Parker .... Sociology
Emily Pas .... Economics
Shawn Rohlin .... Economics
Cynthia Searcy .... Public Administration
Talisha Searcy .... Public Administration
Jeff Thompson .... Economics
Wen Wang .... Economics
Yoshikun Yamamoto .... Public Administration

STAFF

Kelly Bogart .... Administrative Secretary
Martha Bonney .... Publications/Events Coordinator
Karen Cimilluca .... Administrative Secretary
Kati Foley .... Administrative Assistant
Kitty Nasto .... Administrative Secretary
Candi Patterson .... Computer Consultant
Mary Santy .... Administrative Secretary
Tammy Tanner .... Librarian/Office Coordinator

THE ASYMPTOTICS FOR PANEL MODELS WITH COMMON SHOCKS*

Chihwa Kao
Syracuse University
cdkao@maxwell.syr.edu

Lorenzo Trapani
Cass Business School and Bergamo University
L.Trapani@city.ac.uk

Giovanni Urga
Cass Business School
G.Urga@city.ac.uk

January 23, 2006

Abstract

This paper develops a novel asymptotic theory for panel models with common shocks. We assume that contemporaneous correlation can be generated by both the presence of common regressors among units and weak spatial dependence among the error terms. Several characteristics of the panel are considered: cross sectional and time series dimensions can either be fixed or large; factors can either be observable or unobservable; the factor model can describe either cointegration relationship or a spurious regression, and we also consider the stationary case. We derive the rate of convergence and the distribution limits for the ordinary least squares (OLS) estimates of the model parameters under all the aforementioned cases.

JEL Classification: C13; C23

Keywords: Cross-sectional dependence; Common shocks; Nonstationary panel.

*We wish to thank participants in the "1st Italian Congress of Econometrics and Empirical Economics" (Universita' Ca' Foscari, Venice, 24-25 January 2005), and the 9th World Congress of the Econometric Society (London, 19-24 August 2005), in particular Peter Pedroni and Yixiao Sun, for useful comments. The usual disclaimer applies. This paper originated while Chihwa Kao was visiting the Centre for Econometric Analysis at Cass (CEA@Cass). Financial support from City University 2005 Pump Priming Fund and CEA@Cass is gratefully acknowledged. Lorenzo Trapani ackowledges financial support from Cass Business School (RAE Development Funds Scheme) and ESRC Postdoctoral Fellowship Scheme (A.N.: PTA-026-27-1107).

1 INTRODUCTION

THERE IS A GROWING BODY of literature dealing with limit theory for nonstationary panels. While the first generation of these contributions assumed independence across units (see for instance Phillips and Moon (1999), Kao (1999)), in the second generation this assumption is relaxed, and hypothesis testing and estimation methods are evaluated assuming alternative degrees of cross dependence (see, Bai (2003, 2004), Bai and Ng (2002, 2004), Stock and Watson (2002)). We can distinguish the case where regressors are cross-sectionally dependent (see Donald and Lang (2004), Moulton (1990)) from that where it is the error terms across unit to be dependent (see for instance Bai and Kao, 2005; Moon and Perron, 2004) or both (see for instance Ahn, Lee and Schmidt (2001), Pesaran (2005a)).

In this paper we develop a new inferential framework which extends existing limit theory for panel data models. Phillips and Moon (1999) analyze what happens in nonstationary panels when both n and T are large. They consider both cointegrated relationship and spurious regression, getting the seminal result that as $n \to \infty$ a long-run average relationship between two nonstationary panel vectors exists even when the single units do not cointegrate. A similar result is also reported in Kao (1999). However, the asymptotics derived in Phillips and Moon (1999) is based on the assumption of cross section independence, albeit the authors point out that their results still hold when some weak dependence among panel units is allowed. Thus, the case of Phillips and Moon (1999) with any degree of dependence amongst units has remained largely unexplored, and it is likely to lead to different asymptotics. Asymptotic normality may not hold, for example when all or part of the regressors are aggregates, and may result in mixed asymptotic normality, as Andrews (2005) has demonstrated in a cross-sectional context. See also the discussion in Moon and Perron (2004).

Recently, Bai (2003, 2004) and Bai and Ng (2004) have developed an inferential theory for panels where cross sectional dependence is explicitly considered via factor-loadings representation. In these contributions, however, common factors are considered as part of the panel covariance structure rather than as explanatory variables in the regression model. The theory developed in the aforementioned papers works only when n and T tend to infinity along certain paths. From standard factor analysis (see e.g., Anderson (1981)) it is well known that consistent estimation of factors is not possible for a fixed n and consistent estimation of the loading is not possible for a fixed T.

The main aim of this paper is to propose a novel asymptotic theory for panels with common shocks. We generalize the limit theory developed by Phillips and Moon (1999) by employing and extending the theory for factor models in Bai (2003, 2004) and Bai and Ng (2004). Our asymptotics considers several features of the underlying model.

First, we assume that contemporaneous correlation can be generated by both the presence of common regressors (e.g. macro shocks, aggregate fiscal and monetary policies) among units and weak spatial dependence among the error terms.

Second, the common shocks can either be known or unobservable. Common shocks are likely to be only seldom observable, classical examples being the capital asset pricing model (CAPM) or index models. Most often, they are unknown. Classical examples are the cases of index extraction and indicators aggregation in economics (Quah and Sargent (1993), Forni and Reichlin (1998), Bernanke and Boivin (2000)), while in finance the seminal multifactor framework of the arbitrage pricing theory (APT) has generated huge number of contributions in the attempt to identifying the unobserved factors underlying the behavior of asset returns. Factor models are useful for forecasting purposes, as well documented in Stock and Watson (1999, 2005). Bai (2003, 2004), Bai and Ng (2002, 2005) and Boivin and Ng (2005) discuss numerous areas of research where factor models could be employed and some applications in macro and finance.

Third, in our framework the factor model may describe either a cointegration relationship or a spurious regression. We also consider the stationary case, arising e.g. when estimating models using first differenced data.

Fourth, the time series dimension T and the cross-sectional dimension n can be either fixed or large. We develop our limit theory by considering cases where the time series dimension T and the number of units n are large and we also include the case of when either n or T is fixed[1].

A short overview of the results we find under the conditions mentioned above is reported in Table 1 here.

[1] It is important to notice that the notion of fixed or "small" n or T is not well specified. Pesaran (2005b) cites $n < 10$ as the case when the number of cross sectional units is small. More generally, one could think as fixed n or T a number of cross sectional units or time series observations such that the cross sectional or the time series average is still faraway from the asymptotic limit, but such definition depends on the degree of cross sectional dependence or serial correlation in the panel and is therefore of scarce operational use.

Table 1: Consistency and limiting distribution of $\hat{\beta}_{OLS}$: $\quad y_{it} = \alpha_i + \beta' F_t + u_{it}$,

F_t known			F_t unknown		
(n, T)	Consistent	Limiting Distribution	(n, T)	Consistent	Limiting Distribution
Cointegration: $u_{it} \sim I(0)$					
Fixed n $T \to \infty$	Yes	Mixed Normal (Eq.10)		Yes	Non Standard (Eq. 47)
Fixed T $n \to \infty$	Yes	Mixed Normal (Eq.14)		Yes	Mixed Normal (Eq.14)
$(n, T) \to \infty$	Yes	Mixed Normal (Eq.18)			
			$\sqrt{n}/T \to 0$ $\sqrt{T}/n \to 0$	Yes Yes	Mixed Normal (Eq. 32) Non Standard (Eq. 34)
Spurious Regression: $u_{it} \sim I(1)$					
Fixed n $T \to \infty$	No	Non Standard (Eq. 12)		No	Non Standard (Eq. 49)
Fixed T $n \to \infty$	Yes	Non Standard (Eq. 16)		Yes	Non Standard (Eq. 16)
$(n, T) \to \infty$	Yes	Non Standard (Eq. 20)			
			$\sqrt{n}/T \to 0$ $T/\sqrt{n} \to 0$ $\sqrt{n}/T \to 0$	Yes	Non Standard (Eq. 36)
			$T^2/\sqrt{n} \to 0$	Yes	Non Standard (Eq. 34)
First Differences: $\Delta y_{it} = \beta' \Delta F_t + \Delta u_{it}$.					
Fixed n $T \to \infty$	Yes	Normal (Eq. 22)		No	Degenerate (Eq. 51)
Fixed T $n \to \infty$	Yes	Mixed Normal (Eq. 24)		Yes	Mixed Normal (Eq. 24)
$(n, T) \to \infty$	Yes	Normal (Eq. 26)			
			$n/T \to 0$ $T/n \to 0$	Yes Yes	Degenerate (Eq. 38) Degenerate (Eq. 40)

The remainder of the paper is organized as follows. Section 2 introduces the model, including a discussion about the relationship with existing models in the literature, and the main assumptions. The main contribution is reported in Section 3, where we analyze both the cases of known factors (section 3.1) and unknown factors (section 3.2), distinguishing the cases of large n and T, finite T and large n and finite n and large T . Section 4 concludes. Proofs are reported in the Appendix.

Notation is fairly standard. Throughout we use $\|A\|$ to denote the Euclidean norm $\sqrt{tr\,(A'A)}$, $\otimes$ for the Kronecker product, "$\to$" to indicate the ordinary limit, "$\Rightarrow$" to denote weak convergence, "$\gg$" to denote much greater, "$\xrightarrow{p}$" to denote convergence in probability, "$=$" to denote definitional equivalence. Stochastic processes such as $B\,(r)$ on $[0, 1]$ are usually written as B, integrals such as $\int_0^1 B\,(r)\,dr$ as $\int B$ and stochastic integrals such as $\int_0^1 B\,(r)\,dB\,(r)$ as

$\int BdB$.

2 MODEL SPECIFICATION AND ASSUMPTIONS

Consider the following panel regression model with common and idiosyncratic shocks

$$y_{it} = \alpha_i + \beta' F_t + \gamma' x_{it} + u_{it} \tag{1}$$

$i = 1, ..., n$, $t = 1, ..., T$, where β and γ are $(k \times 1)$ and $(p \times 1)$ vectors of slope parameters, respectively, $F_t = (F_{1t}, ..., F_{kt})'$ is a $k \times 1$ vector of common shocks

$$F_t = F_{t-1} + \varepsilon_t.$$

x_{it} is a $(p \times 1)$ vector of observable $I(1)$ individual-specific regressors,

$$x_{it} = x_{it-1} + \epsilon_{it}$$

and u_{it} and ϵ_{it} are the error terms.

The main interest of this paper is on the estimation of the common slope coefficients, β, and thus we do not lose in generality if we restrict our analysis to the model

$$y_{it} = \alpha_i + \beta' F_t + u_{it}, \tag{2}$$

Equation (2) could be either a spurious regression or a cointegration relationship depending on whether u_{it} is $I(1)$ or $I(0)$, respectively. In this paper, we analyze both cases. When common shocks are not observable, we assume that a set of exogenous variables, z_{it}, is observable such that

$$z_{it} = \lambda_i' F_t + e_{it} \tag{3}$$

where λ_i is a vector of factor loadings and e_{it} is the idiosyncratic component. It is important to point out the link between the model consider in (2) here and those in the literature. Model (2) is in the class of nonstationary panel models (see Baltagi and Kao (2000) and Breitung and Pesaran (2005) for a survey) and may also be motivated by Bai (2003, 2004) and Stock and Watson (1999, 2002, 2005). Bai (2003, 2004) assumes β different across i (β_i) where β_i is the loading and F_t is the common factor. In our set up, equation (2) represents a panel regression with common shocks F_t (e.g., macroeconomic variables, latent factors), as opposed to factor-loading specifications such as in Bai (2003, 2004). In Stock and Watson's (2002) setup, y_{it} in (2) (with $n = 1$) is the time series variable to be forecasted and $z_i = (z_{i1}, z_{i2}, ..., z_{iT})'$ is a n-dimensional multiple time series of candidate predictors.

The limiting theory for the OLS estimator for β that we use here depends both on cross-sectional (n) and the time series (T) dimensions, where in the factor analysis setup (e.g., Bai (2003, 2004)), the estimates of β_i only depends on the time series dimension T. More importantly, the common shocks F_t in (2), the same across units, induces cross-sectional dependence. This issue of contemporaneous correlation in panels with common regressors has not fully explored in the panel literature, the only exception being Bai (2005), who considers a concentrated least-squares estimator, and Jin (2005), who considers a maximum likelihood estimator for a discrete choice nonstationary panels.

Finally, it is worth mentioning that Bai and Kao (2005) study panel cointegration with a factor structure in the error terms (not in the regressors). Pesaran (2005a) proposes an estimator that allows for multiple factor error structure. Andrews (2005) studies the ordinary least squares (OLS) estimator with cross-sectional dependence though only in the context of cross section data.

For estimation and inference purposes, model (2) may be also rewritten in first-differenced form:

$$\Delta y_{it} = \beta' \Delta F_t + \Delta u_{it}. \tag{4}$$

The OLS estimator for β in equation (2) is given by:

$$\hat{\beta} = \left[\sum_{i=1}^{n} \sum_{t=1}^{T} \left(F_t - \bar{F}\right) \left(F_t - \bar{F}\right)' \right]^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(F_t - \bar{F}\right) y_{it} \tag{5}$$

where $\bar{F} = T^{-1} \sum_{t=1}^{T} F_t$, or, when using equation (4), by:

$$\hat{\beta}^{FD} = \left[\sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta F_t' \right]^{-1} \left[\sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta y_{it} \right]. \tag{6}$$

The following set of assumptions are used throughout the paper:

Assumption 1: (a) either (i) (*cointegration case*) $u_{it} = D_i(L)\eta_{it}$, or (ii) (*spurious regression case*) $\Delta u_{it} = F_i(L)\eta_{it}$ with $F_i(1) \neq 0$ and such that $\sum_i u_{it} \sim I(1)$; for both cases, $\eta_{it} \sim iid\left(0, \sigma_{\eta i}^2\right)$, with $E\left|\eta_{it}\right|^8 < M$, $\sum_{j=0}^{\infty} j\left|D_{ij}\right| < M$, $\sum_{j=0}^{\infty} j\left|F_{ij}\right| < M$ and $D_i^2(1)\sigma_{\eta i}^2 > 0$, $F_i^2(1)\sigma_{\eta i}^2 > 0$; (b) (time series and cross sectional correlation) $E\left(u_{it}u_{js}\right) = \tau_{ij,ts} = \tau_{ij,|t-s|}$ and $E\left(\Delta u_{it}\Delta u_{js}\right) = \gamma_{ij,ts} = \gamma_{ij,|t-s|}$, with

$$\frac{1}{nT} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{t=1}^{T} \sum_{s=1}^{T} \left|\tau_{ij,ts}\right| < M$$

and $(nT)^{-1} \sum_i \sum_j \sum_t \sum_s \left|\gamma_{ij,ts}\right| < M$.

Assumption 2: $\varepsilon_t = C(L) w_t$ where $C(L) = \sum_{j=0}^{\infty} C_j L^j$; (a) $w_t \sim iid\left(0, \Sigma_u\right)$ with $E\left\|w_t\right\|^{4+\delta} \leq M$ for some $\delta > 0$; (b) $Var\left(\Delta F_t\right) = \Sigma_{\Delta F} = \sum_{j=0}^{\infty} C_j \Sigma_u C_j'$ is a positive definite matrix; (c) the eigenvalues of $\Sigma_{\Delta F}$ are distinct; (d) $\sum_{j=0}^{\infty} j\left\|C_j\right\| < M$ and (e) $C(1)$ has full rank.

Assumption 3: (a) $E\|F_0\|^4 \leq M$ and $E|u_{i0}|^4 \leq M$, and (b) let B_ε be the Brownian motion associated with the partial sums of ε_t with covariance matrix $\Omega_{\varepsilon\varepsilon}$; we assume that the eigenvalues of the random matrix $\int B_\varepsilon B_\varepsilon'$ are distinct with probability 1.

Assumption 4: The loadings λ_i are non random quantities such that *(a)* $\|\lambda_i\| \leq M$; *(b)* either $n^{-1}\sum_{i=1}^{n}\lambda_i\lambda_i' = \Sigma_\Lambda$ if n is finite, or $\lim_{n\to\infty} n^{-1}\sum_{i=1}^{n}\lambda_i\lambda_i' = \Sigma_\Lambda$, if $n \to \infty$; in both cases, the matrix Σ_Λ is positive definite and such that the eigenvalues of $\Sigma_\Lambda\Sigma_{\Delta F}$ are distinct.

Assumption 5: $e_{it} = G_i(L)\nu_{it}$ where (a) $\nu_{it} \sim iid\left(0, \sigma_{vi}^2\right)$, $E|v_{it}|^8 < M$, $\sum_{j=0}^{\infty} j|G_{ij}| < M$ and $G_i^2(1)\sigma_{vi}^2 > 0$; (b) $E(\nu_{it}\nu_{jt}) = \tau_{ij}$ with $\sum_{i=1}^{n}|\tau_{ij}| \leq M$ for all j; (c) $E\left|n^{-1/2}\sum_{i=1}^{n}\left[e_{is}e_{it} - E(e_{is}e_{it})\right]\right|^4 \leq M$ for every (t,s); (d) $E\left[n^{-1}\sum_{i=1}^{n}e_{it}e_{is}\right] = \gamma_{s-t}$, $\left|\gamma_{s-t}\right| \leq M$ for all s and $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}\left|\gamma_{s-t}\right| \leq M$; (e) $E|e_{i0}|^4 \leq M$.

Assumption 6: $\{\varepsilon_t\}$, $\{u_{it}\}$ and $\{e_{it}\}$ are three independent groups.

Assumption 1(a) considers the possibility that equation (2) is either a cointegration or a spurious regression. Processes u_{it} and Δu_{it} are assumed to be invertible MA processes as in Bai (2004) and Bai and Ng (2004), in a similar fashion to processes ε_t and e_{it}. Assumption 1(b) also considers the presence of some, limited, cross sectional dependence among the u_{it}s or the Δu_{it}s and therefore it rules out the possibility that all the cross sectional dependence is taken into account by the common factors F_t - see the related work by Conley (1999).

Even if it refers to a different framework (panel data with common shocks as opposed to factor models), we take a position similar to that in Bai (2003, 2004) and Bai and Ng (2002, 2004). Using the factor models terminology, this means having a model with an "approximate factor structure", e.g., see the discussion in Chamberlain and Rotschild (1983) and Onatski (2005) - which differs from a strict common factor model where the u_{it}s are assumed to be independent across i.

The amount of cross sectional dependence we allow for in Assumption 1(b) is anyway limited, since we have that both $\sum_i\sum_j|\tau_{ij,ts}|$ and $\sum_i\sum_j\left|\gamma_{ij,ts}\right|$ are $O(n)$ for all pairs (s,t); such requirement is conceptually analogous to the condition of absolute summability of autocovariances in the time series framework, and it allows for both a Law of Large Numbers and a Central Limit Theorem to hold for the (rescaled) sequences $\sum_{i=1}^{n}u_{it}$ and $\sum_{i=1}^{n}\Delta u_{it}$, since the variances of both sequences are bounded as $n \to \infty$.

Assumption 2 requires that both the short run and the long run variance of ΔF_t are positive definite (Assumptions 2*(b)* and 2*(e)*, respectively), and therefore the possibility of having cointegration among factors is ruled out. Assumption 2 allows for some weak serial correlation in the dynamics of ε_t. This process can be described as invertible MA process, implied by the absolute summability conditions.

Assumption 3*(a)* is a standard initial condition requirement. In Assumption 3*(b)* the covariance matrix of B_ε, $\Omega_{\varepsilon\varepsilon}$, is positive definite, as ensured by Assumption 2*(e)*. Assumption 4 serves to identify the factors, which, merely for the purpose of a concise discussion, are assumed to be non random. This requirement could be relaxed, as in Bai (2003, 2004) and Bai and Ng (2004), assuming that the λ_is are randomly generated and independent of ε_t and e_{it}, and our results would keep holding. Assumption 4*(b)* ensures that the factor structure is identifiable. Note that it would be possible to relax this assumption by constraining the minimum eigenvalue of $\sum_{i=1}^n \lambda_i \lambda_i'$ to tend to infinity as $n \to \infty$, as pointed out by Onatski (2005). This structure would allow factors to be less pervasive than in our framework, thereby allowing the idiosyncratic component e_{it} in equation (3) to have a greater impact in explaining the contemporaneous correlation among the z_{it}s. Nonetheless, this would be made at the price of losing the possibility to model the z_{it} as a serially correlated process, whilst in our framework some limited time series and cross sectional dependence in model (3) is allowed for - as one could realize from Assumption 5. As pointed out in Bai (2003), the conditions in Assumption 5 are fairly general and allow for consistency and distribution results to hold for the principal component estimator.

Assumption 6 also rules out the existence of any form of dependence between factors F_t and u_{it}. Therefore, it is a stronger requirement than the simple lack of correlation, and we need it in order to prove the main results in our paper.

The following definitions are employed throughout the paper. $\bar{B}_\varepsilon(r)$ is the demeaned Brownian motion associated to the partial sums of F_t, i.e., $\bar{B}_\varepsilon(r) = B_\varepsilon(r) - \int_0^1 B_\varepsilon(r)\,dr$. Let h_i (h_i^Δ) and h_{ij} (h_{ij}^Δ) be the long run variance for u_{it} (Δu_{it}) and the long run covariance between processes u_{it} and u_{jt} (Δu_{it} and Δu_{jt}) - we have $h_{ij} = \sum_{t=1}^T \sum_{s=1}^T \tau_{ij,ts}$ and $h_{ij}^\Delta = \sum_{t=1}^T \sum_{s=1}^T \gamma_{ij,ts}$. Also, let $\bar{h} = \lim_{n\to\infty} n^{-1} \sum_{i=1}^n \sum_{j=1}^n h_{ij}$ and $\bar{h}^\Delta = \lim_{n\to\infty} n^{-1} \sum_{i=1}^n \sum_{j=1}^n h_{ij}^\Delta$. Last, the following variances arising from cross sectional aggregation of the u_{it}s and the Δu_{it}s are often used in our results: $\bar{\tau}_{ts} = \lim_{n\to\infty} n^{-1} \sum_{i=1}^n \sum_{j=1}^n \tau_{ij,ts}$, and $\bar{\gamma}_{ts} = \lim_{n\to\infty} n^{-1} \sum_{i=1}^n \sum_{j=1}^n \gamma_{ij,ts}$.

3 ASYMPTOTIC THEORY

The main objective of this paper is to derive the rate of convergence and limiting distribution of $\hat{\beta}$ and $\hat{\beta}^{FD}$, by considering several features of (2) and (4):

1. the factors F_t can either be known or (more likely) unobservable. The asymptotics of $\hat{\beta}$ and $\hat{\beta}^{FD}$ are affected by the estimation errors if we replace F_t by its estimate $\widehat{F}_t$;

2. the relationship described by equation (2) can be either a cointegration relationship or a spurious regression. As pointed out by Kao (1999) and Phillips and Moon (1999), convergence is obtained at rate $\sqrt{n}$ in panel spurious regression models and $\sqrt{n}T$ for panel cointegrated models. In

this paper, we are going to face a similar issue, which is compounded by the presence of common shocks in the panel regression (2);

3. the time series dimension T and the cross-sectional dimension n can be either fixed or large. Asymptotics are likely to change depending on whether one considers either dimension T or n large, keeping the other one fixed, or whether both n and T are allowed to tend to infinity.

We first start with the exploration of the case of known common shocks (Section 3.1) and then move to the case of unknown common shocks (Section 3.2).

3.1 Known F_t

In the case when F_t is known we have:

$$\hat{\beta} - \beta = \left[\sum_{i=1}^{n} \sum_{t=1}^{T} W_t W_t' \right]^{-1} \left[\sum_{i=1}^{n} \sum_{t=1}^{T} W_t u_{it} \right], \tag{7}$$

where $W_t = F_t - \bar{F}$, and

$$\hat{\beta}^{FD} - \beta = \left[\sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta F_t' \right]^{-1} \left[\sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta u_{it} \right]. \tag{8}$$

The convergence rate and the limiting distribution for $\hat{\beta}$ are now stated in the following theorem.

Theorem 1 *Suppose Assumptions 1-6 hold, and let $Z \sim N(0, I_k)$ be independent of the σ-field generated by the common shocks F_t.*
For fixed n and $T \to \infty$

$$\hat{\beta} - \beta = O_p\left(T^{-1}\right) \tag{9}$$

$$T\left(\hat{\beta} - \beta\right) \Rightarrow \frac{1}{n} \left(\int \bar{B}_\varepsilon \bar{B}_\varepsilon' \right)^{-1/2} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij} \right)^{1/2} Z \tag{10}$$

if equation (2) is a cointegration relationship, and

$$\hat{\beta} - \beta = O_p(1), \tag{11}$$

$$\left(\hat{\beta} - \beta\right) \Rightarrow \left(\int \bar{B}_\varepsilon \bar{B}_\varepsilon' \right)^{-1} \left(\int \bar{B}_\varepsilon B_u \right) \left(\sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^{\Delta} \right)^{1/2} \tag{12}$$

if (2) is a spurious regression.
For fixed T and $n \to \infty$, we have

$$\hat{\beta} - \beta = O_p\left(n^{-1/2}\right), \tag{13}$$

$$\sqrt{n}\left(\hat{\beta}-\beta\right) \Rightarrow \left(\sum_{t=1}^{T} W_t W_t'\right)^{-1} \left(\sum_{t=1}^{T}\sum_{s=1}^{T} W_t W_s' \bar{\tau}_{ts}\right)^{1/2} Z, \tag{14}$$

if (2) is a cointegration regression, whilst if it is a spurious relationship we have

$$\hat{\beta}-\beta = O_p\left(n^{-1/2}\right), \tag{15}$$

$$\sqrt{n}\left(\hat{\beta}-\beta\right) \Rightarrow \left(\sum_{t=1}^{T} W_t W_t'\right)^{-1} \left(\sum_{t=1}^{T} W_t \bar{u}_t\right), \tag{16}$$

where $\bar{u}_t = \lim_{n\to\infty} n^{-1/2} \sum_{i=1}^{n} u_{it}$.
When $(n,T)\to\infty$, *one has*

$$\hat{\beta}-\beta = O_p\left(n^{-1/2}T^{-1}\right), \tag{17}$$

$$\sqrt{n}T\left(\hat{\beta}-\beta\right) \Rightarrow \left(\int \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1/2} \sqrt{\bar{h}} Z, \tag{18}$$

if equation (2) is a cointegration relationship and

$$\hat{\beta}-\beta = O_p\left(n^{-1/2}\right), \tag{19}$$

$$\sqrt{n}\left(\hat{\beta}-\beta\right) \Rightarrow \left(\int \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1} \left(\int \bar{B}_\varepsilon B_u\right) \sqrt{\bar{h}^{\Delta}}, \tag{20}$$

if it is a spurious regression.

Proof. See Appendix. ■

Equations (9)-(12) are the standard superconsistency and inconsistency results in the literature. With respect to the speed of convergence, when $(n,T)\to\infty$ our results in equations (17) and (19) lead to the same orders as in Phillips and Moon (1999) and Kao (1999) for both the cointegration and the spurious regression case. Consistency is achieved under the spurious regression case as well, where the rate of convergence is $\sqrt{n}$. This result, which follows the seminal contributions of Kao (1999) and Phillips and Moon (1999), is reinforced for the case when T is fixed and $n\to\infty$. Equations (13) and (15) prove that irrespective of model (2) to be a cointegration regression or a spurious regression, large n allows for consistency to hold. It is worth observing the complicated distribution that arises when T is finite; this is essentially due, as outlined in the proof, to the presence of serial correlation in the u_{it}s.

For the case of n and T large, the rate of convergence for $\hat{\beta}$ is the same as in Phillips and Moon (1999) under the case of contemporaneous independence across units, but the limiting distributions in equations (18) and (20) differ and are mixed normal rather than normal as in the Phillips and Moon (1999) case. The mixed normality is due to both F_t being nonstationary and common across units, as can be seen by considering equation (14) for $T\to\infty$.

The design matrix $\left(nT^2\right)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}F_tF_t'=T^{-2}\sum_{t=1}^{T}F_tF_t'$ converge in distribution to a random matrix, namely $\int\bar{B}_\varepsilon\bar{B}_\varepsilon'$, rather than to a constant. Of course, $\left(nT^2\right)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}F_tF_t'$ would converge to a constant (in probability) if F_t were not common shocks, i.e., if F_t were replaced by, say, F_{it}.

The convergence rates and the limiting distributions for $\hat{\beta}^{FD}$ are reported in the following theorem.

Theorem 2 *Suppose Assumptions 1-6 hold and let $Z\sim N\left(0,I_k\right)$ be independent of the σ-field generated by ΔF_t.*
For fixed n and $T\rightarrow\infty$

$$\hat{\beta}^{FD}-\beta=O_p\left(T^{-1/2}\right), \tag{21}$$

$$\sqrt{T}\left(\hat{\beta}^{FD}-\beta\right)\Rightarrow n^{-1}\Sigma_{\Delta F}^{-1/2}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}h_{ij}^{\Delta}\right)^{1/2}Z. \tag{22}$$

For T fixed and $n\rightarrow\infty$, we have

$$\hat{\beta}^{FD}-\beta=O_p\left(n^{-1/2}\right), \tag{23}$$

$$\sqrt{n}\left(\hat{\beta}^{FD}-\beta\right)\Rightarrow\left(\sum_{t=1}^{T}\Delta F_t\Delta F_t'\right)^{-1}\left(\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta F_t\Delta F_s'\bar{\gamma}_{ts}\right)Z. \tag{24}$$

When $(n,T)\rightarrow\infty$, one has

$$\hat{\beta}^{FD}-\beta=O_p\left(n^{-1/2}T^{-1/2}\right), \tag{25}$$

$$\sqrt{nT}\left(\hat{\beta}^{FD}-\beta\right)\Rightarrow\Sigma_{\Delta F}^{-1/2}\sqrt{\bar{h}^{\Delta}}Z. \tag{26}$$

Proof. See Appendix. ■

The results were derived for the case of no serial correlation. The presence of time dependence in general involves a more complicated expression of the limiting distributions, but rates of convergence would not be affected. Note also that since the first differenced model is always stationary, irrespective of whether equation (2) is a cointegration equation or a spurious regression, one can always apply the CLT to obtain the limiting distribution of $\hat{\beta}^{FD}-\beta$; this is indirectly shown by the rate of convergence for the case when $(n,T)\rightarrow\infty$, equal to $\sqrt{nT}$.

It is worth noticing the remarkable result in equation (26): one would expect the limiting distribution of $\hat{\beta}^{FD}-\beta$ to be mixed normal given the strong dependence across units due to the terms $\Delta F_t\Delta u_{it}$ sharing the common element ΔF_t across i, as we showed in (24) with large n and fixed T. However, the common shocks are found not to play any role in the case of large n and large T. This result is discussed thoroughly in the proofs of Theorems 1 and 2. and can also be seen in equation (24) which gives the limiting distribution for T fixed and $n \to \infty$. The design matrix $T^{-1}\sum_{t=1}^{T}\Delta F_t \Delta F_t'$ is a random matrix for all finite values of T. However, standard application of the LLN (its validity is ensured by Assumption 2) shows that the design matrix converges to a constant matrix as $T \to \infty$. Therefore, the mixed normality arising for finite T is wiped away by the smoothing over time as well. Asymptotic normality is therefore determined merely by design matrix $T^{-1}\sum_{t=1}^{T}\Delta F_t \Delta F_t'$ being constant asymptotically.

3.2 Unknown F_t

We turn now to the case when common shocks are unknown and thus they need to be estimated. The asymptotics of $\hat{\beta}$ and $\hat{\beta}^{FD}$ are affected by the errors in estimating factors F_t.

Let $\hat{F}_t$ be an estimate of the factor. Denote $\hat{W}_t = \hat{F}_t - T^{-1}\sum_{t=1}^{T}\hat{F}_t$. Estimations of β using the model in levels ($\hat{\beta}$) or first differences ($\hat{\beta}^{FD}$) respectively are now given by:

$$\hat{\beta} = \left[\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\hat{W}_t'\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t y_{it}\right], \tag{27}$$

and

$$\hat{\beta}^{FD} = \left[\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta\hat{F}_t'\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta y_{it}\right], \tag{28}$$

with estimation errors:

$$\hat{\beta} - \beta = \left[\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\hat{W}_t'\right]^{-1}\left\{\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\left[\left(W_t - \hat{W}_t\right)'\beta + u_{it}\right]\right\}, \tag{29}$$

$$\hat{\beta}^{FD} - \beta^{FD} = \left[\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta\hat{F}_t'\right]^{-1}\left\{\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\left[\left(\Delta F_t - \Delta\hat{F}_t\right)'\beta + \Delta u_{it}\right]\right\}. \tag{30}$$

In what follows, for the purpose of a concise discussion, we assume the number of factors k to be known[2]. We like to emphasize that this is does not lead to any loss of generality since the distribution of the estimated factors does

[2] An issue of importance that arises within this framework and that needs tackling prior to estimating the common components F_t is to determine their number, k. In light of some recent contributions, e.g., see Bai and Ng (2002) and Onatski (2005), it is natural to refer to model (3) in order to extract both the common factors F_t and their number k. It is worth pointing out though that determining k crucially depends on whether both n and T are large or if either dimension is fixed. Under all cases, the literature provides methodologies to estimate k consistently, i.e. to obtain an estimate $\hat{k}$ such that, as either $(n, T) \to \infty$ or, alternatively,

not depend on whether k is known or estimated, and therefore the estimation error that arises from using $\hat{k}$ instead of k does not play any role as long as $\hat{k}$ is consistent, e.g., see Bai (2003, p. 143, note 5) for an elegant proof of this statement.

3.2.1 The case of n and T large

In this section, we estimate the common shocks F_t using the principal component estimator. This means minimizing either

$$V_b(k) = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(z_{it} - \lambda_i' F_t\right)^2,$$

when considering F_t in levels, or

$$V_a(k) = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\Delta z_{it} - \lambda_i' \Delta F_t\right)^2$$

when estimating factors ΔF_t from the first differenced version of model (3). Consider the $T \times n$ matrix $Z = (z_1, ..., z_T)'$, and the $T \times k$ matrix of factors $F = (F_1, F_2, ..., F_T)'$. Then each objective function $V_a(k)$ or $V_b(k)$ can be minimized by concentrating out λ and using the normalizations $\Delta F' \Delta F / T = I_k$ or $F'F/T^2 = I_k$. The estimated factor matrices, denoted as $\Delta \hat{F}$ and $\hat{F}$, are $\sqrt{T}$ times eigenvectors corresponding to the k largest eigenvalues of the $T \times T$ matrices $\Delta Z \Delta Z'$ or ZZ'. It is well known that the solutions to the above minimization problems are not unique, e.g., when estimating factors ΔF_t and F_t, these are not directly identifiable even though they are up to a transformation. In our setup, the knowledge of $H_1 \Delta F_t$, $H_1 F_t$ and $H_2 \lambda_i$ is as good as knowing ΔF_t, F_t, and λ_i. For sake of notational simplicity, in what follows we shall assume that H_1 $(k \times k)$ and H_2 $(n \times n)$ are identity matrices.

The convergence rate and the limiting distribution for $\hat{\beta}$ are in the following theorem.

Theorem 3 *Suppose Assumptions 1-6 hold.*
Let equation (2) be a cointegration relationship:
if $\sqrt{n}/T \to 0$

$$\hat{\beta} - \beta = O_p\left(n^{-1/2} T^{-1}\right), \tag{31}$$

$\max\{n, T\} \to \infty$ and $\min\{n, T\}$ is fixed, it holds that $P\left[\hat{k} = k\right] = 1$ and $P\left[\hat{k} \neq k\right] = o_p(1)$. Most often these methods treat estimation of k as a either model selection or a rank estimation problem, thereby employing some information criteria. For the case of either n or T fixed, the contributions by Lewbel (1991), Donald (1997) and Cragg and Donald (1997) ensure consistent estimation of the either the rank of the $n \times n$ matrix $\sum_t z_t z_t'$ with $z_t \equiv [z_{1t}, ..., z_{nt}]'$ or of the $T \times T$ matrix $\sum_i z_i z_i'$ with $z_i \equiv [z_{i1}, ..., z_{iT}]'$, depending on whether n or T is fixed. When $(n, T) \to \infty$, the aforementioned procedures are no longer usable to obtain a consistent $\hat{k}$ and Bai and Ng (2002) propose a consistent estimator for k - see also Onatski (2005). Note that assumptions 2-6 in our settings ensure the applicability of these methods to equation (3), as it can be immediately verified.

$$\sqrt{n}T\left(\hat{\beta}-\beta\right)\Rightarrow\left(\int\bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\right)^{-1/2}\left[\sqrt{\bar{h}}Z_1+\sqrt{\beta'\tilde{Q}_B\Gamma\tilde{Q}_B'\beta}Z_2\right]; \tag{32}$$

if $T/\sqrt{n}\rightarrow 0$

$$\hat{\beta}-\beta=O_p\left(T^{-2}\right), \tag{33}$$

$$T^2\left(\hat{\beta}-\beta\right)\Rightarrow\frac{1}{2}\sigma_e^2\left[\int\bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\right]^{-1}\Omega_{\varepsilon\varepsilon}; \tag{34}$$

where $Z_1\sim N_1\left(0,I_k\right)$ *and* $Z_2\sim N_2\left(0,I_k\right)$ *are independent, the random matrix* $\tilde{Q}_B$ *is defined as*

$$T^{-2}\sum_{t=1}^{T}\hat{W}_tW_t'\Rightarrow\tilde{Q}_B,$$

and

$$\Gamma=\lim_{n\rightarrow\infty}n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j'E\left(e_{it}e_{jt}\right),$$

$$\sigma_e^2=\lim_{n\rightarrow\infty}\frac{1}{n}\sum_{i=1}^{n}\sigma_{ei}^2,$$

where σ_{ei}^2 *is the long run variance of process* $\{e_{it}\}$.
Let equation (2) be a spurious regression:
if $\sqrt{n}/T\rightarrow 0$, *or* $T/\sqrt{n}\rightarrow 0$ *and* $\sqrt{n}/T^2\rightarrow 0$

$$\hat{\beta}-\beta=O_p\left(n^{-1/2}\right), \tag{35}$$

$$\sqrt{n}\left(\hat{\beta}-\beta\right)\Rightarrow\left(\int\bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\right)^{-1}\left(\int\bar{B}_{\varepsilon}B_u\right)\sqrt{\bar{h}^{\Delta}}; \tag{36}$$

if $T^2/\sqrt{n}\rightarrow 0$, *then (33) and (34) hold.*

Proof. See Appendix. ■

Consistency is ensured in both cases, even though $T/\sqrt{n}\rightarrow 0$ results in a slower (than in the case of $\sqrt{n}/T\rightarrow 0$) rate of convergence and in a degenerate behavior of the numerator of $\hat{\beta}-\beta$. This is anyway not surprising given that the factors estimation errors (see Bai and Ng (2002), and Bai (2004) can be decomposed in several terms of different asymptotic stochastic magnitude, which have an impact only on the numerator. Notice the consequence of equation (2) being a spurious regression: as long as the number of cross sectional units is not exceedingly large, the classical $\sqrt{n}$ consistency holds, and we have the same limiting distribution as in equation (20). When n is far larger than T, we have the same result as if relationship (2) were a cointegration relationship.

See below for the case when $\sqrt{n}/T$ tends to a constant.

The convergence rate and the limiting distribution for $\hat{\beta}^{FD}$ are in the following theorem.

Theorem 4 *Suppose Assumptions 1-2 and 4-6 hold.*
If $\frac{n}{T} \to 0$

$$\hat{\beta}^{FD} - \beta^{FD} = O_p\left(n^{-1/2}T^{-1/2}\right), \tag{37}$$

$$\sqrt{nT}\left(\hat{\beta}^{FD} - \beta^{FD}\right) \xrightarrow{p} \Sigma_{\Delta F}^{-1} Q V^{-1}\beta \tag{38}$$

where V *is the probability limit of the diagonal matrix consisting of the first* k *eigenvalues of* $(nT)^{-1}\Delta Z \Delta Z'$ *in decreasing order, and*

$$Q = p\lim T^{-3/2}\sum_{s=1}^{T}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta\tilde{F}_s' n^{-1}\left[\sum_{i=1}^{n}\left(e_{it}e_{is} - \gamma_{s-t}\right)\right].$$

If $\frac{T}{n} \to 0$

$$\hat{\beta}^{FD} - \beta^{FD} = O_p\left(T^{-1}\right), \tag{39}$$

$$T\left(\hat{\beta}^{FD} - \beta^{FD}\right) \xrightarrow{p} \bar{h}_e V^{-1}\beta, \tag{40}$$

where $\bar{h}_e$ *is the long-run variance of the process* $\lim_{n\to\infty} n^{-1/2}\sum_{i=1}^{n} e_{it}$.

Proof. See Appendix. ■

Notice that in this case we have degenerate limiting distributions, despite having consistent estimates.

The condition $n/T \to 0$ means that T is much larger than n, which in turn implies a panel where time series observations outnumber the cross sectional units. In such a case, we still have consistency. The condition $T/n \to 0$ implies that the number of units n is far larger than T. In such a case, consistency is ensured, even though at a "slow" rate, given by T. In this case the impact of n becomes ineffective, just as in Bai (2003, 2004) and Bai and Ng (2002, 2004), where consistency depends on the minimum between T and n or some functions of them. Further, the distribution limit is degenerate, and therefore convergence in distribution can be achieved at a slower speed than $O_p\left(T^{-1}\right)$.

Finally, it can be observed that in both Theorems 3 and 4, the boundary cases $\sqrt{n}/T \to \tau$ or $n/T \to \tau'$ (for τ and τ' are constants), are implicitly analyzed. In these cases, the limit distributions are given by the sum of the limit distributions in equations (32)-(34) and (38)-(40), respectively..

3.2.2 The case of T fixed and n large

When T is fixed and n is large, consistent estimation of factors is still possible, see e.g. Connor and Korajzcyk (1986) and Bai (2003). However, the following restriction is necessary:

Assumption 7: $E\left(e_{it}e_{is}\right) = 0$ for all $t \neq s$.

Assumption 7 rules out the possibility of serial correlation in the data generating process of the e_{it}, and therefore this is a constraint on Assumption 5(d).

However, contemporaneous correlation and cross sectional heteroscedasticity are preserved.

Under Assumptions 4-7, we know that factors estimation is $\sqrt{n}$ consistent, i.e. we have both

$$\hat{F}_t - F_t = O_p\left(n^{-1/2}\right)$$

and

$$\Delta\hat{F}_t - \Delta F_t = O_p\left(n^{-1/2}\right)$$

for all t.

Theorems (5) and (6) do not anyway require $\sqrt{n}$ consistency, since they ensure the consistency of the OLS estimates $\hat{\beta}$ and $\hat{\beta}^{FD}$ for *any* consistent estimate of the factors, irrespective of the rate of convergence.

Theorem 5 *Suppose Assumptions 1-7 hold; then for every consistent estimator $\hat{F}_t$ of F_t and for fixed T and $n \to \infty$ we have the same results as in equations (13)-(16).*

Proof. See Appendix. ■

Theorem 6 *Suppose Assumptions 1-7 hold; then for every consistent estimator $\Delta\hat{F}_t$ of ΔF_t and for fixed T and $n \to \infty$ we have the same results as in equations (23) and (24).*

Proof. See Appendix. ■

In both cases we have the same results as we would have if the F_ts were observable. Therefore, when T is fixed, having large n makes it indifferent to use observed or estimated factors as long as factors are estimated consistently.

3.2.3 The case of n fixed and T large

In what follows, we provide a new inferential theory for the case when factors are unknown and the cross-sectional dimension n is finite. This case has not been explored in the literature, the only exception being Gonzalo and Granger (1995). Our contribution is aimed at making the estimated factors usable in a regression framework.

Rewriting model (3) in the vector form, one gets:

$$z_t = \Lambda F_t + e_t, \tag{41}$$

where $z_t = (z_{1t}, ..., z_{nt})'$, $e_t = [e_{1t}, ..., e_{nt}]'$, and $\Lambda = (\lambda_1, \lambda_2, ..., \lambda_n)'$. Here too one can estimate Λ using the principal components estimator. A feasible estimator of Λ, $\hat{\Lambda}$, is given by the $\sqrt{n}$ times the eigenvectors corresponding to the k largest eigenvalues of $Z'Z$. Notice that this estimator exploits the normalization $\hat{\Lambda}'\hat{\Lambda}/n = I_k$, and it turns out to be computationally convenient for the case of $n < T$. For sake of the notation, and without loss of generality, from Assumption 4 we assume henceforth that $n^{-1}\sum_{i=1}^{n}\lambda_i\lambda_i' = I_k$.

The following theorem characterizes consistency and limiting distribution of $\hat{\Lambda}$.

Proposition 1 *Under Assumptions 3-6 we have*

$$\hat{\Lambda} - \Lambda = O_p\left(T^{-1}\right), \tag{42}$$

$$\begin{aligned} T\left(\hat{\Lambda} - \Lambda\right) \quad \Rightarrow \quad & \left[I_n - n^{-1}\Lambda \int B_\varepsilon B_\varepsilon' \Lambda'\right] \left(\int dW_e B_\varepsilon'\right) \left(\int B_\varepsilon B_\varepsilon'\right)^{-1} \\ & - n^{-1}\Lambda \left(\int dW_e B_\varepsilon'\right) \Lambda \\ & + n^{-1}\left[I_n - 2n^{-1}\Lambda \int B_\varepsilon B_\varepsilon' \Lambda'\right] \Omega_e \Lambda, \end{aligned} \tag{43}$$

where W_e is the Wiener process associated to the partial sums of e_t and $\Omega_e = E\left(e_t e_t'\right)$.

Proof. See Appendix. ■

Note that in this case we have a T-consistent estimate of the factor loadings, even though the principal component estimator of F_t is not consistent (see Bai (2004) and Proposition 2 below) when n is finite.

Henceforth, for sake of notation, we refer to the limiting distribution of $T\left(\hat{\Lambda} - \Lambda\right)$ as D_Λ^1 , i.e. $T\left(\hat{\Lambda} - \Lambda\right) \Rightarrow D_\Lambda^1$. Given the restriction $\hat{\Lambda}'\hat{\Lambda}/n = I_k$, the OLS estimator of F_t, obtained regressing the z_ts on the estimated loadings $\hat{\Lambda}$, is

$$\hat{F}_t = n^{-1}\hat{\Lambda}' z_t.$$

The following Proposition characterizes (the inconsistency of) this estimator:

Proposition 2 *Consider $\hat{F}_t = n^{-1}\hat{\Lambda}' z_t$, and also the first difference estimator, $\Delta\hat{F}_t = n^{-1}\hat{\Lambda}'\Delta z_t$. Then*

$$\max_{1 \le t \le T} \left\| \hat{F}_t - F_t \right\| = O_p(1), \tag{44}$$

and

$$\max_{1 \le t \le T} \left\| \Delta\hat{F}_t - \Delta F_t \right\| = O_p(1) \tag{45}$$

uniformly in t.

Proof. See Appendix. ■

From Proposition 2 we note that the estimates of the factors and of their first difference are inconsistent. However this inconsistency has no impact on the consistency of $\hat{\beta}$ and $\hat{\beta}^{FD}$, though it affects their asymptotic law. See the proofs of Theorems 7 and 8.

The convergence rate and the limiting distribution for $\hat{\beta}$ are in the following theorem.

Theorem 7 *For the estimator $\hat{\beta}$, we have:*

$$\hat{\beta} - \beta = O_p\left(T^{-1}\right), \tag{46}$$

$$T\left(\hat{\beta} - \beta\right) \Rightarrow n^{-1}\left[\int \bar{B}_\varepsilon \bar{B}_\varepsilon'\right]^{-1} \left\{ \begin{array}{c} \int \bar{B}_\varepsilon dB_u \left(\sum_{i=1}^n \sum_{j=1}^n h_{ij}\right)^{1/2} - \left(1 - n^{-1}\right) \times \\ \left[\int \bar{B}_\varepsilon \bar{B}_\varepsilon' \beta + T^{-1} D_\Lambda^1 \Lambda \int \bar{B}_\varepsilon \bar{B}_\varepsilon' \beta + \Lambda' \int d\bar{B}_e \bar{B}_\varepsilon' \beta\right] + \\ n^{-1}\left[T^{-1} \int \bar{B}_\varepsilon \bar{B}_\varepsilon' \Lambda' D_\Lambda^{1\prime} \beta + \Lambda' \Sigma_e \Lambda \beta + \int \bar{B}_\varepsilon d\bar{B}_e' \Lambda \beta\right] \end{array} \right\}, \tag{47}$$

where $\bar{B}_e$ is the demeaned Brownian motion associated to the partial sums of e_t and $\Sigma_e = Var\left(e_t\right)$. When this is a spurious relationship, one gets

$$\hat{\beta} - \beta = O_p\left(1\right), \tag{48}$$

$$\hat{\beta} - \beta \Rightarrow \left(\int \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1} \left(\int \bar{B}_\varepsilon B_u\right) \left(\sum_{i=1}^n \sum_{j=1}^n h_{ij}^\Delta\right)^{1/2}. \tag{49}$$

Proof. See Appendix. ■

Note that even though common shocks cannot be estimated consistently, $\hat{\beta}$ is consistent when (2) represents a cointegration relationship but inconsistent when (2) represents a spurious regression. Factor estimation has an impact on the limit distribution of $\hat{\beta} - \beta$ when equation (2) is a cointegration regression - see equation (47) above. On the other hand, it does not affect the asymptotic distribution when equation (2) is a spurious regression - see equation (49). This can be seen comparing the two distribution limits with equations (10) and (12) respectively, where factors are assumed to be known.

Equations (47) and (49) show an important common feature of this theoretical framework. Only the numerators of equation (47) and (49) depend on whether equation (2) is a cointegrating or spurious regression, whilst the denominators are not affected. This is due to the fact (detailed in the proof) that though $\hat{F}_t$ is not a consistent estimator for F_t, the quantity $\sum \hat{F}_t \hat{F}_t'$ is a consistent estimator for $\sum F_t F_t'$ for *any* consistent estimator of the loadings $\hat{\Lambda}$.

The convergence rate and the limiting distribution for $\hat{\beta}^{FD}$ are in the following theorem.

Theorem 8 *For the first difference estimator $\hat{\beta}^{FD}$, we have:*

$$\hat{\beta}^{FD} - \beta^{FD} = O_p\left(1\right), \tag{50}$$

and

$$\hat{\beta}^{FD} - \beta^{FD} \xrightarrow{p} -\beta + n\left[\Lambda' \Sigma_{\Delta z} \Lambda\right]^{-1} \Sigma_{\Delta F} \beta, \tag{51}$$

where $\Sigma_{\Delta e} = Var\left(\Delta e_t\right)$ and $\Sigma_{\Delta z} = \Lambda \Sigma_{\Delta F} \Lambda' + \Sigma_{\Delta e}$.

Proof. See Appendix. ■

The estimator $\hat{\beta}^{FD}$ is inconsistent. As detailed in the proof, this is due to the two terms $\sum \Delta F_t \Delta F_t'$ and $\sum \Delta e_t \Delta e_t'$ having the same asymptotic order, rather than to the factor estimates being inconsistent. Also, this hold for any consistent estimator $\hat{\Lambda}$ (see discussion in the proof).

Theorems 7 and 8 hold if equation (3) represents a cointegration relationship. We now turn to evaluate the case of $e_{it} \sim I(1)$.

Extension to the case where (3) is a spurious regression

If equation (3) is a spurious regression, then factors have to be estimated using the approach in Bai (2003). Let us consider the first differenced version of equation (3), i.e.$\Delta z_t = \Lambda \Delta F_t + \Delta e_t$, and let $\Delta Z = [\Delta z_1, ..., \Delta z_t]'$. Then Λ can be estimated via principal components obtaining $\hat{\Lambda}^{FD}$, with normalization $\hat{\Lambda}^{FD\prime}\hat{\Lambda}^{FD}/n = I_k$. Under Assumptions 2 and 4-6, the estimated loadings are of asymptotic magnitude $\hat{\Lambda}^{FD} - \Lambda = O_p(1)$ - see Theorem 2 in Bai (2003).

The factors are estimated as

$$\hat{F}_t = n^{-1}\hat{\Lambda}^{FD\prime} z_t.$$

The rate of convergence and the limiting distribution of $\hat{\beta}$ and $\hat{\beta}^{FD}$ are reported in the following theorem.

Theorem 9 *Let equation (3) be a spurious regression and let D^2_Λ be the distribution limit of $\hat{\Lambda}^{FD}$, i.e. $\hat{\Lambda}^{FD} \Rightarrow D^2_\Lambda$. Then*

$$\hat{\beta} - \beta = O_p(1), \tag{52}$$

irrespective of whether equation (2) is a cointegration equation or spurious regression; let $\bar{B}_z$ be the demeaned Brownian motion associated to the partial sums of z_t; the distribution limit is equal to

$$\hat{\beta} - \beta \Rightarrow -\beta + n\left(D^{2\prime}_\Lambda \int \bar{B}_z \bar{B}_z' D^2_\Lambda\right)^{-1} D^{2\prime}_\Lambda \int \bar{B}_z \bar{B}_\varepsilon' \beta \tag{53}$$

when equation (2) is a cointegration relationship, whilst if it represents a spurious regression it holds that

$$\hat{\beta}-\beta \Rightarrow -\beta+n\left(D^{2\prime}_\Lambda \int \bar{B}_z \bar{B}_z' D^2_\Lambda\right)^{-1} D^{2\prime}_\Lambda \left[\int \bar{B}_z \bar{B}_\varepsilon' \beta + \int \bar{B}_z B_u \left(\sum_{i=1}^{n}\sum_{i=1}^{n} h^\Delta_{ij}\right)^{1/2}\right]. \tag{54}$$

With respect to $\hat{\beta}^{FD}$, we have:

$$\hat{\beta}^{FD} - \beta^{FD} = O_p(1), \tag{55}$$

$$\hat{\beta}^{FD} - \beta^{FD} \Rightarrow -\beta + n\left(D^{2\prime}_\Lambda \Sigma_{\Delta z} D^2_\Lambda\right)^{-1} D^{2\prime}_\Lambda \Lambda \Sigma_{\Delta F} \beta. \tag{56}$$

Proof. See Appendix. ■

When factors are estimated from a spurious regression, consistency of the OLS estimates of β is lost irrespective of whether model (2) is stationary or cointegrating or spurious relationship. $e_{it} \sim I(1)$ has also an impact on the limiting distributions, since in this case we have to take account the asymptotic law of $\hat{\Lambda}$ as well.

The rate of convergence of both $\hat{\beta}^{FD} - \beta^{FD}$ and $\hat{\beta} - \beta$ is $O_p(1)$, even though the estimated factors have different stochastic magnitudes as shown in Lemma 4 in Appendix.

4 CONCLUSION

This paper develops limiting theory for the OLS estimator for panel models with common shocks, where contemporaneous correlation is generated by both the presence of common regressors (e.g. macro shocks, aggregate fiscal and monetary policies) among units and weak spatial dependence among the error terms. We derived rates of convergence and limiting distributions under a comprehensive set of alternative characteristics of panels: different combinations of the cross-sectional dimension n and the time series dimension T ; factors being either observable or unobservable; and the main model representing either a cointegrating equation or a spurious regression.

When the common factors are observable, the OLS estimator always provides consistency. This result holds for all possible combinations of the dimensions of n and T, including the case of n fixed, which so far has not been addressed in the literature on non stationary panel factor models. The only exception being the case of spurious regression with fixed n. We extend the study of consistency of OLS estimators to the case when the factors are unobservable and we prove that consistency always holds, the only exceptions being the cases of spurious regression and stationary regression when n is fixed .

A central result is represented by the limiting distributions derived under the strong cross-sectional dependence induced by the presence of common shocks. In this case, we obtained mixed normality as consequence of the common shock being non stationarity, while when shocks are stationary genuinely normal distributions are obtained.

Appendix

Proof of Theorem 1. To prove the theorem, we refer to equation (7) that contains the estimation error $\hat{\beta}-\beta=\left[\sum_i\sum_t W_tW_t'\right]^{-1}\left[\sum_i\sum_t W_tu_{it}\right]$. The proof be derived splitting this quantity into the denominator $\sum_i\sum_t W_tW_t'$ and the numerator $\sum_i\sum_t W_tu_{it}$, and analyzing the asymptotic behavior of both quantities separately.

Let us start considering the denominator $\sum_i\sum_t W_tW_t'$. When $T\to\infty$ and n is fixed, we have from Assumptions 2 and 3 that under both the cases that equation (2) is a spurious regression or a cointegrating one it holds that $\sum_i\sum_t W_tW_t'=O_p\left(T^2\right)$ and

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}W_tW_t'\Rightarrow\int\bar{B}_\varepsilon\bar{B}_\varepsilon'. \tag{57}$$

As $n\to\infty$, and for fixed T, we have $\sum_i\sum_t W_tW_t'=O_p\left(n\right)$

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}W_tW_t'=\frac{1}{T^2}\sum_{t=1}^{T}W_tW_t', \tag{58}$$

whilst as both n and T are large we have $\sum_i\sum_t W_tW_t'=O_p\left(nT^2\right)$

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}W_tW_t'\Rightarrow\int\bar{B}_\varepsilon\bar{B}_\varepsilon'. \tag{59}$$

As far as the numerator is concerned, the proof be derived with respect to three separate cases, following the same structure as in the theorem. We firstly derive the rate of convergence and the limiting distribution of $\sum_i\sum_t W_tu_{it}$ for the case when T is large and n is fixed; we then study the opposite case, when T is fixed and n is large; last, we analyze the case when both T and n are large.

Case 1: large T and fixed n

We firstly focus our attention to the case where equation (2) is a cointegration relationship.

Denote

$$\xi_{nt}=T^{-1}W_t\left(\sum_{i=1}^{n}u_{it}\right)$$

and

$$\xi_{nT}=\sum_{t=1}^{T}\xi_{nt}.$$

Assumption 6 ensures that F_t and the u_{it}s are independent. Also, according to Assumption 1(a), the process $\sum_i u_{it}$ has covariance structure given by

$$E\left[\left(\sum_{i=1}^{n}u_{it}\right)\left(\sum_{i=1}^{n}u_{is}\right)\right]=\sum_{i=1}^{n}\sum_{j=1}^{n}\tau_{ij,ts}.$$

Then the absolutely summability condition on $\tau_{ij,ts}$ over time implied in Assumption 1(b), and Assumptions 2 and 3 ensure that a FCLT holds such that

$$\xi_{nT} \Rightarrow \int \bar{B}_{\varepsilon} dW,$$

where W is a Brownian motion with variance

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{n} \tau_{ij,ts} = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}.$$

An alternative way to write the limiting distribution of ξ_{nT} is

$$\xi_{nT} \Rightarrow \left(\sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij} \right)^{1/2} \left(\int \bar{B}_{\varepsilon} \bar{B}_{\varepsilon}' \right)^{1/2} Z,$$

where $Z \sim N(0, I_k)$.

Hence we have a twofold result. First, the rate of convergence of the numerator of $\hat{\beta} - \beta$ is $O_p(T)$; therefore, given equation (57) that ensures that the denominator of $\hat{\beta} - \beta$ is $O_p(T^2)$, we have that $\hat{\beta} - \beta = O_p(T^{-1})$, proving equation (9). As far as the distribution limit is concerned, we know, combining the asymptotic law of ξ_{nT} with equation (57), we have that

$$\left[\frac{1}{T^2} \sum_{i=1}^{n} \sum_{t=1}^{T} W_t W_t' \right]^{-1} \left[\frac{1}{T} \sum_{i=1}^{n} \sum_{t=1}^{T} W_t u_{it} \right] \Rightarrow \frac{1}{n} \left(\int \bar{B}_{\varepsilon} \bar{B}_{\varepsilon}' \right)^{-1/2} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij} \right)^{1/2} Z,$$

which proves equation (10). Note that independence between Z and $\bar{B}_{\varepsilon}$ is ensured by Assumption 6.

We can now consider the case when equation (2) is a spurious regression and therefore $u_{it} \sim I(1)$.

Define first $\xi_{nt}^{S} = T^{-2} W_t \left(\sum_{i=1}^{n} u_{it} \right)$ and $\xi_{nT}^{S} = \sum_{t=1}^{T} \xi_{nt}^{S}$. The process $\sum_{i=1}^{n} u_{it}$ is still a unit root process with long run variance given by $\sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^{\Delta}$. Therefore, a FCLT, which follows from Assumptions 1(a), 2 and 3, ensures that $\xi_{nT}^{S} = O_p(1)$. Together with equation (57), this proves that $\hat{\beta} - \beta = O_p(1)$, as reported in equation (11). As far as the limiting distribution is concerned, here the asymptotic law of the numerator of $\hat{\beta} - \beta$ is given by

$$\xi_{nT}^{S} = \frac{1}{T^2} \sum_{t=1}^{T} W_t \left(\sum_{i=1}^{n} u_{it} \right) \Rightarrow \left(\int \bar{B}_{\varepsilon} B_u \right) \left(\sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^{\Delta} \right)^{1/2}.$$

Combining this with the asymptotic law of the denominator given in equation (57), we get equation (12).

Case 2: large n and fixed T.

In this case the same approach as in the previous case be followed to prove the main results in the theorem.

Consider first the cointegration case. Define $\tilde{\xi}_{nt} = W_t \left(n^{-1/2} \sum_{i=1}^{n} u_{it}\right)$ and

$$\tilde{\xi}_{nT} = \sum_{t=1}^{T} W_t \left(n^{-1/2} \sum_{i=1}^{n} u_{it}\right).$$

Assumption 1(a) ensures that a CLT holds for $n^{-1/2} \sum_{i=1}^{n} u_{it}$, so that as $n \to \infty$ we have that, for every t, $n^{-1/2} \sum_{i=1}^{n} u_{it} \Rightarrow \bar{u}_t$, where $\bar{u}_t$ is a normally distributed, zero mean random variable with, after Assumption 1(b)

$$E\left[\bar{u}_t \bar{u}_s\right] = \bar{\tau}_{ts}.$$

Therefore, the quantities $W_t \bar{u}_t$ are mixed normals random variables (due to the randomness of W_t) and it ultimately holds that

$$\tilde{\xi}_{nT} \sim N\left[0, \sum_{t=1}^{T} \sum_{s=1}^{T} W_t W_s' \bar{\tau}_{ts}\right] = \left(\sum_{t=1}^{T} \sum_{s=1}^{T} W_t W_s' \bar{\tau}_{ts}\right)^{1/2} Z,$$

where $Z \sim N\left(0, I_k\right)$; Assumption 6 ensures independence between Z and the random variable $\sum_{t=1}^{T} \sum_{s=1}^{T} W_t W_s' \bar{\tau}_{ts}$.

Therefore, in this case the rate of convergence of the numerator of $\hat{\beta} - \beta$ is $O_p\left(\sqrt{n}\right)$. Combining this with the rate of convergence of the denominator, given by equation (58), we have that $\hat{\beta} - \beta = O_p\left(n^{-1/2}\right)$, thereby proving equation (13). As far as the distribution limit is concerned, combining the asymptotic law of $\tilde{\xi}_{nT}$ with equation (58), we ultimately obtain (14).

Under the spurious regression case, define $\tilde{\xi}_{nt}^{S} = W_t \left(n^{-1/2} \sum_{i=1}^{n} u_{it}\right)$ and $\tilde{\xi}_{nT}^{S} = \sum_{t=1}^{T} \tilde{\xi}_{nt}^{S}$. Assumption 1(a) ensures the validity of the CLT for $n^{-1/2} \sum_{i=1}^{n} u_{it}$, so that uniformly in t we have, as $n \to \infty$, $n^{-1/2} \sum_{i=1}^{n} u_{it} \Rightarrow \bar{u}_t$. The process $\bar{u}_t$ is an aggregation of unit root processes, and in light of Assumption 1(a) it is a unit root process with long run variance which by definition is equal to $\bar{h}^{\Delta}$. From this we have that $\tilde{\xi}_{nT}^{S} = O_p(1)$, and combining this with equation (58), we obtain $\hat{\beta} - \beta = O_p(1)$ as reported in equation (15). As far as the limiting distribution is concerned, since $\tilde{\xi}_{nT}^{S}$ is a finite sum, we have $\tilde{\xi}_{nT}^{S} \Rightarrow \sum_{t=1}^{T} W_t \bar{u}_t$ as $n \to \infty$. Combining this with equation (58), we prove the validity of equation (16).

Case 3: large n and large T.

Let us start with the case where equation (2) is a cointegration relationship.

Define $\check{\xi}_{nt} = T^{-1} W_t \left(n^{-1/2} \sum_{i=1}^{n} u_{it}\right)$, and let $\check{\xi}_{nT} = \sum_{t=1}^{T} \check{\xi}_{nt}$. After Assumption 1(b), we know that the process $n^{-1/2} \sum_{i=1}^{n} u_{it}$ has zero mean and covariance structure given by

$$E\left[\left(n^{-1/2} \sum_{i=1}^{n} u_{it}\right)\left(n^{-1/2} \sum_{i=1}^{n} u_{is}\right)\right] = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \tau_{ij,ts}.$$

Moreover, Assumption 6 ensures that $n^{-1/2}\sum_{i=1}^{n}u_{it}$ and W_t are independent. Hence, in light of Assumptions 1(a), 2 and 3, the FCLT ensures that

$$\check{\xi}_{nT} \Rightarrow \int \bar{B}_{\varepsilon} dW,$$

where the Brownian motion W has variance equal to $n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}h_{ij}$. An alternative formulation for the asymptotic distribution of $\check{\xi}_{nT}$, as it is well known,

$$\check{\xi}_{nT} \Rightarrow \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}h_{ij}\right)^{1/2}\left(\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\right)^{1/2} Z,$$

where $Z \sim N(0, I_k)$ and $\bar{B}_{\varepsilon}$ and Z are independent. As $n \to \infty$ we have

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}h_{ij} = \bar{h},$$

and therefore, as $n \to \infty$

$$\check{\xi}_{nT} \Rightarrow \sqrt{\bar{h}}\left(\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\right)^{1/2} Z.$$

Hence, as far as the rate of convergence of $\check{\xi}_{nT}$ is concerned, we have $\check{\xi}_{nT} = O_p(1)$. Combining this with equation (59), we get that $\hat{\beta}-\beta = O_p\left(n^{-1/2}T^{-1}\right)$, proving equation.(17). As far as the distribution limit is concerned, combining the asymptotic law of $\check{\xi}_{nT}$ with equation (59), we have:

$$\left[\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}W_tW_t'\right]\left[\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}W_tu_{it}\right] \Rightarrow \sqrt{\bar{h}}\left(\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\right)^{-1/2} Z,$$

which corresponds to equation (18).

We now turn to the case when equation (2) is a spurious regression. Let $\check{\xi}_{nt}^{S} = T^{-2}W_t\left(n^{-1/2}\sum_{i=1}^{n}u_{it}\right)$ and $\check{\xi}_{nT}^{S} = \sum_{t=1}^{T}\check{\xi}_{nt}^{S}$. For fixed n the process $n^{-1/2}\sum_{i=1}^{n}u_{it}$ is still a unit root process with long run variance given by $n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}h_{ij}^{\Delta}$. Therefore, for fixed n and as $T \to \infty$, a FCLT, which follows from Assumptions 1(a), 2 and 3, ensures that $\xi_{nT}^{S} = O_p(1)$. This result, together with equation (59), proves that $\hat{\beta}-\beta = O_p\left(n^{-1/2}\right)$, as reported in equation (19). As far as the limiting distribution is concerned as $T \to \infty$ we have

$$\check{\xi}_{nT}^{S} = \frac{1}{T^2}\sum_{t=1}^{T}W_t\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}u_{it}\right) \Rightarrow \left(\int \bar{B}_{\varepsilon}B_u\right)\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}h_{ij}^{\Delta}\right)^{1/2};$$

taking the limit for $n \to \infty$ leads to

$$\left(\int \bar{B}_\varepsilon B_u\right)\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} h_{ij}^{\Delta}\right)^{1/2} \to \sqrt{\bar{h}^{\Delta}}\left(\int \bar{B}_\varepsilon B_u\right). \tag{60}$$

Combining this result with the one reported in equation (59), we ultimately get equation (20). ■

Proof of Theorem 2. To prove the theorem, we refer to equation (8) that contains the estimation error $\hat{\beta}^{FD} - \beta = \left[\sum_i \sum_t \Delta F_t \Delta F_t'\right]^{-1}\left[\sum_i \sum_t \Delta F_t \Delta u_{it}\right]$. The proof be derived splitting this quantity into the denominator $\sum_i \sum_t \Delta F_t \Delta F_t'$ and the numerator $\sum_i \sum_t \Delta F_t \Delta u_{it}$, and analyzing the asymptotic behavior of both quantities separately.

Let us start considering the denominator $\sum_i \sum_t \Delta F_t \Delta F_t'$. When $T \to \infty$ and n is fixed, we have from Assumption 2 and the Law of Large Numbers that under both the cases that equation (2) is a spurious regression or a cointegrating one it holds that $\sum_i \sum_t \Delta F_t \Delta F_t' = O_p(T)$ and

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} \Delta F_t \Delta F_t' \xrightarrow{p} \Sigma_{\Delta F}. \tag{61}$$

As $n \to \infty$, and for fixed T, we have

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T} \Delta F_t \Delta F_t' = \sum_{t=1}^{T} \Delta F_t \Delta F_t', \tag{62}$$

whilst as both n and T are large we have

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} \Delta F_t \Delta F_t' \xrightarrow{p} \Sigma_{\Delta F}, \tag{63}$$

with $\sum_{i=1}^{n}\sum_{t=1}^{T} \Delta F_t \Delta F_t' = O_p(nT)$.

As far as the numerator is concerned, as in the case of Theorem 1 the proof be derived with respect to three separate cases, following the same structure as in the theorem. We firstly derive the rate of convergence and the limiting distribution of $\sum_i \sum_t \Delta F_t \Delta u_{it}$ for the case when T is large and n is fixed; we then study the opposite case, when T is fixed and n is large; last, we analyze the case when both T and n are large. The proofs for each of the three cases are along the same lines as in Theorem 1. It is worth noticing though that both under the case when equation (2) is a cointegration relationship and when it is a spurious regression, Δu_{it} is a stationary process. Therefore, there is no need to distinguish between the two cases unlike in Theorem 1.

Case 1: large T and fixed n

Denote

$$\xi_{nt}^{\Delta} = T^{-1/2}\Delta F_t\left(\sum_{i=1}^{n} \Delta u_{it}\right)$$

and

$$\xi_{nT}^{\Delta} = \sum_{t=1}^{T} \xi_{nt}^{\Delta}.$$

Assumption 6 ensures that ΔF_t and the Δu_{it}s are independent. Also, according to Assumption 1(b), the process $\sum_i \Delta u_{it}$ has zero mean and covariance structure

$$E\left[\left(\sum_{i=1}^{n} \Delta u_{it}\right)\left(\sum_{i=1}^{n} \Delta u_{is}\right)\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij,ts}.$$

Therefore the process ξ_{nt}^{Δ} has zero mean and covariance structure given by

$$E\left[\xi_{nt}^{\Delta}\xi_{nt}^{\Delta\prime}\right] = \frac{1}{T}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij,ts}\right) E\left(\Delta F_t \Delta F_s'\right).$$

After Assumption 1(b) and 2, that ensure weak dependence over time a CLT holds. Therefore, as $T \to \infty$, we have

$$\begin{aligned} \xi_{nT}^{\Delta} \quad &\Rightarrow \quad \left[\lim_{T\to\infty} \frac{1}{T}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij,ts}\right) E\left(\Delta F_t \Delta F_s'\right)\right]^{1/2} Z \qquad (64) \\ &= \quad \left(\sum_{i=1}^{n}\sum_{j=1}^{n} h_{ij}^{\Delta}\right)^{1/2} \Sigma_{\Delta F}^{1/2} Z, \end{aligned}$$

where $Z \sim N(0, I_k)$. Hence we have a twofold result. First, the rate of convergence of the numerator of $\hat{\beta}^{FD} - \beta$ is $O_p\left(\sqrt{T}\right)$; therefore, given equation (61) that ensures that the denominator of $\hat{\beta}^{FD} - \beta$ is $O_p(T)$, we have that $\hat{\beta}^{FD} - \beta = O_p\left(T^{-1/2}\right)$, proving equation (21). As far as the distribution limit is concerned, combining the asymptotic law of ξ_{nT}^{Δ} with equation (61), we have that

$$\left[\frac{1}{T}\sum_{i=1}^{n}\sum_{t=1}^{T} \Delta F_t \Delta F_t'\right]^{-1}\left[\frac{1}{\sqrt{T}}\sum_{i=1}^{n}\sum_{t=1}^{T} \Delta F_t \Delta u_{it}\right] \Rightarrow \frac{1}{n}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} h_{ij}^{\Delta}\right)^{1/2} \Sigma_{\Delta F}^{-1/2} Z,$$

which proves equation (22).

Case 2: large n and fixed T.

In this case the same approach as in the previous case be followed to prove the main results in the theorem.

Define $\tilde{\xi}_{nt}^{\Delta} = \Delta F_t \left(n^{-1/2} \sum_{i=1}^{n} \Delta u_{it}\right)$ and

$$\tilde{\xi}_{nT}^{\Delta} = \sum_{t=1}^{T} \tilde{\xi}_{nt}.$$

Assumption 1(a) ensures that a CLT holds for $n^{-1/2}\sum_{i=1}^{n}\Delta u_{it}$, so that as $n\to\infty$ we have that, for every t, $n^{-1/2}\sum_{i=1}^{n}\Delta u_{it}\Rightarrow\bar{u}_t$, where $\Delta\bar{u}_t$ is a normally distributed, zero mean random variable with covariance structure

$$E\left[\Delta\bar{u}_t\Delta\bar{u}_s\right]=\sum_{t=1}^{T}\sum_{s=1}^{T}\bar{\gamma}_{ts}.$$

Hence, in light of Assumption 6, $\tilde{\xi}_{nt}^{\Delta}$ is a zero mean random variable whose covariance structure is given by (after Assumption 1(a))

$$E\left[\tilde{\xi}_{nt}^{\Delta}\tilde{\xi}_{ns}^{\Delta\prime}\right]=\sum_{t=1}^{T}\sum_{s=1}^{T}\bar{\gamma}_{ts}E\left(\Delta F_t\Delta F_s^{\prime}\right).$$

Since $\tilde{\xi}_{nT}^{\Delta}$ is a finite sum of normally distributed random variables, we have that

$$\tilde{\xi}_{nT}^{\Delta}\sim\left(\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta F_t\Delta F_s^{\prime}\bar{\gamma}_{ts}\right)^{1/2}Z,$$

where $Z\sim N\left(0,I_k\right)$; Assumption 6 ensures independence between Z and the random variable $\sum_t\sum_s\Delta F_t\Delta F_s^{\prime}\bar{\gamma}_{ts}$. Therefore, in this case the rate of convergence of the numerator of $\hat{\beta}^{FD}-\beta$ is $O_p\left(\sqrt{n}\right)$. Combining this with the rate of convergence of the denominator, given by equation (62), we have that $\hat{\beta}^{FD}-\beta=O_p\left(n^{-1/2}\right)$, thereby proving equation (23). Also, combining this with equation (62), we ultimately obtain (24).

Case 3: large n and large T.

Define $\check{\xi}_{nt}^{\Delta}=T^{-1/2}\Delta F_t\left(n^{-1/2}\sum_{i=1}^{n}\Delta u_{it}\right)$, and let $\check{\xi}_{nT}^{\Delta}=\sum_{t=1}^{T}\check{\xi}_{nt}^{\Delta}$. In light of the passages derived above, the $\check{\xi}_{nt}^{\Delta}$s are random variables with zero mean and covariance structure given by

$$E\left[\check{\xi}_{nt}^{\Delta}\check{\xi}_{ns}^{\Delta\prime}\right]=\frac{1}{T}\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma_{ij,ts}\right)E\left(\Delta F_t\Delta F_s^{\prime}\right).$$

From equation (64) we know that, for fixed n and as $T\to\infty$

$$\check{\xi}_{nT}^{\Delta}\Rightarrow\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}h_{ij}^{\Delta}\right)^{1/2}\Sigma_{\Delta F}^{1/2}Z,$$

with $Z\sim N\left(0,I_k\right)$. As $n\to\infty$ we have

$$\lim_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}h_{ij}\right)^{1/2}\Sigma_{\Delta F}^{1/2}\to\sqrt{\bar{h}^{\Delta}}\Sigma_{\Delta F}^{1/2}.$$

Hence, as far as the rate of convergence of $\check{\xi}_{nT}^{\Delta}$ is concerned, we have $\check{\xi}_{nT}^{\Delta} = O_p(1)$. Combining this with equation (63), we get that $\hat{\beta}^{FD} - \beta = O_p\left(n^{-1/2}T^{-1/2}\right)$, proving equation.(23). As far as the distribution limit is concerned, we know that

$$\check{\xi}_{nT}^{\Delta} \Rightarrow \sqrt{\bar{h}^{\Delta}}\Sigma_{\Delta F}^{1/2}Z,$$

as $(n,T) \to \infty$ with $Z \sim N(0, I_k)$. Combining this result with equation (63), we have:

$$\left[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta F_t \Delta F_t'\right]^{-1}\left[\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta F_t \Delta u_{it}\right] \Rightarrow \sqrt{\bar{h}^{\Delta}}\Sigma_{\Delta F}^{-1/2}Z,$$

which corresponds to equation (24). ■

Lemma 1 *Let Assumptions 1-6 hold. Then the following results hold for the estimated factors $\hat{F}_t$ when $(n,T) \to \infty$:*

1.
$$\begin{aligned} & V_{nT}\left(\hat{F}_t - F_t\right) \\ = \; & T^{-1}\sum_{s=1}^{T}\hat{F}_s\gamma_{s-t} + T^{-1}\sum_{s=1}^{T}\hat{F}_s\zeta_{st} + T^{-1}\sum_{s=1}^{T}\hat{F}_s\eta_{st} + T^{-1}\sum_{s=1}^{T}\hat{F}_s\xi_{st}, \end{aligned}$$

   where $\gamma_{s-t} = E\left[n^{-1}e_t'e_s\right]$,

$$\zeta_{st} = \frac{e_t'e_s}{n} - \gamma_{s-t},$$

$$\eta_{st} = \frac{1}{n}\Delta F_s'\Lambda' e_t,$$

$$\xi_{st} = \frac{1}{n}\Delta F_t'\Lambda' e_s,$$

   and V_{nT} is a diagonal matrix containing the largest k eigenvalues of $(nT)^{-1}ZZ'$ in decreasing order;

2. *Denote $C_{nT} = \min\{\sqrt{n}, T\}$. Consistency of $\hat{F}_t$ is expressed as*

   (a) $\max_{1\le t\le T}\left\|\hat{F}_t - F_t\right\| = O_p\left(T^{-1}\right) + O_p\left(n^{-1/2}T^{1/2}\right)$ *and*

   (b) $\sum_{t=1}^{T}\left\|\hat{F}_t - F_t\right\|^2 = O_p\left(TC_{nT}^{-2}\right);$

3. *It holds that:*

   (a) $\sum_{t=1}^{T}\left(\hat{F}_t - F_t\right)' e_t = O_p\left(TC_{nT}^{-2}\right);$

(b) $\sum_{t=1}^{T}\left(\hat{F}_t - F_t\right)' F_t = O_p(1) + O_p\left(n^{-1/2}T\right) = O_p\left(TC_{nT}^{-2}\right);$

(c) $\sum_{t=1}^{T}\left(\hat{F}_t - F_t\right)' \hat{F}_t = O_p\left(TC_{nT}^{-2}\right).$

4. *When $\frac{\sqrt{n}}{T} \to 0$ as $(n,T) \to \infty$, we have*

$$\sqrt{n}\left(W_t - \hat{W}_t\right) = \frac{1}{T^2}\sum_{s=1}^{T}\hat{W}_s W_s' \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\lambda_i e_{it},$$

with

$$n^{-1/2}\sum_{i=1}^{n}\lambda_i e_{it} \Rightarrow Z_t,$$

where $Z_t \backsim N(0,\Gamma)$ and $T^{-2}\sum_{s=1}^{T}\hat{W}_s W_s \Rightarrow \tilde{Q}_B$; $\tilde{Q}_B$ and Z are independent - see Bai (2004).

Proof. See Bai (2004). ■

Lemma 2 *Lemma 1 ensures that*

1. $T^{-2}\sum_{t=1}^{T}\hat{W}_t\hat{W}_t' = T^{-2}\sum_{t=1}^{T}W_tW_t' + o_p\left(T^{-1/2}C_{nT}^{-1}\right);$
2. $n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it} = n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}W_t u_{it} + O_p\left(C_{nT}^{-1}\right);$
3. $T^{-1}\sum_{t=1}^{T}\hat{W}_t\left(F_t - \hat{F}_t\right) = T^{-1}\sum_{t=1}^{T}W_t'\left(F_t - \hat{F}_t\right) + O_p\left(C_{nT}^{-2}\right);$

Proof. Proof is as follows:

$$\begin{aligned}
\frac{1}{T^2}\sum_{t=1}^{T}\hat{W}_t\hat{W}_t' &= \frac{1}{T^2}\sum_{t=1}^{T}\left(W_t + \hat{W}_t - W_t\right)\left(W_t + \hat{W}_t - W_t\right)' \\
&= \frac{1}{T^2}\sum_{t=1}^{T}W_tW_t' + \frac{1}{T^2}\sum_{t=1}^{T}W_t\left(\hat{W}_t - W_t\right)' \\
&\quad + \frac{1}{T^2}\sum_{t=1}^{T}\left(\hat{W}_t - W_t\right)W_t' + \frac{1}{T^2}\sum_{t=1}^{T}\left(\hat{W}_t - W_t\right)\left(\hat{W}_t - W_t\right)' \\
&= I + II + III + IV.
\end{aligned}$$

Consider II and III. Using the Cauchy-Schwartz inequality we get straightforwardly

$$\frac{1}{T^2}\sum_{t=1}^{T}W_t\left(\hat{W}_t - W_t\right)' = O_p\left(\frac{1}{\sqrt{T}}\right)O_p\left(\frac{1}{C_{nT}}\right)o_p(1) = o_p\left(\frac{1}{\sqrt{T}C_{nT}}\right).$$

Consider now IV. In this case, Lemma 1.2.*(b)* states that

$$T^{-2}\sum_{t=1}^{T}\left(\hat{W}_t - W_t\right)\left(\hat{W}_t - W_t\right)' = O_p\left(T^{-1}C_{nT}^{-2}\right).$$

Then

$$\frac{1}{T^2}\sum_{t=1}^{T}\hat{W}_t\hat{W}_t' = \frac{1}{T^2}\sum_{t=1}^{T}W_tW_t' + o_p\left(\frac{1}{\sqrt{T}C_{nT}}\right) + O_p\left(\frac{1}{TC_{nT}^2}\right),$$

which proves part 1 of the Lemma. Consider now part 2:

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_tu_{it} = \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}W_tu_{it} + \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{W}_t - W_t\right)u_{it} = I + II.$$

For term I, Theorem 1 ensures that $n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}W_tu_{it} = O_p(1)$. As far as II is concerned, using the Cauchy-Schwartz inequality we get

$$\begin{aligned}
& \left\|\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{W}_t - W_t\right)u_{it}\right\| \\
= & \left\|\frac{1}{T}\sum_{t=1}^{T}\left(\hat{W}_t - W_t\right)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}u_{it}\right\| \\
\leq & \left(\frac{1}{T}\sum_{t=1}^{T}\left\|\hat{W}_t - W_t\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{i=1}^{n}\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}u_{it}\right\|^2\right)^{1/2} = O_p\left(\frac{1}{C_{nT}}\right)
\end{aligned}$$

given that $T^{-1}\sum_{t=1}^{T}\left\|\hat{W}_t - W_t\right\|^2 = O_p\left(C_{nT}^{-2}\right)$ and $n^{-1/2}\sum_{i=1}^{n}u_{it} = O_p(1)$. Hence,

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_tu_{it} = \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}W_tu_{it} + O_p\left(\frac{1}{C_{nT}}\right),$$

proving part 2 of the Lemma. To prove part 3, we note that

$$\frac{1}{T}\sum_{t=1}^{T}\hat{W}_t'\left(F_t - \hat{F}_t\right) = \frac{1}{T}\sum_{t=1}^{T}W_t'\left(F_t - \hat{F}_t\right) + \frac{1}{T}\sum_{t=1}^{T}\left(\hat{W}_t - W_t\right)'\left(F_t - \hat{F}_t\right) = I + II.$$

Term I is bounded by $O_p\left(C_{nT}^{-1}\right)$ - see Lemma 1.3.*(c)* - whilst II is bounded by

$$\begin{aligned}
& \left(\frac{1}{T}\sum_{t=1}^{T}\left\|\hat{W}_t - W_t\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|F_t - \hat{F}_t\right\|^2\right)^{1/2} \\
= & O_p\left(\frac{1}{C_{nT}}\right)O_p\left(\frac{1}{C_{nT}}\right) = O_p\left(\frac{1}{C_{nT}^2}\right).
\end{aligned}$$

Hence,

$$\frac{1}{T}\sum_{t=1}^{T}\hat{W}_t\left(F_t - \hat{F}_t\right) = O_p\left(\frac{1}{C_{nT}}\right) + O_p\left(\frac{1}{C_{nT}^2}\right).$$

■

Proof of Theorem 3. According to equation (29)

$$\hat{\beta} - \beta = \left[\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\hat{W}_t'\right]^{-1}\left\{\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\left[\left(W_t - \hat{W}_t\right)'\beta + u_{it}\right]\right\}.$$

Let us first consider the denominator of this expression. Assumption 3 and Lemma 2.1 imply that

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\hat{W}_t' = O_p\left(nT^2\right), \tag{65}$$

and

$$\left(nT^2\right)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\hat{W}_t' \Rightarrow \int \bar{B}_\varepsilon \bar{B}_\varepsilon'; \tag{66}$$

this holds under both the cases of cointegration and spurious regression.

We now prove Theorem 3 for the case when equation (2) is a cointegration relationship. The numerator of $\hat{\beta} - \beta$ is given by

$$n\sum_{t=1}^{T}\hat{W}_t\left(W_t - \hat{W}_t\right)'\beta + \sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it} = I + II.$$

Let us consider II. We know from Theorem 1 and Lemma 2.2 that, as far as II is concerned

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it} = \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}W_t u_{it} + o_p(1) = O_p(1),$$

and therefore

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it} \Rightarrow \left(\int \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{1/2}\sqrt{\bar{h}}Z, \tag{67}$$

where $Z \sim N(0, I_k)$.As far as term I is concerned, two cases are possible:

1. if $\sqrt{n}/T \to 0$, we know from Lemma 1.4 that

$$\sqrt{n}\left(W_t - \hat{W}_t\right) = \frac{1}{T^2}\sum_{s=1}^{T}\hat{W}_s W_s' \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\lambda_i e_{it},$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\lambda_i e_{it} \Rightarrow Z_t,$$

with $Z_t \backsim N(0, \Gamma)$ for every t. Therefore the asymptotic magnitude of term I is the same as that of term II and equal to $O_p(\sqrt{n}T)$. This proves equation (31). As far as the distribution limit is concerned, we can write

$$\begin{aligned}
& \frac{1}{T}\sum_{t=1}^{T} W_t\sqrt{n}\left(W_t-\hat{W}_t\right)'\beta \\
= & \frac{1}{T}\sum_{t=1}^{T} W_t\left[\frac{1}{T^2}\sum_{s=1}^{T}\hat{W}_sW_s'\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\lambda_ie_{it}\right)\right]'\beta+o_p(1),
\end{aligned}$$

and since by definition $T^{-2}\sum_{s=1}^{T}\hat{W}_sW_s'\Rightarrow\tilde{Q}_B$, we have

$$\frac{1}{T}\sum_{t=1}^{T} W_t\left[\frac{1}{T^2}\sum_{s=1}^{T}\hat{W}_sW_s'\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\lambda_ie_{it}\right)\right]'\beta\Rightarrow\left(\int\bar{B}_\varepsilon\bar{B}_\varepsilon'\right)^{1/2}\left(\beta'\tilde{Q}_B\Gamma\tilde{Q}_B'\beta\right)^{1/2}Z.,$$

with $Z\sim N(0,I_k)$. Combining this with the asymptotic law of II and with equation (66), we obtain equation (32);

2. if $T/\sqrt{n}\to 0$, after Lemma 1.3.*(c)*, we have

$$\sum_{t=1}^{T}\hat{W}_t\left(W_t-\hat{W}_t\right)'=O_p\left(n^{-1/2}T\right)+O_p(1),$$

and the term that dominates is the one with asymptotic magnitude $O_p(1)$. Therefore, $I=\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\left(W_t-\hat{W}_t\right)'=O_p(n)$, and term $II=\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_tu_{it}=O_p(\sqrt{n}T)$ is dominated. The order of magnitude of the numerator is now $O_p(n)$, and combining this with equation (65) we have

$$\hat{\beta}-\beta=O_p\left(T^{-2}\right),$$

which proves equation (33). As far as the limiting distribution is concerned, following Bai (2004), we can write

$$\sum_{t=1}^{T}\hat{W}_t\left(W_t-\hat{W}_t\right)'=\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}W_tW_s'\left(\frac{1}{n}\sum_{i=1}^{n}e_{it}e_{is}\right)+o_p(1),$$

and asymptotically we have:

$$\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}W_tW_s'\left(\frac{1}{n}\sum_{i=1}^{n}e_{it}e_{is}\right)=\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T}W_te_{it}\right)\left(\frac{1}{T}\sum_{s=1}^{T}W_s'e_{is}\right).$$

We know that $T^{-1}\sum_{t=1}^{T}W_te_{it}\Rightarrow\int\bar{B}_\varepsilon dB_{ei}$, where $B_{ei}(r)$ is the Brownian motion associated to the partial sums of e_{it} with long run variance σ_{ei}^2; therefore, applying a LLN, the limit for $n\to\infty$ is given by

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}E\left[\left(\int\bar{B}_\varepsilon dB_{ei}\right)\left(\int\bar{B}_\varepsilon' dB_{ei}\right)\right]=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Var\left(\int\bar{B}_\varepsilon dB_{ei}\right).$$

Since we have that $Var\left(\int \bar{B}_{\varepsilon} dB_{ei}\right) = \sigma_{ei}^2 E\left(\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\right)$, it holds that

$$\begin{aligned} & \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} Var\left(\int \bar{B}_{\varepsilon} dB_{ei}\right) \\ = & \left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sigma_{ei}^2\right) E\left(\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\right) = \frac{1}{2}\sigma_e^2\Omega_{\varepsilon\varepsilon}, \end{aligned}$$

given the definition of σ_e^2 and that $E\left(\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\right) = 1/2\Omega_{\varepsilon\varepsilon}$. Combining this equation (66), equation (34) is proved.

We now prove results when we equation (2) is a spurious regression. Here, as far as term $\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it}$ in equation (29) is concerned, we have

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it} = \sum_{i=1}^{n}\sum_{t=1}^{T} W_t u_{it} + \sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{W}_t - W_t\right) u_{it}.$$

After equation (19) we know

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

As per $\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{W}_t - W_t\right) u_{it}$, application of the Cauchy-Schwartz inequality leads to

$$\begin{aligned} & \sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{W}_t - W_t\right) u_{it} \\ = & \sum_{t=1}^{T}\left(\hat{W}_t - W_t\right)\left(\sum_{i=1}^{n} u_{it}\right) \\ \leq & \left[\sum_{t=1}^{T}\left\|\hat{W}_t - W_t\right\|^2\right]^{1/2}\left[\sum_{t=1}^{T}\left\|\sum_{i=1}^{n} u_{it}\right\|^2\right]^{1/2} \\ = & O_p\left(\sqrt{T}C_{nT}^{-1}\right) O_p\left(\sqrt{n}T\right), \end{aligned}$$

and therefore $\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{W}_t - W_t\right) u_{it}$ is always dominated by $\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it}$. Hence, we have

$$\frac{1}{\sqrt{n}T^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it} = \frac{1}{\sqrt{n}T^2}\sum_{i=1}^{n}\sum_{t=1}^{T} W_t u_{it} + o_p(1),$$

and, after equation (60), we have

$$\frac{1}{\sqrt{n}T^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it} \Rightarrow \left(\int \bar{B}_{\varepsilon} B_u\right)\sqrt{\bar{h}^{\Delta}}.$$

Consequently, there are two possibilities for the rate of convergence and the limiting distribution of the numerator:

- when $\sqrt{n}/T^2 \to 0$, two subcases are possible:
  - $\sqrt{n}/T \to 0$, and in such case we have $\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\left(W_t - \hat{W}_t\right)'\beta = O_p\left(\sqrt{n}T\right)$; therefore the term that dominates is $\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it}$; combining this with equations (65) and (66), equations (35) and (36) can be obtained;
  - $T/\sqrt{n} \to 0$ and $\sqrt{n}/T^2 \to 0$, and here $\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\left(W_t - \hat{W}_t\right)'\beta = O_p\left(n\right)$; in this case, again the term that dominates is $\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it}$ combining this with equations (65) and (66), equations (35) and (36) hold;
- when $T^2/\sqrt{n} \to 0$, we have that $\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\left(W_t - \hat{W}_t\right)'\beta = O_p\left(n\right)$, and this is the dominating term. This leads to the same results as in equations (33) and (34).

■

Lemma 3 *Let Assumptions 1-2 and 4-6 hold. Then, for the estimated factors $\Delta\hat{F}_t$, it holds that*

1.
$$\begin{aligned} & V_{nT}\left(\Delta\hat{F}_t - \Delta F_t\right) \\ = \ & T^{-1}\sum_{s=1}^{T}\Delta\hat{F}_s\gamma_{s-t} + T^{-1}\sum_{s=1}^{T}\Delta\hat{F}_s\zeta_{st} + T^{-1}\sum_{s=1}^{T}\Delta\hat{F}_s\eta_{st} + T^{-1}\sum_{s=1}^{T}\Delta\hat{F}_s\xi_{st}, \end{aligned}$$

   where $\gamma_{s-t} = E\left[n^{-1}\sum_{i=1}^{n}e_{it}e_{is}\right]$,
$$\zeta_{st} = n^{-1}\sum_{i=1}^{n}e_{it}e_{is} - \gamma_{s-t},$$
$$\eta_{st} = n^{-1}\Delta F_s'\Lambda' e_t,$$
$$\xi_{st} = n^{-1}\Delta F_t'\Lambda' e_s,$$
   and V_{nT} is a diagonal matrix containing the largest k eigenvalues of $(nT)^{-1}\Delta Z\Delta Z'$ in decreasing order;

2. *Denote $\delta_{nT} = \min\left\{\sqrt{n},\sqrt{T}\right\}$. Consistency of $\Delta\hat{F}_t$ is ensured by*
   (a) $\max_{1\le t\le T}\left\|\Delta\hat{F}_t - \Delta F_t\right\| = O_p\left(T^{-1/2}\right) + O_p\left(n^{-1/2}T^{1/2}\right)$;

(b) $\sum_{t=1}^{T} \left\| \Delta \hat{F}_t - \Delta F_t \right\|^2 = O_p \left(T \delta_{nT}^{-2} \right)$;

3. *It holds that:*

   (a) $\sum_{t=1}^{T} \left(\Delta \hat{F}_t - \Delta F_t \right)' e_t = O_p \left(T \delta_{nT}^{-2} \right)$;

   (b) $\sum_{t=1}^{T} \left(\Delta \hat{F}_t - \Delta F_t \right)' \Delta F_t = O_p \left(T \delta_{nT}^{-2} \right)$;

   (c) $\sum_{t=1}^{T} \left(\Delta \hat{F}_t - \Delta F_t \right)' \Delta \hat{F}_t = O_p \left(T \delta_{nT}^{-2} \right)$;

4. *The relationship between factors and ζ_{st} is given by $\sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \Delta F_s' \zeta_{st} = O_p \left(n^{-1/2} T^{3/2} \right)$.*

Proof. See Bai and Ng (2002) and Bai (2003). ■

Proof of Theorem 4. Recall equation (30):

$$\hat{\beta}^{FD} - \beta^{FD} = \left[\sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_t \Delta \hat{F}_t' \right]^{-1} \left\{ \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_t \left[\left(\Delta F_t - \Delta \hat{F}_t \right)' \beta + \Delta u_{it} \right] \right\}.$$

We firstly study the rate of convergence and the distribution limit of the denominator. The following decomposition holds:

$$\begin{aligned} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_t \Delta \hat{F}_t' &= n \sum_{t=1}^{T} \Delta F_t \Delta F_t' + n \sum_{t=1}^{T} \Delta F_t \left(\Delta \hat{F}_t - \Delta F_t \right)' \\ &\quad + n \sum_{t=1}^{T} \left(\Delta \hat{F}_t - \Delta F_t \right) \Delta F_t' + n \sum_{t=1}^{T} \left(\Delta \hat{F}_t - \Delta F_t \right) \left(\Delta \hat{F}_t - \Delta F_t \right)' \\ &= I + II + III + IV. \end{aligned}$$

After Assumption 2 we have

$$I = n \sum_{t=1}^{T} \Delta F_t \Delta F_t' = O_p (nT).$$

Also

$$II = n \sum_{t=1}^{T} \left(\Delta \hat{F}_t - \Delta F_t \right) \Delta F_t' = O_p \left(nT \delta_{nT}^{-2} \right),$$

and

$$IV = n \sum_{t=1}^{T} \left(\Delta \hat{F}_t - \Delta F_t \right) \left(\Delta \hat{F}_t - \Delta F_t \right)' = O_p \left(nT \delta_{nT}^{-2} \right),$$

using Lemma 2.3.*(b)* and 2.2.*(b)* respectively. Therefore

$$\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_t \Delta \hat{F}_t' = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta F_t' + O_p \left(\delta_{nT}^{-2} \right) \tag{68}$$

and, for $(n,T)\to\infty$

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta\hat{F}_t' \xrightarrow{p} \Sigma_{\Delta F}. \tag{69}$$

Let us now turn to the numerator of $\hat{\beta}^{FD}-\beta^{FD}$. We have

$$\begin{aligned}
&\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\left[\left(\Delta F_t-\Delta\hat{F}_t\right)'\beta+\Delta u_{it}\right]\\
=&\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta F_t\Delta u_{it}+\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\Delta\hat{F}_t-\Delta F_t\right)\Delta u_{it}\\
&+n\sum_{t=1}^{T}\Delta\hat{F}_t\left(\Delta F_t-\Delta\hat{F}_t\right)'\beta=I+II+III.
\end{aligned}$$

We know, from equation (25) in Theorem 2 that:

$$I=\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta F_t\Delta u_{it}=O_p\left(\sqrt{nT}\right).$$

Also, following Bai (2003, pp. 163-164), we could prove

$$II=\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\Delta\hat{F}_t-\Delta F_t\right)\Delta u_{it}=O_p\left(\sqrt{n}T\delta_{nT}^{-2}\right).$$

Lemma 2.3.*(c)* ensures that

$$III=n\sum_{t=1}^{T}\Delta\hat{F}_t\left(\Delta F_t-\Delta\hat{F}_t\right)'=nO_p\left(T\delta_{nT}^{-2}\right)=O_p\left(nT\delta_{nT}^{-2}\right).$$

Note that term III dominates term II by a factor $\sqrt{n}$. Also, III always dominates I since it always holds that $nT\delta_{nT}^{-2}>\sqrt{nT}$; in fact, this is the same as writing

$$\sqrt{n}\sqrt{T}=\min\left(\sqrt{n},\sqrt{T}\right)\max\left(\sqrt{n},\sqrt{T}\right)>\delta_{nT}^2=\left[\min\left(\sqrt{n},\sqrt{T}\right)\right]^2.$$

Therefore, term III in the numerator always dominates. According to Lemma 3.1, III can be decomposed into four terms of magnitude

$$\begin{aligned}
\sum_{t=1}^{T}\Delta\hat{F}_t\left(\Delta F_t-\Delta\hat{F}_t\right)'\beta &= O_p\left(\sqrt{T}\delta_{nT}^{-1}\right)+O_p\left(Tn^{-1/2}\delta_{nT}^{-1}\right)+O_p\left(\sqrt{\frac{T}{n}}\right)+O_p\left(\sqrt{\frac{T}{n}}\right)\\
&= a+b+c+d.
\end{aligned}$$

Two cases may occur:

1. $\frac{n}{T} \to 0$; in this case, $\delta_{nT} = \sqrt{n}$. The dominating term is b and

$$b = n \sum_{t=1}^{T} \Delta \hat{F}_t \left(\Delta F_t - \Delta \hat{F}_t \right)' \beta = \frac{n}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta \hat{F}_t \Delta \tilde{F}_s' \zeta_{st} V^{-1} \beta + o_p(1),$$

where

$$\zeta_{st} = \frac{1}{n} \sum_{i=1}^{n} \left(e_{it} e_{is} - \gamma_{s-t} \right) = O_p\left(n^{-1/2} \right).$$

After similar passages as above, we have

$$\sum_{t=1}^{T} \sum_{s=1}^{T} \Delta \hat{F}_t \Delta \tilde{F}_s' \zeta_{st} V^{-1} \beta = \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \Delta F_s' \zeta_{st} V^{-1} \beta + o_p(1).$$

After Lemma 2.4, we know that $T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \Delta F_s' \zeta_{st} = O_p\left(n^{-1/2} T^{1/2} \right)$. Therefore the order of magnitude of the numerator is $O_p\left(n^{1/2} T^{1/2} \right)$, and combining this with equation (68) we obtain equation (37). As per the limiting distribution, since by definition

$$Q = p \lim \frac{1}{n T^{3/2}} \sum_{s=1}^{T} \sum_{t=1}^{T} \Delta \hat{F}_t \Delta \tilde{F}_s' \zeta_{st},$$

combining this with equation (69), one can derive equation (38);

2. $\frac{T}{n} \to 0$, and in such case, given that $\delta_{nT} = \sqrt{T}$, the dominating term is a. Given its definition, after Lemma 2.2.*(a)*, we have

$$a = \sum_{t=1}^{T} \Delta \hat{F}_t \left(\Delta F_t - \Delta \hat{F}_t \right)' \beta = \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta \hat{F}_t \Delta \tilde{F}_s' \gamma_{s-t} V^{-1} \beta + o_p(1),$$

and

$$\begin{aligned} & \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta \hat{F}_t \Delta \tilde{F}_s' \gamma_{s-t} \\ = \quad & \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \Delta F_s' \gamma_{s-t} + \sum_{t=1}^{T} \sum_{s=1}^{T} \left(\Delta \hat{F}_t - \Delta F_t \right) \Delta F_s' \gamma_{s-t} \\ & + \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \left(\Delta \tilde{F}_s - \Delta F_s \right)' \gamma_{s-t} + \\ & \sum_{t=1}^{T} \sum_{s=1}^{T} \left(\Delta \hat{F}_t - \Delta F_t \right) \left(\Delta \tilde{F}_s - \Delta F_s \right)' \gamma_{s-t}. \end{aligned}$$

Then we can show that:

(a)

$$\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta F_t \Delta F_s' \gamma_{s-t} = O_p\left(T\right);$$

(b)

$$\begin{aligned}\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta F_t\left(\Delta \tilde{F}_s - \Delta F_s\right)'\gamma_{s-t} &\leq \max_t \left\|\Delta F_t\right\| \max_s \left\|\Delta \tilde{F}_s - \Delta F_s\right\| \sum_{t=1}^{T}\sum_{s=1}^{T}\left|\gamma_{s-t}\right| \\ &= O_p\left(1\right) O_p\left(T^{-1/2}\right) O_p\left(T\right) = O_p\left(T^{1/2}\right);\end{aligned}$$

(c)

$$\begin{aligned}&\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\Delta \hat{F}_t - \Delta F_t\right)\left(\Delta \tilde{F}_s - \Delta F_s\right)'\gamma_{s-t} \\ \leq\ & \left(\max_s \left\|\Delta \tilde{F}_s - \Delta F_s\right\|\right)\sum_{t=1}^{T}\sum_{s=1}^{T}\left|\gamma_{s-t}\right| = O_p\left(T^{-1}\right) O_p\left(T\right) = O_p\left(1\right).\end{aligned}$$

Therefore, the dominating term is the first one with

$$n\sum_{t=1}^{T}\Delta \hat{F}_t\left(\Delta F_t - \Delta \hat{F}_t\right)'\beta = O_p\left(nT\right).$$

Combining this with the rate of convergence of the denominator as given in equation (68), we obtain equation (39). As far as the distribution limit is concerned, we have

$$\frac{1}{nT}\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta \hat{F}_t \Delta \hat{F}_s' \gamma_{t-s} V^{-1}\beta \xrightarrow{p} \bar{h}_e \Sigma_{\Delta F} V^{-1}\beta.$$

Combining this with equation (69), and recalling the definition of $\Sigma_{\Delta F}$, we can derive equation (40).

■

Proof of Theorem 5. The results stated in the theorem hold for any consistent estimator of F_t; we therefore consider an estimator, $\breve{F}_t$, such that for all t

$$\breve{F}_t - F_t = O_p\left(n^{-\delta}\right),$$

for some $\delta > 0$. In this case we have

$$\begin{aligned}\sum_{t=1}^{T}\sum_{i=1}^{n}\breve{F}_t u_{it} &= \sum_{t=1}^{T}\sum_{i=1}^{n}F_t u_{it} + \sum_{t=1}^{T}\sum_{i=1}^{n}\left(\breve{F}_t - F_t\right)u_{it} \\ &= O_p\left(n^{1/2}\right) + O_p\left(n^{-\delta}\right) O_p\left(n^{1/2}\right) = O_p\left(n^{1/2}\right),\end{aligned}$$

where the first term is $O_p\left(n^{1/2}\right)$ as proved in Theorem 1 and the second one is always dominated. Note that the summation over t does not play any role since T is fixed. Moreover, in light of the consistency of $\breve{F}_t$ we have

$$\sum_{t=1}^{T}\sum_{i=1}^{n}\breve{F}_t\breve{F}_t' = \sum_{t=1}^{T}\sum_{i=1}^{n}F_tF_t' + o_p(1) = O_p(n).$$

■

Proof of Theorem 6. This theorem can be proved following the same lines as for Theorem 5 and therefore is omitted. ■

Proof of Proposition 1. Equation (42) follows from Lemma 3 in Bai (2004).

As far as equation (43) is concerned, let $\bar{F}_t$ be the principal component estimator for F_t as defined in Bai (2004). Then we know (see e.g. the proof of Lemma 3 in Bai, 2004) that $T\left(\hat{\Lambda}-\Lambda\right)$ can be decomposed as

$$T\left(\hat{\Lambda}-\Lambda\right) = \frac{1}{T}\left[\sum_{t=1}^{T}e_tF_t' + \sum_{t=1}^{T}e_t\left(\bar{F}_t - F_t\right)' + \Lambda\sum_{t=1}^{T}\left(F_t-\bar{F}_t\right)\bar{F}_t'\right]\left[\frac{1}{T^2}\sum_{t=1}^{T}\bar{F}_t\bar{F}_t'\right]^{-1}. \tag{70}$$

As far as the denominator of this expression is concerned, let $\Xi = \int B_\varepsilon B_\varepsilon'$. We have

$$\sum_{t=1}^{T}\bar{F}_t\bar{F}_t' = \sum_{t=1}^{T}F_tF_t' + \sum_{t=1}^{T}\left(\bar{F}_t - F_t\right)\bar{F}_t' + \sum_{t=1}^{T}\bar{F}_t\left(\bar{F}_t - F_t\right)' + \sum_{t=1}^{T}\left(\bar{F}_t - F_t\right)\left(\bar{F}_t - F_t\right)',$$

where

$$\sum_{t=1}^{T}F_tF_t' = O_p\left(T^2\right),$$

$$\sum_{t=1}^{T}\left(\bar{F}_t - F_t\right)\bar{F}_t' = O_p(T),$$

$$\sum_{t=1}^{T}\left(\bar{F}_t - F_t\right)\left(\bar{F}_t - F_t\right)' = O_p(T);$$

the last two equalities come directly from Lemma B.4(ii) and Lemma B.1 in Bai (2004). Therefore

$$T^{-2}\sum_{t=1}^{T}\bar{F}_t\bar{F}_t' = T^{-2}\sum_{t=1}^{T}F_tF_t' + O_p\left(T^{-1}\right)$$

and

$$T^{-2}\sum_{t=1}^{T}\bar{F}_t\bar{F}_t' \Rightarrow \Xi.$$

As far as the numerator of equation (70) is concerned, we study each term. First of all we know that $T^{-1}\sum_{t=1}^{T} e_t F_t' \Rightarrow \int dW_e B_\varepsilon$. The limiting distribution of $\sum_{t=1}^{T} e_t \left(\bar{F}_t - F_t\right)'$ can be obtained from the following decomposition - see Bai (2004, p. 164) for details:

$$\bar{F}_t - F_t = T^{-2}\sum_{s=1}^{T}\tilde{F}_s\gamma_n(s,t) + T^{-2}\sum_{s=1}^{T}\tilde{F}_s\zeta_{st} + T^{-2}\sum_{s=1}^{T}\tilde{F}_s\eta_{st} + T^{-2}\sum_{s=1}^{T}\tilde{F}_s\xi_{st},$$

where (as in Lemma 1) we let $\gamma_n(s,t) = E(e_t'e_s/n)$, $\zeta_{st} = e_t'e_s/n - \gamma_n(s,t)$, $\eta_{st} = F_s'\Lambda' e_t/n$, $\xi_{st} = F_t'\Lambda' e_s/n$. Hence

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} e_t\left(\bar{F}_t - F_t\right)' &= T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_t\tilde{F}_s'\gamma_n(s,t) + T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_t\tilde{F}_s'\zeta_{st} \\
&\quad + T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_t\tilde{F}_s'\eta_{st} + T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_t\tilde{F}_s'\xi_{st} \\
&= I + II + III + IV,
\end{aligned}$$

and

$I = O_p\left(T^{-1}\right)$ since $E\left\|e_t\tilde{F}_s'\gamma_n(s,t)\right\| \leq |\gamma_n(s,t)|\left(\max_{s,t} E\left\|e_t\tilde{F}_s'\right\|\right)$ and $\max_{s,t} E\left\|e_t\tilde{F}_s'\right\| = O_p(T)$;

$II = n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_t\tilde{F}_s'e_t'e_s - T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_t\tilde{F}_s'\gamma_n(s,t)$ and we have

$$\begin{aligned}
n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_t\tilde{F}_s'e_t'e_s &= n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_te_t'e_s\tilde{F}_s' \\
&= n^{-1}T^{-1}\left(T^{-1}\sum_{t=1}^{T} e_te_t'\right)\left(T^{-1}\sum_{s=1}^{T} e_s\tilde{F}_s'\right) = O_p\left(T^{-1}\right);
\end{aligned}$$

$III = n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_t\tilde{F}_s'F_s'\Lambda' e_t$ with

$$\begin{aligned}
n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_t\tilde{F}_s'F_s'\Lambda' e_t &= n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_te_t'\Lambda F_s\tilde{F}_s' \\
&= n^{-1}\left(T^{-1}\sum_{t=1}^{T} e_te_t'\right)\Lambda\left(T^{-2}\sum_{s=1}^{T} F_s\tilde{F}_s'\right) = O_p(1);
\end{aligned}$$

$IV = n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_t\tilde{F}_s'F_t'\Lambda' e_s$ and

$$\begin{aligned}
n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_t\tilde{F}_s'F_t'\Lambda' e_s &= n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_tF_t'\Lambda' e_s\tilde{F}_s' \\
&= n^{-1}T^{-1}\left(T^{-1}\sum_{t=1}^{T} e_tF_t'\right)\Lambda'\left(T^{-1}\sum_{s=1}^{T} e_s\tilde{F}_s'\right) = O_p\left(T^{-1}\right).
\end{aligned}$$

Therefore the term that dominates is III and

$$n^{-1}\left(T^{-1}\sum_{t=1}^{T}e_te_t'\right)\Lambda\left(T^{-2}\sum_{s=1}^{T}F_s\tilde{F}_s'\right)\Rightarrow n^{-1}\Omega_e\Lambda Q.$$

Finally, as far as the term $\Lambda\sum_{t=1}^{T}\left(F_t-\bar{F}_t\right)\bar{F}_t'$ in equation (70) is concerned, we have

$$\begin{aligned}T^{-1}\Lambda\sum_{t=1}^{T}\left(F_t-\bar{F}_t\right)\bar{F}_t' &= -T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T}\tilde{F}_s\bar{F}_t'\gamma_n\left(s,t\right)-T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T}\tilde{F}_s\bar{F}_t'\zeta_{st}\\ &\quad -T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T}\tilde{F}_s\bar{F}_t'\eta_{st}-T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T}\tilde{F}_s\bar{F}_t'\xi_{st}\\ &= a+b+c+d.\end{aligned}$$

We have that the terms a and b follow from the proof of Lemma B.4 in Bai, 2004):

$$\begin{aligned}a&=O_p\left(T^{-1}\right);\\ b&=O_p\left(T^{-1}\right),\end{aligned}$$

the term

$$c=n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T}\tilde{F}_s\bar{F}_t'F_s\Lambda'e_t,$$

with

$$\begin{aligned}n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T}\tilde{F}_s\bar{F}_t'F_s\Lambda'e_t &= n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T}\tilde{F}_sF_s'\Lambda'e_t\bar{F}_t'\\ &= n^{-1}\left(T^{-2}\sum_{s=1}^{T}\tilde{F}_sF_s'\right)\Lambda'\left(T^{-1}\sum_{t=1}^{T}e_t\bar{F}_t'\right)=O_p\left(1\right);\end{aligned}$$

and

$$d=n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T}\tilde{F}_s\bar{F}_t'F_t'\Lambda'e_s,$$

with

$$\begin{aligned}n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T}\tilde{F}_s\bar{F}_t'F_t'\Lambda'e_s &= n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T}\tilde{F}_se_s'\Lambda F_t'\bar{F}_t'\\ &= n^{-1}\left(T^{-1}\sum_{s=1}^{T}\tilde{F}_se_s'\right)\Lambda'\left(T^{-1}\sum_{t=1}^{T}F_t\bar{F}_t'\right)=O_p\left(1\right).\end{aligned}$$

Thus the limiting distribution of $\Lambda\sum_{t=1}^{T}\left(F_t-\bar{F}_t\right)\bar{F}_t'$ is determined by c and d, and we have

$$\begin{aligned}
c &= n^{-1}\left(T^{-2}\sum_{s=1}^{T}F_s\tilde{F}_s'\right)\Lambda'\left(T^{-1}\sum_{t=1}^{T}e_t\bar{F}_t'\right)= \\
& n^{-1}\left(T^{-2}\sum_{s=1}^{T}F_s\tilde{F}_s'\right)\Lambda'\left(T^{-1}\sum_{t=1}^{T}e_tF_t\right) \\
& +n^{-1}\left(T^{-2}\sum_{s=1}^{T}F_s\tilde{F}_s'\right)\Lambda'\left[T^{-1}\sum_{t=1}^{T}e_t\left(\bar{F}_t-F_t\right)'\right] \\
&\Rightarrow n^{-1}Q\Lambda'\left[\int dW_eB_\varepsilon'+n^{-1}\Omega_e\Lambda Q\right],
\end{aligned}$$

and

$$d=n^{-1}\left(T^{-1}\sum_{s=1}^{T}\tilde{F}_se_s'\right)\Lambda'\left(T^{-1}\sum_{t=1}^{T}F_t\bar{F}_t'\right)\Rightarrow n^{-1}\left[\int B_\varepsilon dW_e'+n^{-1}Q\Lambda'\Omega_e\right]\Lambda Q.$$

Combining the limiting distributions of all terms $\sum_{t=1}^{T}\bar{F}_t\bar{F}_t'$, $\sum_{t=1}^{T}e_tF_t'$, $\sum_{t=1}^{T}e_t\left(\bar{F}_t-F_t\right)'$ and $\Lambda\sum_{t=1}^{T}\left(F_t-\bar{F}_t\right)\bar{F}_t'$ in equation (70), we obtain equation (43). ■

Proof of Proposition 2. Consider the estimation error

$$\begin{aligned}
\hat{F}_t-F_t &= n^{-1}\hat{\Lambda}'z_t-F_t \\
&= n^{-1}\hat{\Lambda}'\Lambda F_t+n^{-1}\hat{\Lambda}'e_t-F_t \\
&= n^{-1}\hat{\Lambda}'\hat{\Lambda}F_t+n^{-1}\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)F_t+n^{-1}\hat{\Lambda}'e_t-F_t.
\end{aligned}$$

Since we know that, by construction, $\hat{\Lambda}'\hat{\Lambda}=nI_k$, we have

$$n^{-1}\hat{\Lambda}'z_t-F_t=n^{-1}\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)F_t+n^{-1}\hat{\Lambda}'e_t=I+II.$$

As far as I is concerned, it holds that, omitting n^{-1} for the sake of brevity

$$\max_{1\leq t\leq T}\left\|\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)F_t\right\|\leq\left\|\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)\right\|\max_{1\leq t\leq T}\left\|F_t\right\|;$$

since

$$\left\|\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)\right\|=O_p\left(T^{-1}\right)$$

and

$$\max_{1\leq t\leq T}\left\|F_t\right\|=O_p\left(T^{1/2}\right),$$

we get

$$\max_{1\leq t\leq T}\left\|\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)F_t\right\|=O_p\left(T^{-1/2}\right).$$

Therefore $I = O_p\left(T^{-1/2}\right)$ uniformly in t. As per II, we have

$$\begin{aligned}
\hat{\Lambda}' e_t &= \Lambda' e_t + \left(\hat{\Lambda} - \Lambda\right)' e_t \leq \max_{1 \leq t \leq T} \left\|\Lambda' e_t\right\| + \max_{1 \leq t \leq T} \left\|\left(\hat{\Lambda} - \Lambda\right)' e_t\right\| \\
&\leq \|\Lambda\| \max_{1 \leq t \leq T} \|e_t\| + \left\|\left(\hat{\Lambda} - \Lambda\right)\right\| \max_{1 \leq t \leq T} \|e_t\| = O_p(1) + O_p\left(T^{-1}\right) O_p(1).
\end{aligned}$$

Hence, $II = O_p(1)$. Thus we have

$$\max_{1 \leq t \leq T} \left\|\hat{F}_t - F_t\right\| = O_p(1),$$

which proves equation (44).

Equation (45) can be derived following a similar argument. ■

Proof of Theorem 7. Recall equation (29)

$$\hat{\beta} - \beta = \left[\sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_t \hat{W}_t'\right]^{-1} \left\{\sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_t \left[\left(W_t - \hat{W}_t\right)' \beta + u_{it}\right]\right\}.$$

As far as the denominator of $\hat{\beta} - \beta$ is concerned, we have

$$\sum_{t=1}^{T} \hat{W}_t \hat{W}_t' = \sum_{t=1}^{T} W_t W_t' + o_p(1).$$

We prove this with respect to $\sum_{t=1}^{T} \hat{F}_t \hat{F}_t'$; extension to $\sum_{t=1}^{T} \hat{W}_t \hat{W}_t'$ is straightforward though notationally more involved. First, consider the following decomposition:

$$\begin{aligned}
\sum_{t=1}^{T} \hat{F}_t \hat{F}_t' &= \sum_{t=1}^{T} F_t F_t' + \sum_{t=1}^{T} \hat{F}_t \left(F_t - \hat{F}_t\right)' \\
&\quad + \sum_{t=1}^{T} \left(F_t - \hat{F}_t\right) \hat{F}_t' + \sum_{t=1}^{T} \left(F_t - \hat{F}_t\right) \left(F_t - \hat{F}_t\right)' \\
&\quad + I + II + III + IV.
\end{aligned}$$

We have

$$I = \sum_{t=1}^{T} F_t F_t' = O_p\left(T^2\right).$$

As far as II and III are concerned, it holds that

$$\begin{aligned}
III &= \sum_{t=1}^{T} \left[n^{-1} \hat{\Lambda}' \Lambda F_t + n^{-1} \hat{\Lambda}' e_t - F_t\right] z_t' \hat{\Lambda} n^{-1} \\
&= \sum_{t=1}^{T} \left[n^{-1} \hat{\Lambda}' \hat{\Lambda} F_t - n^{-1} \hat{\Lambda}' \left(\hat{\Lambda} - \Lambda\right) F_t + n^{-1} \hat{\Lambda}' e_t - F_t\right] z_t' \hat{\Lambda} n^{-1} \\
&= -n^{-2} \hat{\Lambda}' \left(\hat{\Lambda} - \Lambda\right) \left[\sum_{t=1}^{T} F_t z_t'\right] \hat{\Lambda} + n^{-2} \hat{\Lambda}' \left[\sum_{t=1}^{T} e_t z_t'\right] \hat{\Lambda},
\end{aligned}$$

with

$$n^{-2}\hat{\Lambda}'\left(\hat{\Lambda}-\Lambda\right)\left[\sum_{t=1}^{T}F_{t}z_{t}'\right]\hat{\Lambda}=O_{p}\left(T^{-1}\right)O_{p}\left(T^{2}\right)=O_{p}\left(T\right),$$

and

$$n^{-2}\hat{\Lambda}'\left[\sum_{t=1}^{T}e_{t}z_{t}'\right]\hat{\Lambda}=O_{p}\left(T\right);$$

therefore $II=O_{p}\left(T\right)$. As far as IV is concerned

$$\begin{aligned}
IV &= n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\left[\left(\Lambda-\hat{\Lambda}\right)F_{t}+e_{t}\right]\left[\left(\Lambda-\hat{\Lambda}\right)F_{t}+e_{t}\right]'\hat{\Lambda} \\
&= n^{-2}\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)\sum_{t=1}^{T}F_{t}F_{t}'\left(\Lambda-\hat{\Lambda}\right)'\hat{\Lambda} \\
&\quad +n^{-2}\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)\sum_{t=1}^{T}F_{t}e_{t}'\hat{\Lambda}+n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}e_{t}F_{t}'\left(\Lambda-\hat{\Lambda}\right)'\hat{\Lambda} \\
&\quad +n^{-2}\hat{\Lambda}'\left(\sum_{t=1}^{T}e_{t}e_{t}'\right)\hat{\Lambda},
\end{aligned}$$

with

$$n^{-2}\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)\sum_{t=1}^{T}F_{t}F_{t}'\left(\Lambda-\hat{\Lambda}\right)'\hat{\Lambda}=O_{p}\left(1\right),$$

$$n^{-2}\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)\sum_{t=1}^{T}F_{t}e_{t}'\hat{\Lambda}=O_{p}\left(1\right),$$

and

$$n^{-2}\hat{\Lambda}'\left(\sum_{t=1}^{T}e_{t}e_{t}'\right)\hat{\Lambda}=O_{p}\left(T\right);$$

therefore, $IV=O_{p}\left(T\right)$. Thus we get

$$T^{-2}\sum_{t=1}^{T}\hat{F}_{t}\hat{F}_{t}'=T^{-2}\sum_{t=1}^{T}F_{t}F_{t}'+O_{p}\left(T^{-1}\right).$$

Note that even if the estimated factors are not consistent, $T^{-2}\sum_{t=1}^{T}\hat{F}_{t}\hat{F}_{t}'$ is a consistent estimator for $T^{-2}\sum_{t=1}^{T}F_{t}F_{t}'$. This holds for any consistent estimator $\hat{\Lambda}$ such that $\hat{\Lambda}-\Lambda=O_{p}\left(T^{-\delta}\right)$; in such case, consistency would be ensured at a rate $\min\left\{1,\delta\right\}$.

With respect to the numerator of equation (29), this is equal to

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_{t}u_{it}+\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_{t}\left(W_{t}-\hat{W}_{t}\right)'\beta=I+II.$$

We have:

$$
\begin{aligned}
I &= \sum_{t=1}^{T} W_t u_{it} + \sum_{t=1}^{T} \left(\hat{W}_t - W_t \right) u_{it} \\
&= \sum_{t=1}^{T} W_t u_{it} + n^{-1} \hat{\Lambda}' \left(\Lambda - \hat{\Lambda} \right) \sum_{t=1}^{T} W_t u_{it} + n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} e_t u_{it},
\end{aligned}
$$

with

$$
\sum_{t=1}^{T} W_t u_{it} = O_p(T),
$$

$$
n^{-1} \hat{\Lambda}' \left(\Lambda - \hat{\Lambda} \right) \sum_{t=1}^{T} W_t u_{it} = O_p\left(T^{-1}\right) O_p(T) = O_p(1),
$$

and

$$
n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} e_t u_{it} = O_p\left(T^{1/2}\right),
$$

which follows from Assumption 6. As far as II is concerned, we have

$$
\begin{aligned}
II &= n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} z_t \left(W_t - n^{-1} \hat{\Lambda}' \bar{z}_t \right)' \\
&= n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} z_t \left[W_t - n^{-1} \hat{\Lambda}' \Lambda W_t - n^{-1} \hat{\Lambda}' e_t - n^{-1} \left(\hat{\Lambda} - \Lambda \right) z_t \right]' \\
&= -n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \bar{z}_t e_t' \Lambda - n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \bar{z}_t \bar{z}_t' \left(\hat{\Lambda} - \Lambda \right)' = O_p(T) + O_p\left(T^{-1}\right) O_p\left(T^2\right).
\end{aligned}
$$

Hence, the numerator of equation (29) is $O_p(T)$. Combining this result with the asymptotic magnitude of the denominator of equation (29), we get

$$
\begin{aligned}
& \left[\sum_{i=1}^{n} \sum_{t=1}^{T} W_t W_t' + o_p(1) \right]^{-1} \left\{ \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_t \left[\left(W_t - \hat{W}_t \right)' \beta + u_{it} \right] \right\} \\
&= O_p\left(T^{-2}\right) O_p(T) = O_p\left(T^{-1}\right).
\end{aligned}
$$

This proves equation (46).

As far as the limiting distribution of the numerator of equation (29) is concerned, we first study the term $\sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_t \left(W_t - \hat{W}_t \right)' \beta$. We have:

$$
\begin{aligned}
\sum_{t=1}^{T} \hat{W}_t \left(W_t - \hat{W}_t \right)' \beta &= n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} \bar{z}_t \left(W_t - n^{-1} \hat{\Lambda}' \bar{z}_t \right)' \beta \\
&= n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} \bar{z}_t W_t' \beta - n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \bar{z}_t \bar{z}_t' \hat{\Lambda} \beta \\
&= I + II.
\end{aligned}
$$

Since it holds that $\bar{z}_t = \Lambda W_t + \bar{e}_t$, we have

$$I = n^{-1}\hat{\Lambda}'\sum_{t=1}^{T}\Lambda W_t W_t'\beta + n^{-1}\hat{\Lambda}'\sum_{t=1}^{T}\bar{e}_t W_t'\beta,$$

and

$$\begin{aligned}
-II &= n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\Lambda W_t W_t'\Lambda'\hat{\Lambda}\beta + n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\bar{e}_t\bar{e}_t'\hat{\Lambda}\beta + \\
&\quad + n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\Lambda W_t\bar{e}_t'\hat{\Lambda}\beta + n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\bar{e}_t W_t'\Lambda'\hat{\Lambda}\beta.
\end{aligned}$$

Given that asymptotically $\hat{\Lambda} = \Lambda + T^{-1}D_{\Lambda}^{1}$, we have, with respect to I:

$$\begin{aligned}
n^{-1}\hat{\Lambda}'\sum_{t=1}^{T}\Lambda W_t W_t'\beta + n^{-1}\hat{\Lambda}'\sum_{t=1}^{T}\bar{e}_t W_t'\beta &= n^{-1}\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\beta + n^{-1}T^{-1}D_{\Lambda}^{1}\Lambda\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\beta \\
&\quad + n^{-1}\Lambda'\int d\bar{B}_e\bar{B}_{\varepsilon}'\beta, \qquad (71)
\end{aligned}$$

and, as far as II is concerned

$$\begin{aligned}
& n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\Lambda W_t W_t'\Lambda'\hat{\Lambda}\beta + n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\bar{e}_t\bar{e}_t'\hat{\Lambda}\beta + n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\Lambda W_t\bar{e}_t'\hat{\Lambda}\beta + n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\bar{e}_t W_t'\Lambda'\hat{\Lambda}\beta \\
= \; & n^{-2}T^{-1}D_{\Lambda}^{1}\Lambda\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\beta + n^{-2}T^{-1}\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\Lambda' D_{\Lambda}^{1\prime}\beta \\
& + n^{-2}\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\beta + n^{-2}\Lambda'\Sigma_e\Lambda\beta + n^{-2}\int \bar{B}_{\varepsilon}d\bar{B}_e'\Lambda\beta \\
& + n^{-2}\Lambda'\int d\bar{B}_e\bar{B}_{\varepsilon}'\beta. \qquad (72)
\end{aligned}$$

Thus, combining equations (71) and (72) we have

$$\begin{aligned}
\hat{\Lambda}'\sum_{t=1}^{T}\bar{z}_t W_t'\beta - \hat{\Lambda}'\sum_{t=1}^{T}\bar{z}_t\bar{z}_t'\hat{\Lambda}\beta &= n^{-1}\left(1-n^{-1}\right)\left[\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\beta + T^{-1}D_{\Lambda}^{1}\Lambda\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\beta + \Lambda'\int d\bar{B}_e\bar{B}_{\varepsilon}'\beta\right] \\
&\quad - n^{-2}\left[T^{-1}\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\Lambda' D_{\Lambda}^{1\prime}\beta + \Lambda'\Sigma_e\Lambda\beta + \int \bar{B}_{\varepsilon}d\bar{B}_e'\Lambda\beta\right].
\end{aligned}$$

As far as the term $\sum_{t=1}^{T}\hat{W}_t\left(\sum_{i=1}^{n}u_{it}\right)$ is concerned, we have

$$\begin{aligned}
\sum_{t=1}^{T}\hat{W}_t\left(\sum_{i=1}^{n}u_{it}\right) &= n^{-1}\hat{\Lambda}'\sum_{t=1}^{T}\bar{z}_t\left(\sum_{i=1}^{n}u_{it}\right) \\
&= n^{-1}\hat{\Lambda}'\sum_{t=1}^{T}\Lambda W_t\left(\sum_{i=1}^{n}u_{it}\right) + n^{-1}\hat{\Lambda}'\sum_{t=1}^{T}\bar{e}_t\left(\sum_{i=1}^{n}u_{it}\right),
\end{aligned}$$

which asymptotically leads to

$$\frac{1}{T}n^{-1}\hat{\Lambda}'\Lambda\sum_{t=1}^{T}W_t\left(\sum_{i=1}^{n}u_{it}\right)\Rightarrow n^{-1}\int\bar{B}_{\varepsilon}dB_u\left(\sum_{i=1}^{n}\sum_{j=1}^{n}h_{ij}\right)^{1/2}.$$

This completes the proof of (47).

Finally, we consider the case when equation (2) is a spurious relationship. Since $u_{it}\sim I(1)$, we have that

$$\sum_{t=1}^{T}\hat{W}_t\left(W_t-\hat{W}_t\right)'\beta=O_p(T),$$

and

$$\begin{aligned}\sum_{t=1}^{T}\hat{W}_tu_{it} &= \sum_{t=1}^{T}W_tu_{it}+n^{-1}\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)\sum_{t=1}^{T}W_tu_{it}+n^{-1}\hat{\Lambda}'\sum_{t=1}^{T}e_tu_{it}\\ &= O_p\left(T^2\right)+O_p(T)+O_p(T),\end{aligned}$$

so that $\sum_{t=1}^{T}\hat{W}_t\left[\left(W_t-\hat{W}_t\right)'\beta+u_{it}\right]=O_p\left(T^2\right)$, which proves (48).

In this case the limiting distribution of the numerator is given by the leading term $\sum_{t=1}^{T}W_t\left(\sum_{i=1}^{n}u_{it}\right)$, so that the same result as in equation (60) holds, namely

$$\frac{1}{nT^2}\sum_{t=1}^{T}W_t\left(\sum_{i=1}^{n}u_{it}\right)\Rightarrow\sqrt{\bar{h}^{\Delta}}\left(\int\bar{B}_{\varepsilon}B_u\right).$$

This proves equation (49). ■

Proof of Theorem 8. Consider equation (30)

$$\hat{\beta}^{FD}-\beta^{FD}=\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta\hat{F}_t'\right]^{-1}\left\{\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\left[\left(\Delta F_t-\Delta\hat{F}_t\right)'\beta+\Delta u_{it}\right]\right\}.$$

As far as the denominator is concerned, we have

$$\sum_{t=1}^{T}\Delta\hat{F}_t\Delta\hat{F}_t'=n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\Delta z_t\Delta z_t'\hat{\Lambda}=O_p(T).$$

As far as the numerator of $\hat{\beta}^{FD}-\beta^{FD}$ is concerned, we have

$$\begin{aligned}\sum_{t=1}^{T}\Delta\hat{F}_t\left(\Delta F_t-\Delta\hat{F}_t\right)' &= n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\Delta z_t\left[\Delta F_t'\left(\hat{\Lambda}-\Lambda\right)'-\Delta e_t'\right]\hat{\Lambda}\\ &= n^{-2}\hat{\Lambda}'\left[\sum_{t=1}^{T}\Delta z_t\Delta F_t'\right]\left(\hat{\Lambda}-\Lambda\right)'\hat{\Lambda}-n^{-2}\hat{\Lambda}'\left[\sum_{t=1}^{T}\Delta z_t\Delta e_t'\right]\hat{\Lambda}\\ &= O_p(T)\,O_p\left(T^{-1}\right)+O_p(T)=O_p(T).\end{aligned}$$

Also we have

$$\begin{aligned}\sum_{t=1}^{T} \Delta \hat{F}_t \Delta u_{it} &= \sum_{t=1}^{T} \Delta F_t \Delta u_{it} + \sum_{t=1}^{T} \left(\Delta \hat{F}_t - \Delta F_t \right) \Delta u_{it} \\ &= O_p\left(\sqrt{T}\right) + O_p\left(\sqrt{T}\right) = O_p\left(\sqrt{T}\right).\end{aligned}$$

This proves (50).

The limiting distribution of $\hat{\beta}^{FD} - \beta^{FD}$ can be obtained as follows. Consider first the denominator of $\hat{\beta}^{FD} - \beta^{FD}$. Given that $T^{-1}\sum_{t=1}^{T} \Delta \hat{F}_t \Delta \hat{F}_t' = n^{-2}T^{-1}\hat{\Lambda}\sum_{t=1}^{T} \Delta z_t \Delta z_t' \hat{\Lambda}'$, and recalling that

$$p\lim \frac{1}{T}\sum_{t=1}^{T} \Delta z_t \Delta z_t' = \Sigma_{\Delta z},$$

we have

$$n^{-2}T^{-1}\hat{\Lambda}'\sum_{t=1}^{T} \Delta z_t \Delta z_t' \hat{\Lambda} \xrightarrow{p} n^{-2}\Lambda' \Sigma_{\Delta z} \Lambda.$$

As far as the numerator of $\hat{\beta}^{FD} - \beta^{FD}$ is concerned, the term that dominates is $\sum_{t=1}^{T} \Delta \hat{F}_t \left(\Delta F_t - \Delta \hat{F}_t \right)' \beta$ and we have:

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T} \Delta \hat{F}_t \left(\Delta F_t - \Delta \hat{F}_t \right)' \beta = \frac{1}{T} n^{-1}\hat{\Lambda}' \sum_{t=1}^{T} \Delta z_t \left(\Delta F_t' - n^{-1}\Delta z_t' \hat{\Lambda}' \right) \beta \\ = \; & \frac{1}{T} n^{-1}\hat{\Lambda}' \sum_{t=1}^{T} \Delta z_t \Delta F_t' \beta - \frac{1}{T} n^{-2}\hat{\Lambda}' \sum_{t=1}^{T} \Delta z_t \Delta z_t' \hat{\Lambda}' \beta \\ = \; & \frac{1}{T} n^{-1}\hat{\Lambda}' \sum_{t=1}^{T} \Lambda \Delta F_t \Delta F_t' \beta + \frac{1}{T} n^{-1}\hat{\Lambda}' \sum_{t=1}^{T} \Delta e_t \Delta F_t' \beta - \frac{1}{T} n^{-2}\hat{\Lambda}' \sum_{t=1}^{T} \Delta z_t \Delta z_t' \hat{\Lambda}' \beta,\end{aligned}$$

where $n^{-1}T^{-1}\hat{\Lambda}'\sum_{t=1}^{T} \Delta e_t \Delta F_t' \beta$ is of order $O_p\left(T^{-1/2}\right)$. Since

$$n^{-1}\frac{1}{T}\hat{\Lambda}'\sum_{t=1}^{T} \Lambda \Delta F_t \Delta F_t' \beta \xrightarrow{p} n^{-1}\Sigma_{\Delta F}\beta,$$

and

$$n^{-2}\frac{1}{T}\hat{\Lambda}'\sum_{t=1}^{T} \Delta z_t \Delta z_t' \hat{\Lambda}' \beta \xrightarrow{p} n^{-2}\Lambda' \Sigma_{\Delta z} \Lambda \beta,$$

we have

$$\frac{1}{T}\sum_{t=1}^{T} \Delta \hat{F}_t \left(\Delta F_t - \Delta \hat{F}_t \right)' \beta \xrightarrow{p} n^{-1}\Sigma_{\Delta F}\beta - n^{-2}\Lambda' \Sigma_{\Delta z} \Lambda \beta.$$

Recalling that the denominator converges to $n^{-2}\Lambda' \Sigma_{\Delta z} \Lambda$ in probability, we finally obtain equation (51). ■

Lemma 4 *Let $\hat{\Lambda}$ be the principal component estimator for Λ in*

$$\Delta z_t = \Lambda \Delta F_t + \Delta e_t,$$

and let $\hat{F}_t = n^{-1}\hat{\Lambda}^{FD\prime} z_t$ and $\Delta\hat{F}_t = n^{-1}\hat{\Lambda}^{FD\prime}\Delta z_t$. It holds that

$$\max_{1\leq t\leq T}\left\|\hat{F}_t - F_t\right\| = O_p\left(T^{1/2}\right), \tag{73}$$

uniformly in t and with respect to the first differenced estimates given by $\Delta\hat{F}_t = \hat{\Lambda}^{FD\prime}\Delta z_t$

$$\max_{1\leq t\leq T}\left\|\Delta\hat{F}_t - \Delta F_t\right\| = O_p\left(1\right), \tag{74}$$

uniformly in t.

Proof. Consider the estimation error

$$\begin{aligned}
\hat{F}_t - F_t &= n^{-1}\hat{\Lambda}^{FD\prime} z_t - F_t \\
&= n^{-1}\hat{\Lambda}^{FD\prime}\Lambda F_t + n^{-1}\hat{\Lambda}^{FD\prime} e_t - F_t \\
&= n^{-1}\hat{\Lambda}^{FD\prime}\hat{\Lambda}^{FD} F_t + n^{-1}\hat{\Lambda}^{FD\prime}\left(\Lambda - \hat{\Lambda}^{FD}\right) F_t + n^{-1}\hat{\Lambda}^{FD\prime} e_t - F_t \\
&= n^{-1}\hat{\Lambda}^{FD\prime}\left(\Lambda - \hat{\Lambda}^{FD}\right) F_t + n^{-1}\hat{\Lambda}^{FD\prime} e_t = I + II.
\end{aligned}$$

As far as I is concerned, we have, omitting the term n^{-1}

$$\max_{1\leq t\leq T}\left\|\hat{\Lambda}^{FD\prime}\left(\Lambda - \hat{\Lambda}^{FD}\right) F_t\right\| \leq \left\|\hat{\Lambda}^{FD\prime}\left(\Lambda - \hat{\Lambda}^{FD}\right)\right\| \max_{1\leq t\leq T}\left\|F_t\right\|;$$

given that $\left\|\hat{\Lambda}^{FD\prime}\left(\Lambda - \hat{\Lambda}^{FD}\right)\right\| = O_p\left(1\right)$ and

$$\max_{1\leq t\leq T}\left\|F_t\right\| = O_p\left(T^{1/2}\right),$$

we get

$$\max_{1\leq t\leq T}\left\|\hat{\Lambda}^{FD\prime}\left(\Lambda - \hat{\Lambda}^{FD}\right) F_t\right\| = O_p\left(T^{1/2}\right).$$

As far as term I is concerned, it holds that:

$$\begin{aligned}
\hat{\Lambda}^{FD\prime} e_t &= \Lambda' e_t + \left(\hat{\Lambda}^{FD} - \Lambda\right)' e_t \leq \max_{1\leq t\leq T}\left\|\Lambda' e_t\right\| \max_{1\leq t\leq T}\left\|\left(\hat{\Lambda}^{FD} - \Lambda\right)' e_t\right\| \\
&\leq \left\|\Lambda\right\| \max_{1\leq t\leq T}\left\|e_t\right\| + \left\|\left(\hat{\Lambda}^{FD} - \Lambda\right)\right\| \max_{1\leq t\leq T}\left\|e_t\right\| = O_p\left(T^{1/2}\right) + O_p\left(1\right) O_p\left(T^{1/2}\right).
\end{aligned}$$

Hence, $\hat{\Lambda}^{FD\prime} e_t = O_p\left(T^{1/2}\right)$. This proves equation (73). Equation (74) can be derived noting that

$$\Delta\hat{F}_t - \Delta F_t = \hat{\Lambda}^{FD\prime}\left(\Lambda - \hat{\Lambda}^{FD}\right)\Delta F_t + \hat{\Lambda}^{FD\prime}\Delta e_t = I + II.$$

Here, as far as I is concerned, we have

$$\max_{1\leq t\leq T}\left\|\hat{\Lambda}^{FD\prime}\left(\Lambda-\hat{\Lambda}^{FD}\right)\Delta F_t\right\|\leq\left\|\hat{\Lambda}^{FD\prime}\left(\Lambda-\hat{\Lambda}^{FD}\right)\right\|\max_{1\leq t\leq T}\left\|\Delta F_t\right\|,$$

where

$$\max_{1\leq t\leq T}\left\|\Delta F_t\right\|=O_p\left(1\right).$$

As far as I is concerned, we have

$$\begin{aligned}\hat{\Lambda}^{FD\prime}\Delta e_t &= \Lambda'\Delta e_t+\left(\hat{\Lambda}^{FD}-\Lambda\right)'\Delta e_t\leq\max_{1\leq t\leq T}\left\|\Lambda'\Delta e_t\right\|\max_{1\leq t\leq T}\left\|\left(\hat{\Lambda}^{FD}-\Lambda\right)'\Delta e_t\right\|\\ &\leq \left\|\Lambda\right\|\max_{1\leq t\leq T}\left\|\Delta e_t\right\|+\left\|\left(\hat{\Lambda}^{FD}-\Lambda\right)\right\|\max_{1\leq t\leq T}\left\|\Delta e_t\right\|=O_p\left(1\right)+O_p\left(1\right)O_p\left(1\right).\end{aligned}$$

■

Proof of Theorem 9. Equation (29) states that

$$\hat{\beta}-\beta=\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\hat{W}_t'\right]^{-1}\left\{\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\left[\left(W_t-\hat{W}_t\right)'\beta+u_{it}\right]\right\}.$$

With respect to the denominator we have

$$\sum_{t=1}^{T}\hat{W}_t\hat{W}_t'=O_p\left(T^2\right),$$

which follows from the proof of theorem 7. The limiting distribution is given by

$$\frac{1}{T^2}\sum_{t=1}^{T}\hat{W}_t\hat{W}_t'=\frac{1}{T^2}n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\bar{z}_t\bar{z}_t'\hat{\Lambda}\Rightarrow D_{\Lambda}^{2\prime}\int\bar{B}_z\bar{B}_z'D_{\Lambda}^2.$$

Note that the denominator is not affected by the u_{it}s, and therefore the results derived for its rate of convergence and its limiting distribution hold irrespective of whether equation (2) is a cointegration or a spurious regression.

As far as the numerator is concerned, this is given by

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\left(W_t-\hat{W}_t\right)'\beta+\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_tu_{it}=I+II.$$

As far as I is concerned we have

$$\sum_{t=1}^{T}\hat{W}_t\left(W_t-\hat{W}_t\right)'\beta=O_p\left(T^2\right),$$

and the asymptotic law of I is given by

$$\begin{aligned}\frac{1}{T^2}\sum_{t=1}^{T}\hat{W}_t\left(W_t-\hat{W}_t\right)'\beta &= \frac{1}{T^2}n^{-1}\hat{\Lambda}'\sum_{t=1}^{T}\bar{z}_tW_t\beta-\frac{1}{T^2}n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\bar{z}_t\bar{z}_t'\hat{\Lambda}\beta\\ &\Rightarrow n^{-1}D_{\Lambda}^{2\prime}\int\bar{B}_z\bar{B}_{\varepsilon}'\beta-n^{-2}D_{\Lambda}^{2\prime}\int\bar{B}_z\bar{B}_z'D_{\Lambda}^2\beta.\end{aligned}$$

Since I does not depend on the u_{it}s, the results derived above for its rate of convergence and its limiting distribution hold independently of whether equation (2) is a cointegration or a spurious regression.

As far as II is concerned, we have that $II = O_p(T)$ - if equation (2) is a cointegrating regression - or $II = O_p\left(T^2\right)$ - if it is a spurious regression; in this case we have

$$\begin{aligned}
\frac{1}{T^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it} &= \frac{1}{T^2} n^{-1}\hat{\Lambda}'\sum_{t=1}^{T}\bar{z}_t\left(\sum_{i=1}^{n}u_{it}\right) \\
&\Rightarrow n^{-1}D_{\Lambda}^{2\prime}\int \bar{B}_z B_u\left(\sum_{i=1}^{n}\sum_{i=1}^{n}h_{ij}^{\Delta}\right)^{1/2}.
\end{aligned}$$

Thus, when equation (2) cointegrates, the term that dominates in the numerator is $I = \sum_{t=1}^{T}\hat{W}_t\left(W_t - \hat{W}_t\right)'\beta$. Combining its rate of convergence and its limiting distribution with the ones of the denominator, we prove equations (52) and (53). When equation (2) is a spurious regression, term II has the same order of magnitude as I. This does not affect the rate of convergence of $\hat{\beta}-\beta$, but it does affect its limiting distribution. Combining the asymptotic law of the denominator of $\hat{\beta}-\beta$ with the limiting distribution of terms I and II, we finally get equation (54).

Consider now $\hat{\beta}^{FD} - \beta^{FD}$. From equation (30) we know that

$$\hat{\beta}^{FD} - \beta^{FD} = \left[\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta\hat{F}_t'\right]^{-1}\left\{\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\left[\left(\Delta F_t - \Delta\hat{F}_t\right)'\beta + \Delta u_{it}\right]\right\}.$$

The denominator of this expression is

$$\sum_{t=1}^{T}\Delta\hat{F}_t\Delta\hat{F}_t' = O_p(T), \tag{75}$$

as proved in Theorem 8, and using the definition of D_{Λ}^2 it holds that

$$\frac{1}{T}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta\hat{F}_t' = \frac{1}{T}n^{-2}\sum_{t=1}^{T}\hat{\Lambda}'\Delta z_t\Delta z_t'\hat{\Lambda} \Rightarrow n^{-2}D_{\Lambda}^{2\prime}\Sigma_{\Delta z}D_{\Lambda}^2.$$

As far as the numerator is concerned, we have

$$\sum_{t=1}^{T}\Delta\hat{F}_t\left(\Delta F_t - \Delta\hat{F}_t\right)'\beta = O_p(T),$$

and

$$\sum_{t=1}^{T}\Delta\hat{F}_t\Delta u_{it} = O_p\left(T^{1/2}\right),$$

so that we have $\sum_{t=1}^{T}\Delta\hat{F}_t\left[\left(\Delta F_t-\Delta\hat{F}_t\right)'\beta+\Delta u_{it}\right]=O_p\left(T\right)$. Combining this with equation (75), we obtain equation (55). As far as the limiting distribution of the numerator is concerned, this can be derived noting that

$$\begin{aligned}\sum_{t=1}^{T}\Delta\hat{F}_t\left(\Delta F_t-\Delta\hat{F}_t\right)'\beta &= n^{-1}\sum_{t=1}^{T}\hat{\Lambda}'\Delta z_t\left(\Delta F_t-n^{-1}\hat{\Lambda}'\Delta z_t\right)'\beta \\ &= n^{-1}\hat{\Lambda}'\sum_{t=1}^{T}\Delta z_t\Delta F_t'\beta-n^{-2}\sum_{t=1}^{T}\hat{\Lambda}'\Delta z_t\Delta z_t'\hat{\Lambda}\beta,\end{aligned}$$

and

$$\begin{aligned}\frac{1}{T}n^{-1}\hat{\Lambda}'\sum_{t=1}^{T}\Delta z_t\Delta F_t'\beta &= n^{-1}\hat{\Lambda}'\Lambda\left(\frac{1}{T}\sum_{t=1}^{T}\Delta F_t\Delta F_t'\right)\beta+n^{-1}\hat{\Lambda}'\left(\frac{1}{T}\sum_{t=1}^{T}\Delta e_t\Delta F_t'\right)\beta \\ &\Rightarrow n^{-1}D_{\Lambda}^{2\prime}\Lambda\Sigma_{\Delta F}\beta,\end{aligned}$$

since $Cov\left(\Delta e_t,\Delta F_t\right)=0$. Combining this with the asymptotic law of the denominator, we get equation (56). ■

References

[1] Ahn, S.G., Lee, Y-H., and Schmidt, P. (2001), "GMM Estimation of Linear Panel Data Models with Time-Varying Individual Effects", *Journal of Econometrics*, 102, 219-255.

[2] Anderson, T. W. (1981), *An Introduction to Multivariate Statistical Analysis*, New York: Wiley.

[3] Andrews, D. W. K. (2005), "Cross-Section Regression with Common Shocks," *Econometrica*, 73, 1551-1586.

[4] Bai, J. (2003), "Inferential Theory for Structural Models of Large Dimensions", *Econometrica,* 71, 135-171.

[5] Bai, J. (2004), "Estimating Cross-Section Common Stochastic Trends in Nonstationary Panel Data", *Journal of Econometrics*, 122, 137-183.

[6] Bai, J. (2005), "Panel Data Models with Interactive Effects," Working Paper.

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

[8] Bai, J., and Ng, S. (2004), "A Panic Attack on Unit Roots and Cointegration", *Econometrica,* 72, 1127-1178.

[9] Bai, J., and Ng, S. (2005), "Evaluating Latent and Observed Factors in Macroeconomics and Finance", *Journal of Econometrics* (article in press).

[10] Bai, J., and Kao, C., (2005), "On the Estimation and Inference of a Cointegrated Regression in Panel Data with Cross-sectional Dependence," unpublished manuscript.

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

[12] Bernanke, B., Boivin, J., (2000), "Monetary Policy in a Data-Rich Environment", *Journal of Monetary Economics*, 50, 525-46.

[13] Boivin, J., Ng, S. (2003), "Are More Data Always Better for Factor Analysis", *Journal of Econometrics* (article in press).

[14] Breitung, J., Pesaran, M.H. (2005), "Unit Roots and Cointegration in Panels", forthcoming in L. Matyas, and P. Sevestre, The Econometrics of Panel Data (Third Edition), Kluwer Academic Publishers.

[15] Chamberlain, G., and Rothschild, M. (1983), "Arbitrage, Factor Structure and Mean-Variance Analysis in Large Asset markets," *Econometrica*, 51, 1305-1324.

[16] Conley, T. G. (1999), "GMM Estimation with Cross Sectional Dependence", *Journal of Econometrics*, 92, 1-45.

[17] Cragg, J., Donald, S. (1997), "Inferring the Rank of a Matrix", *Journal of Econometrics*, 76, 223-250.

[18] Donald, S.G. (1997), "Inference Concerning the Number of Factors in a Multivariate Nonparametric Relationship", *Econometrica*, Vol. 65, No.1, 103-131.

[19] Donald, S.G., Lang, K. (2004), "Inference with Difference in Differences and Other Panel Data", mimeo.

[20] Forni, M., Reichlin, L. (1998), "Let's Get Real: a Factor Analytic Approach to Disaggregated Business Cycle Dynamics", *Review of Economic Studies*, 65, 453-473.

[21] Gonzalo, J., and Granger, C.W.J. (1995), "Estimation of Common Long Memory Components in Cointegrated Systems", *Journal of Business and Economic Statistics*, 13, 27-35.

[22] Jin. S. (2005), "Discrete Choice Modeling with Nonstationary Panels Applied to Exchange Rate Regime Choice," unpublished manuscript.

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

[24] Lewbel, A. (1991), "The Rank of Demand Systems: Theory and Nonparametric Estimation", *Econometrica*, 59, No. 3, 711-730.

[25] Moon, H. R., and Perron, B. (2004), "Testing for a Unit Root in Panels with Dynamic Factors", *Journal of Econometrics*, 122, 81-126.

[26] Moulton, B.R. (1990), "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Unit", *The Review of Economics and Statistics*, 72(2), 334-38.

[27] Onatski, A. (2005), "Determining the Number of Factors form Empirical Distribution of Eigenvalues", unpublished manuscript.

[28] Pesaran, M. H. (2005a), "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure", http://www.econ.cam.ac.uk/faculty/pesaran/wp.htm#cdlarge.

[29] Pesaran, M. H. (2005b), "A Simple Panel Unit Root Test in the Presence of Cross Section Dependence", http://www.econ.cam.ac.uk/faculty/pesaran/PesaranCADFJan05.pdf

[30] Phillips, P. C. B., and Moon, H. R. (1999), "Linear Regression Limit Theory for Nonstationary Panel Data", *Econometrica*, 67, 1057-1112.

[31] Quah, D., and Sargent, T. J. (1993), "A Dynamic Index Model for Large Cross Sections," in J. H. Stock and M. W. Watson, eds, *Business Cycles, Indicators, and Forecasting*, Univesity of Chicago, Chicago.

[32] Stock, J. H., and Watson, M. W. (1999), "Foresting Inflation," *Journal of Monetary Economics*, 44, 293-335.

[33] Stock, J. H., and Watson, M. W. (2002), "Forecasting Using Principal Components From a Large Number of Predictors," *Journal of the American Statistical Association*, 97, 1167-1179.

[34] Stock, J. H., and Watson, M. W. (2005), "Implications of Dynamic Factor Models for Var Analysis," unpublished manuscript.