CENTRE FOR ECONOMETRIC ANALYSIS
CEA@Cass

http://www.cass.city.ac.uk/cea/index.html

Cass Business School
Faculty of Finance
106 Bunhill Row
London EC1Y 8TZ

Asymptotics for Panel Models with Common Shocks (Extended Version).

Chihwa Kao, Lorenzo Trapani and Giovanni Urga

CEA@Cass Working Paper Series

WP–CEA–01-2010

Asymptotics for Panel Models with Common Shocks - Extended Version*†

Chihwa Kao
Syracuse University
cdkao@maxwell.syr.edu

Lorenzo Trapani
Cass Business School
L.Trapani@city.ac.uk

Giovanni Urga
Cass Business School and Universita' di Bergamo
G.Urga@city.ac.uk

July 28, 2010

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 a cointegration relationship or a spurious regression, and we also consider the stationary case. We derive the rate of convergence and the limit distributions for the ordinary least squares (OLS) estimates of the model parameters under all the aforementioned cases.

JEL Classification: C13, C23.

Keywords: Panel data, common shocks, cross-sectional dependence, asymptotics, joint limit, martingale difference sequence.

*This paper is an extended version of Kao, Trapani and Urga (2010, forthcoming in *Econometric Reviews*), and it contains all Lemmas and proofs referred to in the published version.

†This article originated while Chihwa Kao was visiting the Centre for Econometric Analysis at Cass (CEA@Cass). We thank the Editor, Esfandiar Maasoumi, and Associate Editor and a Referee, for extremely constructive comments which have greatly improved the presentation of this paper. We wish to thank participants in the "1st Italian Congress of Econometrics and Empirical Economics" (Universita' Ca' Foscari, Venice, 24-25 January 2005), the 9th World Congress of the Econometric Society (London, 19-24 August 2005), in particular Peter Pedroni and Yixiao Sun, the North American Summer Meeting of the Econometric Society (Minneapolis, MN, 22-25 June 2006) and the 13th Conference on Panel Data (Cambridge, 7-9 July, 2006), the Monday Workshop of the Faculty of Finance at Cass Business School and the Lunch Seminars of the Department of Management and Information Technology at Bergamo University for useful comments. The paper has hugely benefited from conversations with Badi Baltagi and Jan Kiviet. All remaining errors are the authors' sole responsibility.

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 various degrees of cross-sectional dependence, e.g., 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 the case where it is the error terms across unit to be dependent (see for instance Bai and Kao, 2006; Moon and Perron, 2004) or both (see for instance Ahn, Lee, and Schmidt, 2001; Pesaran 2006).

The main aim of this paper is to propose a novel asymptotic theory for panel models where common shocks are present among the regressors, thereby introducing strong cross-sectional dependence. We generalize the asymptotics developed by Phillips and Moon (1999) and Andrews (2005) by employing and extending the theory for factor models in Bai (2003, 2004) and Bai and Ng (2004).

Phillips and Moon (1999) analyze nonstationary panels when both cross-sectional dimension n and time-series dimension T are large. They derive 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-sectional independence though the authors point out that their results still hold when certain degree of weak cross-sectional dependence is allowed. Thus, the derivation of sequential and joint asymptotics with arbitrary dependence amongst units remains 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 common across cross-sectional units, and may result in mixed asymptotic normality, as Andrews (2005) has demonstrated in a cross-sectional context. Andrews (2005, Theorem 4, p. 1567) proves that the presence of common factors among the cross-sectional units makes the limiting distribution of the OLS estimator of the regression slope mixed normal and not normal as in the classical regression analysis. Note that in this case mixed normality of the OLS estimator of the regression slope holds even if regressors are stationary, i.e., $I(0)$, and independent of errors. This finding is also obtained in our paper when studying the distribution limit for the OLS estimator of the regression slope for the fixed T case (see equation 20 in Theorem 2 below), while when we consider the $T \to \infty$ case, not explored by Andrews (2005), we show that in the stationary case as $T \to \infty$ the OLS estimator of the regression slope is normally distributed.

1.1 Basic Model and Extensions

In this paper we consider the following panel regression model with common shocks

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

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

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

Equation (1) could be either a spurious regression or a cointegration model depending on whether u_{it} is $I(1)$ or $I(0)$, respectively. It is important to emphasize that, as far as the presence of F_t is concerned, equation (1) represents a *panel regression model* with a set of regressors, F_t, which is common across units and with common slope coefficient β. Model (1) differs from a factor-loading specification as in Bai (2004) and Bai and Ng (2004), for example. Thus, in our setup F_t is a genuine (observable or unobservable) regressor rather than a "common factor". A framework which is similar in spirit to the one in this paper is in Stock and Watson (1999, 2002, 2005), where y_{it} in (1) (with $n = 1$) is the time-series variable to be predicted and $z_i = (z_{i1}, z_{i2}, ..., z_{iT})'$ is an n-dimensional multiple time-series of candidate predictors; also, model (1) resembles the panel cointegration model with global stochastic trends of Bai, Kao and Ng (2009), although (1) assumes having common β.

When common shocks are not observable, we assume that a set of exogenous variables, z_{it}, is observable such that

$$z_{it} = \lambda_i' F_t + e_{it} \tag{2}$$

where λ_i is a vector of factor loadings and e_{it} is an idiosyncratic component. We assume throughout the paper, for the sake of the simplicity of the notation, that the number of the z_{it}s is the same as that of the y_{it}s. However, the panel dimensions of y_{it} and the z_{it} may be different, for example y_{it} may refer to individuals while z_{it} may index several macro variables.

To extend our results to the stationary panel model case, we also consider the first-differenced form of model (1),

$$\Delta y_{it} = \beta' \Delta F_t + \Delta u_{it}. \tag{3}$$

Model (1) considers a very simple specification. However, it could be argued that a more complete and realistic framework should also embed a set of idiosyncratic shocks, i.e.,

$$y_{it} = \alpha_i + \beta' F_t + \gamma' x_{it} + u_{it}. \tag{4}$$

For the sake of notational simplicity, the main results in the paper, reported in Section 3, are derived under the restrictive assumptions of no idiosyncratic shocks, i.e., under the constrain that $\gamma = 0$. However, in Section 4 we show that our main results concerning the asymptotics of the estimator of β are still useful in presence of a more complicated specification as (4). This is obviously true when the regressors F_t and x_{it} are orthogonal. We also examine the case whereby the x_{it} are allowed to be correlated with F_t via the factor-loadings specification

$$x_{it} = \tau_i G_t + \omega_{it} \tag{5}$$

where G_t is a set of common factors that can be independent of the regressors F_t or (fully or partly) overlap with them, and ω_{it} is a unit specific (stationary or nonstationary) shock. A similar framework that allows for cross-sectional dependence among the idiosyncratic regressors and dependence between the idiosyncratic regressors and the common regressors is in Pesaran (2006) and Kapetanios, Pesaran and Yamagata (2006), even though in our paper F_t is a set of regressors and not nuisance parameters. Note that allowing for x_{it} being dependent upon F_t through some possibly heterogeneous loadings τ_i allows for the response of y_{it} to F_t being (indirectly) heterogeneous across individuals.

Models (1) and (4) are frequently employed for the purpose of forecasting (Stock and Watson, 1999, 2002, 2005), and they encompass a wide set of models in economics and finance. As a general interpretation, such models represent the decision of a microeconomic agent i (y_{it}), being influenced by macroeconomic factors F_t and by a set of individual specific characteristics, α_i and possibly x_{it}. Examples in the literature include, inter alia: demand for household food consumption (see e.g., Dynarski and Sheffrin, 1985, where households are assumed to have the same elasticity to food price, which is the common shock, and to permanent income, which is the idiosyncratic variable); firm size evolving according to a random walk, a case known in the literature as Gibrat's law (see Sutton, 1997; Geroski et al., 2002); other examples can also be found in micro demand for investment, consumption, labor demand. Moreover, the forward rate unbiasedness hypothesis postulates that the forward rate is an unbiased predictor of the corresponding future spot rate. This hypothesis has been extensively tested for exchange rates (Baillie and Bollerslev, 1989; Liu and Maynard, 2005; Westerlund, 2007). Another example in finance are models for default intensity for firm i at time t expressed as function of common factors (such as U.S. 3-month T-bill and the trailing 1-year returns) and idiosyncratic covariates such as distance to default and trailing 1-year stock return of the firm i (see Das, Duffie, Kapadia and Saita, 2007). Relevant is also the literature on output convergence where output for country i at time t depends on a set of common, to all n countries, technological shocks/knowledge and heterogenous degrees of access to the technological knowledge (Pesaran, 2007; Phillips and Sul, 2007). Considering (3), which represents a stationary panel regression with common shocks, the most natural application one may have in mind is to asset pricing models, such as the APT, where asset returns are explained by common factors (such as e.g., market return and powers thereof to represent coskewness and cokurtosis, macro factors, etc.); see Cochrane (2005) for a comprehensive review.

1.2 Main Results

Our asymptotic theory 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. Classical examples of observed common shocks are index models such as those used in international trade, labor economics, urban regional, public economics and finance literature. Most often, shocks are unknown, as in the cases of index extraction and indicators aggregation in economics, e.g., Quah and Sargent (1993), Forni and Reichlin (1998), and Bernanke and Boivin (2000). Third, regression model (1) may describe either a cointegration relationship or a spurious regression. 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.

An overview of the results derived in this paper is reported in Table 1.

[Insert Table 1 somewhere here]

As Table 1 shows, this paper provides a unified framework for the asymptotics of panels with common shocks. Particularly, results for the case of large n and large T with observable F_t are novel and can be thought of as extensions

Table 1: Consistency (C) and Limiting Distribution (LD) of $\hat{\beta}$: $y_{it} = \alpha_i + \beta' F_t + u_{it}$					
F_t known			F_t unknown		
(n, T)	C	LD	(n, T)	C	LD
Cointegration: $u_{it} \sim I(0)$					
Fixed n $T \to \infty$	Yes	Mixed Normal (Eq.11)		Yes	Non Standard (Eq. 37)
Fixed T $n \to \infty$	Yes	Mixed Normal (Eq.13)		Yes	Mixed Normal (Eq.13)
$(n, T) \to \infty$					
$n/T \to 0$	Yes	Mixed Normal (Eq.15)	$n/T \to 0$	Yes	Mixed Normal (Eq. 27)
$T/n \to 0$	Yes	Mixed Normal (Eq.16)	$T/n \to 0$	Yes	Non Standard (Eq. 32)
Spurious Regression: $u_{it} \sim I(1)$					
Fixed n $T \to \infty$	No	Non Standard (Eq. 12)		No	Non Standard (Eq. 38)
Fixed T $n \to \infty$	Yes	Non Standard (Eq. 14)		Yes	Non Standard (Eq. 14)
$(n, T) \to \infty$					
$n/T \to 0$	Yes	Non Standard (Eq. 17)	$n/T \to 0$	Yes	Non Standard (Eq. 29)
$T/n \to 0$	Yes	Non Standard (Eq. 18)	$T/n \to 0$	Yes	Non Standard (Eq. 30)
First Differences: $\hat{\beta}^{FD}$: $\Delta y_{it} = \beta' \Delta F_t + \Delta u_{it}$					
Fixed n $T \to \infty$	Yes	Normal (Eq. 19)		No	Degenerate (Eq. 39)
Fixed T $n \to \infty$	Yes	Mixed Normal (Eq. 20)		Yes	Mixed Normal (Eq. 20)
$(n, T) \to \infty$					
$n/T \to 0$	Yes	Normal (Eq. 21)	$n/T \to 0$	Yes	Normal (Eq. 31)
$T/n \to 0$	Yes	Degenerate (Eq. 22)	$T/n \to 0$	Yes	Degenerate (Eq. 32)

of the asymptotic theory for panels derived by Phillips and Moon (1999) and Kao (1999), who consider a model with cross-sectional independence. Assuming cross-sectional dependence in the panel changes the asymptotic theory, and a typical feature (discussed in greater details hereafter) is the asymptotic distribution of estimates being no longer normal as opposed to the independence case. An important result here is the extension of the joint limit theory to the strong dependence case, and the development of a method of proof for the asymptotics of double sums involving common shocks. Thus, although our results are specific to model (1), the method of proof we follow can be extended to study the asymptotics of estimators and tests for different models. For example, the method of proof developed here extends readily to inferential theory for cointegrated panels with common factors (Westerlund, 2007) or it can be used to show the asymptotics of t-tests for long run parameters in mixed panels (Fuertes, 2008; Ng, 2008); other applications to models where common factors are treated either as nuisance parameters or are genuine observable regressors are possible.

Results obtained for the case whereby the common shocks F_t are not observable are also new. The asymptotic theory for the estimates of the common shocks F_t is based on previous work by Bai (2003, 2004) and Bai and Ng (2002, 2004), and extended to the case of finite n. When common shocks are not observable, the estimated latent variables F_t are used as generated regressors to estimate β. This introduces a new error component in the regression equation. An important contribution of our paper is to study the impact of the estimation error when one needs to use an estimate of F_t in the regression model; see e.g., although in a nonparametric set-up, Connor, Hagmann and Linton (2007). Note that in Table 1, the "non standard" limiting distributions depend also on the assumptions made on the data generating process (DGP). Section 3 provides details on this.

Last, it is important to note that, as far as Theorems 1-4 are concerned, when both n and T are large the limits we derived are *joint* limits, which we obtain for $(n, T) \to \infty$ under, as a restriction on the rate of expansion of n and

T, $n/T \to 0$. Although more details on the method of proof are provided in Theorem 9 in Appendix B and in the remarks and proofs (reported in Appendix C) of the other theorems, the derivation of the joint limit is carried out by conditioning on the σ-field generated by the common shocks F_t. We show, in a similar spirit to Andrews (2005), that this entails that the quantities involved in the derivation of the asymptotics are martingale difference sequences (MDS), conditional on F_t. For each of the cases considered here, we then prove a joint Liapunov condition, under $(n, T) \to \infty$, which allows to apply the MDS central limit theorem (CLT) discussed in Hall and Heyde (1980) as $(n, T) \to \infty$. The restriction on the rate of expansion n/T is derived using similar arguments as in Phillips and Moon (1999), based on the Beveridge-Nelson (BN) decomposition of the series involved in the calculations.

The remainder of the paper is organized as follows. Section 2 introduces and comments on the main assumptions. In Section 3, we report the asymptotic theory of the ordinary least square (OLS) estimators of β in models (1) and (3). We analyze both the cases of known factors (Section 3.1) and unknown factors (Section 3.2), and we distinguish the cases of large n and T, finite T and large n and finite n and large T. Section 4 considers a discussion of the asymptotics for the estimator of when the data are generated by (4). Some Monte Carlo evidence is reported in Section 5. Section 6 concludes. Appendix A reports and discusses a joint MDS CLT. The main proofs are reported in Appendix B, contained in the present paper. Other proofs and preliminary Lemmas are in Appendix C.

Notation is fairly standard. Throughout the paper, $\|A\|$ denotes $\sqrt{tr\left(A'A\right)}$, $\to$ the ordinary limit, $\Rightarrow$ weak convergence, and $\xrightarrow{p}$ convergence in probability. Stochastic processes such as $B\left(r\right)$ on $[0, 1]$ are usually written as B, integrals such as $\int_0^1 B\left(r\right) dr$ as $\int B$ and stochastic integrals such as $\int_0^1 B\left(r\right) dB\left(r\right)$ as $\int BdB$. We let $M < \infty$ be a generic positive number, not depending on T or n.

2 Model and Assumptions

We assume that y_{it} is generated as follows

$$\begin{aligned} y_{it} &= \alpha_i + \beta' F_t + u_{it} \\ F_t &= F_{t-1} + \varepsilon_t \\ z_{it} &= \lambda_i' F_t + e_{it} \end{aligned} \tag{6}$$

$i = 1, ..., n$; $t = 1, ..., T$; β is a $k \times 1$ vector of slope parameters; $F_t = (F_{1t}, ..., F_{kt})'$ is a $k \times 1$ vector of common shocks; u_{it} may be $I(1)$ or $I(0)$ (spurious regression or cointegration relationship); z_{it} is a set of observed exogenous variables.

Define B_ε as the Brownian motion associated with the partial sums of ε_t with covariance matrix $\Omega_{\varepsilon\varepsilon}$ and $\bar{B}_\varepsilon(r)$ as the demeaned Brownian motion associated to the partial sums of F_t, i.e., $\bar{B}_\varepsilon(r) = B_\varepsilon(r) - \int_0^1 B_\varepsilon(r)\,dr$. 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, $\eta_{it} \overset{i.i.d.}{\sim} (0, \sigma_\eta^2)$ over t and i, with $E|\eta_{it}|^8 < M$, $\sum_{j=0}^\infty j|D_{ij}| < M$, $\sum_{j=0}^\infty j|F_{ij}| < M$ and $D_i^2(1)\sigma_\eta^2 > 0$, $F_i^2(1)\sigma_\eta^2 > 0$; the two MA processes $u_{it} = D_i(L)\eta_{it}$ and $\Delta u_{it} = F_i(L)\eta_{it}$ are assumed to be invertible; (b) (time-series and cross-sectional correlation) letting $E(u_{it}u_{js}) = \tau_{ij,ts} = \tau_{ij,|t-s|}$ and $E(\Delta u_{it}\Delta u_{js}) = \gamma_{ij,ts} = \gamma_{ij,|t-s|}$, as $n \to \infty$ a law of large numbers (LLN) and a CLT hold for the quantities $n^{-1/2}\sum_i u_{it}$ and $n^{-1/2}\sum_i \Delta u_{it}$.

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

Assumption 3: $E\|F_0\|^4 \leq M$ and $E|u_{i0}|^4 \leq M$.

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 $\lim_{n\to\infty} n^{-1}\sum_{i=1}^n \lambda_i\lambda_i' =$

Σ_Λ, if $n \to \infty$; in both cases, the matrix Σ_Λ is positive definite and such that the eigenvalues of the matrix $\Sigma_\Lambda^{1/2} \Sigma_{\Delta F} \Sigma_\Lambda^{1/2}$ are distinct, and the eigenvalues of the stochastic matrix $\Sigma_\Lambda^{1/2} \int B_\varepsilon B_\varepsilon' \Sigma_\Lambda^{1/2}$are distinct with probability 1.

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

Assumption 6: $\{\varepsilon_t\}$, $\{u_{it}\}$ and $\{e_{it}\}$ are three independent groups; F_0 is independent of $\{u_{it}\}$ and $\{e_{it}\}$.

Assumption 1(a) considers the possibility that equation (1) 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, e.g., 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", which differs from a strict common factor model where the u_{it} are assumed to be independent across i, e.g., Chamberlain and Rothschild (1983) and Onatski (2005).

The amount of cross-sectional dependence we allow for in Assumption 1(b) is anyway limited, since we require that it allows a LLN and a CLT to hold for the (rescaled) sequences $\sum_{i=1}^{n} u_{it}$ and $\sum_{i=1}^{n} \Delta u_{it}$.

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. Both the short run and the long run variance of ΔF_t are positive definite (Assumptions 2(b) and 2(d), respectively). Note that Assumption 2(d) rules out the possibility that the (common) regressors F_t in model (1) are cointegrated. This requirement is standard in cointegration to have non-degenerate limiting distributions.

Assumption 3 is a standard initial condition requirement. Assumption 4 serves to identify the factor loadings, 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_i 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 (2) to have a greater impact in explaining the contemporaneous correlation among the z_{it}. 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 (2) 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 analysis (PCA).

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. 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 $h_{ij} = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} \tau_{ij,ts}$ and $h_{ij}^{\Delta} = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} \gamma_{ij,ts}$. 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} and the Δu_{it} 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 the OLS, $\hat{\beta}$

$$\hat{\beta} = \left[\sum_{i=1}^{n} \sum_{t=1}^{T} \left(F_t - \bar{F}\right) \left(F_t - \bar{F}\right)' \right]^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(F_t - \bar{F}\right) y_{it} \tag{7}$$

in equation (1), where $\bar{F} = T^{-1} \sum_{t=1}^{T} F_t$, and $\hat{\beta}^{FD}$ when using equation (3),

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

We consider several features of (1) and (3):

1. the shocks F_t can either be known or (more likely) unobservable;

2. the relationship described by equation (1) 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;

3. the time series dimension T and the cross-sectional dimension n can be either fixed or large.

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 Observable F_t

When F_t is known we have:

$$\hat{\beta} - \beta = \left[\sum_{i=1}^{n} \sum_{t=1}^{T} W_t W_t' \right]^{-1} \left[\sum_{i=1}^{n} \sum_{t=1}^{T} W_t u_{it} \right] \tag{9}$$

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

$$\hat{\beta}^{FD} - \beta = \left[\sum_{i=1}^{n}\sum_{t=1}^{T} \Delta F_t \Delta F_t'\right]^{-1} \left[\sum_{i=1}^{n}\sum_{t=1}^{T} \Delta F_t \Delta u_{it}\right]. \tag{10}$$

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

Theorem 1 *Suppose Assumptions 1-6 hold, and define $Z \sim N(0, I_k)$, independent of the F_ts. For fixed n and $T \to \infty$*

$$T\left(\hat{\beta} - \beta\right) \Rightarrow \frac{1}{n}\left(\int \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1/2} \left(\sum_{i=1}^{n}\sum_{j=1}^{n} h_{ij}\right)^{1/2} \times Z, \tag{11}$$

if equation (1) is a cointegration relationship, and

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

if (1) is a spurious regression. For fixed T and $n \to \infty$, we have

$$\sqrt{n}\left(\hat{\beta} - \beta\right) \Rightarrow \left(\sum_{t=1}^{T} W_t W_t'\right)^{-1} \left(\sum_{t=1}^{T}\sum_{s=1}^{T} W_t W_s' \bar{\tau}_{ts}\right)^{1/2} \times Z, \tag{13}$$

if (1) is a cointegration regression, whilst if it is a spurious relationship we have

$$\sqrt{n}\left(\hat{\beta} - \beta\right) \Rightarrow \left(\sum_{t=1}^{T} W_t W_t'\right)^{-1} \left(\sum_{t=1}^{T} W_t \bar{u}_t\right), \tag{14}$$

where $\bar{u}_t = \lim_{n\to\infty} n^{-1/2} \sum_{i=1}^{n} u_{it}$.
Let equation (1) be a cointegration relationship; as $(n, T) \to \infty$ with $n/T \to 0$:

$$\sqrt{n}T\left(\hat{\beta} - \beta\right) \Rightarrow \left[\sqrt{\bar{h}}\left(\int \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1/2}\right] \times Z; \tag{15}$$

as $(n, T) \to \infty$ with $T/n \to 0$:

$$T^{3/2}\left(\hat{\beta} - \beta\right) \Rightarrow \left(\int \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1} \Delta_1, \tag{16}$$

where Δ_1 is defined in the proof - see equation (61).
Let equation (1) be a spurious regression; as $(n, T) \to \infty$ with $n/T \to 0$:

$$\sqrt{n}\left(\hat{\beta} - \beta\right) \Rightarrow \left(\int \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1} \left(\int \bar{B}_\varepsilon B_u\right) \sqrt{\bar{h}^{\Delta}}; \tag{17}$$

as $(n,T)\to\infty$ *with* $T/n\to 0$*:*

$$\sqrt{T}\left(\hat{\beta}-\beta\right)\Rightarrow\left(\int\bar{B}_{\varepsilon}\bar{B}_{\varepsilon}^{\prime}\right)^{-1}\Delta_{2}, \tag{18}$$

where Δ_2 *is defined in the proof - see equation (64).*

Proof. See Appendix B. ■

Equation (12) is a standard results in the literature. With respect to the speed of convergence, when $(n,T)\to\infty$ our results lead to the same rates of convergence 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 $\sqrt{n}$. Note that, irrespective of model (1) being a cointegration regression or a spurious regression, large n allows for consistency to hold.

For the case of $(n,T)\to\infty$ with $n/T\to 0$, the rate of convergence for $\hat{\beta}$ is the same as in Phillips and Moon (1999) under the case of contemporaneous independence across cross-sectional units. The limiting distributions in (15) and (17) are mixed normal rather than normal, as in Phillips and Moon (1999). The mixed normality is due to both F_t being nonstationary and common across cross-sectional units, as can be seen by considering equation (13) for $T\to\infty$. The design matrix $\left(nT^2\right)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}F_tF_t^{\prime}=T^{-2}\sum_{t=1}^{T}F_tF_t^{\prime}$ converge in distribution to a random matrix, namely $\int\bar{B}_{\varepsilon}\bar{B}_{\varepsilon}^{\prime}$, rather than to a constant. Of course, $\left(nT^2\right)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}F_tF_t^{\prime}$ would converge to a constant (in probability) if F_t were not common shocks, i.e., if F_t were replaced by F_{it}. Theorem 1 also explores the case of a "short" panel, where $T/n\to 0$. In this case, $\hat{\beta}$ is still consistent, irrespective of whether (1) is a cointegration or a spurious regression, although consistency is achieved at a "slower" rate than in the case whereby $n/T\to 0$. The limiting distributions, given in (16) and (18), are non standard, and they depend upon Δ_1 and Δ_2. As shown in Appendix C, these terms come from the bias arising from the BN decomposition of F_t and u_{it}, and thus they depend upon the assumptions on the DGP of F_t and u_{it}. A similar argument is discussed in Phillips and Moon (1999); of course, if F_t and u_{it} had initial values set equal to zero, and if they were i.i.d. processes, then (15) and (17) would hold for $(n, T) \to \infty$ for all combinations of n and T.

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 define $Z \sim N(0, I_k)$, independent of the $\Delta F_t s$.*
For fixed n and $T \to \infty$

$$\sqrt{T}\left(\hat{\beta}^{FD} - \beta\right) \Rightarrow n^{-1}\Sigma_{\Delta F}^{-1/2}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} h_{ij}^{\Delta}\right)^{1/2} \times Z. \tag{19}$$

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

$$\sqrt{n}\left(\hat{\beta}^{FD} - \beta\right) \Rightarrow \left(\sum_{t=1}^{T} \Delta F_t \Delta F_t'\right)^{-1}\left(\sum_{t=1}^{T}\sum_{s=1}^{T} \Delta F_t \Delta F_s' \bar{\gamma}_{ts}\right)^{1/2} \times Z. \tag{20}$$

When $(n, T) \to \infty$, under $n/T \to 0$ it holds that

$$\sqrt{nT}\left(\hat{\beta}^{FD} - \beta\right) \Rightarrow \left(\Sigma_{\Delta F}^{-1/2}\sqrt{\bar{h}^{\Delta}}\right) \times Z; \tag{21}$$

as $(n, T) \to \infty$, under $T/n \to 0$, it holds that

$$T\left(\hat{\beta}^{FD} - \beta\right) \xrightarrow{p} \Sigma_{\Delta F}^{-1}\Delta_3, \tag{22}$$

where Δ_3 is defined in the proof - see equation (77).

Proof. See Appendix C. ■

Since the first differenced model is always stationary, irrespective of whether equation (1) is a cointegration equation or a spurious regression, one can always apply the CLT to obtain the limiting distribution of $\hat{\beta}^{FD} - \beta$.

Equation (21) states that the limiting distribution of $\hat{\beta}^{FD} - \beta$ is normal instead of mixed normal, despite the strong dependence across cross-sectional units. This can be seen from equation (20), which gives the limiting distribution for T fixed and $n \to \infty$. The matrix $T^{-1}\sum_{t=1}^{T} \Delta F_t \Delta F_t'$ is random for finite T, but it converges to a constants as $T \to \infty$ due to a LLN.

Equation (22) refers to a panel where $T/n \to 0$, and thus the number of cross-sectional units is "much larger" than the number of time-series observations. This case is similar to that found in Theorem 1: when $T/n \to 0$, the bias in the BN decomposition dominates, thereby making the limiting distribution non standard and depending upon the assumptions on the DGP of F_t and u_{it}. Of course, if one knows $\Sigma_{\Delta F}^{-1}\Delta_3$ (or could estimate it consistently at a rate ϕ_{nT}) then the remainder in the BN decomposition of $\hat{\beta}^{FD} - \left(\beta + \Sigma_{\Delta F}^{-1}\Delta_3\right)$ has mean zero (or of order $O_p\left(\phi_{nT}\right)$). Under additional assumptions, the bias in the BN decomposition for $\Sigma_{\Delta F}^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta F_t \Delta u_{it}$ (see also equation (75) in Appendix C) is of order $O_p\left(\sqrt{n}\right)$ (resp., $O_p\left(n\phi_{nT}\right)$). Therefore, when normalized by $\sqrt{nT}$, the bias is always dominated by the martingale approximation to $\Sigma_{\Delta F}^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta F_t \Delta u_{it}$ - resp., of order $O_p\left(\sqrt{\frac{n}{T}}\phi_{nT}\right)$. The former case implies that (21) holds as $(n, T) \to \infty$, with no restriction on the rate of expansion of n and T as they pass to infinity. This is consistent with Phillips and Moon (1999, p. 1074, Remark (a)), and a similar argument could be in principle applied to Theorem 1.

3.2 Unobservable 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 shocks F_t.

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

$$\hat{\beta} = \left[\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\hat{W}_t'\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t y_{it}\right] \tag{23}$$

and

$$\hat{\beta}^{FD} = \left[\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta\hat{F}_t'\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta y_{it}\right] \tag{24}$$

with estimation errors:

$$\hat{\beta} - \beta = \left[\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\hat{W}_t'\right]^{-1}\left\{\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\left[\left(W_t - \hat{W}_t\right)'\beta + u_{it}\right]\right\} \tag{25}$$

and

$$\hat{\beta}^{FD} - \beta = \left[\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta\hat{F}_t'\right]^{-1}\left\{\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\left[\left(\Delta F_t - \Delta\hat{F}_t\right)'\beta + \Delta u_{it}\right]\right\}. \tag{26}$$

We assume that the number of common shocks k is known. This does not lead to any loss of generality since the distribution of the estimated shocks does 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).

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 shocks ΔF_t from (2). Consider the $T \times n$ matrix $Z = (z_1, ..., z_T)'$, and the $T \times k$ matrix of shocks $F = (F_1, F_2, ..., F_T)'$. Then each objective function $V_a(k)$ or $V_b(k)$ can be minimized by concentrating out λ and obtaining estimates $\Delta\hat{F}$ and $\hat{F}$ using the normalizations $\Delta\hat{F}'\Delta\hat{F}/T = I_k$ or $\hat{F}'\hat{F}/T^2 = I_k$. The estimated shock matrices $\Delta\hat{F}$ and $\hat{F}$ are $\sqrt{T}$ and T times respectively the eigenvectors corresponding to the k largest eigenvalues of the $T \times T$ matrices $\Delta Z\Delta Z'$ or ZZ'.

In the context of unobservable common factors, the problem of identification arises. It is well known (see e.g., Bai, 2003, and Bai, 2004) that the solutions to the above minimization problems are not unique, e.g., when estimating shocks ΔF_t and F_t, these are not directly identifiable even though they are up to a transformation. This entails that whilst it is possible to "consistently" estimate the space spanned by the common factors ΔF_t and F_t, it is not possible to estimate ΔF_t and F_t themselves. Thus, $\hat{F}_t$ and $\Delta \hat{F}_t$ are, respectively, "consistent" estimators for $H_1' F_t$ and $H_1' \Delta F_t$, where H_1 is a non singular $k \times k$ matrix, meaning that $\hat{F}_t - H_1' F_t$ and $\Delta \hat{F}_t - H_1' \Delta F_t$ converge to zero in some sense. This makes it impossible to estimate β consistently, although once again the space spanned by β can be estimated consistently. Due to the rotational indeterminacy of the estimation of F_t and ΔF_t, the estimators $\hat{\beta}$ and $\hat{\beta}^{FD}$ may "consistently" estimate $H_1^{-1}\beta$, so that $\hat{\beta} - H_1^{-1}\beta$ may converge to zero in some sense as n and T (or either) pass to infinity. Whilst there is no direct consistency result for β, being able to estimate the space it spans is sufficient for many purposes. For example, the quantity βF_t can be consistently estimated by $\hat{\beta}\hat{F}_t$, and therefore predicting using (1) is feasible. Also, it is possible to carry out inference on e.g., the significance of the F_ts as regressors in (1) - see also a similar discussion in Bai (2003, p. 145). Likewise, it is impossible to consistently estimate the loadings λ_i in (2), although it is possible to consistently estimate the space spanned, i.e. $\hat{\lambda}_i - H_2 \lambda_i$, where the rotation matrix H_2 is $n \times n$, can be shown to converge to zero in some sense.

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

Theorem 3 *Suppose Assumptions 1-6 hold.*
Let equation (1) be a cointegration relationship; as $(n, T) \to \infty$ *with* $n/T \to 0$*:*

$$\sqrt{n}T\left(\hat{\beta} - H_1^{-1}\beta\right) \Rightarrow \left(H_1' \int \bar{B}_\varepsilon \bar{B}_\varepsilon' H_1\right)^{-1/2} \left[\bar{h} + \beta' H_1'^{-1} \tilde{Q}_B \Gamma \tilde{Q}_B' H_1^{-1}\beta\right]^{1/2} \times Z \tag{27}$$

where $Z \sim N_1(0, I_k)$ *independent of the* σ*-field generated by the common shocks* F_t *and of the random matrix* $\tilde{Q}_B$*. Also*

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

and

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

When $(n,T)\to\infty$ *with* $T/n\to 0$*, it holds that*

$$T^{3/2}\left(\hat{\beta}-H_1^{-1}\beta\right)\Rightarrow\left(H_1'\int\bar{B}_\varepsilon\bar{B}_\varepsilon'H_1\right)^{-1}\Delta_1 \tag{28}$$

where Δ_1 *is defined in (61).*

Let equation (1) be a spurious regression; as $(n,T)\to\infty$ *with* $n/T\to 0$*:*

$$\sqrt{n}\left(\hat{\beta}-H_1^{-1}\beta\right)\Rightarrow\left(H_1'\int\bar{B}_\varepsilon\bar{B}_\varepsilon'H_1\right)^{-1}\left(H_1'\int\bar{B}_\varepsilon B_u\right)\sqrt{\bar{h}^{\Delta}}. \tag{29}$$

As $(n,T)\to\infty$ *with* $T/n\to 0$*, it holds that*

$$\sqrt{T}\left(\hat{\beta}-H_1^{-1}\beta\right)\Rightarrow\left(H_1'\int\bar{B}_\varepsilon\bar{B}_\varepsilon'H_1\right)^{-1}\Delta_2 \tag{30}$$

where Δ_2 *is defined in (64).*

Proof. See Appendix B. ■

The estimator $\hat{\beta}$ is always consistent, even though $T/n\to 0$ results in a slower rate of convergence and in a degenerate behavior of the numerator of $\hat{\beta}-H_1^{-1}\beta$. When $n/T\to\infty$, results for the case of (1) being a cointegrating regression are essentially the same as in equation (15) in Theorem 1. The only difference is that now the variance of $\hat{\beta}-H_1^{-1}\beta$ is "inflated" by the non-negative random variable $\beta'H_1'^{-1}\tilde{Q}_B\Gamma\tilde{Q}_B'H_1^{-1}\beta$, which arises from the estimation error when replacing F_t with $\hat{F}_t$.

Notice the consequence of equation (1) being a spurious regression: as long as the number of cross sectional units n is "smaller" than T, the classical $\sqrt{n}$ consistency holds, and we have the same limiting distribution as in equation (17).

In both cases, the limiting distributions become non standard when $T/n\to 0$.

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

$$n\left(\hat{\beta}^{FD}-H_1^{-1}\beta\right)\xrightarrow{p}\left(H_1'\Sigma_{\Delta F}H_1\right)^{-1}\times Z, \tag{31}$$

where Z is $N(0,Q)$ *with*

$$\begin{aligned} Q &= \lim_{n,T\to\infty} \frac{1}{T^4}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{u=1}^{T}\sum_{v=1}^{T} H_1'\Delta F_t\left(\Delta\hat{F}_s - H_1'\Delta F_s\right)' V^{-1}\beta \times \\ & \beta' V^{-1}\left(\Delta\hat{F}_u - H_1'\Delta F_u\right)' \Delta F_v H_1 \times \\ & \sum_{i=1}^{n}\sum_{j=1}^{n} E\left\{\left[e_{it}e_{is} - E\left(e_{it}e_{is}\right)\right]\left[e_{ju}e_{jv} - E\left(e_{ju}e_{jv}\right)\right]\right\}, \end{aligned}$$

and 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.*

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

$$T\left(\hat{\beta}^{FD} - H_1^{-1}\beta\right) \xrightarrow{p} \bar{h}_e V^{-1}\beta + \left(H_1'\Sigma_{\Delta F}H_1\right)^{-1}\Delta_3 \quad (32)$$

where $\bar{h}_e$ *is the long-run variance of* $\lim_{n\to\infty} n^{-1/2}\sum_{i=1}^{n} e_{it}$ *and* Δ_3 *is defined in (77).*

Proof. See Appendix C. ■

Notice that in this case we have a degenerate limiting distribution when $\frac{T}{n}\to 0$, despite having a consistent estimate. The distribution limit depends on the bias arising in the BN decomposition, Δ_3, but also on the presence of the error term $\Delta\hat{F}_t - H_1'\Delta F_t$.

3.2.2 The case of T fixed and n large

When T is fixed and n is large, consistent estimation of shocks is still possible, see 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 DGP 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 shocks estimation is $\sqrt{n}$ consistent, i.e., we have both

$$\hat{F}_t - H_1'F_t = O_p\left(n^{-1/2}\right)$$

and

$$\Delta\hat{F}_t - H_1'\Delta F_t = O_p\left(n^{-1/2}\right)$$

for all t.

Theorem 5 *Suppose Assumptions 1-7 hold; then for every consistent estimator $\hat{F}_t$ of $H_1'F_t$ and for fixed T and $n \to \infty$ we have the same result as in equation (14).*

Proof. See Appendix C. ■

Theorem 6 *Suppose Assumptions 1-7 hold; then for every consistent estimator $\Delta\hat{F}_t$ of $H_1'\Delta F_t$ and for fixed T and $n \to \infty$ we have the same result as in equation (20).*

Proof. See Appendix C. ■

Theorems (5) and (6) do not anyway require $\sqrt{n}$ consistency, since they ensure the consistency of $\hat{\beta}$ and $\hat{\beta}^{FD}$ for any consistent estimate of the shocks, irrespective of the rate of convergence. 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 shocks as long as shocks are estimated consistently.

3.2.3 The case of n fixed and T large

In what follows, we develop the inferential theory for the case when shocks are unknown and the cross-sectional dimension n is finite. Rewriting model (2) in the vector form, one gets:

$$z_t = \Lambda F_t + e_t \tag{33}$$

where $z_t = (z_{1t}, ..., z_{nt})'$, $e_t = [e_{1t}, ..., e_{nt}]'$, and $\Lambda = (\lambda_1, \lambda_2, ..., \lambda_n)'$. One can estimate Λ using PCA. 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 $n < T$. Henceforth, for sake of the notation and without loss of generality, we modify Assumption 4 by assuming 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*

$$\begin{aligned} T\left(\hat{\Lambda} - H_2\Lambda\right) \;\Rightarrow\; & \left[I_n - n^{-1}H_2\Lambda\left(H_1'\int \bar{B}_\varepsilon\bar{B}_\varepsilon' H_1\right)\Lambda' H_2'\right]\left(\int dW_e\bar{B}_\varepsilon' H_1\right)\left(H_1'\int \bar{B}_\varepsilon\bar{B}_\varepsilon' H_1\right)^{-1} \\ & -n^{-1}\Lambda H_2\left(\int dW_e\bar{B}_\varepsilon' H_1\right)\Lambda H_2' \\ & +n^{-1}\left[I_n - 2n^{-1}\Lambda H_2\left(H_1'\int \bar{B}_\varepsilon\bar{B}_\varepsilon' H_1'\right)\Lambda' H_2'\right]\Omega_e\Lambda H_2 \qquad (34) \end{aligned}$$

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

Proof. See Appendix B. ■

Note that in this case we have a T-consistent estimate of Λ, even though the PCA of F_t is not consistent, e.g., see Bai (2004) and Proposition 2 below, when n is finite.

Define the limiting distribution of $T\left(\hat{\Lambda} - H_2\Lambda\right)$ as D_Λ^1, i.e., $T\left(\hat{\Lambda} - H_2\Lambda\right) \Rightarrow D_\Lambda^1$. Given the restriction $\hat{\Lambda}'\hat{\Lambda}/n = I_k$, the OLS estimator of F_t, obtained regressing the z_t 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*

$$\max_{1\le t\le T}\left\|\hat{F}_t - H_1'F_t\right\| = O_p(1) \qquad (35)$$

and

$$\max_{1\le t\le T}\left\|\Delta\hat{F}_t - H_1'\Delta F_t\right\| = O_p(1) \qquad (36)$$

uniformly in t.

Proof. See Appendix C. ∎

Proposition 2 states that the estimates of the shocks and of their first difference are inconsistent, in that the estimation error does not die out as $T \to \infty$. However this inconsistency has no impact on the consistency of $\hat{\beta}$ and $\hat{\beta}^{FD}$, though it affects their asymptotic distributions. 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:*

$$T\left(\hat{\beta} - H_1^{-1}\beta\right) \Rightarrow \left[H_1' \int \bar{B}_\varepsilon \bar{B}_\varepsilon' H_1\right]^{-1} \left\{ \begin{array}{c} \int H_1' \bar{B}_\varepsilon dB_u \left(\sum_{i=1}^n \sum_{j=1}^n h_{ij}\right)^{1/2} \\ -n^{-1}\left[H_1' \int \bar{B}_\varepsilon d\bar{B}_e' H_2 \Lambda H_1^{-1}\beta + n^{-1}\Lambda' H_2' \Sigma_e H_2 \Lambda H_1^{-1}\beta\right] \\ -n^{-1}\left(H_1' \int \bar{B}_\varepsilon \bar{B}_\varepsilon' H_1\right)\left(D_\Lambda^{1\prime} H_2 \Lambda - \Lambda' H_2' D_\Lambda^1\right) H_1^{-1}\beta \end{array} \right\} \tag{37}$$

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

$$\hat{\beta} - H_1^{-1}\beta \Rightarrow \left(H_1' \int \bar{B}_\varepsilon \bar{B}_\varepsilon' H_1\right)^{-1} \left(H_1 \int \bar{B}_\varepsilon B_u\right) \left(\sum_{i=1}^n \sum_{j=1}^n h_{ij}^\Delta\right)^{1/2}. \tag{38}$$

Proof. See Appendix C. ∎

Note that even though common shocks cannot be estimated consistently, $\hat{\beta}$ is consistent when (1) is a cointegration relationship but inconsistent when (1) instead is a spurious regression. With respect to the case of observable shocks, shock estimation has an impact on the limit distribution of $\hat{\beta} - H_1^{-1}\beta$ when equation (1) is a cointegration regression - see equation (37) above. On the other hand, it does not affect the asymptotic distribution when equation (1) is a spurious regression - see equation (38).

Equations (37) and (38) show an important common feature of this theoretical framework. Only the numerators of equation (37) and (38) depend on whether equation (1) 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 of $\hat{\beta}^{FD}$ are in the following theorem.

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

$$\hat{\beta}^{FD} - H_1^{-1}\beta \xrightarrow{p} -H_1^{-1}\beta + n\left[\Lambda' \Sigma_{\Delta z} \Lambda\right]^{-1} \left[H_1' \Sigma_{\Delta F} H_1\right] H_1^{-1}\beta \tag{39}$$

where $\Sigma_{\Delta e} = Var\left(\Delta e_t\right)$ *and* $\Sigma_{\Delta z} = \Lambda\left(H_1' \Sigma_{\Delta F} H_1\right)\Lambda' + \Sigma_{\Delta e}$.

Proof. See Appendix C. ■

The estimator $\hat{\beta}^{FD}$ is inconsistent. This is due to the two terms $\sum \Delta F_t \Delta F_t'$ and $\sum \Delta e_t \Delta e_t'$ in the denominator having the same asymptotic magnitude, rather than to the common shock estimates being inconsistent. Also, this holds for any consistent estimator $\hat{\Lambda}$ (see proof in Appendix C).

4 Extensions

In this section, we consider two extensions of our basic framework:

(i) the case of model (2), where our basic framework (1) also contains some idiosyncratic shocks;

(ii) the case where model (1) is misspecified, and the common shocks F_t actually have unit specific slopes, say β_i.

4.1 The case of idiosyncratic shocks

Model (1) assumes that the DGP of y_{it} depends only on a set of common shocks. In this section, we briefly consider the case where the model is augmented to take into account the presence of unit-specific variables. Even though the algebra becomes more tedious, all the results derived in the previous section still hold. The only novelty is the covariance between the common shocks F_t and the unit specific regressors.

Recall the augmented model (4)

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

with $i = 1, ..., n$, $t = 1, ..., T$, where β and γ are $(k \times 1)$ and $(p \times 1)$ vectors of slope parameters. The idiosyncratic variables x_{it} are a $(p \times 1)$ vector of observable $I(1)$ individual-specific regressors, defined as

$$x_{it} = \tau_i G_t + \omega_{it} \tag{40}$$

where G_t is an $R \times 1$ vector of $I(1)$ variables that may contain some elements of F_t, and τ_i is a $p \times R$ matrix and

$$\omega_{it} = \omega_{it-1} + \epsilon_{it} \tag{41}$$

with ω_{it} assumed for simplicity i.i.d. across i and such that $\{\epsilon_{it}\}$, $\{\varepsilon_t\}$, $\{F_0\}$ and $\{\omega_{i0}\}$ are independent groups. Equation (40) considers the possible presence of "correlation" between F_t and x_{it}. A similar framework that allows for cross dependence among the idiosyncratic regressors and dependence between the idiosyncratic regressors and the common regressors is considered in Pesaran (2006) and Kapetanios, Pesaran and Yamagata (2006). Thus, cross dependence is accounted for directly, via F_t, and indirectly, via the factor structure in x_{it}.

The impact of the presence of the x_{it}s on the LS estimator of β will be discussed considering (for the purpose of a concise discussion) the case of observable F_t. Let $\bar{x}_{it} = x_{it} - T^{-1} \sum_{t=1}^{T} x_{it}$; then,

$$\begin{bmatrix} \hat{\beta} - \beta \\ \hat{\gamma} - \gamma \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} \sum_{t=1}^{T} W_t W_t' & \sum_{i=1}^{n} \sum_{t=1}^{T} W_t \bar{x}_{it}' \\ \sum_{i=1}^{n} \sum_{t=1}^{T} \bar{x}_{it} W_t' & \sum_{i=1}^{n} \sum_{t=1}^{T} \bar{x}_{it} \bar{x}_{it}' \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{n} \sum_{t=1}^{T} W_t u_{it} \\ \sum_{i=1}^{n} \sum_{t=1}^{T} \bar{x}_{it} u_{it} \end{bmatrix} \tag{42}$$

and

$$\begin{bmatrix} \hat{\beta}^{FD} - \beta \\ \hat{\gamma}^{FD} - \gamma \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta F_t' & \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta x_{it}' \\ \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta x_{it} \Delta F_t' & \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta x_{it} \Delta x_{it}' \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta u_{it} \\ \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta x_{it} \Delta u_{it} \end{bmatrix}. \tag{43}$$

Thus, the asymptotics of $\hat{\beta} - \beta$ (and of $\hat{\beta}^{FD} - \beta$) depends on terms containing the x_{it}s as well, due to the presence of the off-diagonal term $\sum_{i=1}^{n} \sum_{t=1}^{T} \bar{x}_{it} W_t'$ in the denominator of (42) and of $\sum_{i=1}^{n} \sum_{t=1}^{T} \Delta x_{it} \Delta F_t'$ in (43).

Depending on whether or not these quantities are asymptotically zero (i.e. F_t and x_{it} or their first differences are asymptotically orthogonal), the asymptotic distribution of $\hat{\beta}$ and $\hat{\beta}^{FD}$ may or may not change with respect to the results reported in Section 3. In order to investigate the cases whereby $\hat{\beta} - \beta$ (or $\hat{\beta}^{FD} - \beta$) is orthogonal to $\hat{\gamma} - \gamma$ ($\hat{\gamma}^{FD} - \gamma$ respectively), consider the following preliminary assumption, where $\bar{G}_t = G_t - T^{-1} \sum_{t=1}^{T} G_t$ and similarly $\bar{\omega}_{it}$.

Assumption 8: (1) as $n \to \infty$, $n^{-1} \sum_{i=1}^{n} \tau_i = \bar{\tau}$; (2) as $(n, T) \to \infty$, *(i)* $\sum_{i=1}^{n} \sum_{t=1}^{T} \bar{x}_{it} \bar{x}_{it}' = O_p\left(nT^2\right)$, *(ii)* $\sum_{i=1}^{n} \sum_{t=1}^{T} W_t \bar{\omega}_{it}' = O_p\left(\sqrt{n}T^2\right)$ and *(iii)* as $T \to \infty$, $T^{-2} \sum_{t=1}^{T} \bar{G}_t W_t' \Rightarrow \int \bar{B}_G \bar{B}_{\varepsilon}'$; (3) as $(n, T) \to \infty$, *(i)* $\sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \bar{x}_{it} \Delta \bar{x}_{it}' = O_p(nT)$, *(ii)* $\sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \omega_{it} \Delta F_t' = O_p\left(\sqrt{nT}\right)$ and *(iii)* as $T \to \infty$, $\sum_{t=1}^{T} \Delta F_t \Delta G_t' = O_p\left(T^{\varkappa}\right)$ with $\varkappa = 1/2$ or 1.

Assumption 8 requires some asymptotic results to hold with respect to the newly introduced variables x_{it}, G_t and ω_{it}, and it could be expressed using some more primitive assumptions. For example, parts 2*(i)* and 3*(i)* could be shown using the same arguments as in Phillips and Moon (1999); likewise, 2*(ii)* and 3*(ii)* could be proved, under suitable assumptions, using similar derivations as for the proofs of Theorems 1 and 2. Note that a necessary condition in order for 3*(ii)* to hold is that $\Delta \omega_{it}$ and ΔF_t be uncorrelated. The result in 2*(iii)* could be proved using a FCLT argument. Note that 3*(iii)* accommodates both situations whereby F_t and G_t are independent of each other or overlap. The former case holds under $\varkappa = 1/2$, which implies that a CLT is required to hold for the sequence $\sum_{t=1}^{T} \Delta F_t \Delta G_t'$; the latter case is entailed by $\varkappa = 1$, which requires a LLN to hold for $\sum_{t=1}^{T} \Delta F_t \Delta G_t'$.

The following theorems provide a summary of the values of $\sum_{i=1}^{n} \sum_{t=1}^{T} \bar{x}_{it} W_t'$ and $\sum_{i=1}^{n} \sum_{t=1}^{T} \Delta x_{it} \Delta F_t'$ under various combinations of n and T.

Proposition 3 *Let Assumptions 8(1) and 8(2) hold, and define* $\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{x}_{it}W_t' = \sum_{i=1}^{n}\sum_{t=1}^{T}\bar{\omega}_{it}W_t' + \sum_{i=1}^{n}\tau_i\sum_{t=1}^{T}\bar{G}_tW_t' = a+b$.
As $(n,T)\to\infty$, $a = O_p\left(\sqrt{n}T^2\right)$ *and* $b = O_p\left(nT^2\right)$ *with*

$$\frac{1}{nT^2}\sum_{i=1}^{n}\tau_i\sum_{t=1}^{T}\bar{G}_tW_t' \Rightarrow \bar{\tau}\int \bar{B}_G\bar{B}_\varepsilon'.$$

For fixed T *and as* $n\to\infty$, $a = O_p(\sqrt{n})$ *and* $b = O_p(n)$ *with* $n^{-1}b \Rightarrow \bar{\tau}\sum_{t=1}^{T}\bar{G}_tW_t'$. *For fixed* n *and as* $T\to\infty$, $a = b = O_p\left(T^2\right)$.

Proof. See Appendix C. ■

Proposition 4 *Let Assumptions 8(1) and 8(3) hold, and define* $\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta F_t' = \sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\omega_{it}\Delta F_t' + \sum_{i=1}^{n}\tau_i\sum_{t=1}^{T}\Delta G_t\Delta F_t' = a+b$. *As* $(n,T)\to\infty$, $a = O_p\left(\sqrt{nT}\right)$ *and* $b = O_p\left(nT^{\varkappa}\right)$. *When* $\varkappa = 1$, *it holds that*

$$\frac{1}{nT}\sum_{i=1}^{n}\tau_i\sum_{t=1}^{T}\Delta G_t\Delta F_t' \to \bar{\tau}E\left(\Delta G_t\Delta F_t'\right).$$

For fixed T *and as* $n\to\infty$, $a = O_p(\sqrt{n})$ *and* $b = O_p(n)$ *with* $n^{-1}b \Rightarrow \bar{\tau}\sum_{t=1}^{T}\Delta G_t\Delta F_t'$. *For fixed* n *and as* $T\to\infty$, $a = O_p\left(\sqrt{T}\right)$ *and* $b = O_p\left(T^{\varkappa}\right)$.

Proof. See Appendix C. ■

Propositions 3 and 4 illustrate the cases when $\hat{\beta}-\beta$ ($\hat{\beta}^{FD}-\beta$) is orthogonal to $\hat{\gamma}-\gamma$ ($\hat{\gamma}^{FD}-\gamma$), thereby making the presence of the idiosyncratic shocks x_{it} irrelevant for the asymptotics of $\hat{\beta}-\beta$ (or $\hat{\beta}^{FD}-\beta$). Orthogonality between $\hat{\beta}-\beta$ and $\hat{\gamma}-\gamma$ requires two necessary conditions to hold: $n\to\infty$ and $\bar{\tau}=0$. Note that when n is fixed, $\hat{\beta}-\beta$ and $\hat{\gamma}-\gamma$ cannot be orthogonal, irrespective of the presence of the common factors G_t in the DGP of x_{it}. As far as $\hat{\beta}^{FD}-\beta$ and $\hat{\gamma}^{FD}-\gamma$ are concerned, $n\to\infty$ and $\bar{\tau}=0$ are only sufficient conditions but they are not necessary: if the x_{it}s are purely idiosyncratic variables, i.e. with no common factor structure ($\tau_i = 0$), or if the x_{it}s do have a common factor that is unrelated to F_t (i.e. $\varkappa = 1/2$), then this suffices to have asymptotic orthogonality between $\hat{\beta}^{FD}-\beta$ and $\hat{\gamma}^{FD}-\gamma$. Similar results could be proved for the case of unobservable F_t.

4.2 The case of heterogeneous slopes

In this section, we consider the case of heterogeneous slopes, i.e., the case whereby the coefficient of common shocks F_t is unit specific. This entails that (1) now becomes

$$y_{it} = \alpha_i + \beta_i' F_t + u_{it}. \tag{44}$$

Although β_i is different across units, the researcher could however be interested in estimating the average β; a typical case of this is in the literature on convergence (see e.g., Temple, 1999, p. 126). Thus, the model that will be used for estimation is

$$y_{it} = \alpha_i + \beta' F_t + v_{it}$$

where $v_{it} = u_{it} + (\beta_i - \beta)' F_t$; when unobservable common shocks are considered, the model becomes

$$y_{it} = \alpha_i + \beta' (H_1')^{-1} \hat{F}_t + v_{it} \tag{45}$$

where now the error term $v_{it} = u_{it} + (\beta_i - \beta)' F_t - \beta' \left(\hat{F}_t - H_1' F_t\right)$; it is important to note that, as it is well known from Phillips and Moon (1999), neglected heterogeneity introduces a further, nonstationary components in the error term, given by $(\beta_i - \beta)' F_t$. Thus, (45) is always a spurious regression. When first differenced data are used, (45) becomes

$$\Delta y_{it} = \beta' (H_1')^{-1} \Delta \hat{F}_t + \Delta v_{it} \tag{46}$$

where $\Delta v_{it} = \Delta u_{it} + (\beta_i - \beta)' \Delta F_t - \beta' (H_1')^{-1} \left(\Delta \hat{F}_t - H_1' \Delta F_t\right)$.

For the sake of brevity, we will focus our attention to the case whereby F_t is unobservable, thus analyzing the estimates of β from (45) and (46). The estimation errors are, respectively

$$\hat{\beta} - H_1^{-1}\beta = \left[\sum_{i=1}^{n}\sum_{t=1}^{T} \hat{W}_t \hat{W}_t'\right]^{-1} \left\{\sum_{i=1}^{n}\sum_{t=1}^{T} \hat{W}_t \left[u_{it} - \left(\hat{W}_t - H_1' W_t\right)' H_1^{-1}\beta + W_t' (\beta_i - \beta)\right]\right\} \tag{47}$$

$$\begin{aligned} \hat{\beta}^{FD} - H_1^{-1}\beta &= \left[\sum_{i=1}^{n}\sum_{t=1}^{T} \Delta\hat{F}_t \Delta\hat{F}_t'\right]^{-1} \\ &\times \left\{\sum_{i=1}^{n}\sum_{t=1}^{T} \Delta\hat{F}_t \left[\Delta u_{it} - \left(\Delta\hat{F}_t - H_1'\Delta F_t\right)' H_1^{-1}\beta + \Delta F_t' (\beta_i - \beta)\right]\right\}. \end{aligned} \tag{48}$$

Consider also the following assumption on the β_is.

Assumption 9: It holds that *(i)* $\beta_i \overset{iid}{\sim} (\beta, \Sigma_\beta)$ with $E\|\beta_i - \beta\|^{4+\delta} < \infty$ and *(ii)* β_i is independent of all other quantities.

Assumption 9 yields the usual asymptotic results for $\{\beta_i\}_{i=1}^n$, such as a CLT and a LLN. Part *(ii)* entails that the long run average parameter β (see Phillips and Moon, 1999 and 2000) is genuinely $E(\beta_i)$.

To illustrate the main point (summarized in a theorem hereafter), consider (47) - similar arguments hold for (48) and can thus be readily extended. Looking at the numerator, we know from Lemma 2 in Appendix A that, as $(n,T) \to \infty$, $\sum_{i=1}^n \sum_{t=1}^T \hat{W}_t \left(\hat{W}_t - H_1' W_t\right)' H_1^{-1}\beta = O_p\left(nTC_{nT}^{-1}\right)$. Also, as far as $\sum_{i=1}^n \sum_{t=1}^T \hat{W}_t u_{it}$ is concerned, its magnitude depends on whether (44) is a cointegrated or a spurious regression, being of order $O_p(\sqrt{n}T) + O_p\left(n\sqrt{T}\right)$ in the former case and of order $O_p\left(\sqrt{n}T^2\right) + O_p\left(nT^{3/2}\right)$ in the latter - the terms $O_p\left(n\sqrt{T}\right)$ and $O_p\left(nT^{3/2}\right)$ come from the remainders in the BN decomposition. Neglecting the heterogeneity of the β_is entails that a further error term, $\sum_{i=1}^n \sum_{t=1}^T \hat{W}_t W_t' (\beta_i - \beta)$, is present. To evaluate its magnitude, consider the denominator as well

$$\left[\sum_{i=1}^n \sum_{t=1}^T \hat{W}_t \hat{W}_t'\right]^{-1} \sum_{i=1}^n \sum_{t=1}^T \hat{W}_t W_t' (\beta_i - \beta) \tag{49}$$
$$= \frac{1}{n} H_1^{-1} \sum_{i=1}^n (\beta_i - \beta) + \left[\sum_{t=1}^T \hat{W}_t \hat{W}_t'\right]^{-1} \left[\sum_{t=1}^T \hat{W}_t \left(W_t - \hat{W}_t\right)'\right] \left[\frac{1}{n}\sum_{i=1}^n (\beta_i - \beta)\right]$$
$$= I + II.$$

Assumption 9 yields $I = O_p\left(n^{-1/2}\right)$. Also, $II = O_p\left(T^{-2}\right) O_p\left(TC_{nT}^{-1}\right) O_p\left(n^{-1/2}\right)$, and thus II is always dominated. Therefore, $\sum_{i=1}^n \sum_{t=1}^T \hat{W}_t W_t' (\beta_i - \beta) = O_p\left(\sqrt{n}T^2\right)$, thereby making the estimation error $\hat{\beta} - \beta = O_p\left(n^{-1/2}\right)$. This result is in line with the findings in Phillips and Moon (1999) and Moon and Phillips (2000), and it is consistent with the idea that neglecting heterogeneity creates an extra, nonstationary, error term, thereby making the model equivalent to a spurious regression. Thus, when (44) is a cointegrating equation, as

$(n, T) \to \infty$ with $n/T \to 0$, we have $\hat{\beta} - H_1^{-1}\beta = O_p\left(n^{-1/2}\right)$ and the limiting distribution of $\hat{\beta} - H_1^{-1}\beta$ is driven by $n^{-1/2}H_1^{-1}\sum_{i=1}^{n}(\beta_i - \beta)$. When (44) is a spurious regression, the numerator is driven by both $n^{-1/2}H_1^{-1}\sum_{i=1}^{n}(\beta_i - \beta)$ and by $n^{-1/2}T^{-2}\sum_{i=1}^{n}\sum_{t=1}^{T}H_1^{-1}W_t u_{it}$. However, Assumption 9*(ii)* entails independence (conditional on the common shocks F_t) between the two terms.

The following propositions summarize the asymptotics of $\hat{\beta} - H_1^{-1}\beta$ and $\hat{\beta}^{FD} - H_1^{-1}\beta$; results are presented in two cases, namely $n \to \infty$ and fixed n.

Proposition 5 *Let Assumptions 1-6 and 9 hold and assume* $(n, T) \to \infty$ *with* $n/T^2 \to 0$. *Then,*

$$\sqrt{n}\left(\hat{\beta} - H_1^{-1}\beta\right) \Rightarrow H_1^{-1}\Sigma_\beta^{1/2} \times Z \tag{50}$$

when (45) is a cointegration relationship, where $Z \sim N(0, I_k)$, *and*

$$\sqrt{n}\left(\hat{\beta}^{FD} - H_1^{-1}\beta\right) \Rightarrow H_1^{-1}\Sigma_\beta^{1/2} \times Z. \tag{51}$$

Also, when (45) is a spurious regression under $n/T \to 0$

$$\sqrt{n}\left(\hat{\beta} - H_1^{-1}\beta\right) \Rightarrow \left(H_1'\int \bar{B}_\varepsilon \bar{B}_\varepsilon' H_1\right)^{-1}\left(H_1^{-1}\int \bar{B}_\varepsilon B_u\right)\sqrt{\bar{h}^\Delta} + H_1^{-1}\Sigma_\beta^{1/2} \times Z \tag{52}$$

where the two random variables are independent.

Proof. See Appendix C. ■

Proposition 6 *Let Assumptions 1-6 and 9 hold and assume* $T \to \infty$. *Then* $\hat{\beta} - H_1^{-1}\beta = O_p(1)$ *with* $E\left(\hat{\beta} - H_1^{-1}\beta\right) = 0$, *Also,* $\hat{\beta}^{FD} - H_1^{-1}\beta = O_p(1)$.

Proof. See Appendix C. ■

Propositions 5 and 6 characterize the asymptotics of $\hat{\beta} - H_1^{-1}\beta$ and $\hat{\beta}^{FD} - H_1^{-1}\beta$, under the cases of n passing to infinity and being fixed respectively.

As Proposition 5 shows, neglecting heterogeneity always results in $\sqrt{n}$-consistency; this result is already well known in the nonstationary case, as proved by Phillips and Moon (1999), and it essentially follows form the fact that in the numerators of both (47) and (48), the terms that dominate are, respectively, $\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t W_t'(\beta_i - \beta) = O_p\left(\sqrt{n}T^2\right)$ and $\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta F_t'(\beta_i - \beta) = O_p(\sqrt{n}T)$. This is also shown in greater detail in the proofs. The main result is that the error arising from neglected heterogeneity dominates the common shock estimation error.

Proposition 6 states that, when n is fixed, β cannot be estimated consistently. Whilst this is in line with Theorems 7 and 8 for the case of spurious regression and for the first-differenced model respectively, it is now also the case for (44) being a cointegrated equation. However, albeit inconsistent, $\hat{\beta}$ is unbiased.

5 Conclusion

This paper developed 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 cross-sectional units and weak dependence among the error terms. We derived rates of convergence and limiting distributions under a comprehensive set of alternative characteristics of panels: several combinations of the cross-sectional dimension n and the time series dimension T; shocks being either observable or unobservable; and stationary and nonstationary panel models, the latter representing either a cointegrating equation or a spurious regression.

When the common shocks are observable, the OLS estimator always provides consistent estimates of the β, the case of spurious regression with fixed n being the only exception. Consistency 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 nonstationary panel factor models. We extend the study of consistency of OLS estimators to the case when the shocks are unobservable and we prove that consistency always holds, the cases of spurious regression and stationary regression when n is fixed being the only exceptions. All the asymptotics for $(n, T) \to \infty$ has been derived in the joint limit, using an approach based on the application of a conditional MDS CLT.

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 shocks being nonstationary; when shocks are stationary, normal distributions are obtained.

In this paper, we primarily consider a panel regression model with only latent shocks F_t as regressors. As we discuss in Section 4, this formulation can be extended to a more general framework containing also idiosyncratic regressors, i.e., $y_{it} = \alpha_i + \beta' F_t + \gamma' x_{it} + u_{it}$, and with heterogeneous slopes, i.e., $y_{it} = \alpha_i + \beta_i' F_t + u_{it}$.

The results derived in this article are asymptotic, and therefore it would be important to assess the finite sample behavior of the estimates via Monte Carlo exercises. Another important extension is to relax the exogeneity hypothesis in Assumption 6(a). In this case, fully modified OLS (Phillips and Hansen, 1990) and/or instrumental variable estimators may be employed. These interesting issues are beyond the scope of the present paper, and we leave them for future studies.

Appendix A: Central Limit Theorem for Multi-index Martingale Difference Sequences

In this Appendix, we provide a joint CLT for MDS, based upon the theory in Hall and Heyde (1980). This is the building block employed to prove the joint limits in Theorems 1 to 4.

Theorem 9 *Consider the sequence of k-dimensional random variables $\{\xi_{iT}\}_{i=1}^{n}$. Let C be an invariant σ-field such that, conditionally on C, (i) the ξ_{iT} are independent across i; (ii) $E[\xi_{iT}| C] = 0$ for all i; (iii) for some $\delta > 0$ it holds that*

$$E\|\xi_{iT}| C\|^{2+\delta} < \infty \tag{53}$$

as $T \to \infty$. Then, as $(n, T) \to \infty$, it holds that conditional on C

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_{iT} \Rightarrow V^{1/2} \times Z \tag{54}$$

where $Z \sim N(0, I_k)$ is independent of V a conditional variance and defined as

$$\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\xi_{iT} \xi_{iT}^{'} \middle| C \right) \xrightarrow{p} V. \tag{55}$$

Proof. In order to prove this theorem, we show that all the conditions required in Corollary 3.1 in Hall and Heyde (1980, p. 58) are satisfied, and thus this corollary can be applied here to prove (54). Consider the σ-field defined as $I_{n,i} = \{\xi_{1T}, ..., \xi_{iT}\} \cup C$; as n expands, the σ-fields $I_{n,i}$ are nested since $I_{n,i} = I_i$ for any $i \leq n$, and therefore condition (3.21) in Hall and Heyde (1980, p. 58) holds. Henceforth, we therefore use the simpler notation I_i, thus suppressing the dependence on n. According to Assumptions (i) and (ii), the ξ_{iT} are independent across i conditional on C; this entails that $E[\xi_{iT}| I_{i-1}] = E[\xi_{iT}| C] = 0$, where the last equality holds by assumption. Thus, the ξ_{iT} are (conditional on C) a zero mean martingale difference array. Equation (53) is essentially a conditional Liapunov condition, which requires conditional integrability of order $2 + \delta$. This also implies that $E\|\xi_{iT}| I_{i-1}\|^2 = E\|\xi_{iT}| C\|^2 < \infty$, and therefore the conditional variance of ξ_{iT} is finite. Also, the conditional Liapunov condition (53) is sufficient for a conditional Lindeberg condition to hold, whereby as $n \to \infty$ for some $\varepsilon > 0$ it holds that

$$\sum_{i=1}^{n} \left[\xi_{iT}^2 d \left(|\xi_{iT}| > \varepsilon \right) \middle| I_{i-1} \right] \xrightarrow{p} 0$$

where $d(\cdot)$ is the indicator function. Thus, all the assumptions required for Corollary 3.1 in Hall and Heyde (1980, p. 58) are satisfied, and a (cross-sectional) CLT holds for the sequence $\sum_{i=1}^{n} \xi_{iT}$, and as $n \to \infty$, the sequence $n^{-1/2} \sum_{i=1}^{n} \xi_{iT}$ converges to a normal random variable, whose asymptotic variance is given in (55). ■

Remarks

R1 Theorem 9 is a joint CLT for MDS sequences, of a similar type to those studied in Hall and Heyde (1980). The only difference is that the random sequence $\sum_{i=1}^{n} \xi_{iT}$ depends on two indexes, n and T, both allowed to pass to infinity, which makes the result applicable in a panel data framework. From the technical viewpoint, Theorem 9 lays out some sufficient conditions whereby Corollary 3.1 in Hall and Heyde (1980, p. 58) holds. A key role is played by the conditional Liapunov condition (53), which states that the sequence ξ_{iT} is (conditionally upon C) integrable of order $2+\delta$ as T passes to infinity. This ensures that the MDS is square integrable, and that a Lindeberg condition holds for the ξ_{iT} as $T \to \infty$; see also Phillips and Moon (1999, p. 1071) for the case of i.i.d. panel models.

R2 Theorem 9 can be applied to the panel models when $(n, T) \to \infty$ and in presence of e.g., strong cross-sectional dependence arising from the presence of a common factor structure. Cross sectional independence among ξ_{iT} is not required, unlike in Phillips and Moon (1999), as long as the ξ_{iT}s are cross sectionally independent conditionally on some invariant σ-field C. When this is the case, the joint asymptotic theory follows from the MDS CLT.

R3 A similar approach was employed to derive asymptotic results in cross-sectional regressions with common shocks by Andrews (2005). To the best of our knowledge, this is the first time that the MDS CLT is applied to study the asymptotics of multi-index sequences. Although the MDS CLT is essentially a cross-sectional result, the role played by T in Theorem 9 is still quite evident, e.g., from (53).

R4 Theorem 9 suggests a methodology to derive the joint asymptotics for panels with possibly strong cross-sectional dependence as $(n, T) \to \infty$, with no need to make appeal to sequential limit. To illustrate this, consider the double sum $\sum_i \sum_t \lambda_i f_t x_{it}$, where e.g., x_{it} is i.i.d. across i, λ_i is non random and f_t is a variable common to all i - this term could arise when studying the estimation error in a panel regression where the error term has a factor structure such as $y_{it} = \alpha + \beta x_{it} + u_{it}$ with $u_{it} = \lambda_i f_t$. The asymptotic theory derived in Phillips and Moon (1999) cannot be applied to $\sum_i \sum_t \lambda_i f_t x_{it}$, since the sequence $\lambda_i f_t x_{it}$ is not independent across i. The limiting distribution of $\sum_i \sum_t \lambda_i f_t x_{it}$ can be studied by applying Theorem 9 to the sequence $\xi_{iT} = s_T \sum_t f_t (\lambda_i x_{it})$, where s_T is a suitable normalization, and by considering the σ-field generated by the f_t, say C_f. If $E[\xi_{iT} | C_f] = 0$ for $T \to \infty$ and if it can be proven that as $T \to \infty$ (53) holds, then Theorem 9 ensures that $\left(n^{-1/2} s_T\right) \sum_i \sum_t \lambda_i f_t x_{it}$ converges to a normal random variable with mean zero and variance V.

R5 Equation (55) suggests a method to calculate V. Note that V can be a constant or a random variable, depending on the assumptions on ξ_{iT}, and thus the limiting distribution of $n^{-1/2} \sum_i \xi_{iT}$ can be mixed normal, as already suggested by Andrews (2005) in the cross-sectional case.

Appendix B: Proofs

Henceforth, we define $C_{nT} = \min\{\sqrt{n}, T\}$ and $\delta_{nT} = \min\left\{\sqrt{n}, \sqrt{T}\right\}$. Also, in order to keep the notation simple, in this and in the other appendices, we set the rotation matrices H_1 and H_2 to be identity matrices of dimensions k and n respectively.

Proof of Theorem 1. Consider the estimation error $\hat{\beta}-\beta = \left[\sum_i \sum_t W_t W_t'\right]^{-1} \left[\sum_i \sum_t W_t u_{it}\right]$ as defined in (9).

Let us start with the denominator $\sum_i \sum_t W_t W_t'$. When $T \to \infty$ and n is fixed, it holds that $\sum_i \sum_t W_t W_t' = O_p\left(T^2\right)$ irrespective of whether (1) is a spurious or a cointegrating regression from Assumptions 2 and 3, and

$$\frac{1}{nT^2} \sum_{i=1}^{n} \sum_{t=1}^{T} W_t W_t' \Rightarrow \int \bar{B}_{\varepsilon} \bar{B}_{\varepsilon}'. \tag{56}$$

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

$$\frac{1}{nT^2} \sum_{i=1}^{n} \sum_{t=1}^{T} W_t W_t' = \frac{1}{T^2} \sum_{t=1}^{T} W_t W_t' \tag{57}$$

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

$$\frac{1}{nT^2} \sum_{i=1}^{n} \sum_{t=1}^{T} W_t W_t' \Rightarrow \int \bar{B}_{\varepsilon} \bar{B}_{\varepsilon}'. \tag{58}$$

As far as the numerator is concerned, we derive the asymptotics 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 $\sum_i \sum_t W_t u_{it}$ 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.

Case 1: large T and fixed n.

We firstly focus our attention to the case where equation (1) 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 functional central limit theorem (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

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

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

This entails that the rate of convergence of the numerator of $\hat{\beta}-\beta$ is $O_p(T)$; therefore, given equation (56) that ensures that the denominator of $\hat{\beta}-\beta$ is $O_p(T^2)$, we have that $\hat{\beta}-\beta = O_p(T^{-1})$. As far as the distribution limit is concerned, we know, combining the asymptotics of ξ_{nT} with equation (56), we have that

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

which proves equation (11). Independence between Z and $\bar{B}_{\varepsilon}$ is ensured by Assumption 6.

We now consider the case when equation (1) is a spurious regression, i.e., $u_{it} \sim I(1)$.

Define $\xi_{nt}^S = T^{-2} W_t \left(\sum_{i=1}^n u_{it}\right)$ 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 (56), this proves that $\hat{\beta} - \beta = O_p(1)$. As far as the limiting distribution is concerned, the asymptotics 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 asymptotics of the denominator given in (56), we get equation (12).

Case 2: large n and fixed T.

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

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

where $Z \sim N(0, I_k)$; 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, the rate of convergence of the numerator of $\hat{\beta} - \beta$ is $O_p(\sqrt{n})$. Combining this with the rate of convergence of the denominator, given by (57), we have that $\hat{\beta} - \beta = O_p\left(n^{-1/2}\right)$. As far as the distribution limit is concerned, combining the asymptotic law of $\tilde{\xi}_{nT}$ with (57), we obtain (13).

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 $n \to \infty$, $n^{-1/2}\sum_{i=1}^n u_{it} \Rightarrow \bar{u}_t$. We have that $\tilde{\xi}_{nT}^S = O_p(1)$, and combining this with equation (57), we obtain $\hat{\beta} - \beta = O_p\left(n^{-1/2}\right)$. 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 (57), we prove the validity of equation (14).

Case 3: large n and large T.

The proof is largely based on Theorem 9.

Let us start with the case where equation (1) is a cointegration relationship. Define $\check{\xi}_{iT} = T^{-1}\sum_{t=1}^T W_t u_{it}$, and consider the BN decomposition for W_t and u_{it}, given respectively by

$$W_t \overset{a.s.}{=} W_t^* + W_0 + \tilde{w}_0 - \tilde{w}_t,$$
$$u_{it} \overset{a.s.}{=} u_{it}^* + u_{i0} + \tilde{\eta}_{i0} - \tilde{\eta}_{it},$$

where $W_t^* = C(1)\sum_{j=1}^t w_j$, $u_{it}^* = D_i(1)\eta_{it}$, $\tilde{w}_t = \sum_{j=0}^\infty\left(\sum_{i=j+1}^\infty C_i\right) w_{t-j}$ and $\tilde{\eta}_{it}$ defined similarly. Thus

$$\check{\xi}_{iT} = \frac{1}{T}\sum_{t=1}^T W_t^* u_{it}^* + R_{iT} = \check{\xi}_{iT}^* + R_{iT}, \tag{59}$$

where, as far as the remainder R_{iT} is concerned, it can be proved using similar arguments as Phillips and Moon (1999) that $R_{iT} = O_p\left(T^{-1/2}\right)$. Consider $\check{\xi}_{iT}^*$ and let C be the σ-field generated by the F_ts. Then $E\left[\check{\xi}_{iT}^*\middle| C\right] = T^{-1}\sum_{t=1}^T W_t^* E\left(u_{it}^*\right) = 0$ for all i and T. Defining I_i as the σ-field generated by F_t and $\left(\check{\xi}_{1T}^*, ..., \check{\xi}_{iT}^*\right)$, it holds that $\left\{\check{\xi}_{iT}^*, I_i\right\}$ is an MDS since $E\left[\check{\xi}_{iT}^*\middle| I_{i-1}\right] = E\left[\check{\xi}_{iT}^*\middle| C\right] = 0$. For some constant M_δ and some $\delta > 0$, we have that, uniformly in i

$$\begin{aligned} E\left\| \check{\xi}_{iT}^{*} \right| C \Big\|^{2+\delta} &= E\left\| \frac{1}{T}\sum_{t=1}^{T} W_t^* u_{it}^* \right| C \Bigg\|^{2+\delta} \\ &\leq M_\delta \sum_{t=1}^{T} E\left\| \frac{1}{T} W_t^* u_{it}^* \right| C \Big\|^{2+\delta} \\ &= M_\delta \frac{1}{T^{2+\delta}} \sum_{t=1}^{T} \|W_t^*\|^{2+\delta} E\left|u_{it}^*\right|^{2+\delta}. \end{aligned}$$

Since $E\left|u_{it}^*\right|^{2+\delta} < \infty$ for all i and constant over t by Assumption 1, it holds that $E\left\| \check{\xi}_{iT}^{*} \right| C \Big\|^{2+\delta} < \infty$ if $T^{-(2+\delta)} \sum_{t=1}^{T} \|W_t^*\|^{2+\delta}$ is stochastically bounded as $T \to \infty$. This is ensured by Theorem 5.3 in Park and Phillips (1999), which holds due to Assumption 2. Thus, a Liapunov condition holds whereby $E\left\| \check{\xi}_{iT}^{*} \right| C \Big\|^{2+\delta} < \infty$ for some $\delta > 0$ - in light of Assumptions 1 and 2, this is ensured for up to $\delta > 2$. Thus the joint MDS CLT can be employed to get, for $(n, T) \to \infty$

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \check{\xi}_{iT}^{*} \Rightarrow \left[\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} E\left(\check{\xi}_{iT}^{*} \check{\xi}_{jT}^{*\prime} \right| C \Big) \right]^{1/2} \times Z$$

with $Z \sim N(0, I_k)$ independent of $E\left(\check{\xi}_{iT}^{*} \check{\xi}_{iT}^{*\prime} \right| C \Big)$. Thus

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \check{\xi}_{iT} &\overset{a.s.}{=} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \check{\xi}_{iT}^{*} + O\left(\sqrt{\frac{n}{T}} \right) \qquad (60) \\ \Rightarrow & \left[\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} E\left(\check{\xi}_{iT}^{*} \check{\xi}_{jT}^{*\prime} \right| C \Big) \right]^{1/2} \times Z, \end{aligned}$$

under $n/T \to 0$. Last, $n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} E\left(\check{\xi}_{iT}^{*} \check{\xi}_{iT}^{*\prime} \right| C \Big)$ for $(n, T) \to \infty$ is given by

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} E\left(\check{\xi}_{iT}^{*} \check{\xi}_{jT}^{*\prime} \right| C \Big) &= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} W_t^* W_s^{*\prime} E\left(u_{it}^* u_{js}^*\right) \\ &= \left[\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} E\left(u_{it}^* u_{jt}^*\right) \right] \left[\frac{1}{T^2} \sum_{t=1}^{T} W_t^* W_t^{*\prime} \right] \end{aligned}$$

where the last equality holds because $E\left(u_{it}^{*}u_{is}^{*}\right)=0$ for all $t\neq s$. As $(n,T)\rightarrow\infty$, the FCLT implied by Assumption 2 and the definition of $\bar{h}$ entail $n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(\check{\xi}_{iT}^{*}\check{\xi}_{iT}^{*\prime}\Big| C\right)\Rightarrow\bar{h}\left(\int\bar{B}_{\varepsilon}\bar{B}_{\varepsilon}^{\prime}\right)$, and therefore as $(n,T)\rightarrow\infty$ with $n/T\rightarrow 0$ we have

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}W_{t}u_{it}\Rightarrow\sqrt{\bar{h}}\left(\int\bar{B}_{\varepsilon}\bar{B}_{\varepsilon}^{\prime}\right)^{1/2}\times Z.$$

Note that in this proof the restriction whereby $n/T\rightarrow 0$ arises for the same reason as in Phillips and Moon (1999), i.e., from the fact that (either or both) W_t and u_{it} could be time dependent and the initial conditions W_0 and u_{i0} need not be zero. Combining this with equation (58), we get that $\hat{\beta}-\beta=O_{p}\left(n^{-1/2}T^{-1}\right)$. As far as the limiting distribution is concerned, combining the asymptotic law of $\check{\xi}_{nT}$ with equation (58), we have:

$$\left[\frac{1}{nT^{2}}\sum_{i=1}^{n}\sum_{t=1}^{T}W_{t}W_{t}^{\prime}\right]^{-1}\left[\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}W_{t}u_{it}\right]\Rightarrow\sqrt{\bar{h}}\left(\int\bar{B}_{\varepsilon}\bar{B}_{\varepsilon}^{\prime}\right)^{-1/2}Z$$

which corresponds to equation (15). When $T/n\rightarrow 0$, the term that dominates in (60) is the second one, and thus the order of magnitude of the numerator of $\hat{\beta}-\beta$ is $O_{p}\left(n\sqrt{T}\right)$. To prove (16), define

$$\frac{\sqrt{T}}{n}\sum_{i=1}^{n}R_{iT}\stackrel{c}{\rightarrow}\Delta_{1} \tag{61}$$

where "$\stackrel{c}{\rightarrow}$" denotes here convergence in some sense (e.g., in distribution or in probability) and R_{iT} is defined in (59). When normalizing $\hat{\beta}-\beta$ by $T^{3/2}$, the asymptotic law of the numerator is therefore given by the quantity Δ_1 defined in (61). Combining this with equation (58), (16) follows immediately.

We now turn to the case when equation (1) is a spurious regression. Let $\check{\xi}_{iT}^{S}=T^{-2}\sum_{t=1}^{T}W_{t}u_{it}$, and consider the BN decomposition of W_t and u_{it}, whereby

$$W_{t}\stackrel{a.s.}{=}W_{t}^{*}+W_{0}+\tilde{w}_{0}-\tilde{w}_{t},$$
$$u_{it}\stackrel{a.s.}{=}u_{it}^{*}+u_{i0}+\tilde{\eta}_{i0}-\tilde{\eta}_{it},$$

with $u_{it}^{*}=F_{i}(1)\sum_{j=1}^{t}\eta_{ij}$ and the other variables defined accordingly. It holds that

$$\check{\xi}_{iT}^{S} = \frac{1}{T^2}\sum_{t=1}^{T} W_t^{*} u_{it}^{*} + R_{iT} = \check{\xi}_{iT}^{S*} + R_{iT}^{S}; \tag{62}$$

again, it can be shown that $R_{iT}^{S} = O_p\left(T^{-1/2}\right)$. Conditioning on C, it holds that $E\left[\check{\xi}_{iT}^{S*}\middle| C\right] = T^{-2}\sum_{t=1}^{T} W_t^{*} E\left(u_{it}^{*}\right) = 0$ for all i and T. Defining I_i as the σ-field generated by F_t and $\left(\check{\xi}_{1T}^{S*}, ..., \check{\xi}_{iT}^{S*}\right)$, it therefore holds that $\left\{\check{\xi}_{iT}^{S*}, I_i\right\}$ is an MDS since $E\left[\check{\xi}_{iT}^{S*}\middle| I_{i-1}\right] = E\left[\check{\xi}_{iT}^{S*}\middle| C\right] = 0$. A Liapunov condition can be proved noting that for all i and some $\delta > 0$

$$\begin{aligned} E\left\|\check{\xi}_{iT}^{S*}\middle| C\right\|^{2+\delta} &= E\left\|\frac{1}{T^2}\sum_{t=1}^{T} W_t^{*} u_{it}^{*}\middle| C\right\|^{2+\delta} \\ &\leq E\left\|\left(\frac{1}{T^2}\sum_{t=1}^{T}\left\|W_t^{*}\right\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{t=1}^{T}\left|u_{it}^{*}\right|^2\right)^{1/2}\middle| C\right\|^{2+\delta} \\ &= \left(\frac{1}{T^2}\sum_{t=1}^{T}\left\|W_t^{*}\right\|^2\right)^{1+\delta/2} E\left[\left(\frac{1}{T^2}\sum_{t=1}^{T}\left|u_{it}^{*}\right|^2\right)^{1+\delta/2}\right]. \end{aligned}$$

The FCLT and the Continuous Mapping Theorem (CMT) ensure that as $T \to \infty$, both quantities are stochastically bounded. Thus, $E\left|\check{\xi}_{iT}^{S*}\middle| C\right|^{2+\delta}$ is bounded for all i and the (joint) MDS CLT can be employed to get, for $(n, T) \to \infty$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\check{\xi}_{iT}^{S*} \Rightarrow \left[\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} E\left(\check{\xi}_{iT}^{S*}\check{\xi}_{jT}^{S*\prime}\middle| C\right)\right]^{1/2} \times Z$$

with $Z \sim N\left(0, I_k\right)$ independent of $E\left(\check{\xi}_{iT}^{S*}\check{\xi}_{iT}^{S*\prime}\middle| C\right)$. Thus

$$\begin{aligned} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\check{\xi}_{iT}^{S} &\overset{a.s.}{=} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\check{\xi}_{iT}^{S*} + O\left(\sqrt{\frac{n}{T}}\right) \\ &\Rightarrow \left[\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} E\left(\check{\xi}_{iT}^{S*}\check{\xi}_{jT}^{S*\prime}\middle| C\right)\right]^{1/2} \times Z \end{aligned} \tag{63}$$

under $n/T \to 0$. Note that

$$\frac{1}{n}\sum_{i=1}^{n} E\left(\check{\xi}_{iT}^{S*}\check{\xi}_{iT}^{S*\prime}\middle| C\right) = \frac{1}{T^4}\sum_{t=1}^{T}\sum_{s=1}^{T} W_t^{*} W_s^{*\prime}\left[\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} E\left(u_{it}^{*} u_{js}^{*}\right)\right]$$

and an alternative representation for the random variable $\left[n^{-1}\sum_{i=1}^{n} E\left(\check{\xi}_{iT}^{S*}\check{\xi}_{iT}^{S*\prime}\Big| C\right)\right]^{1/2}\times$ Z is

$$\left[n^{-1}\sum_{i=1}^{n} E\left(\check{\xi}_{iT}^{S*}\check{\xi}_{iT}^{S*\prime}\Big| C\right)\right]^{1/2}\times Z \overset{D}{=} \left(\int \bar{B}_{\varepsilon}B_{u}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} h_{ij}^{\Delta}\right)^{1/2}$$

where "$\overset{D}{=}$" means equality in distribution. Thus, $n^{-1/2}\sum_{i=1}^{n}\check{\xi}_{iT}^{S*} \Rightarrow \left(\int \bar{B}_{\varepsilon}B_{u}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} h_{ij}^{\Delta}\right)^{1/2}$. This result, together with equation (58), proves that $\hat{\beta}-\beta = O_p\left(n^{-1/2}\right)$. As far as the limiting distribution is concerned, combining this result with the one reported in equation (58), we get equation (17). When $T/n \to 0$, the term that dominates in (63) is the second one, and the order of magnitude of the numerator of $\hat{\beta}-\beta$ is $O_p\left(nT^{3/2}\right)$. To prove (18), define

$$\frac{\sqrt{T}}{n}\sum_{i=1}^{n} R_{iT}^{S} \overset{c}{\to} \Delta_2 \tag{64}$$

where again "$\overset{c}{\to}$" denotes convergence in some sense, and R_{iT}^{S} is defined in (62). When normalizing $\hat{\beta}-\beta$ by $\sqrt{T}$, the asymptotic law of the numerator is therefore given by the quantity Δ_2 defined in (64). Combining this with equation (58), (18) follows immediately. ∎

Proof of Theorem 3. According to equation (25)

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

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 (1) is a cointegration relationship. Following Theorem 9, define

$$\begin{aligned}
\hat{\xi}_{iT} &= \frac{1}{T}\sum_{t=1}^{T}\hat{W}_t\left[\left(W_t-\hat{W}_t\right)'\beta+u_{it}\right] \\
&= \frac{1}{T}\sum_{t=1}^{T}W_t\left[\left(W_t-\hat{W}_t\right)'\beta+u_{it}\right]+\frac{1}{T}\sum_{t=1}^{T}\left(\hat{W}_t-W_t\right)\left(W_t-\hat{W}_t\right)'\beta \\
&\quad +\frac{1}{T}\sum_{t=1}^{T}\left(\hat{W}_t-W_t\right)u_{it} \\
&= \bar{\xi}_{iT}+\xi_{iT}^{a}+\xi_{iT}^{b}.
\end{aligned}$$

Consider first $\bar{\xi}_{iT}$; applying the BN decomposition to W_t and u_{it} with $W_t = W_t^* + R_{WT}$ and $u_{it} = u_{it}^* + R_{uiT}$, we have

$$\begin{aligned}
\bar{\xi}_{iT} &= \frac{1}{T}\sum_{t=1}^{T}W_t^*\left[\left(W_t-\hat{W}_t\right)'\beta+u_{it}^*\right]+R_{iT} \\
&= \bar{\xi}_{iT}^*+R_{iT} \qquad (67)
\end{aligned}$$

where following similar arguments as in Phillips and Solo (1992) it follows that $R_{iT} = O_p\left(T^{-1/2}\right)$. Conditioning on the σ-field C generated by the F_t, it holds that

$$E\left[\left.\bar{\xi}_{iT}^*\right|C\right]=\frac{1}{T}\sum_{t=1}^{T}W_t^*E\left(u_{it}^*\right)+\frac{1}{T}\sum_{t=1}^{T}W_t^*E\left[\left.W_t-\hat{W}_t\right|C\right]'\beta=O_p\left(\frac{1}{T\sqrt{T}}\right)$$

because $E(u_{it}) = 0$ and

$$\begin{aligned}
-E\left[\left.W_t-\hat{W}_t\right|C\right] &= \frac{1}{T^2}\sum_{s=1}^{T}\hat{W}_s\gamma_{s-t}\leq\frac{1}{T^2}\max_{1\leq t\leq T}\left|\hat{W}_t\right|\sum_{s=1}^{T}\left|\gamma_{s-t}\right| \\
&= \frac{1}{T^2}O_p\left(\sqrt{T}\right)O(1)=O_p\left(\frac{1}{T\sqrt{T}}\right).
\end{aligned}$$

Thus, letting $\bar{I}_i$ the union between the σ-field generated by $\left\{\bar{\xi}_{1T}^*, ..., \bar{\xi}_{iT}^*\right\}$ and

C, $\{\bar{\xi}^*_{iT}, \bar{I}_i\}$ is an MDS as $T \to \infty$. Also for some constant $M_\delta < \infty$

$$
\begin{aligned}
E \left\| \bar{\xi}^*_{iT} \right| C \|^{2+\delta} &= E \left\| \frac{1}{T} \sum_{t=1}^{T} W^*_t \left[\left(W_t - \hat{W}_t \right)' \beta + u^*_{it} \right] \right| C \Bigg\|^{2+\delta} \\
&\leq M_\delta \frac{1}{T^{2+\delta}} \sum_{t=1}^{T} \|W^*_t\|^{2+\delta} E \left| \left(W_t - \hat{W}_t \right)' \beta + u^*_{it} \right| C \Bigg|^{2+\delta} \\
&\leq M_\delta \frac{\|\beta\|}{T^{2+\delta}} \sum_{t=1}^{T} \|W^*_t\|^{2+\delta} E \left\| \left(W_t - \hat{W}_t \right) \right| C \Bigg\|^{2+\delta} \\
&\quad + M_\delta \frac{1}{T^{2+\delta}} \sum_{t=1}^{T} \|W^*_t\|^{2+\delta} E \left| u^*_{it} \right|^{2+\delta} \\
&= I + II.
\end{aligned}
$$

Considering II, given that $E\left|u^*_{it}\right|^{2+\delta} < \infty$ for all i and for some $\delta > 0$, and since $\sum_{t=1}^{T} \|W^*_t\|^{2+\delta} = O_p\left(T^{2+\delta}\right)$, it follows that $II = O_p(1)$. As far as I is concerned, note that for some $M'_\delta < \infty$

$$
\begin{aligned}
T^{2(2+\delta)} E \left\| \left(W_t - \hat{W}_t \right) \right| C \Bigg\|^{2+\delta} &\leq M'_\delta \sum_{s=1}^{T} \left\| \hat{F}_s \right\|^{2+\delta} E \left| \frac{e'_t e_s}{n} \right|^{2+\delta} \\
&\quad + M'_\delta \sum_{s=1}^{T} \left\| \hat{F}_s \right\|^{2+\delta} E \left| \eta_{st} \right| C|^{2+\delta} + M'_\delta \sum_{s=1}^{T} \left\| \hat{F}_s \right\|^{2+\delta} E \left| \xi_{st} \right| C|^{2+\delta}
\end{aligned}
$$

and Assumption 2 ensures that $E\left|e'_t e_s/n\right|^{2+\delta} < \infty$, $E\left|\eta_{st}\right| C|^{2+\delta} < \infty$ and $E\left|\xi_{st}\right| C|^{2+\delta} < \infty$. Since $\sum_{s=1}^{T} \left\| \hat{F}_s \right\|^{2+\delta} = O_p\left(T^{2+\delta}\right)$, as $T \to \infty$, $II = o_p(1)$. Thus, $E \left\| \bar{\xi}^*_{iT} \right| C \|^{2+\delta} < \infty$ as $T \to \infty$ and therefore an MDS-CLT can be applied to $\bar{\xi}^*_{iT}$. As far as ξ^a_{iT} and ξ^b_{iT} are concerned, note that

$$
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi^a_{iT} = \frac{\sqrt{n}}{T} \sum_{t=1}^{T} \left(\hat{W}_t - W_t \right) \left(W_t - \hat{W}_t \right)' \beta = O_p\left(\frac{\sqrt{n}}{C^2_{nT}} \right)
$$

according to Lemma 1.2(b); Lemma 2.2 ensures that $n^{-1/2} \sum_{i=1}^{n} \xi^b_{iT} = O_p\left(C^{-1}_{nT}\right)$.

Thus, as $(n,T) \to \infty$ with $n/T \to 0$

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\xi}_{iT} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\bar{\xi}_{iT} + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_{iT}^{a} + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_{iT}^{b} \\
&\overset{a.s.}{=} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\bar{\xi}_{iT}^{*} + O_p\left(\sqrt{\frac{n}{T}}\right) + O_p\left(\frac{\sqrt{n}}{C_{nT}^{2}}\right) + O_p\left(\frac{1}{C_{nT}}\right) \qquad (68) \\
\Rightarrow &\left[\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(\bar{\xi}_{iT}^{*}\bar{\xi}_{jT}^{*\prime}\middle| C\right)\right]^{1/2} \times Z
\end{aligned}$$

with $Z \sim N(0, I_k)$ is independent of $E\left(\bar{\xi}_{iT}^{*}\bar{\xi}_{iT}^{*\prime}\middle| C\right)$. It holds that

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(\bar{\xi}_{iT}^{*}\bar{\xi}_{jT}^{*\prime}\middle| C\right) &= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}W_t^{*}W_s^{*\prime}\left[\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(u_{it}^{*}u_{js}^{*}\right)\right] \\
&+\frac{2}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}W_t^{*}W_s^{*\prime}E\left[\left(W_t-\hat{W}_t\right)^{\prime}\beta\middle| C\right]\left[\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(u_{js}^{*}\right)\right] \\
&+\frac{n}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}W_t^{*}W_s^{*\prime}\beta^{\prime}E\left[\left(\hat{W}_t-W_t\right)\left(\hat{W}_s-W_s\right)^{\prime}\middle| C\right]\beta \\
&= I + II + III.
\end{aligned}$$

Note that, as in the proof of Theorem 1, $I \Rightarrow \sqrt{h}\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}^{\prime}$. Also, since $E\left(u_{js}^{*}\right) = 0$ for all j and s, $II = 0$. As far as III is concerned, this is given by the variance of

$$\frac{\sqrt{n}}{T}\sum_{t=1}^{T}W_t\left(\hat{W}_t - W_t\right)^{\prime}\beta.$$

Under $n/T \to 0$, Lemma 1.4 holds and

$$\sqrt{n}\left(\hat{W}_t - W_t\right) = \frac{1}{T^2}\sum_{s=1}^{T}\hat{W}_sW_s^{\prime}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\lambda_i e_{it} + O_p\left(\frac{1}{\sqrt{nT}}\right).$$

Thus,

$$\frac{\sqrt{n}}{T}\sum_{t=1}^{T}W_t\left(\hat{W}_t - W_t\right)^{\prime}\beta = \frac{1}{\sqrt{n}T}\sum_{t=1}^{T}\sum_{i=1}^{n}W_te_{it}\lambda_i^{\prime}\left(\frac{1}{T^2}\sum_{s=1}^{T}\hat{W}_sW_s^{\prime}\right)^{\prime}\beta + o_p(1).$$

As far as $\left(\sqrt{n}T\right)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{n}W_te_{it}\lambda_i^{\prime}$ is concerned, similar arguments as in the proof of Theorem 1 lead, as $(n,T) \to \infty$ under $n/T \to 0$, to

$$\frac{1}{\sqrt{n}T}\sum_{t=1}^{T}\sum_{i=1}^{n}W_te_{it}\lambda_i^{\prime} \Rightarrow \left[\left(\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}^{\prime}\right)\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j^{\prime}E\left(e_{it}e_{jt}\right)\right]^{1/2} \times Z$$

and $\lim_{n\to\infty} n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j' E(e_{it}e_{jt}) = \Gamma$ by definition. Since by definition $T^{-2}\sum_{s=1}^{T}\hat{W}_s W_s' \Rightarrow \tilde{Q}_B$, we have

$$\frac{\sqrt{n}}{T}\sum_{t=1}^{T} W_t\left(\hat{W}_t - W_t\right)'\beta \Rightarrow \left(\int \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{1/2}\left(\beta'\tilde{Q}_B\Gamma\tilde{Q}_B'\beta\right)^{1/2}\times Z.$$

Therefore

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\xi}_{iT} \Rightarrow \left(\int \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{1/2}\sqrt{\bar{h}+\beta'\tilde{Q}_B\Gamma\tilde{Q}_B'\beta}\times Z;$$

combining this with (66), we obtain (27). Note that when $T/n\to\infty$, the term that dominates in (68) is the second one, of magnitude $O_p\left(\sqrt{n/T}\right)$. Thus, the numerator of $\hat{\beta}-\beta$ is of order $O_p\left(n\sqrt{T}\right)$, and since the denominator is $O_p\left(nT^2\right)$ one has $\hat{\beta}-\beta = O_p\left(T^{-3/2}\right)$. When normalizing $\hat{\beta}-\beta$ by $T^{3/2}$, the asymptotic law of the numerator is given by the quantity Δ_1 defined in (61); (28) follows.

We now turn to the case where (1) is a spurious regression. Considering the numerator of $\hat{\beta}-\beta$, we have

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

Lemma 1.3.(c) ensures that $I = O_p\left(nTC_{nT}^{-1}\right)$. Equation (62) in the proof of Theorem 1 states that

$$\begin{aligned} II &= \sum_{i=1}^{n}\sum_{t=1}^{T}W_t^* u_{it}^* + T^2\sum_{i=1}^{n}R_{iT}^S \\ &= O_p\left(\sqrt{n}T^2\right) + O_p\left(nT^{3/2}\right) \\ &= II_a + II_b.\end{aligned}$$

Last, $III = O_p\left(\sqrt{n}T^{3/2}C_{nT}^{-1}\right)$ after Lemma 2.4. Thus, as $(n,T)\to\infty$ with $n/T\to 0$, the asymptotics of the numerator of $\hat{\beta}-\beta$ is driven by II_a, and therefore

$$\sqrt{n}\left(\hat{\beta}-\beta\right) \overset{a.s.}{=} \left[\sum_{t=1}^{T} W_t W_t'\right]^{-1} \left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T} W_t u_{it}\right] + o_p(1)$$

$$\Rightarrow \quad \sqrt{\bar{h}^{\Delta}}\left(\int \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\right)^{-1}\left(\int \bar{B}_{\varepsilon}B_u\right).$$

As far as the case whereby $(n,T)\to\infty$ with $T/n\to 0$, note that the term that dominates the asymptotic law of the numerator of $\hat{\beta}-\beta$ is now II_b; thus, as $(n,T)\to\infty$ with $T/n\to 0$, $\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t\left[\left(W_t-\hat{W}_t\right)'\beta+u_{it}\right] = O_p\left(nT^{3/2}\right)$, and therefore $\hat{\beta}-\beta = \left[O_p\left(nT^2\right)\right]^{-1} O_p\left(nT^{3/2}\right) = O_p\left(T^{-1/2}\right)$. Equation (30) follows from the definition of (64). ■

Proof of Proposition 1. 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

$$T\left(\hat{\Lambda}-\Lambda\right) = \frac{1}{T}\left[\sum_{t=1}^{T} e_t F_t' + \sum_{t=1}^{T} e_t\left(\bar{F}_t-F_t\right)' + \Lambda\sum_{t=1}^{T}\left(F_t-\bar{F}_t\right)\bar{F}_t'\right]\left[\frac{1}{T^2}\sum_{t=1}^{T}\bar{F}_t\bar{F}_t'\right]^{-1}. \tag{69}$$

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

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

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

and

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

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 (69) 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

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

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

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

$IV = n^{-1}T^{-3}\sum_{s=1}^{T}\sum_{t=1}^{T} e_t\tilde{F}_s'F_t'\Lambda' e_s$ 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 (69) is concerned, we have

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

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

$$a = O_p\left(T^{-1}\right);$$

$$b = O_p\left(T^{-1}\right),$$

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

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

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

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

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

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

and

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

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 (69), we obtain equation (34). ■

Appendix C: Supplementary Material

C.1 Useful Lemmas

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

1.

$$V_{nT}\left(\hat{F}_t - F_t\right) = T^{-2}\sum_{s=1}^{T}\hat{F}_s\gamma_{s-t} + T^{-2}\sum_{s=1}^{T}\hat{F}_s\zeta_{st} + T^{-2}\sum_{s=1}^{T}\hat{F}_s\eta_{st} + T^{-2}\sum_{s=1}^{T}\hat{F}_s\xi_{st} \quad (70)$$

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

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

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

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

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

2. *Denote $C_{nT} = \min\{\sqrt{n}, T\}$. Then*

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

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

   (c) $\sum_{t=1}^{T}\left\|\hat{W}_t - W_t\right\|^2 = O_p\left(TC_{nT}^{-2}\right)$.

3. *It holds that:*

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

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

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

4. When $n/T^3 \to 0$ as $(n,T) \to \infty$, we have

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

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.

Proof. Part 1 of the Lemma corresponds to expression (A.1) in Bai (2004, p. 164). Part 1.2*(a)* is Proposition 1 in Bai (2004, p. 143), and Part 1.2*(b)* is Lemma 1 (Bai, 2004, p. 143). To prove 1.2*(c)*, let $a_t = \hat{F}_t - F_t$ and $\bar{a} = T^{-1}\sum_{t=1}^{T} a_t$ and note that

$$\begin{aligned}\sum_{t=1}^{T}\left\|\hat{W}_t - W_t\right\|^2 &= \sum_{t=1}^{T}\|a_t - \bar{a}\|^2 \\ &= \sum_{t=1}^{T} a_t' a_t - T\bar{a}'\bar{a}.\end{aligned}$$

Then we have

$$\bar{a}'\bar{a} = \frac{1}{T^2}\left[\sum_{t=1}^{T}\left(\hat{F}_t - F_t\right)\right]'\left[\sum_{t=1}^{T}\left(\hat{F}_t - F_t\right)\right]$$

so that

$$\begin{aligned}\bar{a}'\bar{a} &= \frac{1}{T^2}\left\|\sum_{t=1}^{T}\left(\hat{F}_t - F_t\right)\right\|^2 \\ &\leq \frac{1}{T^2}\sum_{t=1}^{T}\left\|\hat{F}_t - F_t\right\|^2 = O_p\left(T^{-1}C_{nT}^{-2}\right)\end{aligned}$$

after part 2*(b)* of the Lemma, and thus

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

Part 3 of the Lemma is a direct application of 1.2*(b)* and the Cauchy-Schwartz inequality. Finally, to prove 1.4 note

$$\sqrt{n}\left(\hat{W}_t - W_t\right) = \sqrt{n}\left(\hat{F}_t - F_t\right) - \sqrt{n}\bar{a},$$

and since $\bar{a}$ is a dominated term we can treat $\sqrt{n}\left(\hat{W}_t - W_t\right)$ as $\sqrt{n}\left(\hat{F}_t - F_t\right)$. Lemma 1.4 now corresponds to Theorem 2 in Bai (2004, p. 148). ■

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. *if* $u_{it} \sim I(0)$, *then* $n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it} = n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}W_t u_{it} + O_p\left(C_{nT}^{-1}\right)$;
3. $T^{-1}\sum_{t=1}^{T}\hat{W}_t\left(F_t - \hat{F}_t\right) = T^{-1}\sum_{t=1}^{T}W_t'\left(F_t - \hat{F}_t\right) + O_p\left(C_{nT}^{-2}\right)$;
4. *if* $u_{it} \sim I(1)$, *then* $n^{-1/2}T^{-2}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{W}_t u_{it} = n^{-1/2}T^{-2}\sum_{i=1}^{n}\sum_{t=1}^{T}W_t u_{it} + O_p\left(T^{-1}C_{nT}^{-1}\right)$.

Proof. Consider the following decomposition:

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

Consider II and III. Using the Cauchy-Schwartz inequality

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

Consider now IV. In this case, Lemma 1.2.*(c)* 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:

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

Using the Cauchy-Schwartz inequality we get

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

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,

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

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

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

We have

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

Last, as far as part 4 is concerned, we have

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

also,

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

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

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

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

Lemma 3 *Let Assumptions 1-2 and 4-6 hold. Then, for the estimated shocks* $\Delta\hat{F}_t$, *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 $\delta_{nT} = \min\left\{\sqrt{n}, \sqrt{T}\right\}$. *Then*

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

(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. *It holds that* $\sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \Delta F_s' \zeta_{st} = O_p \left(n^{-1/2} T^{3/2} \right)$.

Proof. Part 1 is equation (A.1) in Bai (2003, p. 158). Results *(a)* and *(b)* in part 2 corresponds to Proposition 2 and Lemma A.1 in Bai (2003, p. 147 and 159 respectively). The three statements in part 3 of the Lemma are, respectively, Lemmas B.1, B.2 and B.3 in Bai (2003, p. 163-165). Finally, part 4 is discussed in Bai (2003, p. 165). ■

C.2 Proofs

Proof of Theorem 2. Consider the estimation error $\hat{\beta}^{FD} - \beta = \left[\sum_i \sum_t \Delta F_t \Delta F_t'\right]^{-1} \left[\sum_i \sum_t \Delta F_t \Delta u_{it}\right]$, defined in (10).

Let us start with the denominator $\sum_i \sum_t \Delta F_t \Delta F_t'$. When $T \to \infty$ and n is fixed, under both the cases that equation (1) is a spurious or a cointegrating regression, Assumption 2 and the LLN entail $\sum_i \sum_t \Delta F_t \Delta F_t' = O_p(T)$ and

$$\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta F_t' \xrightarrow{p} \Sigma_{\Delta F}. \tag{71}$$

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

$$\frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta F_t' = \sum_{t=1}^{T} \Delta F_t \Delta F_t' \tag{72}$$

whilst as both n and T are large we have

$$\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta F_t' \xrightarrow{p} \Sigma_{\Delta F} \tag{73}$$

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

Consider the numerator. 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, although here there is no distinction between cointegration and spurious regression.

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

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

Assumptions 1(b) and 2 entail that a CLT holds. Therefore, as $T\to\infty$, we have

$$\begin{aligned}\xi_{nT}^{\Delta} &\Rightarrow \left[\lim_{T\to\infty}\frac{1}{T}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma_{ij,ts}\right)E\left(\Delta F_t\Delta F_s'\right)\right]^{1/2}Z \qquad (74)\\ &= \left(\sum_{i=1}^{n}\sum_{j=1}^{n}h_{ij}^{\Delta}\right)^{1/2}\Sigma_{\Delta F}^{1/2}Z,\end{aligned}$$

where $Z\sim N\left(0,I_k\right)$. Thus, the rate of convergence of the numerator of $\hat{\beta}^{FD}-\beta$ is $O_p\left(\sqrt{T}\right)$; combining this with (71), we have that $\hat{\beta}^{FD}-\beta=O_p\left(T^{-1/2}\right)$. As far as the distribution limit is concerned, combining the asymptotics of ξ_{nT}^{Δ} with (71), 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 (19).

Case 2: large n and fixed T.

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))

$$E\left[\tilde{\xi}_{nt}^{\Delta}\tilde{\xi}_{ns}^{\Delta\prime}\right]=\sum_{t=1}^{T}\sum_{s=1}^{T}\bar{\gamma}_{ts}E\left(\Delta F_t\Delta F_s'\right).$$

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

$$\tilde{\xi}^{\Delta}_{nT} \sim \left(\sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \Delta F_s' \bar{\gamma}_{ts} \right)^{1/2} Z$$

where $Z \sim N(0, I_k)$; Assumption 6 ensures independence between Z and the random variable $\sum_t \sum_s \Delta F_t \Delta F_s' \bar{\gamma}_{ts}$. Therefore, in this case the rate of convergence of the numerator of $\hat{\beta}^{FD} - \beta$ is $O_p(\sqrt{n})$. Combining this with the rate of convergence of the denominator, given by equation (72), we have that $\hat{\beta}^{FD} - \beta = O_p\left(n^{-1/2}\right)$. Also, combining this with (72), we obtain (20).

Case 3: large n and large T.

The proof is, as in the case of Theorem 1, an application of Theorem 9. Define $\check{\xi}^{\Delta}_{iT} = T^{-1/2} \sum_{t=1}^{T} \Delta F_t \Delta u_{it}$, and consider the BN decomposition

$$\begin{aligned} \Delta F_t &= \Delta F_t^* + R_{Ft}, \\ \Delta u_{it} &= \Delta u_{it}^* + R_{uit}, \end{aligned}$$

where the expression for the remainders R_{Ft} and R_{uit} can be found in Phillips and Solo (1992), $\Delta F_t^* = C(1)\, w_t$ and Δu_{it}^* is i.i.d. across t. Then we may write

$$\check{\xi}^{\Delta}_{iT} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \Delta F_t^* \Delta u_{it}^* + R^{\Delta}_{iT} = \check{\xi}^{\Delta*}_{iT} + R^{\Delta}_{iT} \tag{75}$$

with $R^{\Delta}_{iT} = O_p\left(T^{-1/2}\right)$ - see Phillips and Solo (1992). Then letting I_i^{Δ} be the union of the σ-field generated by the F_t (referred to as C) and $\left\{\check{\xi}^{\Delta*}_{1T}, ..., \check{\xi}^{\Delta*}_{iT}\right\}$, it holds that $\left\{\check{\xi}^{\Delta*}_{iT}, I_i\right\}$ is an MDS since $E\left[\check{\xi}^{\Delta*}_{iT} \middle| I_{i-1}\right] = E\left[\check{\xi}^{\Delta*}_{iT} \middle| C\right] = T^{-1/2} \sum_{t=1}^{T} \Delta F_t^* E\left(\Delta u_{it}^*\right) = 0$. Also, for all $\delta > 0$ and for some constant $M_\delta < \infty$

$$E\left\| \check{\xi}^{\Delta*}_{iT} \middle| C \right\|^{2+\delta} \le M_\delta \frac{1}{T^{1+\delta/2}} \sum_{t=1}^{T} \|\Delta F_t^*\|^{2+\delta} E\left|\Delta u_{it}^*\right|^{2+\delta}$$

and since $E\left|\Delta u_{it}^*\right|^{2+\delta} < \infty$ and constant over t and $E\|\Delta F_t^*\|^{2+\delta} < \infty$, $E\left\| \check{\xi}^{\Delta*}_{iT} \middle| C \right\|^{2+\delta} = O_p\left(T^{-\delta/2}\right)$ for all i. Thus, an MDS-CLT can be applied and

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \check{\xi}^{\Delta}_{iT} &\overset{a.s.}{=} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \check{\xi}^{\Delta*}_{iT} + O\left(\sqrt{\frac{n}{T}}\right) \qquad (76) \\ &\Rightarrow \left[\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} E\left(\check{\xi}^{\Delta*}_{iT} \check{\xi}^{\Delta*\prime}_{iT} \middle| C \right) \right]^{1/2} \times Z \end{aligned}$$

under $n/T \to 0$, with $Z \sim N(0, I_k)$ independent of $E\left(\check{\xi}_{iT}^{\Delta*}\check{\xi}_{iT}^{\Delta*\prime}\middle| C\right)$. It holds that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(\check{\xi}_{iT}^{\Delta*}\check{\xi}_{iT}^{\Delta*\prime}\middle| C\right) &= \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta F_t^*\Delta F_s^{*\prime}\left[\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(\Delta u_{it}^*\Delta u_{js}^*\right)\right] \\ &= \left[\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(\Delta u_{it}^*\Delta u_{jt}^*\right)\right]\left(\frac{1}{T}\sum_{t=1}^{T}\Delta F_t^*\Delta F_t^{*\prime}\right)\end{aligned}$$

and by definition $\left[n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(\Delta u_{it}^*\Delta u_{jt}^*\right)\right]\left(T^{-1}\sum_{t=1}^{T}\Delta F_t^*\Delta F_t^{*\prime}\right) \to \bar{h}^{\Delta}\Sigma_{\Delta F}$. Combining this with (73), we get that $\hat{\beta}^{FD} - \beta = O_p\left(n^{-1/2}T^{-1/2}\right)$, and also, as $(n,T) \to \infty$ with $n/T \to 0$,

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

which corresponds to equation (20). When $(n,T) \to \infty$ with $T/n \to 0$, the term that dominates in (76) is the second one, and thus order of magnitude of the numerator of $\hat{\beta}^{FD} - \beta$ is $O_p(n)$. Define

$$\frac{\sqrt{T}}{n}\sum_{i=1}^{n}R_{iT}^{\Delta} \xrightarrow{c} \Delta \tag{77}$$

where "$\xrightarrow{c}$" denotes convergence in some sense and R_{iT}^{Δ} is defined in (75). When normalizing $\hat{\beta} - \beta$ by n, the asymptotic law of the numerator is therefore given by the quantity Δ_3 defined in (77). Combining this with equation (73), (22) follows. ■

Proof of Theorem 4. Recall equation (26):

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

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

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

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

$$\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_t \Delta \hat{F}_t' = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta F_t' + O_p\left(\delta_{nT}^{-2}\right) \tag{78}$$

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

$$\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_t \Delta \hat{F}_t' \xrightarrow{p} \Sigma_{\Delta F}. \tag{79}$$

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

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

As far as I is concerned, (75) ensures that

$$I = \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t^* \Delta u_{it}^* + \sqrt{T} \sum_{i=1}^{n} R_{iT}^{\Delta} = I_a + I_b$$

where $I_a = O_p\left(\sqrt{nT}\right)$ and $I_b = O_p\left(n\right)$ - see also (76). Also, following Bai (2003, p 163-164), we could prove

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

last, Lemma 2.3.*(c)* ensures that

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

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

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

Last, note that as $(n,T)\to\infty$ with $n/T\to 0$, term I_b is dominated by I_a and thus by III as well; conversely, as $(n,T)\to\infty$ with $T/n\to 0$, $III = O_p(n)$ and therefore the asymptotics of the numerator of $\hat{\beta}^{FD} - \beta$ is driven by I_b and III. According to Lemma 3.1, III can be decomposed into four terms of magnitude

$$\begin{aligned} n\sum_{t=1}^{T}\Delta\hat{F}_t\left(\Delta F_t - \Delta\hat{F}_t\right)'\beta &= O_p\left(n\sqrt{T}\delta_{nT}^{-1}\right) + O_p\left(\sqrt{n}T\delta_{nT}^{-1}\right) + O_p\left(\sqrt{nT}\right) + O_p\left(\sqrt{nT}\right) \\ &= a+b+c+d. \end{aligned} \tag{80}$$

Thus, two cases may occur:

1. $\frac{n}{T}\to 0$; in this case, $\delta_{nT} = \sqrt{n}$. The asymptotics of the numerator of $\hat{\beta}^{FD} - \beta$ is driven by III and in (80) the dominating term is b with $b = O_p(T)$. Thus, $\hat{\beta}^{FD} - \beta = O_p\left(n^{-1}\right)$; the limiting distribution of $n\left(\hat{\beta}^{FD} - \beta\right)$ is driven by

   $$n\left(\hat{\beta}^{FD} - \beta\right) = \left[\frac{1}{T}\sum_{s=1}^{T}\Delta F_t\Delta F_t'\right]^{-1}\times\left[\frac{n}{T}\sum_{t=1}^{T}\Delta\hat{F}_t\left(\Delta F_t - \Delta\hat{F}_t\right)'\beta\right],$$

   and since

   $$\frac{n}{T}\sum_{t=1}^{T}\Delta\hat{F}_t\left(\Delta F_t - \Delta\hat{F}_t\right)'\beta = -\frac{n}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta F_t\left(\Delta\hat{F}_s - \Delta F_s\right)'\zeta_{st}V^{-1}\beta + o_p(1),$$

   with

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

after rearranging we have:

$$\begin{aligned} b &= -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\frac{\sqrt{n}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta F_t\left(\Delta\hat{F}_s-\Delta F_s\right)'\left[e_{it}e_{is}-E\left(e_{it}e_{is}\right)\right]V^{-1}\beta\right\} \\ &\quad +o_p\left(1\right) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\chi_{iT}+o_p\left(1\right). \end{aligned}$$

Conditional on C, χ_{iT} is a zero mean MDS, and therefore the MDS CLT ensures that

$$b\xrightarrow{d}\left[\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\chi_{iT}\chi_{jT}'\right]^{1/2}\times N\left(0,I_k\right).$$

Equation (31) follows directly from the definition of χ_{iT};

2. $\frac{T}{n}\to 0$, and in such case, given that $\delta_{nT}=\sqrt{T}$, asymptotics of the numerator of $\hat{\beta}^{FD}-\beta$ is driven by III and by I_b as well. Let us consider first the term III; the asymptotics, according to (80), is given by a. Given its definition, after Lemma 2.2*(a)*, we have

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

Consider $\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta\hat{F}_t\Delta\tilde{F}_s'\gamma_{s-t}$; we have

$$\begin{aligned} &\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta\hat{F}_t\Delta\tilde{F}_s'\gamma_{s-t} \\ =\;&\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta F_t\Delta F_s'\gamma_{s-t}+\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\Delta\hat{F}_t-\Delta F_t\right)\Delta F_s'\gamma_{s-t} \\ &+\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta F_t\left(\Delta\tilde{F}_s-\Delta F_s\right)'\gamma_{s-t}+ \\ &\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\Delta\hat{F}_t-\Delta F_t\right)\left(\Delta\tilde{F}_s-\Delta F_s\right)'\gamma_{s-t}. \\ =\;&a_1+a_2+a_3+a_4. \end{aligned}$$

It holds that $a_1 = O_p(T)$; also

$$\begin{aligned} a_2 &\leq \max_{1\leq t\leq T} \|\Delta F_t\| \max_{1\leq s\leq T} \left\|\Delta \tilde{F}_s - \Delta F_s\right\| \sum_{t=1}^{T}\sum_{s=1}^{T} |\gamma_{s-t}| \\ &= O_p(1)\, O_p\left(T^{-1/2}\right) O_p(T) = O_p\left(T^{1/2}\right), \end{aligned}$$

and a similar argument holds for a_3. Last

$$a_4 \leq \left(\max_{1\leq t\leq T} \left\|\Delta \tilde{F}_t - \Delta F_t\right\|\right)^2 \sum_{t=1}^{T}\sum_{s=1}^{T} |\gamma_{s-t}| = O_p\left(T^{-1}\right) O_p(T) = O_p(1).$$

Thus, a_1 dominates in III as $(n,T) \to \infty$ with $T/n \to 0$ and therefore

$$III \overset{a.s.}{=} \frac{n}{T} \sum_{t=1}^{T}\sum_{s=1}^{T} \Delta F_t \Delta F_s' \gamma_{s-t} + o_p(1) = O_p(n).$$

Since $I_b = O_p(n)$ as well, the order of magnitude of the numerator is $O_p(n)$. To derive the distribution limit, as far as III is concerned, after normalizing by $1/n$, we have as $(n,T) \to \infty$ with $T/n \to 0$

$$\frac{1}{T} \sum_{t=1}^{T}\sum_{s=1}^{T} \Delta F_t \Delta F_s' \gamma_{t-s} V^{-1}\beta \xrightarrow{p} \bar{h}_e \Sigma_{\Delta F} V^{-1}\beta;$$

as far as I_b is concerned, this follows from the definition of Δ_3 in (77). Combining this with equation (79), and recalling the definition of $\Sigma_{\Delta F}$, we can derive equation (32).

■

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

$$\frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\breve{F}_t\breve{F}_t' = \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}F_tF_t' + o_p\left(1\right) = O_p\left(1\right).$$

■

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

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

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

since

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

and

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

we get

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

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

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

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

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

which proves equation (35). Equation (36) can be derived following a similar argument. ■

Proof of Theorem 7. Recall equation (25)

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

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

$$\frac{1}{T^2} \sum_{t=1}^{T} \hat{W}_t \hat{W}_t' = \frac{1}{T^2} \sum_{t=1}^{T} W_t W_t' + o_p(1).$$

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

$$\begin{aligned}
IV &= n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\left[\left(\Lambda-\hat{\Lambda}\right)F_t+e_t\right]\left[\left(\Lambda-\hat{\Lambda}\right)F_t+e_t\right]'\hat{\Lambda} \\
&= n^{-2}\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)\sum_{t=1}^{T}F_tF_t'\left(\Lambda-\hat{\Lambda}\right)'\hat{\Lambda} \\
&\quad +n^{-2}\hat{\Lambda}'\left(\Lambda-\hat{\Lambda}\right)\sum_{t=1}^{T}F_te_t'\hat{\Lambda}+n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}e_tF_t'\left(\Lambda-\hat{\Lambda}\right)'\hat{\Lambda} \\
&\quad +n^{-2}\hat{\Lambda}'\left(\sum_{t=1}^{T}e_te_t'\right)\hat{\Lambda}
\end{aligned}$$

with

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

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

and

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

therefore, $IV=O_p\left(T\right)$. Thus we get

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

Note that even if the estimated shocks are not consistent, $T^{-2}\sum_{t=1}^{T}\hat{F}_t\hat{F}_t'$ is a consistent estimator for $T^{-2}\sum_{t=1}^{T}F_tF_t'$. 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 $\min\{1,\delta\}$.

With respect to the numerator of equation (25), 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:

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

with

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

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

and

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

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

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

Hence, the numerator of equation (25) is $O_p(T)$. Combining this result with the asymptotic magnitude of the denominator of equation (25), we get

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

To find the limiting distribution of the numerator of (25) 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:

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

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

$$\begin{aligned}I &= -n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\left(\Lambda W_t+\bar{e}_t\right)e_t'\Lambda\beta \\ &= -n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\Lambda W_te_t'\Lambda\beta-n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\bar{e}_te_t'\Lambda\beta \\ &\Rightarrow -n^{-1}\int\bar{B}_\varepsilon d\bar{B}_e'\Lambda\beta-n^{-2}\Lambda'\Sigma_e\Lambda\beta. \qquad (81)\end{aligned}$$

As far as II is concerned, recalling that $T\left(\hat{\Lambda}-\Lambda\right)\Rightarrow D_\Lambda^1$, we have

$$\begin{aligned}II &= -n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\left(\Lambda W_t+\bar{e}_t\right)\left(\Lambda W_t+\bar{e}_t\right)'\left(\hat{\Lambda}-\Lambda\right)\beta \\ &= -n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\Lambda W_tW_t'\Lambda'\left(\hat{\Lambda}-\Lambda\right)\beta-n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\Lambda W_t\bar{e}_t'\left(\hat{\Lambda}-\Lambda\right)\beta \\ &\quad -n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\bar{e}_tW_t'\Lambda'\left(\hat{\Lambda}-\Lambda\right)\beta-n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\bar{e}_t\bar{e}_t'\left(\hat{\Lambda}-\Lambda\right)\beta \\ &\Rightarrow -n^{-1}\int\bar{B}_\varepsilon\bar{B}_\varepsilon'\Lambda'D_\Lambda^1\beta. \qquad (82)\end{aligned}$$

Likewise

$$\begin{aligned}III &= n^{-2}\hat{\Lambda}'\sum_{t=1}^{T}\bar{z}_tW_t'\left(\hat{\Lambda}-\Lambda\right)'\hat{\Lambda} \\ &\Rightarrow n^{-1}\int\bar{B}_\varepsilon\bar{B}_\varepsilon'D_\Lambda^{1\prime}\Lambda\beta \qquad (83)\end{aligned}$$

Thus, combining equations (81), (82) and (83) we have

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

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

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

This completes the proof of (37).

Finally, we consider the case when equation (1) 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

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

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

In this case the limiting distribution of the numerator is given by the leading term $\sum_{t=1}^{T} W_t \left(\sum_{i=1}^{n} u_{it}\right)$, which converges to $\sqrt{\bar{h}^{\Delta}} \left(\int \bar{B}_{\varepsilon} B_u\right)$. This proves equation (38). ∎

Proof of Theorem 8. Consider equation (26)

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

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 $\hat{\beta}^{FD} - \beta$ 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

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

The limiting distribution of $\hat{\beta}^{FD} - \beta$ can be obtained as follows. Consider first the denominator of $\hat{\beta}^{FD} - \beta$. 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 $\hat{\beta}^{FD} - \beta$ 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 (39). ■

Proof of Proposition 3. The results follow directly from Assumption 8, parts (1) and (2). ■

Proof of Proposition 4. The results follow directly from Assumption 8, parts (1) and (3). ■

Proof of Proposition 5. Equation (47) follows from Assumption 9 and (49). Similarly, (52), and the restriction $n/T \to 0$, follow from (49) and the proof of Theorem 3. Finally, consider (51). As far as the denominator is concerned, we have $(nT)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta\hat{F}_t' \xrightarrow{p} \Sigma_{\Delta F}$. As far as the numerator is concerned, it holds that that $\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta u_{it} = O_p\left(\sqrt{nT}\right)$, and $\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\left(\Delta\hat{F}_t-\Delta F_t\right)' = O_p\left(nT\delta_{nT}^{-2}\right)$ from Lemma 3.3*(b)*. This, together with Assumption 9, also yields $\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta\hat{F}_t\Delta F_t'(\beta_i-\beta) = \sum_{t=1}^{T}\Delta\hat{F}_t\Delta\hat{F}_t'\sum_{i=1}^{n}(\beta_i-\beta)+ o_p(1) = O_p(\sqrt{n}T)$. Thus, $\hat{\beta}^{FD}-\beta = O_p\left(n^{-1/2}\right)$ with $\sqrt{n}\left(\hat{\beta}^{FD}-\beta\right) = n^{-1/2}\sum_{i=1}^{n}(\beta_i-\beta)+o_p(1)$, whence (51). The proof of Proposition 6 is similar and thus omitted. ■

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

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

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

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

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

[7] Bai, J., and Kao, C. (2006), "On the Estimation and Inference of a Cointegrated Regression in Panel Data with Cross-sectional Dependence," in Panel Data Econometrics: Theoretical Contributions and Empirical Applications, Edited by Badi Baltagi, Elsevier Science.

[8] Bai, J., Kao, C., and Ng, S. (2009), "Panel Cointegration with Global Stochastic Trends," *Journal of Econometrics*, 149, 82-99.

[9] Baillie, R. T., and Bollerslev, T. (1989), "Common Stochastic Trends in a Sysyem of Exchange Rates," *Journal of Finance*, 44, 167-181.

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

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

[12] Cochrane, J. H. (2005), *Asset Pricing*, Princeton University Press, Princeton.

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

[14] Connor, G., Hagmann, M., and Linton, O. B. (2007) "Efficient Estimation of a Semiparametric Characteristic-Based Factor Model of Security Returns," LSE STICERD Research Paper No. EM524.

[15] Connor G., and Korajczyk, R. A. (1986), "Performance Measurement with the Arbitrage Pricing Theory," *Journal of Financial Economics*, 15, 373-394.

[16] Das, S. R., Duffie, D., Kapadia, N., and Saita, L. (2007), "Common Failings: How Corporate Defaults Are Correlated," *Journal of Finance*, 57, 93-117.

[17] Dynarski, M., and Sheffrin, S. M. (1985), "Housing Purchases and Transitory Income: A Study With Panel Data,", *Review of Economics and Statistics*, 67, 195-204

[18] Donald, S. G., and Lang, K. (2007), "Inference with Difference-in-Differences and Other Panel Data," *Review of Economics and Statistics*, 3, 221-233.

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

[20] Fuertes, A. M. (2008), "Sieve Bootstrap t-test on Long-Run Average Parameters," *Computational Statistics and Data Analysis*, 52, 3354-70.

[21] Geroski, P., Lazarova, S., Urga, G., and Walters, C. F. (2002), "Are Differences in Firm Size Transitory or Permanent," *Journal of Applied Econometrics*, 18, 47-59.

[22] Hall, P., and Heyde, C. C. (1980), *Martingale Limit Theory and its Applications*, New York, University Press.

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

[24] Kapetanios, G., Pesaran, M. H., and Yamagata, T. (2006), "Panels with Nonstationary Multifactor Error Structures," manuscript.

[25] Liu, W., and Maynard, A. (2005), "Testing Forward Rate Unbiasedness Allowing for Persistent Regressors," *Journal of Empirical Finance*, 12, 613-628.

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

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

[28] Ng, S. (2008), "A Simple Test for Non-Stationarity in Mixed Panels," *Journal of Business and Economics Statistics*, 26, 113-128.

[29] Onatski, A. (2005), "Determining the Number of Factors From Empirical Distribution of Eigenvalues," manuscript.

[30] Park, J. Y., and Phillips, P. C. B. (1999), "Asymptotics for Nonlinear Transformations of Integrated Time Series," *Econometric Theory,* 15, 269-298.

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

[32] Pesaran, M. H. (2007), "A Pair-Wise Approach to Testing for Output and Growth Convergence," *Journal of Econometrics*, 138, 312-355.

[33] Phillips, P. C. B., and Hansen, B. E. (1990), "Statistical Inference in Instrumental Variables Regression with I(1) Processes," *Review of Economic Studies*, 57, 99–125.

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

[35] Phillips, P. C. B., and Solo, V. (1992), "Asymptotics for Linear Processes," *Annals of Statistics*, 20, 971–1001.

[36] Phillips, P. C. B., and Sul, D. (2007), "Transition Modeling and Econometric Convergence Tests," *Econometrica*, 75, 1771-1855.

[37] 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*, University of Chicago, Chicago.

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

[39] 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.

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

[41] Sutton, J. (1997), "Gibrat's Legacy," *Journal of Economic Literature*, 35, 40-59.

[42] Temple, J. (1999), "The New Growth Evidence," *Journal of Economic Literature*, 37, 112-156.

[43] Westerlund, J. (2007), "Estimating Cointegrated Panels with Common Factors and the Forward Rate Unbiasedness Hypothesis," *Journal of Financial Econometrics*, 5, 491-522.