Department of Management and Information Technology

Working Paper

Series “Economics and Management”

n. 15/EM – 2006

Asymptotics for panel models with common shocks

by

Chihwa Kao, Lorenzo Trapani and Giovanni Urga
COMITATO DI REDAZIONE\textsuperscript{s}

Series Economics and Management (EM): Stefano Paleari, Andrea Salanti
Series Information Technology (IT): Stefano Paraboschi
Series Mathematics and Statistics (MS): Luca Brandolini, Alessandro Fassò

\textsuperscript{s} L’accesso alle \textit{Series} è approvato dal Comitato di Redazione. I \textit{Working Papers} ed i \textit{Technical Reports} della Collana dei Quaderni del Dipartimento di Ingegneria Gestionale e dell’Informazione costituiscono un servizio atto a fornire la tempestiva divulgazione dei risultati dell’attività di ricerca, siano essi in forma provvisoria o definitiva.
ASYMPTOTICS FOR PANEL MODELS WITH COMMON SHOCKS

Chihwa Kao
Syracuse University
cdkao@maxwell.syr.edu

Lorenzo Trapani
Cass Business School and Universita’ di Bergamo
L.Trapani@city.ac.uk

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

November 3, 2006

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

*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 usual disclaimer applies. This paper originated while Chihwa Kao was visiting the Centre for Econometric Analysis at Cass (CEA@Cass). Financial support from City University 2005 Pump Priming Fund and CEA@Cass is gratefully acknowledged. Lorenzo Trapani acknowledges financial support from Cass Business School (RAE Development Funds Scheme) and ESRC Postdoctoral Fellowship Scheme (PTA-026-27-1107).
**JEL Classification:** C13, C23.

**Keywords:** Panel data, common shocks, cross-sectional dependence, asymptotics.
1 INTRODUCTION

There is a growing body of literature dealing with limit theory for nonstationary panels. While the first generation of these contributions assumed independence across units (see for instance Phillips and Moon (1999), Kao (1999)), in the second generation this assumption is relaxed, and hypothesis testing and estimation methods are evaluated assuming alternative degrees of cross dependence (see Bai (2003, 2004), Bai and Ng (2002, 2004), Stock and Watson (2002)). We can distinguish the case where regressors are cross-sectionally dependent (see Donald and Lang (2004), Moulton (1990)) from the case where it is the error terms across unit to be dependent (see for instance Bai and Kao, 2005; Moon and Perron, 2004) or both (see for instance Ahn, Lee and Schmidt (2001), Pesaran (2006)).

The main aim of this paper is to propose a novel asymptotic theory for panels with common shocks. We generalize the limit theory developed by Phillips and Moon (1999) 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 $n$ and $T$ 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 section independence though the authors point out that their results still hold when certain degree of weak dependence among panel units is allowed. Thus, the case of Phillips and Moon (1999) 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 aggregates, and may result in mixed asymptotic normality, as Andrews (2005) has demonstrated in a cross-sectional context.  

Andrews (2005, see Theorem 4, p. 1567) proves that the presence of common

\footnote{See also the discussion in Moon and Perron (2004).}
factors among the cross-sectional units makes the limiting distribution of the least squares estimator of $\beta$ mixed normal and not normal as in the classical regression analysis. Note that in this case mixed normality of the least squares estimator of $\beta$ holds even if regressors are $I(0)$ and independent of errors. This finding is also obtained in our paper when studying the distribution limit for the least square estimator of $\beta$ for the $T$ fixed case (see equation 24 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 least squares estimator of $\beta$ is normally distributed.

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

$$y_{it} = \alpha_i + \beta' F_t + u_{it}, \quad (1)$$

$i = 1, \ldots, n, \ t = 1, \ldots, T$, where $\beta$ is a $k \times 1$ vector of slope parameters and $F_t = (F_{1t}, \ldots, F_{kt})'$ 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 relationship depending on whether $u_{it}$ is $I(1)$ or $I(0)$, respectively. When common shocks are not observable, we assume that a set of exogenous variables, $z_{it}$, is observable such that

$$z_{it} = \lambda'_i F_t + \epsilon_{it} \quad (2)$$

where $\lambda_i$ is a vector of factor loadings and $\epsilon_{it}$ is an idiosyncratic component. For the sake of the simplicity of the notation, we assume throughout the paper that the number of the $z_{it}$s is the same as that of the $y_{it}$s. However, the panel dimensions of $y_{it}$s and the $z_{it}$s may be different, for example $y_{it}$s may refer to individuals while $z_{it}$s 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),

$$\Delta y_{it} = \beta' \Delta F_t + \Delta u_{it}. \quad (3)$$
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 national 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 (Quah and Sargent (1993), Forni and Reichlin (1998), Bernanke and Boivin (2000)), while in finance the seminal multifactor framework of the arbitrage pricing theory has generated huge number of contributions in the attempt to identify the unobserved factors underlying the behavior of asset returns. Factor models are useful for forecasting purposes, as is well documented by Stock and Watson (1999, 2005). Bai (2003, 2004), Bai and Ng (2002, 2005) and Boivin and Ng (2005) discuss numerous areas of research where factor models could be employed and some applications in macro and finance. Third, 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.

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

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

2It is important to notice that the notion of fixed or "small" $n$ or $T$ is not well specified. Pesaran (2005) cites $n<10$ as the case when the number of cross sectional units is small. More generally, one could think as fixed $n$ or $T$ a number of cross sectional units or time series observations such that the cross-sectional or the time series average is still far away from the asymptotic limit, but such definition depends on the degree of cross sectional dependence or serial correlation in the panel and is therefore of scarce operational use.
Table 1: Consistency (C) and Limiting Distribution (LD) of $\hat{\beta}_{\text{OLS}}$: $y_{it} = \alpha_i + \beta F_t + u_{it}$.

<table>
<thead>
<tr>
<th></th>
<th>$F_t$ known</th>
<th>$F_t$ unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(n, T)$</td>
<td>C</td>
<td>LD</td>
</tr>
<tr>
<td>Cointegration: $u_{it} \sim I(0)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Fixed n T \to \infty$</td>
<td>Yes</td>
<td>Mixed Normal (Eq.10)</td>
</tr>
<tr>
<td>$Fixed T n \to \infty$</td>
<td>Yes</td>
<td>Mixed Normal (Eq.14)</td>
</tr>
<tr>
<td>$(n, T) \to \infty$</td>
<td>Yes</td>
<td>Mixed Normal (Eq.18)</td>
</tr>
<tr>
<td>&amp; $\sqrt{T/n} \to 0$</td>
<td>Yes</td>
<td>Non Standard (Eq. 34)</td>
</tr>
<tr>
<td>Spurious Regression: $u_{it} \sim I(1)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Fixed n T \to \infty$</td>
<td>No</td>
<td>Non Standard (Eq. 12)</td>
</tr>
<tr>
<td>$Fixed T n \to \infty$</td>
<td>Yes</td>
<td>Non Standard (Eq. 16)</td>
</tr>
<tr>
<td>$(n, T) \to \infty$</td>
<td>Yes</td>
<td>Non Standard (Eq. 20)</td>
</tr>
<tr>
<td>&amp; $T/\sqrt{n} \to 0$</td>
<td>Yes</td>
<td>Non Standard (Eq. 34)</td>
</tr>
<tr>
<td>&amp; $\sqrt{n}/T^2 \to 0$</td>
<td>Yes</td>
<td>Non Standard (Eq. 34)</td>
</tr>
<tr>
<td>&amp; $T^2/\sqrt{n} \to 0$</td>
<td>Yes</td>
<td>Non Standard (Eq. 34)</td>
</tr>
<tr>
<td>First Differences: $\hat{\beta}<em>{\text{OLS}}^{FD}$: $\Delta y</em>{it} = \beta' \Delta F_t + \Delta u_{it}$,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Fixed n T \to \infty$</td>
<td>Yes</td>
<td>Normal (Eq. 22)</td>
</tr>
<tr>
<td>$Fixed T n \to \infty$</td>
<td>Yes</td>
<td>Mixed Normal (Eq. 24)</td>
</tr>
<tr>
<td>$(n, T) \to \infty$</td>
<td>Yes</td>
<td>Normal (Eq. 26)</td>
</tr>
<tr>
<td>&amp; $T/n \to 0$</td>
<td>Yes</td>
<td>Degenerate (Eq. 40)</td>
</tr>
</tbody>
</table>
theory of the least square estimators of $\beta$ 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 concludes. Proofs are reported in the Appendix.

Notation is fairly standard. Throughout the paper, $\|A\|$ denotes the Euclidean norm of matrix $A$, $\sqrt{\text{tr}(AA)}$, ”$\to$” the ordinary limit, ”$\Rightarrow$” weak convergence, ”$\mathbb{P}$” convergence in probability. Stochastic processes such as $B(r)$ on $[0,1]$ are usually written as $B$, integrals such as $\int_0^1 B(r)\, dr$ as $\int B$ and stochastic integrals such as $\int_0^1 B(r)\, dB(r)$ as $\int BdB$.

2 MODEL AND ASSUMPTIONS

We assume that $y_{it}$ is generated as follows

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

$$
F_t = F_{t-1} + \epsilon_t,
$$

$$
z_{it} = \lambda_0' i + F_t + e_{it}
$$

$i = 1,\ldots, n; t = 1,\ldots, T$; $\beta$ is a $(k \times 1)$ vector of slope parameters; $F_t = (F_{1t},\ldots,F_{kt})'$ 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.

The following set of assumptions are used throughout the paper:

**Assumption 1:** (a) either (i) (cointegration case) $u_{it} = D_i(L) \eta_{it}$, or (ii) (spurious regression case) $\Delta u_{it} = F_i(L) \eta_{it}$ with $F_i(1) \neq 0$ and such that $\sum_i u_{it} \sim I(1)$; for both cases, $\eta_{it} \sim iid (0, \sigma^2_{\eta})$ over $t$ and $i$, with $E |\eta_{it}|^8 < M$, $\sum_{j=0}^{\infty} |D_{ij}| < M$, $\sum_{j=0}^{\infty} |F_{ij}| < M$ and $D_i^2(1) \sigma^2_{\eta} > 0$, $F_i^2(1) \sigma^2_{\eta} > 0$; (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 and a Central Limit Theorem hold for the quantities $n^{-1/2} \sum_i u_{it}$ and $n^{-1/2} \sum_i \Delta u_{it}$.

**Assumption 2:** $\epsilon_t = C(L) w_t$ where $C(L) = \sum_{j=0}^{\infty} C_j L^j$; (a) $w_t \sim iid (0, \Sigma_w)$
with $E\|u_t\|^{4+\delta} \leq M$ 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 $E|u_0|^4 \leq M$.

**Assumption 4:** The loadings $\lambda_i$ are non random quantities such that (a) $\|\lambda_i\| \leq M$; (b) either $n^{-1} \sum_{i=1}^{n} \lambda_i \lambda_i' = \Sigma_\Lambda$ if $n$ is finite, or $\lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} \lambda_i \lambda_i' = \Sigma_\Lambda$; if $n \to \infty$; in both cases, the matrix $\Sigma_\Lambda$ is positive definite and such that the eigenvalues of the matrix $\Sigma_\Lambda^{1/2} \Sigma_{\Delta F} \Sigma_\Lambda^{1/2}$ and of the stochastic matrix $\Sigma_\Lambda^{1/2} \int B_t B_t' \Sigma_\Lambda^{1/2}$ are distinct with probability 1.

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

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

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

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

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

Assumption 2 allows for some weak serial correlation in the dynamics of \( \varepsilon_t \). 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 \( \Delta F_t \) 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 analysis to have non-degenerate limiting distributions.

Assumption 3 is a standard initial condition requirement. Assumption 4 serves to identify the factors, which, merely for the purpose of a concise discussion, are assumed to be non random. This requirement could be relaxed, as in Bai (2003, 2004) and Bai and Ng (2004), assuming that the \( \lambda_i \)s are randomly generated and independent of \( \varepsilon_t \) 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 \( z_{it} \)s. Nonetheless, this would be made at the price of losing the possibility to model the \( z_{it} \) as a serially correlated process, whilst in our framework some limited time series and cross sectional dependence in model (2) is allowed for — as one could realize from Assumption 5. As pointed out in Bai (2003),

\footnote{Note that Bai and Ng (2004) allow for factors to be cointegrated given that they consider an approximate factor model with \( F_t \) as common factors. In our paper \( F_t \)s denote instead a set of regressors. Therefore, we need to rule out cointegration among regressors to have non-degenerate limiting distributions.}
the conditions in Assumption 5 are fairly general and allow for consistency and distribution results to hold for the principal component estimator.

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

The following definitions are employed throughout the paper. $B_\varepsilon$ is the Brownian motion associated with the partial sums of $\varepsilon_t$ with covariance matrix $\Omega_{\varepsilon\varepsilon}$, $\bar{B}_\varepsilon (r)$ is the demeaned Brownian motion associated to the partial sums of $F_t$, i.e., $\bar{B}_\varepsilon (r) = B_\varepsilon (r) - \int_0^1 B_\varepsilon (r) \, dr$. Let $h_i \ (h^\Delta_i)$ and $h_{ij} \ (h^\Delta_{ij})$ be the long run variance for $u_{it}$ ($\Delta u_{it}$) and the long run covariance between processes $u_{it}$ and $u_{jt}$ ($\Delta u_{it}$ and $\Delta u_{jt}$). We have $h_{ij} = \sum_{t=1}^T \sum_{s=1}^T \tau_{ij,ts}$ and $h^\Delta_{ij} = \sum_{t=1}^T \sum_{s=1}^T \gamma_{ij,ts}$.

Also, let $\bar{h} = \lim_{n \to \infty} n^{-1} \sum_{i=1}^n \sum_{j=1}^n h_{ij}$ and $\bar{h}^\Delta = \lim_{n \to \infty} n^{-1} \sum_{i=1}^n \sum_{j=1}^n h^\Delta_{ij}$.

Last, the following variances arising from cross sectional aggregation of the $u_{is}$ and the $\Delta u_{is}$ are 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 ASYMPTOTICS

The main objective of this paper is to derive the rate of convergence and limiting distribution of $\hat{\beta}$ and $\hat{\beta}^{FD}$ defined where the OLS estimator for $\beta$ in equation (1) is given by:

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

(5)

where $\bar{F} = T^{-1} \sum_{t=1}^T F_t$, or, when using equation (3), by:

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

(6)

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

1. the shocks $F_t$ can either be known or (more likely) unobservable. The asymptotics of $\hat{\beta}$ and $\hat{\beta}^{FD}$ are affected by the estimation errors if we replace $F_t$ by its estimate $\hat{F}_t$;
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 $\sqrt{nT}$ for panel cointegrated models. In this paper, we are going to face a similar issue, which is compounded by the presence of common shocks in the panel regression (1);

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

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

### 3.1 Observable $F_t$

In the case when $F_t$ is known we have:

$$\hat{\beta} - \beta = \left[ \sum_{i=1}^{n} \sum_{t=1}^{T} W_i W_i^t \right]^{-1} \left[ \sum_{i=1}^{n} \sum_{t=1}^{T} W_i u_{it} \right], \quad (7)$$

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

$$\hat{\beta}^{FD} - \beta = \left[ \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_i \Delta F_i^t \right]^{-1} \left[ \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_i \Delta u_{it} \right]. \quad (8)$$

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

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

$$\hat{\beta} - \beta = O_p(T^{-1}) \quad (9)$$
\( T (\hat{\beta} - \beta) \Rightarrow \frac{1}{n} \left( \int B_t B_t' \right)^{-1/2} \left( \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij} \right)^{1/2} Z \) \hspace{1cm} (10)

if equation (1) is a cointegration relationship, and

\( \hat{\beta} - \beta = O_p(1) \), \hspace{1cm} (11)

\( (\hat{\beta} - \beta) \Rightarrow \left( \int B_t B_t' \right)^{-1} \left( \int B_t B_u \right) \left( \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^{\Delta} \right)^{1/2} \) \hspace{1cm} (12)

if (1) is a spurious regression.

For fixed \( T \) and \( n \to \infty \), we have

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

\[ \sqrt{n} (\hat{\beta} - \beta) \Rightarrow \left( \sum_{t=1}^{T} W_t W_t' \right)^{-1} \left( \sum_{t=1}^{T} \sum_{s=1}^{T} W_t W_s' \bar{u}_{ts} \right)^{1/2} Z, \] \hspace{1cm} (14)

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

\[ \hat{\beta} - \beta = O_p \left( n^{-1/2} \right) \], \hspace{1cm} (15)

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

where \( \bar{u}_t = \lim_{n \to \infty} n^{-1/2} \sum_{i=1}^{n} u_{it} \).

When \( (n, T) \to \infty \), one has

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

\[ \sqrt{nT} (\hat{\beta} - \beta) \Rightarrow \left( \int B_t B_t' \right)^{-1/2} \sqrt{n} Z, \] \hspace{1cm} (18)

if equation (1) is a cointegration relationship and

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

\[ \sqrt{n} (\hat{\beta} - \beta) \Rightarrow \left( \int B_t B_t' \right)^{-1} \left( \int B_t B_u \right) \sqrt{h_{\Delta}}, \] \hspace{1cm} (20)

if it is a spurious regression.
Proof. See Appendix.

Equations (9)-(12) are the standard superconsistency and inconsistency results in the literature. With respect to the speed of convergence, when \((n, T) \to \infty\) our results in equations (17) and (19) lead to the same orders as in Phillips and Moon (1999) and Kao (1999) for both the cointegration and the spurious regression case. Consistency is achieved under the spurious regression case as well, where the rate of convergence is \(\sqrt{n}\). This result, which follows the seminal contributions of Kao (1999) and Phillips and Moon (1999), is reinforced for the case when \(T\) is fixed and \(n \to \infty\). Equations (13) and (15) prove that irrespective of model (1) to be a cointegration regression or a spurious regression, large \(n\) allows for consistency to hold. It is worth observing the complicated distribution that arises when \(T\) is finite; this is essentially due, as outlined in the proof, to the presence of serial correlation in the \(u_{it}\)’s.

For the case of \(n\) and \(T\) large, the rate of convergence for \(\hat{\beta}\) is the same as in Phillips and Moon (1999) under the case of contemporaneous independence across units, but the limiting distributions in equations (18) and (20) differ and are mixed normal rather than normal as in the Phillips and Moon (1999) case. The mixed normality is due to both \(F_t\) being nonstationary and common across units, as can be seen by considering equation (14) for \(T \to \infty\). The design matrix \((nT^2)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} F_i F'_t = T^{-2} \sum_{i=1}^{T} F_i F'_t\) converge in distribution to a random matrix, namely \(\tilde{\beta}_r \tilde{\beta}'_r\), rather than to a constant. Of course, \((nT^2)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} F_i F'_t\) would converge to a constant (in probability) if \(F_t\) were not common shocks, i.e., if \(F_t\) were replaced by, say, \(F_{it}\).

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

**Theorem 2** Suppose Assumptions 1-6 hold and let \(Z \sim N(0, I_k)\) be independent of the \(\sigma\)-field generated by \(\Delta F_t\).

For fixed \(n\) and \(T \to \infty\)

\[
\hat{\beta}^{FD} - \beta = O_{p}\left(T^{-1/2}\right),
\]  

(21)
\[
\sqrt{T} \left( \hat{\beta}^{FD} - \beta \right) \Rightarrow n^{-1} \Sigma_{\Delta F}^{-1/2} \left( \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^2 \right)^{1/2} Z. \tag{22}
\]

For \( T \) fixed and \( n \to \infty \), we have
\[
\hat{\beta}^{FD} - \beta = O_p \left( n^{-1/2} \right), \tag{23}
\]
\[
\sqrt{n} \left( \hat{\beta}^{FD} - \beta \right) \Rightarrow \left( \sum_{t=1}^{T} \Delta F_t \Delta F'_t \right)^{-1} \left( \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \Delta F'_s \bar{r}_{ts} \right)^{1/2} Z. \tag{24}
\]

When \( (n,T) \to \infty \), one has
\[
\hat{\beta}^{FD} - \beta = O_p \left( n^{-1/2} T^{-1/2} \right), \tag{25}
\]
\[
\sqrt{nT} \left( \hat{\beta}^{FD} - \beta \right) \Rightarrow \Sigma_{\Delta F}^{-1/2} \sqrt{h^2} Z. \tag{26}
\]

**Proof.** See Appendix. □

The results were derived for the case of no serial correlation. The presence of time dependence in general involves a more complicated expression of the limiting distributions, but rates of convergence would not be affected. Note also that since the first differenced model is always stationary, irrespective of whether equation (1) is a cointegration equation or a spurious regression, one can always apply the CLT to obtain the limiting distribution of \( \hat{\beta}^{FD} - \beta \); this is indirectly shown by the rate of convergence for the case when \( (n,T) \to \infty \), equal to \( \sqrt{nT} \).

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

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

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

and

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

with estimation errors:

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

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

In what follows, for the purpose of a concise discussion, we assume the number of shocks $k$ to be known.$^4$ We like to emphasize that this is does not

---

$^4$An issue of importance that arises within this framework and that needs tackling prior to estimating the common components $F_t$ is to determine their number, $k$. In light of some recent contributions, e.g., see Bai and Ng (2002) and Onatski (2005), it is natural to refer to model (2) in order to extract both the common factors $F_t$ and their number $k$. It is worth pointing out though that determining $k$ crucially depends on whether both $n$ and $T$ are large or if either dimension is fixed. Under all cases, the literature provides methodologies to estimate $k$ consistently, i.e. to obtain an estimate $\hat{k}$ such that, as either $\{n,T\} \to \infty$ or, alternatively, max $\{n,T\} \to \infty$ and min $\{n,T\}$ is fixed, it holds that $P \left[ \hat{k} = k \right] = 1$ and $P \left[ \hat{k} \neq k \right] = o_p(1)$. Most often these methods treat estimation of $k$ as either model selection or a rank estimation
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) for an elegant proof of this statement.

3.2.1 The case of \( n \) and \( T \) large

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

\[
V_b(k) = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (z_{it} - \lambda_i F_t)^2 ,
\]

when considering \( F_t \) in levels, or

\[
V_a(k) = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (\Delta z_{it} - \lambda_i \Delta F_t)^2
\]

when estimating shocks \( \Delta F_t \) from the first differenced version of model (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 \( \lambda \) and obtaining estimates \( \Delta \hat{F} \) and \( \hat{F} \) using the normalizations \( \Delta \hat{F}/\hat{F}/T = I_k \) or \( \hat{F}^2/T^2 = I_k \). The estimated shock matrices \( \Delta \hat{F} \) and \( \hat{F} \) are \( \sqrt{T} \) times eigenvectors corresponding to the \( k \) largest eigenvalues of the \( T \times T \) matrices \( \Delta Z \Delta Z' \) or \( ZZ' \). It is well known that the solutions to the above minimization problems are not unique, e.g., when estimating shocks \( \Delta F_t \) and \( F_t \), these are not directly identifiable even though they are up to a transformation. In our setup, the knowledge of \( H_1 \Delta F_t, H_1 F_t \) and \( H_2 \lambda_i \) is as good as knowing \( \Delta F_t, F_t \), and \( \lambda_i \). For sake of notational simplicity, in what follows we shall assume that \( H_1 (k \times k) \) and \( H_2 (n \times n) \) are identity matrices.

problem, thereby employing some information criteria. For the case of either \( n \) or \( T \) fixed, the contributions by Lewbel (1991), Donald (1997) and Cragg and Donald (1997) ensure consistent estimation of the either the rank of the \( n \times n \) matrix \( \sum_i z_i z_i' \) with \( z_i \equiv [z_{i1}, ..., z_{iT}]' \) or of the \( T \times T \) matrix \( \sum_t \Delta z_t \Delta z_t' \) with \( \Delta z_t \equiv [z_{t1}, ..., z_{Tt}]' \), depending on whether \( n \) or \( T \) is fixed. When \( (n,T) \to \infty \), the aforementioned procedures are no longer usable to obtain a consistent \( \hat{k} \) and Bai and Ng (2002) propose a consistent estimator for \( k \) - see also Onatski (2005). Note that assumptions 2-6 in our settings ensure the applicability of these methods to equation (2), as it can be immediately verified.
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:

**if** \( \sqrt{n}/T \to 0 \)

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

\[
\sqrt{n} \left( \hat{\beta} - \beta \right) \Rightarrow \left( \int B_z \hat{B}_z' \right)^{-1/2} \left[ \sqrt{n} Z_1 + \sqrt{T} Q_B \hat{Q}_B' \beta Z_2 \right]; \tag{32}
\]

**if** \( T/\sqrt{n} \to 0 \)

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

\[
T^2 \left( \hat{\beta} - \beta \right) \Rightarrow \frac{1}{2} \sigma^2_e \left[ \int B_z \hat{B}_z' \right]^{-1} \Omega_{ee}; \tag{34}
\]

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

\[
T^{-2} \sum_{t=1}^{T} \hat{W}_t W_t' \Rightarrow \hat{Q}_B,
\]

and

\[
\Gamma = \lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j' E(e_{it} e_{jt}),
\]

\[
\sigma^2_e = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \sigma^2_{e_i},
\]

where \( \sigma^2_{e_i} \) is the long run variance of process \( \{ e_{it} \} \).

Let equation (1) be a spurious regression:

**if** \( \sqrt{n}/T \to 0 \), or \( T/\sqrt{n} \to 0 \) and \( \sqrt{n}/T^2 \to 0 \)

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

\[
\sqrt{n} \left( \hat{\beta} - \beta \right) \Rightarrow \left( \int B_z \hat{B}_z' \right)^{-1} \left( \int B_z B_u \right) \sqrt{h_{\Delta}}; \tag{36}
\]

**if** \( T^2/\sqrt{n} \to 0 \), then (33) and (34) hold.
Proof. See Appendix. ■

Consistency is ensured in both cases, even though $T/\sqrt{n} \to 0$ results in a slower (than in the case of $\sqrt{n}/T \to 0$) rate of convergence and in a degenerate behavior of the numerator of $\hat{\beta} - \beta$. This is anyway not surprising given that the shocks estimation errors (see Bai and Ng (2002), and Bai (2004) can be decomposed in several terms of different asymptotic stochastic magnitude, which have an impact only on the numerator. Notice the consequence of equation (1) being a spurious regression: as long as the number of cross sectional units is not exceedingly large, the classical $\sqrt{n}$ consistency holds, and we have the same limiting distribution as in equation (20). When $n$ is far larger than $T$, we have the same result as if relationship (1) were a cointegration relationship.

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

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

**Theorem 4** Suppose Assumptions 1-2 and 4-6 hold.

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

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

$$\frac{\sqrt{nT}}{\beta^{FD} - \beta} \xrightarrow{p} \Sigma_{DF} Q V^{-1} \beta,$$

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

$$Q = p \lim T^{-3/2} \sum_{s=1}^{T} \sum_{t=1}^{T} \Delta \hat{F}_{it} \Delta \tilde{F}_{st} n^{-1} \left[ \sum_{i=1}^{n} (e_{it} e_{is} - \gamma_{s,t}) \right].$$

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

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

$$\frac{T}{\beta^{FD} - \beta} \xrightarrow{p} \bar{h} e V^{-1} \beta,$$

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

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

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

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

### 3.2.2 The case of \( T \) fixed and \( n \) large

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

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

Assumption 7 rules out the possibility of serial correlation in the data generating process of the \( e_{it} \), and therefore this is a constraint on Assumption 5(d). However, contemporaneous correlation and cross sectional heteroscedasticity are preserved.

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

\[
\hat{F}_t - F_t = O_p\left(n^{-1/2}\right)
\]
and
\[ \Delta \hat{F}_t - \Delta F_t = O_p \left( n^{-1/2} \right) \]
for all \( t \).

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

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

**Proof.** See Appendix.

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

**Proof.** See Appendix.

In both cases we have the same results as we would have if the \( F_t \)'s 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 provide a new inferential theory for the case when shocks are unknown and the cross-sectional dimension \( n \) is finite. This case has not been explored in the literature, the only exception being Gonzalo and Granger (1995). Our contribution is aimed at making the estimated shocks usable in a regression framework.

Rewriting model (2) in the vector form, one gets:
\[ z_t = \Lambda F_t + e_t, \]
where \( z_t = (z_{1t}, ..., z_{nt})' \), \( e_t = [e_{1t}, ..., e_{nt}]' \), and \( \Lambda = (\lambda_1, \lambda_2, ..., \lambda_n)' \). Here too one can estimate \( \Lambda \) using the principal components estimator. A feasible
estimator of $\Lambda$, $\hat{\Lambda}$, is given by the $\sqrt{n}$ times the eigenvectors corresponding to the $k$ largest eigenvalues of $Z'Z$. Notice that this estimator exploits the normalization $\hat{\Lambda}'\hat{\Lambda}/n = I_k$, and it turns out to be computationally convenient for the case of $n < T$. For sake of the notation, and without loss of generality, from Assumption 4 we assume henceforth that $n^{-1}\sum_{i=1}^{n} \lambda_i \lambda_i' = I_k$.

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

**Proposition 1** Under Assumptions 3-6 we have

$$\hat{\Lambda} - \Lambda = O_p(T^{-1}),$$

(42)

$$T \left( \hat{\Lambda} - \Lambda \right) \Rightarrow \left[ I_n - n^{-1} \Lambda \left( \int dB_{\epsilon} B_{\epsilon}' \Lambda' \right) \left( \int dB_{\epsilon} B_{\epsilon}' \right)^{-1} \right. - n^{-1} \Lambda \left( \int dB_{\epsilon} B_{\epsilon}' \right) \Lambda \left. + n^{-1} \left[ I_n - 2n^{-1} \Lambda \int dB_{\epsilon} B_{\epsilon}' \Lambda' \right] \Omega_{\epsilon} \Lambda \right],$$

(43)

where $W_{\epsilon}$ is the Wiener process associated to the partial sums of $e_t$ and $\Omega_{\epsilon} = E(e_t e_t')$.

**Proof.** See Appendix. □

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

Henceforth, for sake of notation, we refer to the limiting distribution of

$T \left( \hat{\Lambda} - \Lambda \right)$ as $D_{\Lambda}$, i.e. $T \left( \hat{\Lambda} - \Lambda \right) \Rightarrow D_{\Lambda}$. Given the restriction $\hat{\Lambda}'\hat{\Lambda}/n = I_k$, the OLS estimator of $F_t$, obtained regressing the $z_t$s 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 \( \hat{F}_t = n^{-1} \hat{N}' z_t \), and also the first difference estimator, \( \Delta \hat{F}_t = n^{-1} \hat{N}' \Delta z_t \). Then

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

and

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

uniformly in \( t \).

Proof. See Appendix. □

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

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

Theorem 7 For the estimator \( \hat{\beta} \), we have:

\[
\hat{\beta} - \beta = O_p(T^{-1}),
\]

\[
T \left( \hat{\beta} - \beta \right) \Rightarrow \left[ \int B_\varepsilon B_\varepsilon' \right]^{-1} \left\{ \int B_\varepsilon dB_u \left( \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij} \right)^{1/2} \right\}
\]

\[
- n^{-1} \left[ \int B_\varepsilon dB_\varepsilon' \Lambda \beta + n^{-1} \Lambda' \Sigma e \Lambda \beta \right]
\]

\[
- n^{-1} \left[ \int B_\varepsilon^2 (\Sigma e^2 \Lambda - \Lambda' \Sigma e \Lambda) \right] \beta
\]

where \( B_\varepsilon \) is the demeaned Brownian motion associated to the partial sums of \( e_t \) and \( \Sigma e = \text{Var}(e_t) \). When this is a spurious relationship, one gets

\[
\hat{\beta} - \beta = O_p(1),
\]

\[
\hat{\beta} - \beta \Rightarrow \left( \int B_\varepsilon B_\varepsilon' \right)^{-1} \left( \int B_\varepsilon B_u \right) \left( \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^2 \right)^{1/2}.
\]

Proof. See Appendix. □

Note that even though common shocks cannot be estimated consistently, \( \hat{\beta} \) is consistent when (1) represents a cointegration relationship but inconsistent
when (1) represents a spurious regression. shock estimation has an impact on
the limit distribution of \( \hat{\beta} - \beta \) when equation (1) is a cointegration regression -
see equation (47) above. On the other hand, it does not affect the asymptotic
distribution when equation (1) is a spurious regression - see equation (49). This
can be seen comparing the two distribution limits with equations (10) and (12)
respectively, where shocks are assumed to be known.

Equations (47) and (49) show an important common feature of this theo-
retical framework. Only the numerators of equation (47) and (49) depend on
whether equation (1) is a cointegrating or spurious regression, whilst the de-
nominators are not affected. This is due to the fact (detailed in the proof)
that though \( \hat{F}_t \) is not a consistent estimator for \( F_t \), the quantity \( \sum \hat{F}_t \hat{F}_t' \) is a
consistent estimator for \( \sum F_t F_t' \) for any consistent estimator of the loadings \( \hat{\Lambda} \).

The convergence rate and the limiting distribution for \( \hat{\beta}^{FD} \) are in the fol-
lowing theorem.

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

\[
\hat{\beta}^{FD} - \beta = O_p(1),
\]

and

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

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

**Proof.** See Appendix. ■

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

Theorems 7 and 8 hold if equation (2) represents a cointegration relationship.
We now turn to evaluate the case of \( e_{it} \sim I(1) \).
4 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 units and weak spatial dependence among the error terms. We derived rates of convergence and limiting distributions under a comprehensive set of alternative characteristics of panels: 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 $\beta$, 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.

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 a mixed normality as consequence of the common shocks being nonstationary, while when shocks are stationary, normal distributions are obtained.

In this paper, we consider a panel regression model with only latent shocks $F_t$ as regressors. This formulation can be extended to a more general framework containing also idiosyncratic regressors, i.e. $y_{it} = \alpha_i + \beta^t F_t + \gamma^t x_{it} + u_{it}$.

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

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

Let us start considering the denominator \( \sum_i \sum_t W_i W_i' \). When \( T \to \infty \) and \( n \) is fixed, we have from Assumptions 2 and 3 that under both the cases that equation (1) is a spurious regression or a cointegrating one it holds that

\[
\sum_i \sum_t W_i W_i' = O_p(T^2)
\]

and

\[
\frac{1}{nT^2} \sum_{i=1}^n \sum_{t=1}^T W_i W_i' = \int B_x B_x'.
\]

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

\[
\frac{1}{nT^2} \sum_{i=1}^n \sum_{t=1}^T W_i W_i' = \frac{1}{T^2} \sum_{t=1}^T W_i W_i',
\]

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

\[
\frac{1}{nT^2} \sum_{i=1}^n \sum_{t=1}^T W_i W_i' \Rightarrow \int B_x B_x'.
\]

As far as the numerator is concerned, the proof be derived with respect to three separate cases, following the same structure as in the theorem. We firstly derive the rate of convergence and the limiting distribution of \( \sum_i \sum_t W_i u_{it} \) 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 FCLT holds such that

\[ \xi_{nT} \Rightarrow \int \dot{B}_x dW, \]

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 \( \xi_{nT} \) is

\[ \xi_{nT} \Rightarrow \left( \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij} \right)^{1/2} \left( \int \dot{B}_x \dot{B}_x' \right)^{1/2} Z, \]

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

Hence we have a twofold result. First, the rate of convergence of the numerator of \( \hat{\beta} - \beta \) is \( O_p(T) \); therefore, given equation (52) that ensures that the denominator of \( \hat{\beta} - \beta \) is \( O_p(T^2) \), we have that \( \hat{\beta} - \beta = O_p(T^{-1}) \), proving equation (9). As far as the distribution limit is concerned, we know, combining the asymptotic law of \( \xi_{nT} \) with equation (52), we have that

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

which proves equation (10). Note that independence between \( Z \) and \( \dot{B}_x \) is ensured by Assumption 6.

We can now consider the case when equation (1) is a spurious regression and therefore \( u_{it} \sim I(1) \).
Define first $\xi_n^S = T^{-2} W_t \left( \sum_{i=1}^n u_{it} \right)$ and $\xi_n^S = \sum_{i=1}^T \xi_{nt}^S$. The process $\sum_{t=1}^n u_{it}$ is still a unit root process with long run variance given by $\sum_{i=1}^n \sum_{j=1}^n h_{ij}^2$. Therefore, a FCLT, which follows from Assumptions 1(a), 2 and 3, ensures that $\xi_n^S = O_p(1)$. Together with equation (52), this proves that $\hat{\beta} - \beta = O_p(1)$, as reported in equation (11). As far as the limiting distribution is concerned, here the asymptotic law of the numerator of $\hat{\beta} - \beta$ is given by

$$\xi_n^S = \frac{1}{T^2} \sum_{t=1}^T W_t \left( \sum_{i=1}^n u_{it} \right) \Rightarrow \left( \int B_t B_u \right) \left( \sum_{i=1}^n \sum_{j=1}^n h_{ij}^2 \right)^{1/2}.$$

Combining this with the asymptotic law of the denominator given in equation (52), we get equation (12).

**Case 2: large $n$ and fixed $T$.**

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

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

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

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)

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

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

$$\tilde{\xi}_n^S \sim N \left[ 0, \sum_{t=1}^T \sum_{s=1}^T W_t W_s^T \tau_{ts} \right] = \left( \sum_{t=1}^T \sum_{s=1}^T W_t W_s^T \tau_{ts} \right)^{1/2} Z,$$

where $Z \sim N(0, I_k)$; Assumption 6 ensures independence between $Z$ and the random variable $\sum_{t=1}^T \sum_{s=1}^T W_t W_s^T \tau_{ts}$.

Therefore, in this case 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 equation (53), we have that \( \hat{\beta} - \beta = O_p(n^{-1/2}) \), thereby proving equation (13). As far as the distribution limit is concerned, combining the asymptotic law of \( \xi_{nt} \) with equation (53), we ultimately obtain (14).

Under the spurious regression case, define \( \tilde{\xi}^S_{nt} = W_t \left( n^{-1/2} \sum_{i=1}^n u_{it} \right) \) and \( \tilde{\xi}^S_{nT} = \sum_{t=1}^T \tilde{\xi}^S_{nt} \). Assumption 1(a) ensures the validity of the CLT for \( n^{-1/2} \sum_{i=1}^n u_{it} \), so that uniformly in \( t \) we have, as \( n \to \infty \), \( n^{-1/2} \sum_{i=1}^n u_{it} \Rightarrow \bar{u}_t \). The process \( \bar{u}_t \) is an aggregation of unit root processes, and in light of Assumption 1(a) it is a unit root process with long run variance which by definition is equal to \( \bar{h} \).

From this we have that \( \tilde{\xi}^S_{nT} = O_p(1) \), and combining this with equation (53), we obtain \( \hat{\beta} = \beta = O_p(1) \) as reported in equation (15). As far as the limiting distribution is concerned, since \( \tilde{\xi}^S_{nT} \) is a finite sum, we have \( \tilde{\xi}^S_{nT} \Rightarrow \sum_{t=1}^T W_t \bar{u}_t \) as \( n \to \infty \). Combining this with equation (53), we prove the validity of equation (16).

**Case 3: large \( n \) and large \( T \).**

Let us start with the case where equation (1) is a cointegration relationship.

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

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

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

\[
\tilde{\xi}_{nT} \Rightarrow \int \bar{B}_\varepsilon dW,
\]

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

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

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

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

29
and therefore, as \( n \to \infty \)

\[
\hat{\xi}_{nT} \Rightarrow \sqrt{h} \left( \int \hat{B}_c B_c^\top \right)^{1/2} Z.
\]

Hence, as far as the rate of convergence of \( \hat{\xi}_{nT} \) is concerned, we have \( \hat{\xi}_{nT} = O_p(1) \). Combining this with equation (54), we get that \( \hat{\beta} - \beta = O_p\left( n^{-1/2}T^{-1} \right) \), proving equation (17). As far as the distribution limit is concerned, combining the asymptotic law of \( \hat{\xi}_{nT} \) with equation (54), we have:

\[
\left[ \frac{1}{nT^2} \sum_{i=1}^{n} \sum_{t=1}^{T} W_i W_i^\top \right] \left[ \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} W_t u_{it} \right] \Rightarrow \sqrt{h} \left( \int \hat{B}_c B_c^\top \right)^{-1/2} Z,
\]

which corresponds to equation (18).

We now turn to the case when equation (1) is a spurious regression. Let \( \hat{\xi}_{nT}^S = T^{-2} W_t \left( n^{-1/2} \sum_{i=1}^{n} u_{it} \right) \) and \( \hat{\xi}_{nT} = \sum_{t=1}^{T} \hat{\xi}_{nT}^S \). For fixed \( n \) the process \( n^{-1/2} \sum_{i=1}^{n} u_{it} \) is still a unit root process with long run variance given by \( n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^\Delta \). Therefore, for fixed \( n \) and as \( T \to \infty \), a FCLT, which follows from Assumptions 1(a), 2 and 3, ensures that \( \hat{\xi}_{nT}^S = O_p(1) \). This result, together with equation (54), proves that \( \hat{\beta} - \beta = O_p\left( n^{-1/2} \right) \), as reported in equation (19). As far as the limiting distribution is concerned as \( T \to \infty \) we have

\[
\hat{\xi}_{nT}^S = \frac{1}{T^2} \sum_{t=1}^{T} W_t \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} u_{it} \right) \Rightarrow \left( \int \hat{B}_c B_c^\top \right) \left( \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^\Delta \right)^{1/2};
\]

taking the limit for \( n \to \infty \) leads to

\[
\left( \int \hat{B}_c B_c^\top \right) \left( \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^\Delta \right)^{1/2} \to \sqrt{h^\Delta} \left( \int \hat{B}_c B_c^\top \right).
\]

Combining this result with the one reported in equation (54), we ultimately get equation (20).

**Proof of Theorem 2.** To prove the theorem, we refer to equation (8) that contains the estimation error

\[
\beta^{FD} - \beta = \left[ \sum_i \sum_t \Delta F_i \Delta F_i^\top \right]^{-1} \left[ \sum_i \sum_t \Delta F_i \Delta u_{it} \right].
\]

The proof be derived splitting this quantity into the denominator \( \sum_i \sum_t \Delta F_i \Delta F_i^\top \)

30
and the numerator $\sum_i \sum_t \Delta F_i \Delta u_{it}$, and analyzing the asymptotic behavior of both quantities separately.

Let us start considering the denominator $\sum_i \sum_t \Delta F_i \Delta F'_t$. When $T \to \infty$ and $n$ is fixed, we have from Assumption 2 and the Law of Large Numbers that under both the cases that equation (1) is a spurious regression or a cointegrating one it holds that $\sum_i \sum_t \Delta F_i \Delta F'_t = O_p(T)$ and

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

(56)

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

$$\frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_i \Delta F'_t = \sum_{t=1}^{T} \Delta F_i \Delta F'_t,$$

(57)

whilst as both $n$ and $T$ are large we have

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

(58)

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

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

**Case 1: large $T$ and fixed $n$**

Denote

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

and

$$\xi_{nT}^\Delta = \sum_{t=1}^{T} \xi_{nt}^\Delta.$$
Assumption 6 ensures that $\Delta F_t$ and the $\Delta u_{it}$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 $\xi_{nT}$ has zero mean and covariance structure given by

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

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

$$\xi_{nT} \Rightarrow \left[ \lim_{T \to \infty} \frac{1}{T} \left( \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij,ts} \right) E (\Delta F_t \Delta F_t') \right]^{1/2} Z,$$

where $Z \sim N(0, I_k)$. Hence we have a twofold result. First, the rate of convergence of the numerator of $\hat{\beta}_{FD}^F - \beta$ is $O_p \left( \sqrt{T} \right)$; therefore, given equation (56) that ensures that the denominator of $\hat{\beta}_{FD}^F - \beta$ is $O_p \left( T \right)$, we have that $\hat{\beta}_{FD}^F - \beta = O_p \left( T^{-1/2} \right)$, proving equation (21). As far as the distribution limit is concerned, combining the asymptotic law of $\xi_{nT}$ with equation (56), we have that

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

which proves equation (22).

Case 2: large $n$ and fixed $T$.

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

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

$$\tilde{\xi}_{nT} = \sum_{t=1}^{T} \tilde{\xi}_{nt}.$$
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 \( t \), \( n^{-1/2} \sum_{i=1}^{n} \Delta u_{it} \Rightarrow \bar{u}_{it} \), where \( \Delta u_{it} \) is a normally distributed, zero mean random variable with covariance structure 

\[
E[\Delta \bar{u}_{t} \Delta \bar{u}_{s}] = \sum_{t=1}^{T} \sum_{s=1}^{T} \gamma_{ts}.
\]

Hence, in light of Assumption 6, \( \xi_{nt}^{\Delta} \) is a zero mean random variable whose

covariance structure is given by (after Assumption 1(a))

\[
E[\xi_{nt}^{\Delta} \xi_{ns}^{\Delta'}] = \sum_{t=1}^{T} \sum_{s=1}^{T} \gamma_{ts} E(\Delta F_{t} \Delta F_{s}').
\]

Since \( \xi_{nt}^{\Delta} \) is a finite sum of normally distributed random variables, we have that

\[
\xi_{nt}^{\Delta} \sim \left( \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_{t} \Delta F_{s}' \gamma_{ts} \right)^{1/2} Z,
\]

where \( Z \sim N(0, I_{k}) \); Assumption 6 ensures independence between \( Z \) and the random variable \( \sum_{t} \sum_{s} \Delta F_{t} \Delta F_{s}' \gamma_{ts} \). 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 (57), we have that

\[
\hat{\beta}^{FD} - \beta = O_{p}(n^{-1/2}),
\]

thereby proving equation (23). Also, combining this with equation (57), we ultimately obtain (24).

Case 3: large \( n \) and large \( T \).

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

\[
E[\xi_{nt}^{\Delta} \xi_{ns}^{\Delta'}] = \frac{1}{T} \left( \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij,ts} \right) E(\Delta F_{t} \Delta F_{s}').
\]

From equation (59) we know that, for fixed \( n \) and as \( T \to \infty \)

\[
\xi_{nt}^{\Delta} \Rightarrow \left( \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^{\Delta} \right)^{1/2} \Sigma_{\Delta F}^{1/2} Z,
\]
with \( Z \sim N(0, I_k) \). As \( n \to \infty \) we have

\[
\lim_{n \to \infty} \left( \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij} \right)^{1/2} \delta \Delta \delta \Delta \to \sqrt{h} \Sigma^{1/2} \Delta \delta \Delta.
\]

Hence, as far as the rate of convergence of \( \delta \Delta nT \) is concerned, we have \( \delta \Delta nT = O_p(1) \). Combining this with equation (58), we get that \( \delta \Delta F - \delta \Delta F = O_p \left( n^{-1/2} T^{-1/2} \right) \), proving equation (23). As far as the distribution limit is concerned, we know that

\[
\delta \Delta nT \Rightarrow \sqrt{h} \Sigma^{1/2} \delta \Delta F Z,
\]

as \((n,T) \to \infty\) with \( Z \sim N(0, I_k) \). Combining this result with equation (58), we have:

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

which corresponds to equation (24).

**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. \[
V_{nT} \left( \hat{F}_t - F_t \right) = T^{-1} \sum_{s=1}^{T} \hat{F}_s \gamma_{s-t} + T^{-1} \sum_{s=1}^{T} \hat{F}_s \zeta_{st} + T^{-1} \sum_{s=1}^{T} \hat{F}_s \eta_{st} + T^{-1} \sum_{s=1}^{T} \hat{F}_s \xi_{st},
\]

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

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

\[
\eta_{st} = \frac{1}{n} \Delta F_i' \Lambda_i' e_t,
\]

\[
\xi_{st} = \frac{1}{n} \Delta F_i' \Lambda_i' e_s,
\]

and \( V_{nT} \) is a diagonal matrix containing the largest \( k \) eigenvalues of \((nT)^{-1} ZZ'\) in decreasing order;
Denote $C_{nT} = \min\{\sqrt{n}, T\}$. Consistency of $\hat{F}$ is expressed as

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

(b) $\sum_{t=1}^{T} \|\hat{F}_t - F_t\|^2 = O_p(1) + O_p(n^{-1/2}T^{1/2})$.

It holds that:

(a) $\sum_{t=1}^{T} (\hat{F}_t - F_t) e_t = O_p(TC_{nT}^{-2})$;

(b) $\sum_{t=1}^{T} (\hat{F}_t - F_t) F_t = O_p(1) + O_p(n^{-1/2}T) = O_p(TC_{nT}^{-2})$;

(c) $\sum_{t=1}^{T} (\hat{F}_t - F_t) \hat{F}_t = O_p(TC_{nT}^{-2})$.

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

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

with

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

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

Proof. See Bai (2004). □

**Lemma 2** Lemma 1 ensures that

1. $T^{-2} \sum_{t=1}^{T} \hat{W}_t \hat{W}_t' = T^{-2} \sum_{t=1}^{T} W_t W_t' + o_p(T^{-1/2}C_{nT}^{-1})$;

2. $n^{-1/2}T^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_t u_{it} = n^{-1/2}T^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} W_t u_{it} + O_p(C_{nT}^{-1})$;

3. $T^{-1} \sum_{t=1}^{T} \hat{W}_t (F_t - \hat{F}_t) = T^{-1} \sum_{t=1}^{T} W_t' (F_t - \hat{F}_t) + O_p(C_{nT}^{-2})$.
Proof. Proof is as follows:

\[
\frac{1}{T^2} \sum_{t=1}^{T} \tilde{W}_t \tilde{W}'_t = \frac{1}{T^2} \sum_{t=1}^{T} \left( W_t + \tilde{W}_t - W_t \right) \left( W_t + \tilde{W}_t - W_t \right)'
\]

\[
= \frac{1}{T^2} \sum_{t=1}^{T} W_t W'_t + \frac{1}{T^2} \sum_{t=1}^{T} W_t \left( \tilde{W}_t - W_t \right)'
\]

\[+ \frac{1}{T^2} \sum_{t=1}^{T} \left( \tilde{W}_t - W_t \right) W'_t + \frac{1}{T^2} \sum_{t=1}^{T} \left( \tilde{W}_t - W_t \right) \left( \tilde{W}_t - W_t \right)'
\]

\[= I + II + III + IV.
\]

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

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

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

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

Then

\[
\frac{1}{T^2} \sum_{t=1}^{T} \tilde{W}_t \tilde{W}''_t = \frac{1}{T^2} \sum_{t=1}^{T} W_t W'_t + o_p \left( \frac{1}{\sqrt{TC_{nT}}} \right) + O_p \left( \frac{1}{T C_{nT}} \right),
\]

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

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

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

\[
\left\| \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \left( \tilde{W}_t - W_t \right) u_{it} \right\|
\]

\[= \left\| \frac{1}{T} \sum_{t=1}^{T} \left( \tilde{W}_t - W_t \right) \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} u_{it} \right\|
\]

\[\leq \left( \frac{1}{T} \sum_{t=1}^{T} \left\| \tilde{W}_t - W_t \right\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{i=1}^{n} \left\| \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} u_{it} \right\|^2 \right)^{1/2} = O_p \left( \frac{1}{C_{nT}} \right)
\]

36
given that $T^{-1}\sum_{t=1}^{T} ||\hat{W}_t - W_t||^2 = O_p\left(C_{nT}^{-2}\right)$ and $n^{-1/2}\sum_{i=1}^{n} u_{it} = O_p(1)$.

Hence,
\[
\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_{it} u_{it} = \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} W_{it} u_{it} + O_p\left(\frac{1}{C_{nT}}\right),
\]
proving part 2 of the Lemma. To prove part 3, we note that
\[
\frac{1}{T} \sum_{t=1}^{T} \hat{W}_t^2 (F_t - \hat{F}_t) = \frac{1}{T} \sum_{t=1}^{T} W_t^2 (F_t - \hat{F}_t) + \frac{1}{T} \sum_{t=1}^{T} (W_t - \hat{W}_t)' (F_t - \hat{F}_t) = I + II.
\]

Term $I$ is bounded by $O_p\left(\frac{1}{C_{nT}}\right)$ - see Lemma 1.3.(c) - whilst $II$ is bounded by
\[
\left(\frac{1}{T} \sum_{t=1}^{T} ||\hat{W}_t - W_t||^2\right)^{1/2} \left(\frac{1}{T} \sum_{t=1}^{T} ||F_t - \hat{F}_t||^2\right)^{1/2} = O_p\left(\frac{1}{C_{nT}}\right) O_p\left(\frac{1}{C_{nT}}\right) = O_p\left(\frac{1}{C_{nT}^2}\right).
\]

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

**Proof of Theorem 3.** According to equation (29)
\[
\hat{\beta} - \beta = \left[\sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_i \hat{W}_t'\right]^{-1} \left\{\sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_t \left[\left(W_t - \hat{W}_t\right)' \beta + u_{it}\right]\right\}.
\]

Let us first consider the denominator of this expression. Assumption 3 and Lemma 2.1 imply that
\[
\sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_i \hat{W}_t' = O_p\left(nT^2\right),
\]
and
\[
(nT^2)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_i \hat{W}_t' \Rightarrow \int B z B' z;
\]
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. The numerator of $\hat{\beta} - \beta$ is given by
\[
\frac{1}{T} \sum_{t=1}^{T} \hat{W}_t (W_t - \hat{W}_t)' \beta + \frac{1}{T} \sum_{t=1}^{T} \hat{W}_t u_{it} = I + II.
\]
Let us consider $II$. We know from Theorem 1 and Lemma 2.2 that, as far as $II$ is concerned

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

and therefore

$$\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} W_{it} u_{it} \Rightarrow \left( \int B_x B_x' \right)^{1/2} \sqrt{kZ}, \tag{62}$$

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

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

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

and

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

with $Z_t \sim N(0, \Gamma')$ for every $t$. Therefore the asymptotic magnitude of term $I$ is the same as that of term $II$ and equal to $O_p(\sqrt{nT})$. This proves equation (31). As far as the distribution limit is concerned, we can write

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

$$= \frac{1}{T} \sum_{t=1}^{T} W_t \left[ \frac{1}{T} \sum_{s=1}^{T} \hat{W}_s W'_s \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \lambda_i e_{it} \right) \right]' \beta + o_p(1),$$

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

$$\frac{1}{T} \sum_{t=1}^{T} \hat{W}_t \left[ \frac{1}{T} \sum_{s=1}^{T} \hat{W}_s W'_s \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \lambda_i e_{it} \right) \right]' \beta \Rightarrow \left( \int B_x B_x' \right)^{1/2} \left( \beta' \hat{Q}_B \Gamma \hat{Q}_B' \beta \right)^{1/2} Z,$$

with $Z \sim N(0, I_k)$. Combining this with the asymptotic law of $II$ and with equation (61), we obtain equation (32);

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

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

38
and the term that dominates is the one with asymptotic magnitude $O_p(1)$.

Therefore, $I = \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_t \left(W_t - \hat{W}_t\right)' = O_p(n)$, and term $II = \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_t u_{it} = O_p(\sqrt{nt})$ is dominated. The order of magnitude of the numerator is now $O_p(n)$, and combining this with equation (60) we have

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

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

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

and asymptotically we have:

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

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

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} E \left[\left(\int B_t dB_{ei}\right) \left(\int B_t dB_{ei}\right)\right] = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \text{Var} \left(\int B_t dB_{ei}\right).$$

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

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \text{Var} \left(\int B_t dB_{ei}\right) = \left(\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \sigma_{ei}^2\right) E \left(\int B_t B_t'\right) = \frac{1}{2} \sigma_e^2 \Omega_{xx},$$

given the definition of $\sigma_e^2$ and that $E \left(\int B_t B_t'\right) = 1/2 \Omega_{xx}$. Combining this equation (61), equation (34) is proved.

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

$$\sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_t u_{it} = \sum_{i=1}^{n} \sum_{t=1}^{T} W_t u_{it} + \sum_{i=1}^{n} \sum_{t=1}^{T} \left(\hat{W}_t - W_t\right) u_{it}.$$
After equation (19) we know
\[ \sum_{i=1}^{n} \sum_{t=1}^{T} W_{it} u_{it} = O_{p} (\sqrt{nT^2}) . \]

As per \( \sum_{i=1}^{n} \sum_{t=1}^{T} (\hat{W}_{t} - W_{t}) u_{it} \), application of the Cauchy-Schwartz inequality leads to
\[
\sum_{i=1}^{n} \sum_{t=1}^{T} (\hat{W}_{t} - W_{t}) u_{it} \leq \left[ \sum_{t=1}^{T} \left\| \hat{W}_{t} - W_{t} \right\|^2 \right]^{1/2} \left[ \sum_{i=1}^{n} \sum_{t=1}^{T} u_{it} \right]^{1/2} = O_{p} \left( \sqrt{T C_{nT}} \right) O_{p} (\sqrt{nT}) ,
\]
and therefore \( \sum_{i=1}^{n} \sum_{t=1}^{T} (\hat{W}_{t} - W_{t}) u_{it} \) is always dominated by \( \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_{t} u_{it} \).

Hence, we have
\[
\frac{1}{\sqrt{nT^2}} \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_{t} u_{it} = \frac{1}{\sqrt{nT^2}} \sum_{i=1}^{n} \sum_{t=1}^{T} W_{t} u_{it} + o_{p} (1) ,
\]
and, after equation (55), we have
\[
\frac{1}{\sqrt{nT^2}} \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_{t} u_{it} \Rightarrow \left( \int \hat{B}_{t} B_{u} \right) \sqrt{h^2} .
\]

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

- when \( \sqrt{n/T^2} \to 0 \), two subcases are possible:
  - \( \sqrt{n/T} \to 0 \), and in such case we have \( \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_{t} (W_{t} - \hat{W}_{t}) \beta = O_{p} (\sqrt{nT}) \); therefore the term that dominates is \( \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_{t} u_{it} \); combining this with equations (60) and (61), equations (35) and (36) can be obtained;
\[ T / \sqrt{n} \to 0 \text{ and } \sqrt{n}/T^2 \to 0, \text{ and here } \sum_{i=1}^{n} \sum_{t=1}^{T} W_t \left( W_t - \hat{W}_t \right)^{'} \beta = O_p(n); \text{ in this case, again the term that dominates is } \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_t u_{it} \]

\[ \text{combining this with equations (60) and (61), equations (35) and (36) hold;} \]

- when \( T^2 / \sqrt{n} \to 0 \), we have that \( \sum_{i=1}^{n} \sum_{t=1}^{T} W_t \left( W_t - \hat{W}_t \right)^{'} \beta = O_p(n), \)
  and this is the dominating term. This leads to the same results as in equations (33) and (34).

Lemma 3 Let Assumptions 1-2 and 4-6 hold. Then, for the estimated shocks \( \hat{F}_t \), it holds that

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

where \( \gamma_{s-t} = E \left[ n^{-1} \sum_{i=1}^{n} e_{it} e_{is} \right], \)
\[ \zeta_{st} = n^{-1} \sum_{i=1}^{n} e_{it} e_{is} - \gamma_{s-t}, \]
\[ \eta_{st} = n^{-1} \Delta F_s \Lambda t \epsilon_t, \]
\[ \xi_{st} = n^{-1} \Delta F_s \Lambda s \epsilon_s, \]

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

2. Denote \( \delta_{nT} = \min \left\{ \sqrt{n}, \sqrt{T} \right\} \). Consistency of \( \Delta \hat{F}_t \) is ensured by

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

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

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

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

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

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

Proof of Theorem 4. Recall equation (30):

$$\hat{\beta}^{FD} - \beta = \left[ \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_t \Delta \hat{F}_t' \right]^{-1} \left\{ \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_t \left[ \left( \Delta F_t - \Delta \hat{F}_t \right)' \beta + \Delta u_t \right] \right\}.$$ We firstly study the rate of convergence and the distribution limit of the denominator. The following decomposition holds:

$$\sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_t \Delta \hat{F}_t' = n \sum_{i=1}^{T} \Delta F_t' \Delta F_t + \sum_{t=1}^{T} \left( \Delta \hat{F}_t - \Delta F_t \right)' \Delta \hat{F}_t + \sum_{t=1}^{T} \left( \Delta \hat{F}_t - \Delta F_t \right)' \Delta F_t + n \sum_{t=1}^{T} \left( \Delta \hat{F}_t - \Delta F_t \right)' \left( \Delta \hat{F}_t - \Delta F_t \right)'$$

$$= I + II + III + IV.$$ 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

$$\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_t \Delta \hat{F}_t' = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_t \Delta F_t' + O_p \left( \delta_{nT}^{-2} \right) \quad (63)$$
and, for \((n, T) \to \infty\)

\[
\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_i \Delta \hat{F}_t^p \to \Sigma F.
\]  

(64)

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

\[
\sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_i \left[ \left( \Delta F_i - \Delta \hat{F}_i \right)' \beta + \Delta u_{it} \right] = \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta F_i \Delta u_{it} + \sum_{i=1}^{n} \sum_{t=1}^{T} \left( \Delta \hat{F}_i - \Delta F_i \right) \Delta u_{it} + \sum_{i=1}^{n} \sum_{t=1}^{T} \left( \Delta \hat{F}_i - \Delta F_i \right)' \beta = I + II + III.
\]

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

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

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

\[
II = \sum_{i=1}^{n} \sum_{t=1}^{T} \left( \Delta \hat{F}_i - \Delta F_i \right) \Delta u_{it} = O_p \left( \sqrt{nT\delta^{-2}} \right).
\]

Lemma 2.3.(c) ensures that

\[
III = n \sum_{i=1}^{n} \Delta \hat{F}_i \left( \Delta F_i - \Delta \hat{F}_i \right)' = nO_p \left( T\delta^{-2} \right) = O_p \left( nT\delta^{-2} \right).
\]

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

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

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

\[
\sum_{i=1}^{T} \Delta \hat{F}_i \left( \Delta F_i - \Delta \hat{F}_i \right)' \beta = O_p \left( \sqrt{T\delta^{-1}} \right) + O_p \left( Tn^{-1/2}\delta^{-1} \right) + O_p \left( \sqrt{T/n} \right) + O_p \left( \sqrt{T/n} \right) = a + b + c + d.
\]

Two cases may occur:
1. \( \frac{\hat{\beta}}{n} \to 0 \); in this case, \( \delta_{nT} = \sqrt{n} \). The dominating term is \( b \) and

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

where

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

After similar passages as above, we have

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

After Lemma 2.4, we know that \( T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \Delta F_s' \zeta_{st} = O_p \left( n^{-1/2} T^{1/2} \right) \).

Therefore the order of magnitude of the numerator is \( O_p \left( n^{-1/2} T^{1/2} \right) \), and combining this with equation (63) we obtain equation (37). As per the limiting distribution, since by definiton

\[
Q = p \lim \frac{1}{n^{5/2}} \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta \hat{F}_t \Delta \tilde{F}_s \zeta_{st},
\]

combining this with equation (64), one can derive equation (38);  

2. \( \frac{T}{n} \to 0 \), and in such case, given that \( \delta_{nT} = \sqrt{T} \), the dominating term is \( a \).

Given its definition, after Lemma 2.2.(a), we have

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

and

\[
\sum_{t=1}^{T} \sum_{s=1}^{T} \Delta \hat{F}_t \Delta \tilde{F}_s' \gamma_{s-t} = \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \Delta F_s' \gamma_{s-t} + \sum_{t=1}^{T} \sum_{s=1}^{T} \left( \Delta \hat{F}_t - \Delta F_t \right) \Delta F_s' \gamma_{s-t}
\]

\[
+ \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \left( \Delta \hat{F}_s - \Delta F_s \right)' \gamma_{s-t} + \sum_{t=1}^{T} \sum_{s=1}^{T} \left( \Delta \hat{F}_t - \Delta F_t \right) \left( \Delta \hat{F}_s - \Delta F_s \right)' \gamma_{s-t}.
\]

Then we can show that:
\( \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \Delta F'_s \gamma_{s-t} = O_p(T); \)

\[
\sum_{t=1}^{T} \sum_{s=1}^{T} \Delta F_t \left( \Delta \hat{F}_s - \Delta F_s \right)' \gamma_{s-t} \leq \max_{t} \| \Delta F_t \| \max_{s} \| \Delta \hat{F}_s - \Delta F_s \| \sum_{t=1}^{T} \sum_{s=1}^{T} | \gamma_{s-t} | = O_p \left( T^{-1/2} \right) O_p(T) = O_p \left( T^{1/2} \right);
\]

\[
\sum_{t=1}^{T} \sum_{s=1}^{T} \left( \Delta \hat{F}_t - \Delta F_t \right) \left( \Delta \hat{F}_s - \Delta F_s \right)' \gamma_{s-t} \leq \left( \max_{s} \| \Delta \hat{F}_s - \Delta F_s \| \right) \sum_{t=1}^{T} \sum_{s=1}^{T} | \gamma_{s-t} | = O_p \left( T^{-1} \right) O_p(T) = O_p(1).
\]

Therefore, the dominating term is the first one with

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

Combining this with the rate of convergence of the denominator as given in equation (63), we obtain equation (39). As far as the distribution limit is concerned, we have

\[ \frac{1}{nT} \sum_{t=1}^{T} \sum_{s=1}^{T} \Delta \hat{F}_t \Delta \hat{F}'_{s} \gamma_{s-t} V^{-1} \beta \overset{p}{\to} \Sigma_{\Delta F} \Sigma_{\Delta F} V^{-1} \beta. \]

Combining this with equation (64), and recalling the definition of \( \Sigma_{\Delta F} \), we can derive equation (40).

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

\[ \hat{F}_t - F_t = O_p \left( n^{-\delta} \right), \]
for some $\delta > 0$. In this case we have
\[
\sum_{t=1}^{T} \sum_{i=1}^{n} \tilde{F}_{t} u_{it} = \sum_{t=1}^{T} \sum_{i=1}^{n} F_{t} u_{it} + \sum_{t=1}^{T} \sum_{i=1}^{n} \left( \tilde{F}_{t} - F_{t} \right) u_{it} = O_p \left( n^{1/2} \right) + O_p \left( n^{-\delta} \right) O_p \left( n^{1/2} \right) = O_p \left( n^{1/2} \right),
\]
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 $\tilde{F}_{t}$ we have
\[
\sum_{t=1}^{T} \sum_{i=1}^{n} \tilde{F}_{t} \bar{F}_{t} = \sum_{t=1}^{T} \sum_{i=1}^{n} F_{t} \bar{F}_{t} + o_p(1) = O_p(1).
\]

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

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

As far as equation (43) is concerned, let $\tilde{F}_{t}$ 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 $T \left( \tilde{\Lambda} - \Lambda \right)$ can be decomposed as
\[
T \left( \tilde{\Lambda} - \Lambda \right) = \frac{1}{T} \left[ \sum_{t=1}^{T} e_{t} F_{t} \left( F_{t} - \bar{F}_{t} \right) + \Lambda \sum_{t=1}^{T} \left( F_{t} - \bar{F}_{t} \right) \bar{F}_{t} \right]^{-1}
\]
(65)

As far as the denominator of this expression is concerned, let $\Xi = \int B_{t} B_{t}'.$ We have
\[
\sum_{t=1}^{T} \tilde{F}_{t} \bar{F}_{t} = \sum_{t=1}^{T} F_{t} F_{t} + \sum_{t=1}^{T} \left( \tilde{F}_{t} - F_{t} \right) \tilde{F}_{t} + \sum_{t=1}^{T} \left( \tilde{F}_{t} - F_{t} \right) \left( F_{t} - \bar{F}_{t} \right) + \sum_{t=1}^{T} \left( F_{t} - \bar{F}_{t} \right) \left( \tilde{F}_{t} - F_{t} \right),
\]
where
\[
\sum_{t=1}^{T} F_{t} F_{t} = O_p \left( T^2 \right),
\]
\[
\sum_{t=1}^{T} \left( \tilde{F}_{t} - F_{t} \right) \tilde{F}_{t} = O_p \left( T \right),
\]
\[
\sum_{t=1}^{T} \left( F_{t} - \bar{F}_{t} \right) \left( \tilde{F}_{t} - F_{t} \right) = O_p \left( T \right),
\]
and
\[
\sum_{t=1}^{T} \left( \tilde{F}_{t} - F_{t} \right) \left( F_{t} - \bar{F}_{t} \right) = O_p \left( T^{3/2} \right).
\]
\[
\sum_{t=1}^{T} (\bar{F}_t - F_t) (\bar{F}_t - F_t)' = O_p(T);
\]
the last two equalities come directly from Lemma B.4(ii) and Lemma B.1 in Bai (2004). Therefore
\[
T^{-2} \sum_{t=1}^{T} F_t F_t' = T^{-2} \sum_{t=1}^{T} F_t F_t' + O_p(T^{-1})
\]
and
\[
T^{-2} \sum_{t=1}^{T} \bar{F}_t \bar{F}_t' \Rightarrow \Xi.
\]
As far as the numerator of equation (65) is concerned, we study each term. First of all we know that \( T^{-1} \sum_{t=1}^{T} \eta_t F_t' \Rightarrow \int dW_t B_t \). The limiting distribution of \( \sum_{t=1}^{T} \eta_t (F_t - F_t)' \) can be obtained from the following decomposition - see Bai (2004, p. 164) for details:
\[
\bar{F}_t - F_t = T^{-2} \sum_{s=1}^{T} \bar{F}_s \gamma_n (s, t) + T^{-2} \sum_{s=1}^{T} \bar{F}_s \zeta_{st} + T^{-2} \sum_{s=1}^{T} \bar{F}_s \eta_{st} + T^{-2} \sum_{s=1}^{T} \bar{F}_s \xi_{st},
\]
where (as in Lemma 1) we let \( \gamma_n (s, t) = E(e_t' e_s/n), \zeta_{st} = e_t' e_s/n - \gamma_n (s, t), \eta_{st} = F_t' \mathcal{N} e_t/n, \xi_{st} = F_t' \mathcal{N} e_s/n \). Hence
\[
\quad \frac{1}{T} \sum_{t=1}^{T} \eta_t (\bar{F}_t - F_t)' = T^{-3} \sum_{s=1}^{T} T^{-3} \sum_{t=1}^{T} \eta_t \bar{F}_s' \gamma_n (s, t) + T^{-3} \sum_{s=1}^{T} T^{-3} \sum_{t=1}^{T} \eta_t \bar{F}_s' \zeta_{st}
\]
\[
\quad \quad + T^{-3} \sum_{s=1}^{T} T^{-3} \sum_{t=1}^{T} \eta_t \bar{F}_s' \eta_{st} + T^{-3} \sum_{s=1}^{T} T^{-3} \sum_{t=1}^{T} \eta_t \bar{F}_s' \xi_{st}
\]
\[
\quad = I + II + III + IV;
\]
and
\[
I = O_p(T^{-1}) \text{ since } E \left| e_t \bar{F}_s' \gamma_n (s, t) \right| \leq |\gamma_n (s, t)| \left( \max_{s,t} E \left| e_t \bar{F}_s' \right| \right) \text{ and } \max_{s,t} E \left| e_t \bar{F}_s' \right| = O_p(T);
\]
\[
II = n^{-1} T^{-3} \sum_{s=1}^{T} T^{-3} \sum_{t=1}^{T} e_t \bar{F}_s' e_s - T^{-3} \sum_{s=1}^{T} T^{-3} \sum_{t=1}^{T} e_t \bar{F}_s' \gamma_n (s, t) \text{ and we have}
\]
\[
n^{-1} T^{-3} \sum_{s=1}^{T} T^{-3} \sum_{t=1}^{T} e_t \bar{F}_s' e_s = n^{-1} T^{-3} \sum_{s=1}^{T} T^{-3} \sum_{t=1}^{T} e_t e_s \bar{F}_s'
\]
\[
= n^{-1} T^{-1} \left( T^{-1} \sum_{t=1}^{T} e_t e_t' \right) \left( T^{-1} \sum_{s=1}^{T} \bar{F}_s' \right) = O_p(T^{-1});
\]
47
$$III = n^{-1}T^{-3} \sum_{s=1}^{T} \sum_{t=1}^{T} e_t \tilde{F}_s^t F_s^t N^e_t$$ with

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

$$= n^{-1} \left( T^{-1} \sum_{t=1}^{T} e_t e'_t \right) \Lambda \left( T^{-2} \sum_{s=1}^{T} F_s \tilde{F}_s^t \right) = O_p(1);$$

$$IV = n^{-1}T^{-3} \sum_{s=1}^{T} \sum_{t=1}^{T} e_t \tilde{F}_s^t F_s^t N^e_s$$ and

$$n^{-1}T^{-3} \sum_{s=1}^{T} \sum_{t=1}^{T} e_t \tilde{F}_s^t F_s^t N^e_s = n^{-1}T^{-3} \sum_{s=1}^{T} \sum_{t=1}^{T} e_t F_s^t N^e_s \tilde{F}_s^t$$

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

Therefore the term that dominates is $$III$$ and

$$n^{-1} \left( T^{-1} \sum_{t=1}^{T} e_t e'_t \right) \Lambda \left( T^{-2} \sum_{s=1}^{T} F_s \tilde{F}_s^t \right) \Rightarrow n^{-3} \Omega \Lambda Q.$$ 

Finally, as far as the term $$\Lambda \sum_{t=1}^{T} (F_t - \tilde{F}_t) \tilde{F}_t^t$$ in equation (65) is concerned, we have

$$T^{-1} \Lambda \sum_{t=1}^{T} (F_t - \tilde{F}_t) \tilde{F}_t^t = -T^{-3} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{F}_s \tilde{F}_t^t \gamma_n(s,t) - T^{-3} \sum_{t=1}^{T} \sum_{s=1}^{T} \tilde{F}_s F_t^t \zeta_{st}$$

$$- T^{-3} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{F}_s \tilde{F}_t^t \eta_{st} - T^{-3} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{F}_s F_t^t \xi_{st}$$

$$= a + b + c + d.$$ 

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

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

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

the term

$$c = n^{-1}T^{-3} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{F}_s \tilde{F}_t^t F_s N^e_t,$$
Combining the limiting distributions of all terms and with

\[ n^{-1}T^{-3} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{F}_s \tilde{F}_t^\prime \Lambda' e_t = n^{-1}T^{-3} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{F}_s \tilde{F}_t^\prime \Lambda' \tilde{F}_t^\prime = n^{-1} \left( T^{-2} \sum_{s=1}^{T} \tilde{F}_s \tilde{F}_s^\prime \right) \Lambda' \left( T^{-1} \sum_{t=1}^{T} e_t \tilde{F}_t^\prime \right) = O_p(1); \]

and

\[ d = n^{-1}T^{-3} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{F}_s \tilde{F}_t^\prime \Lambda' e_s, \]

with

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

Thus the limiting distribution of \( \sum_{t=1}^{T} (F_t - \hat{F}_t) \tilde{F}_t^\prime \) is determined by \( c \) and \( d \), and we have

\[ c = n^{-1} \left( T^{-2} \sum_{s=1}^{T} F_s \tilde{F}_s^\prime \right) \Lambda' \left( T^{-1} \sum_{t=1}^{T} e_t \tilde{F}_t^\prime \right) = n^{-1} \left( T^{-2} \sum_{s=1}^{T} F_s \tilde{F}_s^\prime \right) \Lambda' \left( T^{-1} \sum_{t=1}^{T} e_t F_t \right) + n^{-1} \left( T^{-2} \sum_{s=1}^{T} F_s \tilde{F}_s^\prime \right) \Lambda' \left( T^{-1} \sum_{t=1}^{T} e_t (\hat{F}_t - F_t) \right) \]

\[ \Rightarrow n^{-1}QA' \left[ \int dW_c B_c' + n^{-1} \Omega_c Q \right], \]

and

\[ d = n^{-1} \left( T^{-1} \sum_{s=1}^{T} \tilde{F}_s e_s' \right) \Lambda' \left( T^{-1} \sum_{t=1}^{T} F_t \tilde{F}_t^\prime \right) \Rightarrow n^{-1} \left[ \int B_c dW_c' + n^{-1} QA' \Omega_c \right] \Lambda Q. \]

Combining the limiting distributions of all terms \( \sum_{t=1}^{T} \tilde{F}_t \tilde{F}_t^\prime, \sum_{t=1}^{T} e_t \tilde{F}_t^\prime, \sum_{t=1}^{T} e_t (\hat{F}_t - F_t) \prime \) and \( \Lambda \sum_{t=1}^{T} (F_t - \hat{F}_t) \tilde{F}_t^\prime \) in equation (65), we obtain equation (43). ■

**Proof of Proposition 2.** Consider the estimation error

\[ \hat{F}_t - F_t = n^{-1} \tilde{A}' \tilde{z}_t - F_t = n^{-1} \tilde{A}' \tilde{F}_t + n^{-1} \tilde{A}' e_t - F_t = n^{-1} \tilde{A}' \tilde{F}_t + n^{-1} \tilde{A}' \left( \Lambda - \tilde{A} \right) F_t + n^{-1} \tilde{A}' e_t - F_t. \]
Since we know that, by construction, \( \hat{\Lambda}' \hat{\Lambda} = nI_k \), we have
\[
n^{-1} \hat{\Lambda}' z_t - F_t = n^{-1} \hat{\Lambda}' \left( \Lambda - \hat{\Lambda} \right) F_t + n^{-1} \hat{\Lambda}' e_t = I + II.
\]
As far as \( I \) is concerned, it holds that, omitting \( n^{-1} \) for the sake of brevity
\[
\max_{1 \leq t \leq T} \left\| \hat{\Lambda}' \left( \Lambda - \hat{\Lambda} \right) F_t \right\| \leq \left\| \hat{\Lambda}' \left( \Lambda - \hat{\Lambda} \right) \right\| \max_{1 \leq t \leq T} \| F_t \|;
\]
since
\[
\left\| \hat{\Lambda}' \left( \Lambda - \hat{\Lambda} \right) \right\| = O_p \left( T^{-1} \right)
\]
and
\[
\max_{1 \leq t \leq T} \| F_t \| = O_p \left( T^{1/2} \right),
\]
we get
\[
\max_{1 \leq t \leq T} \left\| \hat{\Lambda}' \left( \Lambda - \hat{\Lambda} \right) F_t \right\| = O_p \left( T^{-1/2} \right).
\]
Therefore \( I = O_p \left( T^{-1/2} \right) \) uniformly in \( t \). As per \( II \), we have
\[
\hat{\Lambda}' e_t = \Lambda' e_t + \left( \hat{\Lambda}' - \Lambda' \right) e_t \leq \max_{1 \leq t \leq T} \| \Lambda' e_t \| + \max_{1 \leq t \leq T} \left\| \left( \hat{\Lambda}' - \Lambda' \right) e_t \right\|
\leq \| \Lambda \| \max_{1 \leq t \leq T} \| e_t \| + \left\| \left( \hat{\Lambda}' - \Lambda' \right) \right\| \max_{1 \leq t \leq T} \| e_t \| = O_p \left( 1 \right) + O_p \left( T^{-1} \right) O_p \left( 1 \right).
\]
Hence, \( II = O_p \left( 1 \right) \). Thus we have
\[
\max_{1 \leq t \leq T} \left\| \hat{F}_t - F_t \right\| = O_p \left( 1 \right),
\]
which proves equation (44).

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

**Proof of Theorem 7.** Recall equation (29)
\[
\hat{\beta} - \beta = \left\{ \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_t \hat{W}_t' \right\}^{-1} \left\{ \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{W}_t \left( (W_t - \hat{W}_t)' \beta + u_{it} \right) \right\}.
\]
As far as the denominator of \( \hat{\beta} - \beta \) is concerned, we have
\[
\frac{1}{T^2} \sum_{t=1}^{T} \hat{W}_t \hat{W}_t' = \frac{1}{T^2} \sum_{t=1}^{T} W_t W_t' + o_p \left( 1 \right).
\]
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:

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

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

$$III = \sum_{t=1}^{T} \left[ n^{-1} \hat{\Lambda}' \hat{F}_t + n^{-1} \hat{\Lambda}' e_t - F_t \right] z_t' \hat{\Lambda} n^{-1}$$

$$= \sum_{t=1}^{T} \left[ n^{-1} \hat{\Lambda}' \hat{F}_t - n^{-1} \hat{\Lambda}' \left( \hat{\Lambda} - \Lambda \right) F_t + n^{-1} \hat{\Lambda}' e_t - F_t \right] z_t' \hat{\Lambda} n^{-1}$$

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

with

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

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

$$IV = n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \left[ \left( \Lambda - \hat{\Lambda} \right) F_t + e_t \right] \left[ \left( \Lambda - \hat{\Lambda} \right) F_t + e_t \right]' \hat{\Lambda}$$

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

$$+ n^{-2} \hat{\Lambda}' \left( \Lambda - \hat{\Lambda} \right) \sum_{t=1}^{T} F_t e_t' \hat{\Lambda} + n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} e_t F_t' \left( \Lambda - \hat{\Lambda} \right)' \hat{\Lambda}$$

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

51
with
\[ n^{-2} \hat{\Lambda}' \left( \Lambda - \hat{\Lambda} \right) \sum_{t=1}^{T} F_t' F_t \left( \Lambda - \hat{\Lambda} \right)' \hat{\Lambda} = O_p(1), \]
\[ n^{-2} \hat{\Lambda}' \left( \Lambda - \hat{\Lambda} \right) \sum_{t=1}^{T} F_t' \hat{e}_t \hat{\Lambda} = O_p(1), \]
and
\[ n^{-2} \hat{\Lambda}' \left( \sum_{t=1}^{T} e_t \hat{e}_t' \right) \hat{\Lambda} = O_p(T); \]
therefore, \( IV = O_p(T) \). Thus we get
\[ T^{-2} \sum_{t=1}^{T} \hat{F}_t \hat{F}_t' = T^{-2} \sum_{t=1}^{T} F_t' F_t' + O_p(T^{-1}). \]
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 \( T^{-2} \sum_{t=1}^{T} F_t F_t' \). This holds for any consistent estimator \( \hat{\Lambda} \) such that \( \hat{\Lambda} - \Lambda = O_p(T^{-\delta}) \); in such case, consistency would be ensured at a rate \( \min \{ 1, \delta \} \).

With respect to the numerator of equation (29), this is equal to
\[ \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{w}_t u_{it} + \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{w}_t \left( \tilde{w}_t \right)' \beta = I + II. \]
We have:
\[ I = \sum_{i=1}^{n} \sum_{t=1}^{T} W_t u_{it} + \sum_{i=1}^{n} \sum_{t=1}^{T} \left( \tilde{w}_t - W_t \right) u_{it} \]
\[ = \sum_{t=1}^{T} W_t \left( \sum_{i=1}^{n} u_{it} \right) + n^{-1} \hat{\Lambda}' \left( \Lambda - \hat{\Lambda} \right) \sum_{t=1}^{T} W_t \left( \sum_{i=1}^{n} u_{it} \right) + n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} e_t \left( \sum_{i=1}^{n} u_{it} \right), \]
with
\[ \sum_{t=1}^{T} W_t \left( \sum_{i=1}^{n} u_{it} \right) = O_p(T), \]
\[ n^{-1} \hat{\Lambda}' \left( \Lambda - \hat{\Lambda} \right) \sum_{t=1}^{T} W_t \left( \sum_{i=1}^{n} u_{it} \right) = O_p(T^{-1}) O_p(T) = O_p(1), \]
and
\[ n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} e_t \left( \sum_{i=1}^{n} u_{it} \right) = O_p(T^{1/2}), \]
which follows from Assumption 6. As far as \( II \) is concerned, we have

\[
II = n^{-1} \hat{\Lambda}' \sum_{i=1}^{n} \sum_{t=1}^{T} z_t \left( W_t - n^{-1} \hat{\Lambda}' \bar{z}_t \right) \beta
\]

\[
= n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} z_t \left[ W_t - n^{-1} \hat{\Lambda}' \hat{\Lambda} W_t + n^{-1} \hat{\Lambda}' \left( \hat{\Lambda} - \Lambda \right) W_t - n^{-1} \hat{\Lambda}' \bar{e}_t - n^{-1} \left( \hat{\Lambda} - \Lambda \right)' z_t \right] \beta
\]

\[
= -n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \bar{z}_t \bar{e}'_t \Lambda \beta + n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \bar{z}_t \bar{W}_t' \left( \hat{\Lambda} - \Lambda \right)' \hat{\Lambda} - n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \bar{z}_t \bar{W}_t' \left( \hat{\Lambda} - \Lambda \right) \beta
\]

\[
= O_p(T) + O_p(T^{-1}) O_p(T^2) + O_p(T^{-1}) O_p(T^2) = O_p(T).
\]

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

\[
\left[ \sum_{i=1}^{n} \sum_{t=1}^{T} W_i W_t' + o_p(1) \right]^{-1} \left\{ \sum_{i=1}^{n} \sum_{t=1}^{T} \bar{W}_t \left( W_t - \bar{W}_t \right)' \beta + \bar{u}_t \right\}
\]

\[
= O_p(T^{-2}) O_p(T) = O_p(T^{-1}).
\]

This proves equation (46).

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

\[
\sum_{i=1}^{n} \sum_{t=1}^{T} \bar{W}_t \left( W_t - \bar{W}_t \right)' \beta = -n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \bar{z}_t \bar{e}'_t \Lambda \beta - n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \bar{z}_t \bar{z}'_t \left( \hat{\Lambda} - \Lambda \right) \beta
\]

\[
+ n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \bar{z}_t \bar{W}_t' \left( \hat{\Lambda} - \Lambda \right)' \hat{\Lambda}
\]

\[
= I + II + III.
\]

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

\[
I = -n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \left( \Lambda W_t + \bar{e}_t \right) \bar{e}'_t \Lambda \beta
\]

\[
= -n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \Lambda W_t \bar{e}'_t \Lambda \beta - n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \bar{e}_t \bar{e}'_t \Lambda \beta
\]

\[
\Rightarrow -n^{-1} \int B_t dB_t \Lambda \beta - n^{-2} \Lambda' \Sigma_e \Lambda \beta.
\]

(66)
As far as II is concerned, recalling that $T \left( \tilde{\Lambda} - \Lambda \right) \Rightarrow D^1_\Lambda$, we have

\[
II = -n^{-2} \tilde{\Lambda}' \sum_{t=1}^{T} (\Lambda W_t + \hat{e}_t) (\Lambda W_t + \hat{e}_t)' \left( \tilde{\Lambda} - \Lambda \right) \beta \\
= -n^{-2} \tilde{\Lambda}' \sum_{t=1}^{T} \Lambda W_t W_t' \Lambda' (\tilde{\Lambda} - \Lambda) \beta - n^{-2} \tilde{\Lambda}' \sum_{t=1}^{T} \Lambda W_t e_t' (\tilde{\Lambda} - \Lambda) \beta \\
- n^{-2} \tilde{\Lambda}' \sum_{t=1}^{T} e_t W_t' \Lambda' (\tilde{\Lambda} - \Lambda) \beta - n^{-2} \tilde{\Lambda}' \sum_{t=1}^{T} e_t e_t' (\tilde{\Lambda} - \Lambda) \beta \\
\Rightarrow -n^{-1} \int B_x B_x' \Lambda' D^1_\Lambda \beta. \quad (67)
\]

Likewise

\[
III = n^{-2} \tilde{\Lambda}' \sum_{t=1}^{T} \hat{e}_t W_t' \left( \Lambda - \tilde{\Lambda} \right) \beta \\
\Rightarrow -n^{-1} \int B_x B_x' \Lambda' \beta. \quad (68)
\]

Thus, combining equations (66), (67) and (68) we have

\[
\sum_{i=1}^{n} \sum_{t=1}^{T} W_t \left( W_t - \hat{W}_t \right)' \beta \Rightarrow -n^{-1} \int B_x dB_x' \Lambda \beta - n^{-2} \Lambda' \Sigma_x \Lambda \beta + n^{-1} \int B_x B_x' [D^1_\Lambda - \Lambda' D^1_\Lambda] \beta.
\]

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

\[
\sum_{t=1}^{T} \hat{W}_t \left( \sum_{i=1}^{n} u_{it} \right) = n^{-1} \tilde{\Lambda}' \sum_{t=1}^{T} \hat{e}_t \left( \sum_{i=1}^{n} u_{it} \right) \\
= n^{-1} \tilde{\Lambda}' \sum_{t=1}^{T} \Lambda W_t \left( \sum_{i=1}^{n} u_{it} \right) + n^{-1} \tilde{\Lambda}' \sum_{t=1}^{T} \hat{e}_t \left( \sum_{i=1}^{n} u_{it} \right),
\]

which asymptotically leads to

\[
\frac{1}{T} n^{-1} \tilde{\Lambda}' A \sum_{t=1}^{T} W_t \left( \sum_{i=1}^{n} u_{it} \right) \Rightarrow \int B_x dB_x \left( \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij} \right)^{1/2}.
\]

This completes the proof of (47).

Finally, we consider the case when equation (1) is a spurious relationship.

Since $u_{it} \sim I(1)$, we have that

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

54
and
\[
\sum_{t=1}^{T} \tilde{W}_t u_{it} = \sum_{t=1}^{T} W_t u_{it} + n^{-1} \hat{\Lambda}' \left( \Lambda - \hat{\Lambda} \right) \sum_{t=1}^{T} W_t u_{it} + n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} e_t u_{it} = O_p(T^2) + O_p(T) + O_p(T),
\]
so that \( \sum_{t=1}^{T} \tilde{W}_t \left( (W_t - \tilde{W}_t)' \beta + u_{it} \right) = O_p(T^2) \), which proves (48).

In this case the limiting distribution of the numerator is given by the leading term \( \sum_{i=1}^{n} W_t (\sum_{i=1}^{n} u_{it}) \), so that the same result as in equation (55) holds, namely
\[
\frac{1}{nT^2} \sum_{t=1}^{T} W_t \left( \sum_{i=1}^{n} u_{it} \right) \Rightarrow \sqrt{\lambda} \left( \int B_x B_u \right).
\]
This proves equation (49).

**Proof of Theorem 8.** Consider equation (30)
\[
\beta^{FD} - \beta = \left[ \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_t \Delta \hat{F}'_t \right]^{-1} \left\{ \sum_{i=1}^{n} \sum_{t=1}^{T} \Delta \hat{F}_t \left[ (\Delta F_t - \Delta \hat{F}_t)' \beta + \Delta u_{it} \right] \right\}.
\]
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 \( \beta^{FD} - \beta \) is concerned, we have
\[
\sum_{t=1}^{T} \Delta \hat{F}_t \left( \Delta F_t - \Delta \hat{F}_t \right)' = n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \Delta z_t \left[ \Delta F'_t \left( \hat{\Lambda} - \hat{\Lambda}' \right) - \Delta \hat{F}'_t \right] \hat{\Lambda}
\]
\[
= n^{-2} \hat{\Lambda}' \left[ \sum_{t=1}^{T} \Delta z_t \Delta F'_t \right] \left( \hat{\Lambda} - \hat{\Lambda}' \right) \hat{\Lambda} - n^{-2} \hat{\Lambda}' \left[ \sum_{t=1}^{T} \Delta z_t \Delta \hat{F}'_t \right] \hat{\Lambda}
\]
\[
= O_p(T) O_p(T^{-1}) + O_p(T) = O_p(T).
\]
Also we have
\[
\sum_{t=1}^{T} \Delta \hat{F}_t \Delta u_{it} = \sum_{t=1}^{T} \Delta F_t \Delta u_{it} + \sum_{t=1}^{T} \left( \Delta \hat{F}_t - \Delta F_t \right) \Delta u_{it}
\]
\[
= O_p(\sqrt{T}) + O_p(\sqrt{T}) = O_p(\sqrt{T}).
\]
This proves (50).
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

\[
p \lim \frac{1}{T} \sum_{t=1}^{T} \Delta z_t \Delta z_t' = \Sigma_{zz},
\]

we have

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

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:

\[
\frac{1}{T} \sum_{t=1}^{T} \Delta \hat{F}_t \left( \Delta F_t - \Delta \hat{F}_t \right)' \beta = \frac{1}{T} n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} \Delta z_t \left( \Delta F_t' - n^{-1} \Delta z_t' \hat{\Lambda}' \right) \beta
\]

\[
= \frac{1}{T} n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} \Delta \Delta F_t \Delta F_t' \beta + \frac{1}{T} n^{-1} \hat{\Lambda}' \sum_{t=1}^{T} \Delta \epsilon_t \Delta F_t' \beta - \frac{1}{T} n^{-2} \hat{\Lambda}' \sum_{t=1}^{T} \Delta z_t \Delta z_t' \hat{\Lambda}' \beta,
\]

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

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

and

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

we have

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

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


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


