# CENTRE FOR ECONOMETRIC ANALYSIS
CEA@Cass

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

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

---

***On Bootstrapping Panel Factor Series***

*Lorenzo Trapani*

---

CEA@Cass Working Paper Series

WP–CEA–04-2011

# On Bootstrapping Panel Factor Series

Lorenzo Trapani*

Cass Business School

October 16, 2011

**Abstract**

This paper studies the asymptotic validity of the bootstrap for nonstationary panel factor series. The analysis assumes a linear process for serial dependence, and sieve bootstrap is proposed to approximate the autocorrelation structure of the processes involved in the model. Two main results are shown. Firstly, a bootstrap Invariance Principle is derived pointwise in $i$, obtaining an upper bound for the order of truncation of the AR polynomial that depends on $n$ and $T$. Consistent estimation of the long run variances is also studied for $(n, T) \rightarrow \infty$. Secondly, joint bootstrap asymptotics is also studied, investigating the conditions under which the bootstrap is valid. Particularly, the extent of cross sectional dependence which can be allowed for is investigated, showing that, in the presence of strong cross dependence, consistent estimation of the long run variance (and therefore validity of the bootstrap) is no longer possible. The paper also considers extensions to the case of a mixture of stationary and nonstationary common factors.

**JEL codes**: C23.

**Keywords**: bootstrap, invariance principle, factor series, Vector AutoRegression, joint asymptotics.

*Cass Business School, Faculty of Finance, 106 Bunhill Row, London EC1Y 8TZ. Tel.: +44 (0) 207 0405260; email: L.Trapani@city.ac.uk. This paper is based on a working paper circulated under the title "Sieve Bootstrap for Large Nonstationary Panel Factor Models". I wish to thank the participants to the 3rd New York Camp Econometrics (Syracuse University, April 4-6, 2008), to the 3rd Italian Congress of Econometrics and Empirical Economics (Universita' Politecnica delle Marche, Ancona, January 30-31, 2009), and to the 4th CSDA International Conference on Computational and Financial Econometrics (CFE'10, University of London, December 10-12, 2010) for valuable feedback; and the participants to Cass Finance Seminar Series and to the Economics Department Seminars at the University of Pavia. The usual disclaimer applies.

# 1 Introduction

In recent years, factor models have achieved great popularity in applied econometrics and statistics. Panel factor series have been extensively used in macroeconometrics to model parsimoniously the presence of cross sectional correlation, e.g. in the analysis of business cycle models (Forni and Reichlin, 1998). Many other applications are possible - see e.g. the discussion in Bai (2004), Bai and Ng (2006a, 2006b, 2010), and the references therein. Nonstationary panel factor series have also been paid noticeable attention in applied statistics, where Lee and Carter's (1992) model for mortality forecasting has generated a huge body of literature (see Girosi and King, 2007).

The literature has recently produced significant developments in the inferential theory. Joint asymptotic theory for $(n, T) \to \infty$ has been studied for the case of stationary and nonstationary data, allowing for serial and cross sectional dependence and heterogeneity - see, *inter alia*, Bai (2003, 2004) and Bai and Ng (2002, 2004). The main focus of this paper is to complement the existing asymptotic theory, by investigating the validity of the bootstrap for nonstationary panel factor series defined as

$$x_{it} = \lambda_i' F_t + u_{it}, \tag{1}$$

with $i = 1, ..., n$ and $t = 1, ..., T$ and

$$F_t = F_{t-1} + \varepsilon_t. \tag{2}$$

Model (1) is a standard nonstationary panel factor model - we also refer to Bai (2004) and the references therein for a discussion of the various possible applications. Bootstrapping (1) could prove useful for at least three reasons. Firstly, as the theory developed in Bai (2004) and Kao, Trapani and Urga (2011) shows, the asymptotics heavily depends on nuisance parameters, and the bootstrap could help deal with this. Moreover, limiting distributions are often complicated and depend on somewhat arbitrary assumptions on the relative speed of divergence of $n$ and $T$. Finally, the common factors $F_t$ are often not observable and need to be estimated, thereby adding a further component to the error term $u_{it}$ in (1). In light of this, and in order to accommodate for serial dependence, this article proposes a sieve bootstrap algorithm (Bühlmann, 1997), building on the theory developed by Park (2002, 2003) and Chang, Park and Song (2006). Whilst this paper moves from a similar research question, namely to show an Invariance Principle (henceforth, IP) for the bootstrap counterpart to $x_{it}$, proving an IP for nonstationary factor models is a different type of exercise to the pure time series case studied by Park (2002) and, in a cointegration framework, by Chang, Park and Song (2006). This is due to two distinctive features of model (1): *(a)* the presence of the latent variables $F_t$, which are replaced by generated regressors, thereby affecting the asymptotics and the bootstrap asymptotics, and *(b)* the fact that the asymptotics, in this framework, depends jointly on two indices, $n$ and $T$.

This article makes two main contributions. In the first part of the paper (Sections 3 and 4), a bootstrap IP is derived and applied to the estimation of loadings, common factors and common components. The resampling algorithm is based on extracting the common factors from (1) by using the Principal Components estimator (PC) and thereafter fitting a Vector AutoRegression (VAR) of order $q$ to the estimated common factors and to the residuals. Two ancillary technical contributions of this section are the asymptotic theory for $\Delta\hat{F}_t = \hat{F}_t - \hat{F}_{t-1}$, and the derivation of an upper bound for $q$, which depends on both $n$ and $T$. These results are based on a "one cross sectional unit at a time" resampling algorithm, and therefore are only pointwise in $i$. Thus, cross sectional dependence among the $u_{it}$s is not taken into account. This is useful in some applications of (1), e.g. when applying bootstrap to approximate the limiting distribution of the estimated $\lambda_i$s, or when cross sectional dependence is negligible - see Section 4 for discussion. In the second part of the paper (Section 5), joint bootstrap asymptotics as $(n, T) \to \infty$ is developed, to also accommodate for the possible presence of cross dependence in the $u_{it}$s. We show that the estimation of the long run variance matrix of the $u_{it}$s is fraught with difficulties, due to its high dimension (growing with $n$). Section 5 contains a negative result, highlighting that consistent estimation of long run covariance matrices is not possible in this context, unless there is very little cross dependence. Finally, the paper also provides some initial results for the extension of bootstrap theory to the case of a mixture of stationary and nonstationary common factors in (1).

The paper is organised as follows. Section 2 introduces the model and discusses the main assumptions. Section 3 contains the resampling algorithm and the relevant asymptotics. Applications to univariate and multivariate problems are in Sections 4 and 5 respectively. The extension to a mixture of stationary and nonstationary common factors in in Section 6. Section 7 concludes. Preliminary lemmas and proofs are in Appendix A and B respectively.

NOTATION Throughout the paper, $\|A\|_p$ denotes the $L_p$-norm of a matrix $A$, i.e. $\max_x \|Ax\|_p / \|x\|_p$ (the Euclidean norm being defined simply as $\|A\|$), "$i_m$" indicates a unit column vector of dimension $m$, "$\to$" the ordinary limit, "$\xrightarrow{d}$" weak convergence, "$\xrightarrow{p}$" convergence in probability, "a.s." stands for "almost surely"; generic finite constants that do not depend on $n$ or $T$ are referred to as $M$. Stochastic processes such as $W(s)$ on $[0,1]$ are usually written as $W$, integrals such as $\int_0^1 W(s)\,ds$ as $\int W$ and stochastic integrals such as $\int_0^1 W(s)\,dW(s)$ as $\int WdW$.

Also, we extensively use the following notation: $\delta_{nT} = \min\left\{\sqrt{n}, \sqrt{T}\right\}$, $C_{nT} = \min\{\sqrt{n}, T\}$, $\varphi_{nT}^F = \min\left\{n, \sqrt{T/\log T}\right\}$ and $\varphi_{nT}^u = \min\left\{\sqrt{n}, \sqrt{T/\log T}\right\}$.

## 2 Model and assumptions

Consider model (1) and the data generating process of $F_t$

$$
\begin{aligned}
x_{it} &= \lambda_i' F_t + u_{it}, \\
F_t &= F_{t-1} + \varepsilon_t,
\end{aligned}
$$

where we assume that the (unobservable) factors $F_t$ are a $k$-dimensional process. We refer to Bai (2004) for the estimation of $k$. All the theory is studied for $(n,T) \to \infty$ jointly - see Phillips and Moon (1999) for definition and discussion.

Consider the following assumptions:

**Assumption 1:** *(time series and cross sectional properties of $u_{it}$)* let $u_t = [u_{1t}, ..., u_{nt}]'$; then $u_t$ admits an invertible $MA(\infty)$ representation

$$
u_t = \Gamma(L)\, e_t^u = \sum_{j=0}^{\infty} \Gamma_j e_{t-j}^u,
$$

where *(i)* $e_t^u$ is i.i.d. across $t$ with $E[e_t^u] = 0$, $E[e_t^u e_t^{u\prime}] = \Sigma_u$; also, letting $e_{it}^u$ be the

$i$-th element of $e_t^u$, $\max_{i,t} E\left|e_{it}^u\right|^8 < \infty$; *(ii)* $\sum_{j=0}^{\infty} \Gamma_j L^j \neq 0$ for all $|L| \leq 1$ and, letting $\Gamma_{i,j}$ be the $i$-th row of $\Gamma_j$, $\max_i \sum_{j=0}^{\infty} j^s \left\|\Gamma_{i,j}\right\| < \infty$ for some $s \geq 1$; *(iii) (cross sectional dependence)* (a) $\left\|\Gamma(1)\right\|_1 \leq M$, $\left\|\Gamma^{-1}(1)\right\|_1 \leq M$, $\left\|\Gamma^{-1}(1)\right\|_{\infty} \leq M$ and $\left\|\Sigma_u\right\|_1 \leq M$; (b) $E\left|n^{-1/2} \sum_{i=1}^{n} \left[u_{is} u_{it} - E\left(u_{is} u_{it}\right)\right]\right|^4 \leq M$ for every $(t, s)$; *(iv) (initial conditions)* $E\left|u_{i0}\right|^4 \leq M$ for all $i$.

**Assumption 2:** *(time series properties of $\varepsilon_t$)* $\varepsilon_t$ is a $k$-dimensional vector random process (with finite $k$) and it admits an invertible $MA(\infty)$ representation where $\varepsilon_t = \alpha(L) e_t^F = \sum_{j=0}^{\infty} \alpha_j L^j e_{t-j}^F$ with *(i)* $e_t^F$ is i.i.d. with $E\left[e_t^F\right] = 0$, $E\left[e_t^F e_t^{F\prime}\right] = \Sigma_e$ and $E\left\|e_t^F\right\|^r < \infty$ for some $r > 4$; *(ii)* $\sum_{j=0}^{\infty} \alpha_j L^j \neq 0$ for all $|L| \leq 1$ and $\sum_{j=0}^{\infty} j^s \left\|\alpha_j\right\| < \infty$ for some $s \geq 1$; *(iii)* the matrix $\Sigma_{\Delta F} = \sum_{j=0}^{\infty} \alpha_j \Sigma_e \alpha_j'$ is positive definite; *(iv) (initial conditions)* $E\left\|F_0\right\|^4 \leq M$.

**Assumption 3:** *(identifiability)* the loadings $\lambda_i$ are *(i)* either nonrandom quantities such that $\left\|\lambda_i\right\| \leq M$, or random quantities such that $E\left\|\lambda_i\right\|^4 < \infty$; *(ii)* 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 $n \to \infty$ with $\Sigma_\Lambda$ positive definite; *(iii)* the eigenvalues of the matrix $\Sigma_\Lambda^{1/2} \Sigma_{\Delta F} \Sigma_\Lambda^{1/2}$ are distinct, and the eigenvalues of the stochastic matrix $\Sigma_\Lambda^{1/2} \left(T^{-2} \sum_{t=1}^{T} F_t F_t'\right) \Sigma_\Lambda^{1/2}$ are a.s. distinct as $T \to \infty$.

**Assumption 4:** *(i)* $\{\varepsilon_t\}$, $\{u_{it}\}$ and $\{\lambda_i\}$ are three mutually independent groups; *(ii)* $F_0$ is independent of $\{u_{it}\}$ and $\{\varepsilon_t\}$.

Parts *(i)* and *(ii)* of Assumption 1 allow to establish an IP for the of the bootstrap value from the general linear process $u_{it}$. Part *(i)* is slightly more stringent than Assumption 3.1 in Park (2002, p. 474), where the existence of the fourth moment suffices. In this context, assuming $r > 4$ is needed for the validity of inferential theory for factor models; see also Assumption C(1) in Bai (2004). Part *(ii)* of the assumption is needed in order to approximate the $AR(\infty)$ polynomial with a finite autoregressive representation - see e.g. Hannan and Kavalieris (1986). Letting $E\left(u_{it} u_{jt}\right) = \tau_{ij}$, part *(iii)* entails that $\sum_{i=1}^{n} \left|\tau_{ij}\right| \leq M$ for all $j$, since $E\left(u_t u_t'\right) = \Gamma(1) \Sigma_u \Gamma'(1)$ and $\left\|E\left(u_t u_t'\right)\right\|_1 \leq \left\|\Gamma(1)\right\|_1^2 \left\|\Sigma_u\right\|_1$. Similar requirements on the (weak) cross dependence of the error term are in Pesaran and Tosetti (2011) and Chudik and Pesaran (2011). That $\left\|\Gamma^{-1}(1)\right\|_1$ be finite could be derived in principle from more primitive assumptions on $\Gamma(1)$ - see e.g. Kolotilina (2009). Part *(iii)* allows for some cross sectional dependence in the error term $u_{it}$; part *(iii)*(b) is the same as Assumption C(4) in Bai (2004). Note that parts *(i)*-*(iii)* entail that $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}\left|\gamma_{s-t}\right| \leq M$, where $\gamma_{s-t} = n^{-1}\sum_{i=1}^{n}\gamma_{i,s-t}$ and $\gamma_{i,s-t} = E\left(u_{it}u_{is}\right)$, which is part (2) of Assumption C in Bai (2004, p. 141).

Assumption 2 mimics Assumption A in Bai (2004) and is required in order for the dimension of the factor space to be estimated consistently, and also to derive the asymptotic theory for the estimated factors. Part *(i)* is enough for both purposes, and it is equivalent to Assumption 3.1(a) in Park (2002, p. 474); part *(ii)* plays the same role as Assumption 1*(ii)*. Note that part *(iii)* rules out cointegration among the $F_t$s, which is the same as Assumption A(2) in Bai (2004, p. 140). Also, Assumption 2 entails a Law of the Iterated Logarithm for $F_t$ (see Phillips and Solo, 1992, Theorem 3.3) to hold, whence $\liminf_{T\to\infty}$ $(\log\log T)$ $T^{-2}\sum_{t=1}^{T}F_tF_t' = D$ where $D$ is a nonrandom positive definite matrix; this is part (3) of Assumption 2 in Bai (2004). Assumptions 3 and 4 are standard requirements for the asymptotics of the estimates of $\lambda_i$ and $F_t$. See Bai (2004) for further details.

PC based inference on $\lambda_i$ and $F_t$ is studied in Bai (2004). The common factors $F_t$ are estimated by $\hat{F}_t$ under the restriction that $T^{-2}\sum_{t=1}^{T}\hat{F}_t\hat{F}_t' = I_k$. The estimated common factor $\hat{F}_t$ is $T$ times the eigenvectors corresponding to the $k$ largest eigenvalues of matrix $XX'$ where $X = [x_1, ..., x_n]'$ with $x_i = [x_{i1}, ..., x_{iT}]'$. Then $\lambda_i$ can be estimated applying OLS to

$$x_{it} = \lambda_i'\hat{F}_t + v_{it}, \tag{3}$$

thus $\hat{\lambda}_i = \left[\sum_{t=1}^{T}\hat{F}_t\hat{F}_t'\right]^{-1}\left[\sum_{t=1}^{T}\hat{F}_t x_{it}\right]$. It is well known that $\lambda_i$ and $F_t$ are identifiable only up to a transformation. Therefore, PC estimates the space spanned by the factors $F_t$ (and the loadings $\lambda_i$), thereby finding $H'F_t$ instead of $F_t$ and $H^{-1}\lambda_i$ instead of $\lambda_i$, where $H$ is a $k \times k$ invertible matrix given by

$$H = \left(\frac{1}{n}\sum_{i=1}^{n}\lambda_i\lambda_i'\right)\left(\frac{1}{T^2}\sum_{t=1}^{T}F_t\hat{F}_t'\right)V_{nT}^{-1}, \tag{4}$$

with $V_{nT}$ a $k \times k$ diagonal matrix containing the eigenvalues of $\frac{1}{nT^2}XX'$ in descending order. The effect of replacing the true, unobservable factors $F_t$ with their estimates $\hat{F}_t$ is to inflate the error term $u_{it}$ in (1):

$$v_{it} = u_{it} + \lambda_i' \left( H' F_t - \hat{F}_t \right). \tag{5}$$

Consider the following notation, which is henceforth used throughout the paper. Let $W_\varepsilon$ be a $k$-dimensional Brownian motion with covariance matrix $\Sigma_{\Delta F}$; also, $W_{u,i}$ denotes a scalar Brownian motion independent of $W_\varepsilon$ with variance $\sigma_{u,i}^2 = \Gamma_i(1) \Sigma_u \Gamma_i'(1)$, where $\Gamma_i(1) = \sum_{j=0}^{\infty} \Gamma_i$ is the $i$-th row of $\Gamma(1)$.

**Proposition 1** *Let Assumptions 1-4 hold. As $T \to \infty$ for every $i$*

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor Ts \rfloor} \begin{bmatrix} \Delta \hat{F}_t \\ u_{it} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} H' W_\varepsilon(s) \\ W_{u,i}(s) \end{bmatrix}, \tag{6}$$

*uniformly in $s$. Also*

$$\frac{1}{T^2} \sum_{t=1}^{T} \hat{F}_t \hat{F}_t' \xrightarrow{d} H' \left( \int W_\varepsilon W_\varepsilon' \right) H, \tag{7}$$

$$\frac{1}{T} \sum_{t=1}^{T} \hat{F}_t u_{it} \xrightarrow{d} H' \int W_\varepsilon dW_{u,i}. \tag{8}$$

*As $(n, T) \to \infty$ with $\frac{n}{T} \to 0$*

$$\frac{1}{\sqrt{nT}} \sum_{t=1}^{\lfloor Ts \rfloor} \sum_{i=1}^{n} u_{it} \xrightarrow{d} \sigma_u W_u(s), \tag{9}$$

$$\frac{1}{\sqrt{nT}} \sum_{t=1}^{T} \sum_{i=1}^{n} \hat{F}_t u_{it} \xrightarrow{d} \sigma_u H' \int W_\varepsilon dW_u, \tag{10}$$

*uniformly in $s$, where $W_u$ is a standard Brownian motion independent of $W_\varepsilon$ and $\sigma_u = \lim_{n\to\infty} n^{-1} i_n' \Gamma(1) \Sigma_u \Gamma'(1) i_n$.*

Proposition 1 contains two types of results: equations (6), (7) and (8), which are univariate, pointwise in $i$, and (9) and (10), which are joint limits. These results, used in conjunction with the Continuous Mapping Theorem (CMT), are the building blocks to prove the validity of bootstrap approximations.

In the remainder of the paper, we show bootstrap analogues to (6)-(8) - Sections 3 and 4. We also show bootstrap counterparts to (9) and (10) - Section 5.

# 3 Univariate sieve bootstrap: algorithm and IP

This section contains the algorithm to generate the bootstrap sample using a "one cross sectional unit at a time" resampling scheme. Asymptotic theory (pointwise in $i$) is reported in Section 3.2. The main output of this section are bootstrap analogues to (6)-(8).

Since (1) is a cointegrating regression, one may apply the algorithm of Chang, Park and Song (2006) to its observable counterpart (3). This would impose a unit root in the bootstrap counterpart to $\hat{F}_t$, which is needed in order for the bootstrap to be consistent - see Park (2003). Henceforth, we define $\xi_{it} = [\Delta F_t', u_{it}]'$, with $\xi_{it} = \sum_{j=1}^{\infty} \beta_{ij} \xi_{it-j} + e_{it}$, also denoting $1 - \sum_{j=1}^{\infty} \beta_{ij}$ as $\beta_i(1)$.

## 3.1 The generation of the bootstrap sample

The presence of serial dependence in $\Delta F_t$ and $u_{it}$ requires a bootstrap algorithm that preserves the autocorrelation structure over time. This can be accomplished by approximating the infinite $AR$ polynomials $\alpha(L)$ and $\Gamma(L)$ by truncating them at lags $q_F$ and $q_{u,i}$ respectively:

$$\Delta F_t = \sum_{j=1}^{q_F} \alpha_{q,j} \Delta F_{t-j} + e_{t,q}^F, \tag{11}$$

$$u_{it} = \sum_{j=1}^{q_{u,i}} \gamma_{q,j}^{(i)} u_{it-j} + e_{it,q}^u. \tag{12}$$

The values of $q_F$ and $q_{u,i}$ depend on $n$ and $T$, as discussed in the following assumption.

**Assumption 5:** As $(n, T) \to \infty$, $q_F \to \infty$ and $q_{u,i} \to \infty$ for each $i$, with $q_F = o\left(\varphi_{nT}^F\right)$ and $q_{u,i} = o\left(\varphi_{nT}^u\right)$ for each $i$.

Assumption 5 contains an upper bound on $q_F$ and $q_{u,i}$. In order for $q_F$ and $q_{u,i}$ to pass to infinity, it is necessary that $(n, T) \to \infty$; no assumptions are needed on the relative speed of divergence of $n$ and $T$. No lower bounds are required for $q_F$ and $q_{u,i}$, as long as they pass to infinity. Using Assumption 5, one could think of selecting $q_F$ and $q_{u,i}$ by using some information criteria such as e.g. AIC or BIC, under the restriction that the maximum lag allowed for be of order $o\left(\varphi_{nT}^{F}\right)$ and $o\left(\varphi_{nT}^{u}\right)$ respectively.

We propose an algorithm similar to Chang, Park and Song (2006) for the case of a cointegration regression (where $n$ is fixed). The main differences here are the presence of unobservable variables and the double-indexed asymptotics.

**Step 1.** (PC estimation)

(1.1) Estimate $\lambda_i$ and $F_t$ in (1) using the PC estimator.

(1.2) Generate the residuals $\hat{u}_{it} = x_{it} - \hat{\lambda}_i'\hat{F}_t$ and define $\hat{\xi}_{it} = \left[\Delta\hat{F}_t', \hat{u}_{it}\right]'$.

**Step 2.** (estimation)

(2.1) Estimate $\alpha_{q,j}$ and $\gamma_{q,j}^{(i)}$ (obtaining $\hat{\alpha}_{q,j}$ and $\hat{\gamma}_{q,j}^{(i)}$ respectively) by applying OLS (or some other estimator, e.g. the Yule-Walker estimator) to $\Delta\hat{F}_t = \sum_{j=1}^{q_F} \alpha_{q,j}\ \Delta\hat{F}_{t-j} + e_{t,q}^{F}$ and $\hat{u}_{it} = \sum_{j=1}^{q_{u,i}} \gamma_{q,j}^{(i)}\hat{u}_{it-j} + e_{it,q}^{u}$.

(2.2) Compute the residuals $\hat{e}_{t,q}^{F} = \Delta\hat{F}_t - \sum_{j=1}^{q_F} \hat{\alpha}_{q,j}\ \Delta\hat{F}_{t-j}$ and $\hat{e}_{it,q}^{u} = \hat{u}_{it} - \sum_{j=1}^{q_{u,i}} \hat{\gamma}_{q,j}^{(i)}\hat{u}_{it-j}$. Define $\hat{e}_{it,q} = \left[\hat{e}_{t,b}^{F\prime}, \hat{e}_{it,q}^{u}\right]'$.

**Step 3.** (bootstrap) for $b = 1, ..., \beth$ iterations

(3.1) (resampling)

(3.1.a) Center the residuals $\hat{e}_{it,q}$ around their mean, as $\bar{e}_{it,q} = \hat{e}_{it,q} - T^{-1}\sum_{t=1}^{T} \hat{e}_{it,q}$.

(3.1.b) Draw (with replacement) $T$ values from $\{\bar{e}_{it,q}\}_{t=1}^{T}$ to obtain the bootstrap sample $\{e_{it,b}\}_{t=1}^{T}$, defining also $e_{it,b} = \left[e_{t,b}^{F\prime}, e_{it,b}^{u}\right]'$.

(3.2) (generation of the bootstrap sample)

(3.2.a) Generate recursively the pseudo sample $\xi_{it,b} = \left[\Delta F_{t,b}', u_{it,b}\right]'$ as $\Delta F_{t,b} = \sum_{j=1}^{q_F} \hat{\alpha}_{q,j}$ $\Delta F_{t-j,b} + e_{t,b}^{F}$ and $u_{it,b} = \sum_{j=1}^{q_{u,i}} \hat{\gamma}_{q,j}^{(i)} u_{it-j,b} + e_{it,b}^{u}$, using as initialization $\left\{\xi_{iq,b}, ..., \xi_{i1,b}\right\} = \left\{\xi_{iq}, ..., \xi_{i1}\right\}$.

(3.2.b) Generate $F_{t,b}$ as $F_{t,b} = F_{0,b} + \sum_{j=1}^{t} \Delta F_{j,b}$, with initialization is $F_{0,b} = \hat{F}_0$, or alternatively $T^{-1}\sum_{t=1}^{T} \hat{F}_t$.

(3.2.c) Generate the pseudo sample $\{x_{it,b}\}_{t=1}^{T}$.

Consider Step 2.1. Since $\Delta \hat{F}_t$ estimates $\Delta F_t$ up to a rotation, $\hat{\alpha}_{q,j}$ estimates a rotation of $\alpha_j$; this however suffices for our purposes (see Lemma 2 below).

As a comment to Step 3.2.c, te possible schemes to generate $x_{it,b}$ are discussed in Section 4. The output of the algorithm above is therefore the bootstrap sample $\left\{\xi_{it,b}\right\}_{t=1}^{T}$. In the next section, an IP for $\left\{\xi_{it,b}\right\}_{t=1}^{T}$ is shown.

### 3.2 Bootstrap asymptotics

Based on a typical approach to prove the validity of the bootstrap, the main purpose of this section is to show that $T^{-1/2}\sum_{t=1}^{\lfloor Ts \rfloor}\xi_{it,b}$ converges (in probability) to the same limit as $T^{-1/2}\sum_{t=1}^{\lfloor Ts \rfloor}\xi_{it}$, uniformly in $s$.

Define the partial sums of $e_{it} = \left[e_t^{F\prime}, e_{it}^{u}\right]'$ as $W_{iT}(s) = T^{-1/2}\sum_{t=1}^{\lfloor Ts \rfloor} e_{it}$. Assumptions 1 and 2 ensure that an IP holds whereby $W_{iT}(s) \xrightarrow{d} W_i(s)$ where $W_i(s)$ is a $(k+1)$-dimensional Brownian motion. This convergence is in the weak form, and it holds in the space of *cadlag* functions $D[0,1]$ endowed with the supremum norm. Weak convergence can be strengthened by defining, on the probability space $(\Omega, \mathcal{F}, P)$, a copy of $W_{iT}(s)$, say $W_{iT}'(s)$, which has the same distribution as $W_{iT}(s)$ and can be chosen such that

$$P\left\{\sup_{0\leq s\leq 1}\left\|W_{iT}'(s) - W_i(s)\right\| \geq \delta\right\} \leq MT^{1-r/2}E\left\|e_{it}\right\|^r, \tag{13}$$

where $\delta > 0$, $r > 2$ and $M$ depends only on $r$. Such results are known as "strong (weak) approximations" (see e.g. Sakhanenko, 1980) and they ensure that $W_{iT}'(s)$ converges a.s. (in probability) to $W_i(s)$. Assumptions 1 and 2, where $r > 4$, entail that (13) holds. In essence, (13) states that, as long as $T^{1-r/2}E\left\|e_{it}\right\|^r \to 0$ either in probability or a.s. for some $r > 2$, an IP holds (in probability or a.s. respectively).

Consider the bootstrap sample $\{e_{it,b}\}_{t=1}^{T}$. This is an i.i.d. sample conditional on $\{\hat{e}_{it}\}_{t=1}^{T}$, on the probability space induced by the bootstrap, say $\left(\Omega^b, \mathcal{F}^b, P^b\right)$. Henceforth, we denote convergence in probability and in distribution in the bootstrap space (with respect to $P^b$) as "$\xrightarrow{p^b}$" and "$\xrightarrow{d^b}$" respectively.

Moments existence for the bootstrap sample is granted by the following Lemma.

**Lemma 1** *Let Assumptions 1-5 hold. As* $(n, T) \to \infty$*, for all* $(i, t)$ *and* $r > 4$

$$
\begin{aligned}
E^b \left\| e_{t,b}^F \right\|^r &= E \left\| e_t^F \right\|^r + O_p\left(q_F^{-rs}\right) + O_p\left(C_{nT}^{-r}\right) + O_p\left[\left(\frac{q_F}{\varphi_{nT}^F}\right)^r\right] && (14) \\
\max_{i,t} E^b \left| e_{it,b}^u \right|^r &= \max_{i,t} E \left| e_{it}^u \right|^r + O_p\left(q_{u,i}^{-rs}\right) + O_p\left(\delta_{nT}^{-r}\right) + O_p\left[\left(\frac{q_{u,i}}{\varphi_{nT}^u}\right)^r\right]. && (15)
\end{aligned}
$$

This result is useful to prove an IP for $e_{it,b}$ using (13). The type of IP that we are able to prove is in the weak form, since (14) and (15) hold in probability. Note that having $q_F$, $q_{u,i} \to \infty$ with upper bounds $\varphi_{nT}^F$ and $\varphi_{nT}^u$ is necessary for the higher order moments of the bootstrap sample to converge to the population values.

Lemma 1 and (13) yield $T^{-1/2} \sum_{t=1}^{\lfloor Ts \rfloor} e_{it,b} \xrightarrow{d^b} W_i(s)$ in $P$. In order to extend this result to the bootstrap sample $\left\{\xi_{it,b}\right\}_{t=1}^T$ generated in Step 3.2*(a)* above, we need the following result as well.

**Lemma 2** *Let Assumptions 1-5 hold. As* $(n, T) \to \infty$*, for all* $i$

$$
\begin{aligned}
\max_{1 \le j \le q_F} \left\| \hat{\alpha}_{q,j} - H' \alpha_j \left(H'\right)^{-1} \right\| &= O_p\left(\sqrt{\frac{\log T}{T}}\right) + O_p\left(\frac{1}{n}\right) + o_p\left(\frac{1}{q_F^s}\right), && (16) \\
\max_{1 \le j \le q_{u,i}} \left| \hat{\gamma}_{q,j}^{(i)} - \gamma_j^{(i)} \right| &= O_p\left(\sqrt{\frac{\log T}{T}}\right) + O_p\left(\frac{1}{\sqrt{n}}\right) + o_p\left(\frac{1}{q_{u,i}^s}\right). && (17)
\end{aligned}
$$

Lemma 2 states that $\hat{\beta}_{q,j}^{(i)}$ is a consistent estimator of the space spanned by $\beta_{ij}$; the issue of identifying $F_t$ affects the estimation of the $\alpha_j$s, which are estimated up to a rotation. The rate $O_p\left(\sqrt{\log T / T}\right)$ is a well-known result in time series analysis (see e.g. Theorem 2.1 in Hannan and Kavalieris, 1986). The rates $O_p(1/n)$ and $O_p\left(\sqrt{1/n}\right)$ are due to the use of generated regressors, $\hat{F}_t$ and $\Delta \hat{F}_t$.

Using Lemmas 1 and 2, a bootstrap IP for $T^{-1/2} \sum_{t=1}^{\lfloor Ts \rfloor} \xi_{it,b}$ can be proved. Consider the partial sums of $\xi_{it}$, $W_{\xi iT}(s) = T^{-1/2} \sum_{t=1}^{\lfloor Ts \rfloor} \xi_{it}$. Assumptions 1 and 2 entail $W_{\xi iT}(s) \xrightarrow{d} W_{\xi i}(s) = \beta_i^{-1}(1) W_i(s)$. Proving the bootstrap IP requires showing that $W_{\xi iT,b}(s) = T^{-1/2} \sum_{t=1}^{\lfloor Ts \rfloor} \xi_{it,b} \xrightarrow{d^b} W_{\xi i}(s)$ as $(n, T) \to \infty$. This can be done by noting that, using the Beveridge-Nelson decomposition

$$
\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor Ts \rfloor} \xi_{it,b} = \hat{\beta}_{i,q}^{-1}(1) \left[\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor Ts \rfloor} e_{it,b}\right] + \frac{\hat{\beta}_{i,q}^{-1}(1)}{\sqrt{T}} \left[\bar{\xi}_{i0,b} - \bar{\xi}_{i\lfloor Ts \rfloor,b}\right], \qquad (18)
$$

where $\bar{\xi}_{it,b} = \sum_{j=1}^{q} \left[ \sum_{k=j}^{q} \hat{\beta}_{q,k}^{(i)} \right] \xi_{it-j+1,b}$. It holds that:

**Lemma 3** *Let Assumptions 1-5 hold. As* $(n, T) \to \infty$*, it holds that* $\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor Ts \rfloor} \xi_{it,b} \stackrel{d^b}{\to} W_{\xi i}(s)$ *in* $P$*, for all* $i$*.*

Lemma 3 entails that the partial sums of the bootstrap process $\left\{ \xi_{it,b} \right\}_{t=1}^{T}$ have the same limiting distribution as the partial sums of $\{\xi_{it}\}_{t=1}^{T}$. In order for this to hold, two results are needed. First, an IP for $\{e_{it,b}\}_{t=1}^{T}$ is needed; this follows from Lemma 1. Also, it must hold that $\hat{\beta}_{i,q}^{-1}(1) \stackrel{p}{\to} \beta_i^{-1}(1)$; as shown in the proof, this is a consequence of Lemma 2. Lemma 3 is the bootstrap counterpart to (6).

Lemmas 1-3 yield a bootstrap analogue to Proposition 1.

**Theorem 1** *Let Assumptions 1-5 hold. Then, as* $(n, T) \to \infty$ *and for all* $i$

$$\frac{1}{T^2} \sum_{t=1}^{T} F_{t,b} F_{t,b}' \stackrel{d^b}{\to} H' \left( \int W_\varepsilon W_\varepsilon' \right) H, \tag{19}$$

$$\frac{1}{T} \sum_{t=1}^{T} F_{t,b} u_{it,b} \stackrel{d^b}{\to} H' \int W_\varepsilon dW_{u,i}, \tag{20}$$

*in* $P$*, with* $W_\varepsilon$ *and* $W_{u,i}$ *defined in Proposition 1.*

Theorem 1 is a similar result to Lemma 3.4 in Chang, Park and Song (2006), and it is the bootstrap counterpart to equations (7) and (8) in Proposition 1. Results are pointwise in $i$; no joint limit theory is developed here.

# 4 Univariate bootstrap

The algorithm proposed above is applied to each unit separately, thereby imposing cross sectional independence. This approach is valid when cross sectional correlation does not need to be taken into account. This is the case when certain "time series problems" are considered, e.g. the estimation of the loadings; in such cases, the bootstrap boils down to a problem similar to Chang, Park and Song (2006). However, when "cross sectional problems" are considered (such as the estimation of common factors) results that are pointwise in $i$ are sufficient only in presence of little or no cross sectional dependence. In this section we consider three applications: we present validity results for the bootstrap estimates of loadings (Section 4.1), common factors (Section 4.2), and common components (Section 4.3).

Throughout the section, we consider two alternative DGPs for $x_{it,b}$:

$$x_{it,b}^{(1)} = \hat{\lambda}_i' \hat{F}_t + u_{it,b}, \tag{21}$$

$$x_{it,b}^{(2)} = \hat{\lambda}_i' F_{t,b} + u_{it,b}. \tag{22}$$

We show that using either (21) or (22) has a marginal impact on the bootstrap theory. Also, when studying the bootstrap approximation of loadings, factors and common components, we consider two alternative estimation techniques, OLS and PC. With OLS, the loadings are estimated through a time series regression, using $\hat{F}_t$ as observable regressors. Similary, the factors are estimated through a cross sectional regression with $\hat{\lambda}_i$ treated as observable. With PC, loadings and factors are extracted from $x_{it,b}^{(1)}$ or $x_{it,b}^{(2)}$, without treating $\hat{\lambda}_i$ or $\hat{F}_t$ as observed. The same restrictions as for the computation of $\left(\hat{\lambda}_i, \hat{F}_t\right)$ can be used at each bootstrap iteration. It can be expected that this approach is less dependent than OLS on the quality of the first step estimates $\left(\hat{\lambda}_i, \hat{F}_t\right)$; also, the bootstrapped errors are allowed to have an impact on the bootstrapped factors.[1]

However, PC cannot estimate factors and loadings directly, but only up to a rotation. Bai and Ng (2011) study under which restrictions the rotation matrix is (asymptotically) the identity matrix, but these restrictions need not always hold in practice. The issue of rotational indeterminacy affects the bootstrap in two ways.

Firstly, it is possible to provide bootstrap approximations for $\hat{\lambda}_i - H^{-1}\lambda_i$ and for $\hat{F}_t - H'F_t$, but the bootstrap is not able to estimate $H$. Whilst this is a general limitation of PC, in many applications knowing $\left(H^{-1}\lambda_i, H'F_t\right)$ is as good as $(\lambda_i, F_t)$; examples include computing common components; confidence intervals for diffusion index forecast (Bai and Ng, 2006a); IV estimation (Bai and Ng, 2010); and testing whether observable economic variables overlap with estimated latent factors (Bai and Ng, 2006b). In these contexts, the bootstrap can be useful.

Secondly, rotational indeterminacy also affects the bootstrap when PC is applied to

[1] I thank an anonymous Referee for pointing this out to me.

(21) or (22). To illustrate this, we consider the estimation of the loadings as a leading example. When using OLS, $\hat{F}_t$ is treated as observable. Thus, there is no rotational indeterminacy, and the bootstrap estimator estimates $\hat{\lambda}_i$ directly. Conversely, when applying PC to e.g. (21), the loadings are estimated up to a rotation matrix $H_1$, given by

$$H_1 = \left[\frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}_i\hat{\lambda}_i'\right]\left[\frac{1}{T^2}\sum_{t=1}^{T}\hat{F}_t\hat{F}_{t,b}'\right]\left[V_{nT}^{b(1)}\right]^{-1}, \tag{23}$$

where $\hat{F}_{t,b}$ is the PC estimate of the common factors and $V_{nT}^{b(1)}$ contains the first $k$ eigenvalues of $\frac{1}{nT^2}X_b^{(1)}X_b^{(1)\prime}$ in descending order with $X_b^{(1)} = \left[x_{1,b}^{(1)}, ..., x_{n,b}^{(1)}\right]'$ and $x_{i,b}^{(1)} = \left[x_{i1,b}^{(1)}, ..., x_{iT,b}^{(1)}\right]'$. The matrix $H_1$ is computed using $\hat{F}_t$, as in (23), or $F_{t,b}$, according as (21) or (22) is used. Thus, $H_1$ is observable: there is no identification issue here. The bootstrap IP, (25) and Proposition 5 below ensures that $H_1$ is invertible.

## 4.1 Loadings

Consider $\hat{\lambda}_i$. Lemma 4 in Bai (2004, p. 147) states that $\hat{\lambda}_i - H^{-1}\lambda_i = O_p\left(T^{-1}\right)$ with

$$\begin{aligned}\hat{\lambda}_i - H^{-1}\lambda_i &= \left[\sum_{t=1}^{T}\hat{F}_t\hat{F}_t'\right]^{-1}\left[H'\sum_{t=1}^{T}F_t u_{it}\right] + \left[\sum_{t=1}^{T}\hat{F}_t\hat{F}_t'\right]^{-1}\times \\ &\quad \left[\sum_{t=1}^{T}\left(\hat{F}_t - H'F_t\right)u_{it}\right] + \left[\sum_{t=1}^{T}\hat{F}_t\hat{F}_t'\right]^{-1}\left[\sum_{t=1}^{T}\hat{F}_t\left(H'F_t - \hat{F}_t\right)'H^{-1}\lambda_i\right] \\ &= I + II + III.\end{aligned} \tag{24}$$

Lemma B.4 in Bai (2004, p. 171) entails that $II$ and $III$ are negligible. Using (7) and (8)

$$T\left(\hat{\lambda}_i - H^{-1}\lambda_i\right) \xrightarrow{d} H^{-1}\left(\int W_\varepsilon W_\varepsilon'\right)^{-1}\left(\int W_\varepsilon dW_{u,i}\right), \tag{25}$$

which is the same as Theorem 3 in Bai (2004).

As mentioned above, OLS can be applied to (21) and (22), treating $\hat{F}_t$ as observable, obtaining $\hat{\lambda}_{i,b}^{OLS(1)}$ and $\hat{\lambda}_{i,b}^{OLS(2)}$ respectively. Alternatively, PC can be applied to $x_{it,b}^{(1)}$ and $x_{it,b}^{(2)}$, obtaining $\hat{\lambda}_{i,b}^{PC(1)}$ and $\hat{\lambda}_{i,b}^{PC(2)}$.

*OLS estimation*

When using OLS to compute $\hat{\lambda}_{i,b}^{OLS(1)}$ and $\hat{\lambda}_{i,b}^{OLS(2)}$, the estimation errors are given by

$$\begin{aligned} \hat{\lambda}_{i,b}^{OLS(1)} - \hat{\lambda}_i &= \left[\sum_{t=1}^{T} \hat{F}_t \hat{F}_t'\right]^{-1} \left[\sum_{t=1}^{T} \hat{F}_t u_{it,b}\right], & (26) \\ \hat{\lambda}_{i,b}^{OLS(2)} - \hat{\lambda}_i &= \left[\sum_{t=1}^{T} F_{t,b} F_{t,b}'\right]^{-1} \left[\sum_{t=1}^{T} F_{t,b} u_{it,b}\right]. & (27) \end{aligned}$$

**Proposition 2** *Let Assumptions 1-5 hold. As $(n,T) \to \infty$ for some $\delta > 0$*

$$T\left[\hat{\lambda}_{i,b}^{OLS(1),(2)} - \hat{\lambda}_i\right] \xrightarrow{d^b} H^{-1}\left(\int W_\varepsilon W_\varepsilon'\right)^{-1}\left(\int W_\varepsilon dW_{u,i}\right) \text{ in } P, \tag{28}$$

$$E^b \left\| T\left[\hat{\lambda}_{i,b}^{OLS(1),(2)} - \hat{\lambda}_i\right] \right\|^{2+\delta} = O_p(1) \text{ in } P. \tag{29}$$

Equation (28) is a weak convergence result: the bootstrap estimation error, $\hat{\lambda}_{i,b}^{OLS(1),(2)} - \hat{\lambda}_i$, has the same limiting distribution as in (25), which stipulates the validity of $\hat{\lambda}_{i,b}^{OLS(1),(2)} - \hat{\lambda}_i$ in approximating the limiting distribution of $\hat{\lambda}_i - H^{-1}\lambda_i$. Equation (29), in essence, is an application of Lemma 1. It ensures that $T\left[\hat{\lambda}_{i,b}^{(1),(2)} - \hat{\lambda}_i\right]$ is uniformly integrable, which is useful when an approximation of the moments of $T\left[\hat{\lambda}_i - H^{-1}\lambda_i\right]$ is needed.

*PC estimation*

When using PC, $\hat{\lambda}_{i,b}^{PC(1)}$ and $\hat{\lambda}_{i,b}^{PC(2)}$ estimate $H_1^{-1}\hat{\lambda}_i$, with $H_1$ defined in (23).

**Proposition 3** *Let Assumptions 1-5 hold. As $(n,T) \to \infty$ for some $\delta > 0$*

$$T\left[\hat{\lambda}_{i,b}^{PC(1),(2)} - H_1^{-1}\hat{\lambda}_i\right] \xrightarrow{d^b} H_1^{-1}\left[H^{-1}\left(\int W_\varepsilon W_\varepsilon'\right)^{-1}\left(\int W_\varepsilon dW_{u,i}\right)\right] \text{ in } P, \tag{30}$$

$$E^b \left\| T\left[\hat{\lambda}_{i,b}^{PC(1),(2)} - H_1^{-1}\hat{\lambda}_i\right] \right\|^{2+\delta} = O_p(1) \text{ in } P. \tag{31}$$

Equation (30) is, in essence, the same as Theorem 3 in Bai (2004): the limiting distribution of $\hat{\lambda}_{i,b}^{PC(1),(2)} - H_1^{-1}\hat{\lambda}_i$ is the same as the limiting distribution of $\hat{\lambda}_i - H^{-1}\lambda_i$, except for the presence of the rotation matrix $H_1$. This is a consequence of the rotational indeterminacy which is typical of PC estimation. Since $H_1$ is observable, the limiting distribution of $\hat{\lambda}_i - H^{-1}\lambda_i$ is approximated by $H_1\left[\hat{\lambda}_{i,b}^{PC(1),(2)} - H_1^{-1}\hat{\lambda}_i\right]$.

### 4.2 Common factors

The building block of the analysis is Theorem 2 in Bai (2004, p. 148): as $(n, T) \to \infty$ with $\frac{n}{T^3} \to 0$, it holds that

$$\sqrt{n}\left[\hat{F}_t - H'F_t\right] \xrightarrow{d} H'\Sigma_\Lambda \times N\left(0, \Gamma_t\right), \tag{32}$$

with $\Gamma_t = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n E\left[\lambda_i \lambda_j' u_{it} u_{jt}\right]$. Under the "one unit at a time" resampling scheme, cross dependence among the $u_{it}$s is forced to be zero; thus, it can be expected that the bootstrap provides valid inference on factors only under $E\left(u_{it}u_{jt}\right) = 0$ for $i \neq j$, which entails $\Gamma_t = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n E\left[\lambda_i \lambda_i' u_{it}^2\right]$. Indeed, as discussed in Section 5, the "one unit at a time" scheme can provide valid inference when cross correlation is different from zero but "negligible" as $n \to \infty$ - see Theorem 3 and the discussion thereafter. More generally, consistent estimation of $\Gamma_t$ is fraught with difficulties; as Bai (2003) points out, HAC-type estimators are not feasible since, in general, the order of cross correlation is unknown.

As in the case of the loadings, there are two possible ways of estimating the common factors from (21) and (22). A cross sectional OLS estimator can be applied, considering $\hat{\lambda}_i$ observable and computing respectively $\hat{F}_{t,b}^{OLS(1)}$ and $\hat{F}_{t,b}^{OLS(2)}$. Alternatively, PC can be used, obtaining $\hat{F}_{t,b}^{PC(1)}$ and $\hat{F}_{t,b}^{PC(2)}$.

*OLS estimation*

The estimation error is

$$\hat{F}_{t,b}^{OLS(2)} - F_{t,b} = \left(\frac{1}{n}\sum_{i=1}^n \hat{\lambda}_i \hat{\lambda}_i'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^n \hat{\lambda}_i u_{it,b}\right). \tag{33}$$

The same equation holds for $\hat{F}_{t,b}^{OLS(1)} - \hat{F}_t$. No identification issue is present when using OLS.

**Proposition 4** *Let Assumptions 1-5 hold. As* $(n, T) \to \infty$

$$\sqrt{n}\left[\hat{F}_{t,b}^{OLS(2)} - F_{t,b}\right] \xrightarrow{d^b} H'\Sigma_\Lambda \times N\left(0, \Gamma_t\right) \text{ in } P, \tag{34}$$

$$E^b \left\|\sqrt{n}\left[\hat{F}_{t,b}^{OLS(2)} - F_{t,b}\right]\right\|^{2+\delta} = O_p\left(1\right) \text{ in } P, \tag{35}$$

*for some* $\delta > 0$. *The same holds for* $\hat{F}_{t,b}^{OLS(1)} - \hat{F}_t$.

Equation (34) stipulates the validity of $\hat{F}_{t,b}^{OLS(2)} - F_{t,b}$ (and of $\hat{F}_{t,b}^{OLS(1)} - \hat{F}_t$) in approximating the limiting distribution of $\hat{F}_t - H'F_t$ given in (32). Note that, unlike Theorem 2 in Bai (2004), no restrictions are needed on the rate of divergence between $n$ and $T$. This is because the OLS estimator uses $\hat{\lambda}_i$ as an observable regressor, thereby not introducing any extra error terms, unlike PC. However, this does not entail that $\sqrt{n}\left[\hat{F}_{t,b}^{OLS(2)} - F_{t,b}\right]$ can be used to approximate the limiting distribution of $\sqrt{n}\left[\hat{F}_t - H'F_t\right]$ for any combination of $n$ and $T$. When $\frac{n}{T^3} \to c > 0$, the limiting distribution of $\sqrt{n}\left[\hat{F}_t - H'F_t\right]$ is not given by (32), and therefore $\sqrt{n}\left[\hat{F}_{t,b}^{OLS(2)} - F_{t,b}\right]$ is no longer valid.

*PC estimation*

Similarly to the case of the loadings, PC identifies the common factors up to the rotation matrix $H_1$, defined in (23).

**Proposition 5** *Let Assumptions 1-5 hold. As* $(n, T) \to \infty$ *with* $\frac{n}{T^3} \to 0$

$$\sqrt{n}\left[\hat{F}_{t,b}^{PC(2)} - H_1' F_{t,b}\right] \xrightarrow{d^b} H_1'\left[H'\Sigma_\Lambda \times N\left(0, \Gamma_t\right)\right] \text{ in } P, \tag{36}$$

$$E^b\left\|\sqrt{n}\left[\hat{F}_{t,b}^{PC(2)} - H_1' F_{t,b}\right]\right\|^{2+\delta} = O_p(1) \text{ in } P, \tag{37}$$

*for some* $\delta > 0$. *The same holds for* $\hat{F}_{t,b}^{PC(1)} - H_1'\hat{F}_t$.

Proposition 5 states the validity of $\sqrt{n}\left[\hat{F}_{t,b}^{PC(2)} - H_1' F_{t,b}\right]$ to approximate the limiting distribution of $\sqrt{n}\left[\hat{F}_t - H'F_t\right]$, and it is the bootstrap counterpart to Theorem 2 in Bai (2004).

## 4.3 Common components

The estimated common components are given by $\hat{C}_{it} = \hat{\lambda}_i'\hat{F}_t$, with

$$\hat{C}_{it} - C_{it} = \left(\hat{F}_t - H'F_t\right)' H^{-1}\lambda_i + \hat{F}_t'\left(\hat{\lambda}_i - H^{-1}\lambda_i\right) = I + II. \tag{38}$$

Bai (2004, Theorem 4, p. 149) shows that, as $(n,T) \to \infty$ with $\frac{n}{T} \to \pi$, for each $(i,t)$ with $t = \lfloor Ts \rfloor$

$$\sqrt{n}\left(\hat{C}_{it} - C_{it}\right) \xrightarrow{d} \lambda_i' \Sigma_\Lambda N(0, \Gamma_t) + \sqrt{\pi} W_\varepsilon(s) \left(\int W_\varepsilon W_\varepsilon'\right)^{-1} \int W_\varepsilon dW_{u,i}, \tag{39}$$

where the first term on the right hand side comes from $I$ in (38) and the second one from $II$.

Choosing either (21) or (22) for the bootstrap approximation of $\hat{C}_{it} - C_{it}$ does not make a difference. However, the estimation technique employed (i.e. OLS or PC) has profound consequences: OLS should not be employed when computing the bootstrap approximation of $\hat{C}_{it} - C_{it}$. In order to illustrate this, let $\hat{\lambda}_{i,b}^{OLS}$ be the OLS estimator of $\hat{\lambda}_i$ in either (21) or (22), and let $F_t^b$ denote the common factors in either (21) or (22) - we omit superscripts to save space. The bootstrap estimate is $C_{it}^b = \hat{\lambda}_{i,b}^{OLS\prime} F_t^b$. Thus, the estimation error is $\hat{\lambda}_{i,b}^{OLS\prime} F_t^b - \hat{\lambda}_i' F_t^b = \left(\hat{\lambda}_{i,b}^{OLS} - \hat{\lambda}_i\right)' F_t^b$: the asymptotics of $C_{it}^b - \hat{C}_{it}$ is driven only by part $II$ in (38). This is due to $F_t^b$ being treated as observable, so that there is no estimation error of the form $\hat{F}_t - H' F_t$. Therefore, $C_{it}^b - \hat{C}_{it}$ cannot be used to approximate the limiting distribution of $\hat{C}_{it} - C_{it}$, unless $\frac{n}{T} \to 0$, which limits its practical use.

Thus, PC should be used. Let $\hat{\lambda}_{i,b}^{PC}$ and $\hat{F}_{t,b}^{PC}$ be the estimates of $\hat{\lambda}_i$ and $\hat{F}_t$ under either (21) or (22) - superscripts are again omitted to save space. Define $\hat{C}_{it,b} = \hat{\lambda}_{i,b}^{PC\prime} \hat{F}_{t,b}^{PC}$; we have

$$\hat{C}_{it,b} - \hat{C}_{it} = \left(\hat{F}_{t,b}^{PC} - H_1' F_t\right)' H_1^{-1} \hat{\lambda}_i + \hat{F}_{t,b}^{PC\prime} \left(\hat{\lambda}_{i,b}^{PC} - H_1^{-1} \hat{\lambda}_i\right). \tag{40}$$

In view of (40), it can be expected that the limiting distribution of $\sqrt{n}\left(\hat{C}_{it,b} - \hat{C}_{it}\right)$ is the same as that of $\sqrt{n}\left(\hat{C}_{it} - C_{it}\right)$ for all combinations of $n$ and $T$ as they pass to infinity.

**Proposition 6** *Let Assumptions 1-5 hold. Then, for all $(i,t)$ such that $t = \lfloor Ts \rfloor$, as $(n,T) \to \infty$ with $\frac{n}{T} \to \pi$*

$$\sqrt{n}\left(\hat{C}_{it,b} - \hat{C}_{it}\right) \xrightarrow{d^b} \lambda_i' \Sigma_\Lambda N(0, \Gamma_t) + \sqrt{\pi} W_\varepsilon(s) \left(\int W_\varepsilon W_\varepsilon'\right)^{-1} \int W_\varepsilon dW_{u,i} \text{ in } P. \tag{41}$$

*Also, for all $(i,t)$ such that $t = \lfloor Ts \rfloor$, as $(n,T) \to \infty$ with $\frac{n}{T} \to \pi$ for some $\delta > 0$*

$$E^b \left| \sqrt{n} \left( \hat{C}_{it,b} - \hat{C}_{it} \right) \right|^{2+\delta} = O_p(1) \ \ in \ P. \tag{42}$$

Results for $\frac{n}{T} \to 0$ and $\frac{T}{n} \to 0$ are in Appendix. The case $\frac{n}{T} \to \pi$ is, as pointed out in Bai (2004), the most useful one, since $\pi$ can be replaced by $\frac{n}{T}$, thereby making the bootstrap approximation of $\hat{C}_{it} - C_{it}$ usable for all combinations of $n$ and $T$.

# 5 Multivariate bootstrap

Results in Sections 3 and 4 are pointwise in $i$, and only consider the time series dimension. This is sufficient for some applications, but in other cases the cross sectional dimension and the presence of cross sectional correlation need to be taken into account. Also, in some applications joint asymptotics results are needed.

The main output of this Section is the derivation of a bootstrap counterpart to equations (9) and (10). It is shown that the moment existence conditions granted by Lemma 1 are sufficient also for joint bootstrap asymptotics. However, consistent estimation of the long run variance of $u_t$ is fraught with difficulties, due to its growing dimension (see Theorem 3).

To study multivariate bootstrap, the algorithm in Section 3.1 is modified by resampling the whole vector $\hat{e}^u_t = \left[\hat{e}^u_{1t,q}, ..., \hat{e}^u_{nt,q}\right]'$, and estimating an $n$-dimensional VAR of order $q$ for $\hat{u}_t = [\hat{u}_{1t}, ..., \hat{u}_{nt}]'$. In order to prove bootstrap analogues to equations (9) and (10), let the $VAR(\infty)$ representation for $u_t$ be $u_t = \sum_{j=1}^{\infty} B_j u_{t-j} + e^u_t$, truncated at lag $q$ as

$$u_t = \sum_{j=1}^{q} B_j u_{t-j} + e^u_t, \tag{43}$$

and let $B(1) = 1 - \sum_{j=1}^{\infty} B_j$; by definition, $B(1) = \Gamma^{-1}(1)$. Also, define the bootstrap counterpart to $e^u_t$, $e^u_{t,b}$, and let $B^*_j$ be some estimator of $B_j$; thus, $B^*(1) = 1 - \sum_{j=1}^{q} B^*_j$ is an estimator for $B(1)$. Note that the number of parameters to be estimated is $qn^2$. Thus, we require that $qn^2 < T$. This constraint on the relative speed of divergence of $n$ and $T$ is stronger than the typical requirement that $\frac{n}{T} \to 0$. The bootstrap sample $u_{t,b}$

can be generated using $u_{t,b}=\sum_{j=1}^{q}B_j^{*}u_{t-j,b}+e_{t,b}^{u}$. No modifications are required to the algorithm in Section 3.1 as far as the generation of $F_{t,b}$ is concerned.

**Theorem 2** *Let Assumptions 1-5 hold, and assume further that* $\|B^{*}(1)-B(1)\|_1=o_p(1)$*. As* $(n,T)\to\infty$ *with* $qn^2<T$

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Ts\rfloor}u_{it,b}\overset{d^b}{\to}\sigma_u W_u(s), \tag{44}$$

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}F_{t,b}u_{it,b}\overset{d^b}{\to}\sigma_u H'\int W_{\varepsilon}dW_u, \tag{45}$$

*uniformly in* $s$*, in* $P$*, where* $\sigma_u$*,* $W_u$ *and* $W_{\varepsilon}$ *are defined in Proposition 1.*

Theorem 2 is a joint asymptotics result. It shows that the distributions of $n^{-1/2}T^{-1/2}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it,b}$ and $n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}F_{t,b}u_{it,b}$ are asymptotically the same as the distributions of $n^{-1/2}T^{-1/2}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it}$ and $n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}F_tu_{it}$.

Equation (44) could be generalised to study multiparameter partial sum processes such as $(nT)^{-1/2}\sum_{i=1}^{\lfloor np\rfloor}\sum_{t=1}^{\lfloor Ts\rfloor}u_{it}$, with $(p,s)\in[0,1]\times[0,1]$. Having $\max_{i,t}E|u_{it}|^{2+\delta}<\infty$ yields $(nT)^{-1/2}\sum_{i=1}^{\lfloor np\rfloor}\sum_{t=1}^{\lfloor Ts\rfloor}u_{it}\overset{d}{\to}\sigma_u W(p,s)$, where $W(\cdot,\cdot)$ is a standard two-dimensional Brownian sheet. This is a standard result in the random fields literature - see, *inter alia*, Bulinski and Shashkin (2006), and Rio (1993) for strong approximations. Therefore, Lemma 1 is sufficient to prove a multiparameter IP for the partial sums of the bootstrap sample $u_{it,b}$. This could be useful when resampling across $i$ as well as across $t$ (see e.g. Kapetanios, 2008, and Levina and Bickel, 2006), although this postulates the existence of some ordering among the units which is not always obvious - see also Goncalves (2010).

In essence, Theorem 2 states that joint asymptotics can be derived for the bootstrap samples under the same assumptions as univariate results, as long as there exists a consistent (in $L_1$-norm) estimator for $B(1)$. Since $B(1)$ is $n\times n$ (with $n\to\infty$), Lemma 2 is not sufficient for this, as it only grants element-wise consistency for $B^{*}(1)$. Although the details are in the proof, here we give a preview of the rationale of the requirement that $\|B^{*}(1)-B(1)\|_1=o_p(1)$. As an illustrative example, consider showing that $(nT)^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\sum_{i=1}^{n}u_{it,b}$ has the same limiting distribution as $(nT)^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\sum_{i=1}^{n}u_{it}$. Writing this in matrix form, a requirement for this is that $(nT)^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}i_n'[B^{*}(1)]^{-1}e_{t,b}^{u}$

and $(nT)^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor} i_n'\left[B(1)\right]^{-1}e_t^u$ should have the same distribution. An IP holds for the partial sums of $e_t^u$ and $e_{t,b}^u$. Thus, following the same lines as in the proof of Lemma 3, we need $(nT)^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\sum_{j=1}^{n}\sum_{i=1}^{n}\left\{\left[B^*(1)\right]^{-1}-\left[B(1)\right]^{-1}\right\}_{ij}e_{jt,b}^u=o_p(1)$, where $\{A\}_{ij}$ denotes the element in position $(i,j)$ of matrix $A$. Since $e_{jt,b}^u$ has finite variance, it is sufficient that $\sup_j\left|\sum_{i=1}^{n}\left\{\left[B^*(1)\right]^{-1}-\left[B(1)\right]^{-1}\right\}_{ij}\right|=o_p(1)$, for which it is sufficient that $\left\|\left[B^*(1)\right]^{-1}-\left[B(1)\right]^{-1}\right\|_1=o_p(1)$. This holds if $\left\|B^*(1)-B(1)\right\|_1=o_p(1)$, since $\left\|\left[B^*(1)\right]^{-1}-\left[B(1)\right]^{-1}\right\|_1\leq\left\|\Gamma^{-1}(1)\right\|_1\left\|\Gamma^{-1}(1)\right\|_\infty\left\|B^*(1)-B(1)\right\|_1$ and $\left\|\Gamma^{-1}(1)\right\|_1$ and $\left\|\Gamma^{-1}(1)\right\|_\infty$ are finite by Assumption 1*(iii)*.

In order to estimate $B(1)$, consider (43). Defining $u_{qt}=\left[u_{t-1}',...,u_{t-q}'\right]'$ and $B_q=\left[B_{q,1}|...|B_{q,q}\right]$, we have

$$u_t=B_qu_{qt}+e_{qt}^u; \tag{46}$$

the feasible estimator of $B_q$ is

$$\hat{B}_q=\left[\sum_{t=q+1}^{T}\hat{u}_t\hat{u}_{qt}'\right]\left[\sum_{t=q+1}^{T}\hat{u}_{qt}\hat{u}_{qt}'\right]^{-1}. \tag{47}$$

Thus, $B(1)$ can be estimated by $\widehat{B_q(1)}=1-\sum_{j=1}^{q}\hat{B}_{q,j}$ - note that (47) requires the inversion of an $nq\times nq$ matrix. More importantly, the VAR approach introduces further restrictions to the applicability of the bootstrap. As pointed out above, in order to implement this approach we need that $qn^2<T$.

We also consider also an alternative estimator of $B(1)$ which does not take into account the cross sectional correlation among the $u_{it}$s. This can be computed from the $\hat{\gamma}_{q,j}$s estimated from (12), and defined as $\widetilde{B_q(1)}=1-\sum_{j=1}^{q}\tilde{B}_{q,j}$, with $\tilde{B}_{q,j}$ an $n\times n$ diagonal matrix whose elements are given by $\hat{\gamma}_{q,j}^{(i)}$. In this case, no VAR is fitted and thus the restriction that $qn^2<T$ is not necessary.

It holds that:

**Theorem 3** *Let $\Lambda=(\lambda_1,\lambda_2,...,\lambda_n)'$, and let Assumptions 1-4 hold with $\|\Lambda\|_1=O_p(n)$. Then*

$$\left\|\widehat{B_q(1)}-B(1)\right\|_1=O_p\left(q\sqrt{\frac{\log T}{T}}\right)+O_p\left(nq^2C_{nT}^{-1}\right)+o\left(q^{-s}\right)+o_p(1). \tag{48}$$

*Assuming* $\sup_j \sum_{i\neq j} |\tau_{ij}| = O\left(n^{-\phi}\right)$ *with* $\phi \geq 0$*, it holds that*

$$\left\| \widetilde{B_q(1)} - B(1) \right\|_1 = O_p\left(\frac{q}{\varphi_{nT}^u}\right) + O_p\left(qn^{-\phi}\right) + o\left(q^{-s}\right). \tag{49}$$

Theorem 3 states that $\widehat{B_q(1)}$ is inconsistent in $L_1$-norm. This result, somewhat constrained by the choice of the matrix norm, can be compared with the analysis in Fan, Fan and Lv (2008). Theorem 3 is a result of independent interest, even outside the context of bootstrap. As far as sieve bootstrap is concerned, the inconsistency of $\widehat{B_q(1)}$ entails that an IP for $u_{it,b}$ cannot be proved - this can be viewed following the same lines as in the proof of Lemma 3. In spite of Assumption 1*(iii)*, which limits the amount of cross dependence among the $u_{it}$s, inconsistency arises due to the presence of $\left[\hat{\Lambda} - \Lambda\left(H'\right)^{-1}\right] F_t$ in the $\hat{u}_{it}$s (see the proof). This could be compared with the results in Chudik and Pesaran (2011), where an assumption similar to 1*(iii)* is sufficient to ensure consistency of the estimated long run covariance matrix. Other, residual-based estimators of the long run variance would similarly be affected by the presence of $\left[\hat{\Lambda} - \Lambda\left(H'\right)^{-1}\right] F_t$. Intuitively, this result reinforces the well-known fact that PC estimation can accommodate for weak cross dependence only.

Turning to $\widetilde{B_q(1)}$, this is consistent only under $\phi > 0$, as long as Assumption 5 is modified to $q \to \infty$ with $q = o\left(\min\left\{\sqrt{\frac{T}{\log T}}, n^{\phi}\right\}\right)$. In this case, $\left\|\widetilde{B_q(1)} - B(1)\right\|_1 = o_p(1)$, as required by Theorem 2. The first term on the right hand side of (49) represents the rate of convergence of the elements on the main diagonal of $\widetilde{B_q(1)}$, as warranted by Lemma 2. The assumption that $\sup_i \sum_{j\neq i} |\tau_{ij}| = O\left(n^{-\phi}\right)$ poses a limitation on the amount of cross dependence among the $u_{it}$s. Although some dependence is allowed for, this is weaker than in an ordinary approximate factor structure framework (see e.g. Assumption C(1) in Bai, 2004), where it suffices to have $n^{-1}\sum_{i=1}^n \sum_{j=1}^n |\tau_{ij}| = O(1)$, for which it is sufficient that $\sup_i \sum_{j=1}^n |\tau_{ij}| = O(1)$. Conversely, the assumption is more general than in classical Principal Component Analysis, where $\tau_{ij} = 0$ for all $i \neq j$. Thus, in essence (49) states that neglecting cross dependence is harmless (and, in fact, advantageous over $\widehat{B_q(1)}$), as long as there is "very little" cross dependence. This result illustrates the fact that the "classical" assumptions of Principal Component Analysis can be relaxed when $n$ is large, but only up to a certain extent. Finally, note that, as long as $\sup_i \sum_{j\neq i} |\tau_{ij}| = O\left(n^{-\phi}\right)$, any consistent estimation technique (e.g. a nonparametric one) for the long run variances of the $u_{it}$s would yield a consistent estimator for $B(1)$.

The results in Theorem 3 extend to other, more general contexts, whenever the long run variance of $u_t$ needs to be estimated. As an example, consider the asymptotics of $\hat{F}_t$ and $\hat{C}_{it}$, reported in equations (32) and (39) respectively. In both cases, the quantity $\Gamma_t = \lim_{n\to\infty} n^{-1}\Lambda'\Gamma(1)\,\Sigma_u\,\Gamma'(1)\,\Lambda$ is present. Unless cross sectional independence is assumed, it is necessary to estimate $\Gamma(1)$, with the difficulties highlighted in Theorem 3.

# 6 Bootstrapping mixed panel factor series

In this section, we discuss the extension of the bootstrap theory derived above to the case in which both $I(1)$ and $I(0)$ common factors are present in the DGP of $x_{it}$, viz.

$$\begin{aligned} x_{it} &= \lambda_i^{F\prime}F_t + \lambda_i^{G\prime}G_t + u_{it} \\ &= \lambda_i^{K\prime}K_t + u_{it}, \end{aligned} \tag{50}$$

with $\lambda_i^K = \left[\lambda_i^{F\prime}, \lambda_i^{G\prime}\right]'$ and $K_t = [F_t', G_t']'$. We assume that, as before, $F_t$ is a $k$-dimensional nonstationary process, and $G_t$ is an $h$-dimensional stationary process.

Model (50) is a useful extension for at least two reasons. Firstly, the presence of common $I(0)$ factors in (50) can accommodate for dynamic factors, i.e. $\lambda_i$ in (1) replaced by $\sum_{j=1}^{\infty}\lambda_i^k L^k$; and for cointegrated factors, i.e. $F_t$ in (1) being cointegrated. Secondly, the presence of stationary common factors also means that, in (1), strong cross dependence among the $u_{it}$s can be accommodated for - as far as weak dependence is concerned, the considerations in Section 5 still hold.

The purpose of this section is to provide some initial results for the extension of bootstrap theory to (50). The relevant inferential theory is in Bai (2004, section 5); see also Maciejowska (2010). In particular, we present extensions of Lemmas 1 and 2. This provides the theoretical framework from which more specialised results, such as validity results for the estimates of $\lambda_i^K$ and $K_t$ can be derived in a similar way as in Sections 4 and 5.

The bootstrap algorithm requires a modification of Section 3.1. More specifically, after estimating $H'K_t$ (for some invertible matrix $H$ defined analogously to (4)), with

$\hat{K}_t = \left[\hat{F}_t', \hat{G}_t\right]'$, define ${}_\Delta\hat{K}_t = \left[\Delta\hat{F}_t', \hat{G}_t\right]'$. As well as estimating $\hat{u}_{it} = \sum_{j=1}^{q_{u,i}} \gamma_{q,j}^{(i)} \hat{u}_{it-j}$ $+e_{it,q}^u$, the algorithm is based on fitting the $(k+h)$-dimensional VAR

$$ {}_\Delta\hat{K}_t = \sum_{j=1}^{q_K} A_{q,j} {}_\Delta\hat{K}_{t-j} + e_{t,q}^K, $$

generating the residuals $\hat{e}_{t,q}^K$. After recentering, the sequence $\left\{\hat{e}_{t,q}^K\right\}$ is resampled as in Step 3.1.*(b)*, obtaining $\left[e_{t,b}^{K\prime}, e_{it,b}^u\right]'$. After generating the pseudo sample ${}_K\xi_{it,b} = \left[{}_\Delta K_{t,b}', u_{it,b}\right]' = \left[\Delta F_{t,b}', G_{t,b}', u_{it,b}\right]'$ similarly to Step 3.2.*(a)*, the first $k$ elements can be integrated as in Step 3.2.*(b)*, thereby getting the bootstrap sample of the nonstationary common factors, $F_{t,b}$.

The following Assumptions are extensions/variations of Assumptions 1-4 and of Assumption 5 respectively, reported here for convenience.

**Assumption 6.** *(a)* Assumption 1 holds; *(b)* the $(k+h)$-dimensional process $[\Delta F_t', G_t']'$ satisfies Assumption 2; *(c)* the loadings $\lambda_i^K$ satifsy Assumption 3; *(d)* $\left\{[\Delta F_t', G_t']'\right\}$, $\{u_{it}\}$ and $\left\{\lambda_i^K\right\}$ are three mutually independent groups, and $K_0$ is independent of $\{u_{it}\}$ and $\left\{\lambda_i^K\right\}$.

**Assumption 7**. As $(n,T) \to \infty$, $q_K \to \infty$ and $q_{u,i} \to \infty$ for each $i$, with $q_K$ and $q_{u,i}$ both $o\left(\varphi_{nT}^u\right)$ for each $i$.

The two Assumptions are very similar to Assumptions 1-5. One difference, in Assumption 7, is that the upper bound for $q_K$ is given by $\min\left\{\sqrt{n}, \sqrt{\log T/T}\right\}$, whereas in Assumption 5 the order of the VAR fitted to $\Delta\hat{F}_t$ is $\min\left\{n, \sqrt{\log T/T}\right\}$.

It holds that

**Lemma 4** *Let Assumptions 6 and 7 hold. As* $(n, T) \to \infty$*, for all* $(i, t)$ *and for* $r > 4$

$$
\begin{aligned}
E^b \left\| e_{t,b}^K \right\|^r &= E \left\| e_t^K \right\|^r + O_p\left(q_K^{-rs}\right) + O_p\left(\delta_{nT}^{-r}\right) && (51)\\
&\quad + O_p\left[\left(\frac{q_K}{\varphi_{nT}^u}\right)^r\right], \\
\max_{i,t} E^b \left| e_{it,b}^u \right|^r &= \max_{i,t} E \left| e_{it}^u \right|^r + O_p\left(q_{u,i}^{-rs}\right) + O_p\left(\delta_{nT}^{-r}\right) && (52)\\
&\quad + O_p\left[\left(\frac{q_{u,i}}{\varphi_{nT}^u}\right)^r\right], \\
\max_{1 \le j \le q_K} \left\| \hat{A}_{q,j\Delta} - H' A_{q,j\Delta} \left(H'\right)^{-1} \right\| &= O_p\left(\sqrt{\frac{\log T}{T}}\right) + O_p\left(\frac{1}{\sqrt{n}}\right) + o_p\left(\frac{1}{q_K^s}\right), && (53)\\
\max_{1 \le j \le q_{u,i}} \left| \hat{\gamma}_{q,j}^{(i)} - \gamma_j^{(i)} \right| &= O_p\left(\sqrt{\frac{\log T}{T}}\right) + O_p\left(\frac{1}{\sqrt{n}}\right) + o_p\left(\frac{1}{q_{u,i}^s}\right). && (54)
\end{aligned}
$$

Lemma 4 is the building block to extend the theory developed in Sections 4 and 5 to (50). Results such as Lemma 3 and Theorem 1 can be proved directly using (51)-(54). The main feature of the Lemma, in terms of rates of convergence, is that according to (53), $\hat{A}_{q,j\Delta}$ is still consistent but at a slower rate than in Lemma 2. In essence, instead of being consistent at a rate $\min\left\{n, \sqrt{\log T / T}\right\}$, $\hat{A}_{q,j\Delta}$ is consistent at a rate $\min\left\{\sqrt{n}, \sqrt{\log T / T}\right\}$. This result is directly related to the findings in Bai (2004) and Maciejowska (2010).

# 7 Concluding remarks

This paper contains results on the validity of sieve bootstrap applied to large, nonstationary panel factor series. Building on a similar research question as in Chang, Park and Song (2006) in the context of cointegrated, finite dimensional VARs, an IP is proved for the bootstrap sample which, together with results on the consistent estimation of long run variances and on the convergence to stochastic integrals of transformations of the bootstrap sample, provides a formal justification to the use of the bootstrap in the context of panel factor series. Whilst the first results are only pointwise, in order to extend the applicability of the sieve bootstrap, joint bootstrap asymptotics is also studied. In this case, the findings are ambiguous: the presence of cross sectional dependence makes bootstrapping invalid, unless cross dependence is very weak. Although this is a negative result, it illustrates the pitfalls and limitations of bootstrapping panel factor models and, more generally, of large panels with cross dependence. As an ancillary result, the paper contains an investigation on the consistency in $L_1$-norm of the estimated long run variance of panel factor models, showing that, whilst element-wise consistency holds, matrix-type consistency is in general hampered by the presence and the extent of cross dependence. These results are of independent interest, and the issue remains as to the consistent estimation of large covariance matrices under general forms of cross dependence. Finally, the paper considers the extension to the case of stationary and nonstationary common factors. This issue is currently under investigation by the author.

## References


[1] Bai, J. 2003. Inferential theory for structural models of large dimensions. Econometrica. 71, 135-171.

[2] Bai, J. 2004. Estimating cross-section common stochastic trends in nonstationary panel data. Journal of Econometrics 122, 137-183

[3] Bai, J., Ng, S. 2002. Determining the number of factors in approximate factor models. Econometrica 70, 191-221

[4] Bai, J., Ng, S. 2004. A PANIC attack on unit roots and cointegration. Econometrica 72, 1127-1178

[5] Bai, J., Ng, S. 2006a. Evaluating latent and observed factors in macroeconomics and finance. Journal of Econometrics 131, 507-537

[6] Bai, J., Ng, S. 2006b. Confidence intervals for diffusion index forecasts and inference with factor-augmented regressions. Econometrica 74, 1133-1150

[7] Bai, J., Ng, S. 2010. Instrumental variables estimation in a data rich environment. Econometric Theory 26, 1607-1637

[8] Bai, J., Ng, S. 2011. Principal components estimation and identification of the factors. Mimeo

[9] Bühlmann, P. 1997. Sieve bootstrap for time series. Bernoulli 3, 123-148

[10] Bulinski, A., Shashkin, A. 2006. Strong invariance principles for dependent random fields. IMS Lecture Notes - Monograph Series Dynamics and Stochastics

[11] Chang, Y., Park, J.Y., Song, K. 2006. Bootstrapping cointegrating regressions. Journal of Econometrics 133, 703-739

[12] Chudik, Pesaran, M.H. 2011. Infinite dimensional VARs and factor models. Journal of Econometrics 163, 4-22

[13] Fan, J., Fan, Y., Lv, J. 2008. High dimensional covariance matrix estimation using a factor model. Journal of Econometrics 147, 186-197

[14] Forni, M., Reichlin, L. 1998. Let's get real: a factor analytic approach to disaggregated business cycle dynamics. Review of Economic Studies 65, 453-473

[15] Girosi, F, King, G. 2007. Understanding the Lee-Carter mortality forecasting method. Mimeo

[16] Goncalves, S. 2011. The moving blocks bootstrap for panel linear regression models with individual fixed effects. Econometric Theory 27, 1048-1082

[17] Hall, P., Heyde, C.C. 1980. Martingale limit theory and its applications. New York, University Press.

[18] Hannan, E.T., Kavalieris, L. 1986. Regression; autoregression models. Journal of Time Series Analysis 7, 27-49

[19] Kao, C., Trapani, L., Urga, G. 2011. Asymptotics for panel models with common shocks. Econometric Reviews (forthcoming).

[20] Kapetanios, G. 2008. A bootstrap procedure for panel datasets with many cross-sectional units. Econometrics Journal 11, 377-395

[21] Kolotilina, L.Y. 2009. Bounds for the infinity norm of the inverse for certain $M$- and $H$-matrices. Linear Algebra and its Applications 430, 692-702

[22] Lee, R.D., Carter L. 1992. Modeling and forecasting the time series of U.S. mortality. Journal of the American Statistical Association 87, 659-71

[23] Levina, E., Bickel, P.J. 2006. Texture synthesis and nonparametric resampling of random fields. Annals of Statistics 34, 1751-1773

[24] Maciejowska, K. 2010. Common factors in nonstationary panel data with a deterministic trend - Estimation and Distribution Theory. Mimeo

[25] Park, J.Y. 2002. An invariance principle for sieve bootstrap in time series. Econometric Theory 18, 469–490

[26] Park, J.Y. 2003. Bootstrap unit root tests. Econometrica 71, 1845-1895

[27] Park, J.Y., Phillips, P.C.B. 1999. Asymptotics for nonlinear transformations of integrated time series. Econometric Theory 15, 269-298

[28] Pesaran, M.H., Tosetti, E. 2011. Large panels with common factors and spatial correlations. Journal of Econometrics 161, 182-202

[29] Phillips, P.C.B. 1988. Weak convergence of sample covariance matrices to stochastic integrals via martingale approximations. Econometric Theory 4, 528-533

[30] Phillips, P.C.B., Moon, H. R. 1999. Linear regression limit theory for nonstationary panel data. Econometrica 67, 1057-1112

[31] Phillips, P.C.B., Solo, V. 1992. Asymptotics for linear processes. Annals of Statistics 20, 971–1001

[32] Rio, E. 1993. Strong approximations for set-indexed partial sums processes via KMT constructions I. Annals of Probability 21, 759-790

[33] Sakhanenko, A.I. 1980. On unimprovable estimates of the rate of convergence in invariance principle. In: Nonparametric Statistical Inference. Colloquia Mathematica Societatis Janos Bolyai, 32, 779-783. Budapest, Hungary

# Appendix A: useful Lemmas

**Lemma A.1** *Let Assumptions 1-4 hold. Then*

**A.1***(i)* $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t - H'\Delta F_t\right\|^2 = O_p\left(C_{nT}^{-2}\right)$,

**A.1***(ii)* $T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{F}_t - H'\Delta F_t\right)'\Delta F_t = O_p\left(T^{-1/2}C_{nT}^{-1}\right)$,

**A.1***(iii)* $T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{F}_t - H'\Delta F_t\right)'\Delta\hat{F}_t = O_p\left(\frac{1}{n}\right) + O_p\left(\frac{1}{T^{3/2}}\right)$.

**Proof.** For the sake of the notation, we omit $H$ whenever possible. Consider A.1*(i)*. It holds that $\Delta\hat{F}_t - \Delta F_t = \left(\hat{F}_t - F_t\right) - \left(\hat{F}_{t-1} - F_{t-1}\right)$; using equation (B.1) in Bai (2004, p. 167) we have

$$V_{nT}\left(\hat{F}_t - H'F_t\right) = T^{-2}\sum_{s=1}^{T}\hat{F}_s\gamma_{s-t} + T^{-2}\sum_{s=1}^{T}\hat{F}_s\zeta_{st} + T^{-2}\sum_{s=1}^{T}\hat{F}_s\eta_{st} + T^{-2}\sum_{s=1}^{T}\hat{F}_s\xi_{st},$$

where $\gamma_{s-t} = n^{-1}E\left(u_t'u_s\right)$, $\zeta_{st} = n^{-1}\left(u_t'u_s\right) - \gamma_{s-t}$, $\eta_{st} = n^{-1}\left(F_s'\Lambda'u_t\right)$ and $\xi_{st} = n^{-1}\left(F_t'\Lambda'u_s\right)$. We omit $H$ and $V_{nT}$ whenever possible; note that they are both full rank matrices, with $\|H\| = O_p(1)$. We have

$$\begin{aligned}
\Delta\hat{F}_t - \Delta F_t &= T^{-2}\sum_{s=1}^{T}\hat{F}_sE\left(\frac{\Delta u_t'u_s}{n}\right) + T^{-2}\sum_{s=1}^{T}\hat{F}_s\left[\frac{\Delta u_t'u_s}{n} - E\left(\frac{\Delta u_t'u_s}{n}\right)\right] \\
&\quad + T^{-2}\sum_{s=1}^{T}\hat{F}_s\left(\frac{F_s'\Lambda'\Delta u_t}{n}\right) + T^{-2}\sum_{s=1}^{T}\hat{F}_s\left(\frac{\Delta F_t'\Lambda'u_s}{n}\right) \\
&= I + II + III + IV.
\end{aligned}$$

It holds that $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t - \Delta F_t\right\|^2 \leq MT^{-1}\sum_{t=1}^{T}\left(\|I\|^2 + \|II\|^2 + \|III\|^2 + \|IV\|^2\right)$. Let $\gamma_{\Delta s-t} = n^{-1}E\left(\Delta u_t'u_s\right)$; then $T^{-1}\sum_{t=1}^{T}\|I\|^2 \leq T^{-2}\left(T^{-2}\sum_{s=1}^{T}\left\|\hat{F}_s\right\|^2\right)\left(T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}\gamma_{\Delta s-t}^2\right) = O_p\left(T^{-2}\right)$ using Lemma B.1 in Bai (2004) and Assumption 1. Also, letting $\zeta_{\Delta st} = n^{-1}\left[\Delta u_t'u_s - E\left(\Delta u_t'u_s\right)\right]$, we have $T^{-1}\sum_{t=1}^{T}\|II\|^2 \leq T^{-5}\sum_{u=1}^{T}\sum_{s=1}^{T}\hat{F}_s'\hat{F}_u\left(\sum_{t=1}^{T}\zeta_{\Delta st}\zeta_{\Delta ut}\right) \leq T^{-2}\left(T^{-2}\sum_{s=1}^{T}\left\|\hat{F}_s\right\|^2\right)\left[T^{-2}\sum_{u=1}^{T}\sum_{s=1}^{T}\left(\sum_{t=1}^{T}\zeta_{\Delta st}\zeta_{\Delta ut}\right)^2\right]^{1/2} = O_p\left(n^{-1}T^{-2}\right)$ using the fact that $E\left(\sum_{t=1}^{T}\zeta_{\Delta st}\zeta_{\Delta ut}\right)^2 \leq T^2\max_{s,t}E\left|\zeta_{\Delta st}\right|^4$ and Assumption 1*(i)*. As far as $III$ is concerned, $T^{-1}\sum_{t=1}^{T}\|III\|^2 \leq n^{-1}\left(T^{-1}\sum_{t=1}^{T}\left\|\frac{\Lambda'\Delta u_t}{\sqrt{n}}\right\|^2\right)\left(T^{-2}\sum_{s=1}^{T}\left\|\hat{F}_s\right\|^2\right)\left(T^{-2}\sum_{s=1}^{T}\|F_s\|^2\right) = O_p\left(n^{-1}\right)$. Similar passages as above yield $T^{-1}\sum_{t=1}^{T}\|IV\|^2 \leq$

$n^{-1}T^{-1}\left(T^{-1}\sum_{t=1}^{T}\|\Delta F_t\|^2\right)\left(T^{-2}\sum_{s=1}^{T}\left\|\hat{F}_s\right\|^2\right)\left(T^{-1}\sum_{s=1}^{T}\left\|\frac{\Lambda' u_s}{\sqrt{n}}\right\|^2\right) = O_p\left(n^{-1}T^{-1}\right).$ Putting all together, we get A.1*(i)*.

Part A.1*(ii)* can be proved in a similar way. We can write

$$\begin{aligned}
& \frac{1}{T}\sum_{t=1}^{T}\left(\Delta\hat{F}_t - \Delta F_t\right)'\Delta F_t \\
= & \; T^{-3}\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{F}_s'\Delta F_t\left(\frac{\Delta u_t' u_s}{n}\right) + T^{-3}\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{F}_s'\Delta F_t\left(\frac{F_s'\Lambda'\Delta u_t}{n}\right) \\
& +T^{-3}\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{F}_s'\Delta F_t\left(\frac{\Delta F_t'\Lambda' u_s}{n}\right) \\
= & \; I + II + III.
\end{aligned}$$

Using Lemmas B.1 and B.4 in Bai (2004), we have $\|I\| \leq n^{-1}T^{-3/2}\left\|\sum_{i=1}^{n}\left(T^{-1}\sum_{s=1}^{T}\hat{F}_s u_{is}\right)\left(T^{-1/2}\sum_{t=1}^{T}\Delta F_t\Delta u_{it}\right)'\right\| = O_p\left(T^{-3/2}\right)$. Also, $\|II\| \leq n^{-1/2}T^{-1/2}\left\|\left(T^{-2}\sum_{s=1}^{T}\hat{F}_s F_s'\right)\left(n^{-1/2}T^{-1/2}\sum_{i=1}^{n}\sum_{t=1}^{T}\lambda_i'\Delta F_t\Delta u_{it}\right)'\right\| = O_p\left(n^{-1/2}T^{-1/2}\right)$. Last, $\|III\| \leq n^{-1/2}T^{-1}\left(n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\lambda_i'\hat{F}_s\Delta u_{is}\right)\left(T^{-1}\sum_{t=1}^{T}\Delta F_t\Delta F_t'\right) = O_p\left(n^{-1/2}T^{-1}\right)$. Combining these results, A.1*(ii)* follows. Equation A.1*(iii)* follows from A.1*(i)*-A.1*(ii)* upon noting that $T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{F}_t - \Delta F_t\right)'\Delta\hat{F}_t = T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{F}_t - \Delta F_t\right)'\Delta F_t + T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t - \Delta F_t\right\|^2$ $= O_p\left(T^{-1/2}C_{nT}^{-1}\right) + O_p\left(C_{nT}^{-2}\right)$. ■

**Lemma A.2** *Let Assumptions 1-4 hold. Then*

**A.2*(i)*** $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t - H'\Delta F_t\right\|^r = O_p\left(C_{nT}^{-r}\right)$,

**A.2*(ii)*** $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t\right\|^r = O_p(1)$,

**A.2*(iii)*** $T^{-1}\sum_{t=1}^{T}|\hat{u}_{it} - u_{it}|^r = O_p\left(\delta_{nT}^{-r}\right)$,

**A.2*(iv)*** $T^{-1}\sum_{t=1}^{T}|\hat{u}_{it}|^r = O_p(1)$ for $r \geq 2$.

**Proof.** We omit $H$ whenever possible. To prove part A.2*(i)*, note that

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t-\Delta F_t\right\|^r &\leq \frac{1}{T}\sum_{t=1}^{T}\left[\left\|\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_s\gamma_{\Delta s-t}\right\|^2\right]^{r/2}+\frac{1}{T}\sum_{t=1}^{T}\left[\left\|\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_s\zeta_{st}\right\|^2\right]^{r/2} \\
&+\frac{1}{T}\sum_{t=1}^{T}\left[\left\|\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_s\left(\frac{F_s'\Lambda'\Delta u_t}{n}\right)\right\|^2\right]^{r/2} \\
&+\frac{1}{T}\sum_{t=1}^{T}\left[\left\|\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_s\left(\frac{\Delta F_t'\Lambda' u_s}{n}\right)\right\|^2\right]^{r/2} \\
&= I+II+III+IV.
\end{aligned}$$

Consider $I$. The Cauchy-Schwartz inequality yields $I\leq T^{-r}\,\frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{T^2}\sum_{s=1}^{T}\left\|\hat{F}_s\right\|^2\right]^{r/2}$ $\left[\sum_{s=1}^{T}\gamma_{\Delta s-t}^2\right]^{r/2}$. Assumption 1 ensures that $\sum_{s=1}^{T}\gamma_{\Delta s-t}^2=O(1)$. Also, $T^{-2}\sum_{s=1}^{T}\left\|\hat{F}_s\right\|^2\leq$ $T^{-2}\sum_{s=1}^{T}\left\|F_s\right\|^2+T^{-2}\sum_{s=1}^{T}\left\|\hat{F}_s-F_s\right\|^2$, with $T^{-2}\sum_{s=1}^{T}\left\|\Delta\hat{F}_s-\Delta F_s\right\|^2=O_p\left(T^{-1}C_{nT}^{-2}\right)$ according to Lemma B.1 in Bai (2004). Given that $T^{-2}\sum_{s=1}^{T}\left\|F_s\right\|^2=O_p(1)$, it holds that $I=O_p\left(T^{-r}\right)$. As far as $II$ is concerned, note $II\leq\left[\frac{1}{T^2}\sum_{s=1}^{T}\left\|\hat{F}_s\right\|^2\right]^{r/2}T^{-1}\sum_{t=1}^{T}\left[\frac{1}{T}\sum_{s=1}^{T}\zeta_{\Delta st}^2\right]^{r/2}$. Since $T^{-1}\sum_{s=1}^{T}\zeta_{\Delta st}^2=O_p\left(n^{-1}\right)$ - see Bai (2003, p. 159) - we have $II=O_p\left(n^{-r/2}\right)$. Considering $III$, it holds that

$$\left\|\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_s\left(\frac{F_s'\Lambda'\Delta u_t}{n}\right)\right\|^r=n^{-r/2}\left\|\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_sF_s'\frac{\Lambda' u_t}{\sqrt{n}}\right\|^r=n^{-r/2}\left\|\frac{\Lambda' u_t}{\sqrt{n}}\right\|^r\left\|\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_sF_s'\right\|^r.$$

Note that $T^{-2}\sum_{s=1}^{T}\hat{F}_sF_s'=T^{-2}\sum_{s=1}^{T}F_sF_s'+T^{-2}\sum_{s=1}^{T}\left(\hat{F}_s-F_s\right)F_s'=O_p(1)+O_p\left(T^{-1}C_{nT}^{-1}\right)$ from Lemma B.1 in Bai (2004). Also, $n^{-r/2}\left\|n^{-1/2}\Lambda' u_t\right\|^r=n^{-r/2}\left\|n^{-1/2}\sum_{i=1}^{n}\lambda_iu_{it}\right\|^r=$ $O_p\left(n^{-r/2}\right)$ after Assumptions 2*(i)* and 3. Thus, $III=O_p\left(n^{-r/2}\right)$. Finally, $IV$ can be rearranged as

$$\begin{aligned}
\left\|\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_s\left(\frac{\Delta F_t'\Lambda' u_s}{n}\right)\right\|^r &= \left[\left\|\frac{1}{nT^2}\sum_{s=1}^{T}\hat{F}_su_s'\Lambda\Delta F_t\right\|^2\right]^{r/2} \\
&= n^{-r/2}\left[\left\|\frac{1}{T^2}\sum_{s=1}^{T}\sum_{i=1}^{n}\hat{F}_s\frac{u_{is}\lambda_i'}{\sqrt{n}}\Delta F_t\right\|^2\right]^{r/2} \\
&\leq n^{-r/2}T^{-r}\left[\left\|\Delta F_t\right\|^2\left\|\frac{1}{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\hat{F}_s\frac{u_{is}\lambda_i'}{\sqrt{n}}\right\|^2\right]^{r/2},
\end{aligned}$$

with $\|\Delta F_t\|^2 = O_p(1)$ and $\left\|(T)^{-1}\sum_{s=1}^{T}\hat{F}_s\left(n^{-1/2}\sum_{i=1}^{n}u_{is}\lambda_i'\right)\right\|^2 = O_p(1) + O_p\left(C_{nT}^{-1}\right)$ from Lemma B.4 in Bai (2004). Hence, $IV = O_p\left(n^{-r/2}T^{-r}\right)$. Thus, we have $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t - \Delta F_t\right\|^r = O_p(T^{-r}) + O_p(n^{-r}) + O_p(n^{-r}) + O_p\left(n^{-r/2}T^{-r}\right) = O_p\left(C_{nT}^{-r}\right)$. Part A.2*(ii)* follows from $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t\right\|^r \leq T^{-1}\sum_{t=1}^{T}\|\Delta F_t\|^r + T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t - \Delta F_t\right\|^r = O_p(1) + O_p\left(C_{nT}^{-r}\right)$. Finally, consider A.2*(iii)*. Since $\hat{u}_{it} = x_{it} - \hat{\lambda}_i'\hat{F}_t$, in light of (1) we have $\hat{u}_{it} - u_{it} = \lambda_i'F_t - \hat{\lambda}_i'\hat{F}_t$, and therefore

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}|\hat{u}_{it} - u_{it}|^r &= \frac{1}{T}\sum_{t=1}^{T}\left|\left(\lambda_i - \hat{\lambda}_i\right)' F_t - \hat{\lambda}_i'\left(\hat{F}_t - F_t\right)\right|^r \\
&\leq M\frac{1}{T}\sum_{t=1}^{T}\left|\left(\hat{\lambda}_i - \lambda_i\right)' F_t\right|^r + M\frac{1}{T}\sum_{t=1}^{T}\left|\hat{\lambda}_i'\left(\hat{F}_t - F_t\right)\right|^r \\
&= I + II.
\end{aligned}
$$

Consider $I$; we have $I \leq \left\|\hat{\lambda}_i - \lambda_i\right\|^r T^{-1}\sum_{t=1}^{T}\|F_t\|^r$. Note that $\hat{\lambda}_i - \lambda_i = O_p\left(T^{-1}\right)$ - see Lemma 3 in Bai (2004, p. 148). Also, Assumptions 1*(i)*, 1*(ii)*, 2*(i)* and 2*(ii)* ensure that $\sum_{t=1}^{T}\|F_t\|^r = O_p\left(T^{1+\frac{1}{2}r}\right)$ - see Park and Phillips (1999, Theorem 5.3). Thus, $I = O_p\left(T^{-\frac{1}{2}r}\right)$. As far as $II$ is concerned, $\sum_{t=1}^{T}\left|\hat{\lambda}_i'\left(\hat{F}_t - F_t\right)\right|^r \leq \left\|\hat{\lambda}_i\right\|^r\sum_{t=1}^{T}\left\|\hat{F}_t - F_t\right\|^r$. Since $\left\|\hat{\lambda}_i\right\|^r = \|\lambda_i + o_p(1)\|^r = O(1)$, similar calculations as before (based on the theory developed in Bai, 2004) would lead to $\sum_{t=1}^{T}\left\|\hat{F}_t - F_t\right\|^r = O_p\left(TC_{nT}^{-r}\right)$. Thus, $II = O_p\left(C_{nT}^{-r}\right)$ and therefore $T^{-1}\sum_{t=1}^{T}|\hat{u}_{it} - u_{it}|^r = O_p\left(\delta_{nT}^{-r}\right)$. Part A.2*(iv)* follows from similar calculations as for the proof of A.2*(ii)*. ■

**Lemma A.3** *Let Assumptions 1-4 hold. Then* $T^{-1}\sum_{t=1}^{T}\hat{u}_{it}^2 = T^{-1}\sum_{t=1}^{T}u_{it}^2 + O_p\left(C_{nT}^{-1}\right) + O_p\left(C_{nT}^{-2}\right)$.

**Proof.** It holds that

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}^2 &= \frac{1}{T}\sum_{t=1}^{T}u_{it}^2 + \frac{2}{T}\sum_{t=1}^{T}u_{it}(\hat{u}_{it} - u_{it}) + \frac{1}{T}\sum_{t=1}^{T}(\hat{u}_{it} - u_{it})^2 \\
&= \frac{1}{T}\sum_{t=1}^{T}u_{it}^2 + I + II.
\end{aligned}
$$

We have $I \leq 2\left[T^{-1}\sum_{t=1}^{T}u_{it}^2\right]^{1/2}\left[T^{-1}\sum_{t=1}^{T}|\hat{u}_{it} - u_{it}|^2\right]^{1/2} = O_p(1)\,O_p\left(C_{nT}^{-1}\right)$, in view of Assumption 1 and Lemma A.2*(iii)*. Also, A.2*(iii)* yields $II = O_p\left(C_{nT}^{-2}\right)$. Putting all together, the Lemma follows. ■

**Lemma A.4** *Let Assumptions 1-5 hold. Then, for some* $r \geq 2$

**A.4*(i)*** $T^{-1}\sum_{t=1}^{T}\left\|\hat{F}_{t,b}-H_1'F_{t,b}\right\|^r = O_p\left(C_{nT}^{-r}\right)$,

**A.4*(ii)*** $T^{-1}\sum_{t=1}^{T}\left(\hat{F}_{t,b}-H_1'F_{t,b}\right)F_{t,b}' = O_p\left(C_{nT}^{-1}\right)$,

**A.4*(iii)*** $T^{-1}\sum_{t=1}^{T}\left(\hat{F}_{t,b}-H_1'F_{t,b}\right)\hat{F}_{t,b}' = O_p\left(C_{nT}^{-1}\right)$,

**A.4*(iv)*** $\hat{F}_{t,b}-H_1'F_{t,b} = O_p\left(n^{-1/2}\right)+O_p\left(T^{-3/2}\right)$.

**Proof.** Prior to starting the proof, recall that $\hat{\lambda}_i - H^{-1}\lambda_i = O_p\left(T^{-1}\right)$. Also, $E\left\|\hat{F}_{t,b}\right\| = O\left(\sqrt{T}\right)$ by construction. Defining $\gamma_{n,|t-s|}^b = n^{-1}\sum_{i=1}^n E\left(u_{it,b}u_{is,b}\right)$, note that, by construction, $u_{it,b}$ is a stationary AR process of finite order; thus, $T^{-1}\sum_{t=1}^T\sum_{s=1}^T \gamma_{n,|t-s|}^b \leq M < \infty$. Also, under the "one unit at a time" bootstrap, the $u_{it,b}$s are independent across $i$ and $E\left|u_{it,b}u_{is,b}\right|^{2+\delta} \leq E\left|u_{it,b}\right|^{4+\delta}$, which is finite (uniformly in $i$) according to Lemma 1. Thus, $n^{-1/2}\sum_{i=1}^n\left(u_{it,b}u_{is,b}-\gamma_{n,|t-s|}^b\right) = O_p(1)$. Let $\hat{\Lambda} = [\lambda_1, ..., \lambda_n]$ and $u_{t,b} = [u_{1t,b}, ..., u_{nt,b}]'$. It holds that

$$V_{nT}^b\left(\hat{F}_{t,b}-H_1'F_{t,b}\right) = \frac{1}{T^2}\sum_{s=1}^T \hat{F}_{s,b}\gamma_{n,|t-s|}^b + \frac{1}{nT^2}\sum_{s=1}^T\sum_{i=1}^n \hat{F}_{s,b}\left(u_{t,b}'u_{s,b}-\gamma_{n,|t-s|}^b\right) + \frac{1}{nT^2}\sum_{s=1}^T \hat{F}_{s,b}F_{s,b}'\hat{\Lambda}'u_{t,b} + \frac{1}{nT^2}\sum_{s=1}^T \hat{F}_{s,b}u_{s,b}'\hat{\Lambda}F_{t,b}. \quad (55)$$

Consider A.4*(i)*: we show it for $r=2$; similar passages as in the proof of Lemma A.2*(i)* yield the result in the general case. It holds that

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^T\left\|\hat{F}_{t,b}-H_1'F_{t,b}\right\|^2 &\leq M\frac{1}{T}\sum_{t=1}^T\left\|\frac{1}{T^2}\sum_{s=1}^T\hat{F}_{s,b}\gamma_{n,|t-s|}^b\right\|^2 \\
&+M\frac{1}{T}\sum_{t=1}^T\left\|\frac{1}{nT^2}\sum_{s=1}^T\sum_{i=1}^n\hat{F}_{s,b}\left(u_{t,b}'u_{s,b}-\gamma_{n,|t-s|}^b\right)\right\|^2 \\
&+M\frac{1}{T}\sum_{t=1}^T\left\|\frac{1}{nT^2}\sum_{s=1}^T\hat{F}_{s,b}F_{s,b}'\hat{\Lambda}'u_{t,b}\right\|^2 + M\frac{1}{T}\sum_{t=1}^T\left\|\frac{1}{nT^2}\sum_{s=1}^T\hat{F}_{s,b}u_{s,b}'\hat{\Lambda}F_{t,b}\right\|^2 \\
&= M\left(I+II+III+IV\right),
\end{aligned}$$

for some $M<\infty$. Terms $I$ and $II$ are $O_p\left(T^{-2}\right)$ and $O_p\left(n^{-1}\right)$ respectively in light of Bai (2004), since Bai's assumptions on the summability of the $\gamma_{n,|t-s|}^b$s and on $n^{-1/2}\sum_{i=1}^n \left(u_{it,b}u_{is,b}-\gamma_{n,|t-s|}^b\right)$ being $O_p(1)$ hold here. Turning to $III$, we have $III \leq M\ \frac{1}{n^2T}$

$\sum_{t=1}^{T} \left\|\hat{\Lambda}' u_{t,b}\right\|^2 \left[\frac{1}{T^2}\sum_{s=1}^{T}\left\|\hat{F}_{s,b}\right\|^2\right] \left[\frac{1}{T^2}\sum_{s=1}^{T}\left\|F_{s,b}\right\|^2\right]$. Consider $\left\|\hat{\Lambda}' u_{t,b}\right\|^2$; note $\hat{\Lambda}' u_{t,b}$ $= \sum_{i=1}^{n} \hat{\lambda}_i u_{it,b}$. Conditional on the sample, the sequence $\left\{\hat{\lambda}_i u_{it,b}\right\}_{i=1}^{n}$ has mean zero, and in view of the "one unit at a time" resampling scheme, it is i.i.d. across $i$. Also, due to the independence between $\hat{\lambda}_i$ and $u_{it,b}$, $E\left\|\hat{\lambda}_i u_{it,b}\right\|^{2+\delta} = E\left\|\hat{\lambda}_i\right\|^{2+\delta} E\left|u_{it,b}\right|^{2+\delta}$, which is finite in view of Assumption 3*(i)*, Proposition 2 and Lemma 1. Thus, a CLT yields $\left\|\hat{\Lambda}' u_{t,b}\right\| = O_p(n)$. Hence, applying the bootstrap IP to $\left[\frac{1}{T^2}\sum_{s=1}^{T}\left\|\hat{F}_{s,b}\right\|^2\right]$ and $\left[\frac{1}{T^2}\sum_{s=1}^{T}\left\|F_{s,b}\right\|^2\right]$, we obtain $III = \frac{1}{n^2}O_p(n) = O_p\left(\frac{1}{n}\right)$. Similar passages yield $IV = O_p\left(\frac{1}{n}\right)$.

Consider part A.4*(ii)*. Omitting $V_{nT}^{b}$

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\left(\hat{F}_{t,b} - H_1' F_{t,b}\right) F_{t,b}' &= \frac{1}{T^3}\sum_{s=1}^{T}\sum_{t=1}^{T}\hat{F}_{s,b}F_{t,b}'\gamma_{n,|t-s|}^{b} + \frac{1}{T^3}\sum_{s=1}^{T}\sum_{t=1}^{T}\hat{F}_{s,b}F_{t,b}'\left(u_{t,b}'u_{s,b} - \gamma_{n,|t-s|}^{b}\right) \\
&\quad + \frac{1}{T^3}\sum_{s=1}^{T}\sum_{t=1}^{T}\hat{F}_{s,b}F_{s,b}'\frac{\hat{\Lambda}' u_{t,b}F_{t,b}'}{n} + \frac{1}{T^3}\sum_{s=1}^{T}\sum_{t=1}^{T}\frac{\hat{F}_{s,b}u_{s,b}'\hat{\Lambda}}{n}F_{t,b}F_{t,b}' \\
&= I + II + III + IV.
\end{aligned}
$$

As far as $I$ and $II$ are concerned, the same arguments as in Bai (2004) can be applied: $I$ and $II$ are $O_p\left(T^{-1}\right)$. Consider $III$; neglecting $\hat{F}_{s,b} - H_1' F_{s,b}$ which is dominated (after adding and subtracting), $III = \left(\frac{1}{T^2}\sum_{t=1}^{T}F_{t,b}F_{t,b}'\right)\left(\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\lambda}_i F_{t,b}' u_{it,b}\right)$. The term $\frac{1}{T^2}\sum_{t=1}^{T}F_{t,b}F_{t,b}'$ is $O_p(1)$ in light of the bootstrap IP; also $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}$ $\hat{\lambda}_i F_{t,b}' u_{it,b} = \frac{1}{nT}H^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\lambda_i F_{t,b}' u_{it,b} + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{\lambda}_i - H^{-1}\lambda_i\right)F_{t,b}' u_{it,b}$. The first term is $O_p\left(n^{-1/2}\right)$, see Bai (2004). Turning to the second term, it is bounded by $\left[\frac{1}{n}\sum_{i=1}^{n}\left\|\hat{\lambda}_i - H^{-1}\lambda_i\right\|^2\right]^{1/2}\left[\frac{1}{n}\sum_{i=1}^{n}\left\|\frac{1}{T}\sum_{t=1}^{T}F_{t,b}' u_{it,b}\right\|^2\right]^{1/2} = O_p\left(T^{-1}\right)$. Thus, $III = O_p\left(C_{nT}^{-1}\right)$. Finally, we turn to $IV = \left(\frac{1}{nT}\sum_{t=1}^{T}\hat{F}_{t,b}u_{t,b}'\hat{\Lambda}\right)\left(\frac{1}{T^2}\sum_{t=1}^{T}F_{t,b}F_{t,b}'\right)$. Consider the first term, the second being $O_p(1)$

$$
\begin{aligned}
\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{F}_{t,b}u_{t,b}'\hat{\Lambda} &= \frac{1}{nT}H_1'\sum_{t=1}^{T}F_{t,b}u_{t,b}'\Lambda H^{-1} + \frac{1}{nT}\sum_{t=1}^{T}\left(\hat{F}_{t,b} - H_1' F_{t,b}\right)u_{t,b}'\Lambda H^{-1} \\
&\quad + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}F_{t,b}u_{it,b}\left(\hat{\lambda}_i - H^{-1}\lambda_i\right)' \\
&\quad + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{F}_{t,b} - H_1' F_{t,b}\right)u_{it,b}\left(\hat{\lambda}_i - H^{-1}\lambda_i\right)'.
\end{aligned}
$$

The first two terms are $O_p\left(n^{-1/2}\right)$ and $O_p\left(n^{-1/2}C_{nT}^{-1}\right)$ respectively - see Bai (2004). The third term is bounded by $\left[\frac{1}{n}\sum_{i=1}^{n}\left\|\hat{\lambda}_i - H^{-1}\lambda_i\right\|^2\right]^{1/2}\left[\frac{1}{n}\sum_{i=1}^{n}\left\|\frac{1}{T}\sum_{t=1}^{T}F_{t,b}u_{it,b}\right\|^2\right]^{1/2}$ as above, and thus it is $O_p\left(T^{-1}\right)$. As far as the fourth one is concerned, it is bounded by $\left[\frac{1}{n}\sum_{i=1}^{n}\left\|\hat{\lambda}_i - H^{-1}\lambda_i\right\|^2\right]^{1/2}\left[\frac{1}{n}\sum_{i=1}^{n}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\hat{F}_{t,b} - H_1'F_{t,b}\right)u_{it,b}\right\|^2\right]^{1/2}$, which can be shown to be $O_p\left(T^{-1}C_{nT}^{-1}\right)$ using similar passages as above. Putting all together, part A.4*(ii)* follows. Also, part A.4*(iii)* follows from A.4*(i)* and A.4*(ii)*.

Turning to A.4*(iv)*, consider (55). The terms $I$ and $II$ have the same asymptotics as in Bai (2004), since the assumptions are the same. Thus, $I = O_p\left(T^{-3/2}\right)$ and $II = O_p\left(n^{-1/2}T^{-1/2}\right)$. As far as $III$ is concerned, recall that $\hat{\Lambda}'u_{t,b} = O_p\left(\sqrt{n}\right)$; thence, given that $\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_{s,b}F_{s,b}' = O_p(1)$, $III = O_p\left(n^{-1/2}\right)$. Finally, consider $IV$; the same arguments as in Bai (2004) yield $IV = O_p\left(n^{-1/2}T^{-1/2}\right)$. Putting all together, A.4*(iv)* follows. ■

**Lemma A.5** *Consider (50). Under Assumption 6, it holds that, for some* $r \geq 2$

**A.5*(i)*** $T^{-1}\sum_{t=1}^{T}\left\|\hat{K}_t - H'K_t\right\|^r = O_p\left(\delta_{nT}^{-r}\right)$,

**A.5*(ii)*** $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t - H_F'\Delta F_t\right\|^r = O_p\left(\delta_{nT}^{-r}\right)$,

**A.5*(iii)*** $T^{-1}\sum_{t=1}^{T}\left(\hat{K}_t - H'K_t\right)K_t' = O_p\left(\delta_{nT}^{-1}\right)$,

**A.5*(iv)*** $T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{F}_t - H_F'\Delta F_t\right)\Delta F_t' = O_p\left(T^{-1/2}\delta_{nT}^{-1}\right)$,

**A.5*(v)*** $T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{F}_t - H_F'\Delta F_t\right)\Delta\hat{F}_t' = O_p\left(\delta_{nT}^{-2}\right)$,

**A.5*(vi)*** $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t\right\|^r = O_p(1)$,

**A.5*(vii)*** $T^{-1}\sum_{t=1}^{T}\left\|\hat{G}_t\right\|^r = O_p(1)$,

**A.5*(viii)*** $T^{-1}\sum_{t=1}^{T}\left|\hat{u}_{it} - u_{it}\right|^r = O_p\left(\delta_{nT}^{-r}\right)$,

**A.5*(ix)*** $T^{-1}\sum_{t=1}^{T}\left|\hat{u}_{it}\right|^r = O_p(1)$,

**A.5*(x)*** $T^{-1}\sum_{t=1}^{T}\hat{u}_{it}^2 = \frac{1}{T}\sum_{t=1}^{T}u_{it}^2 + O_p\left(\delta_{nT}^{-1}\right)$.

**Proof.** Most of the passages in the proof are similar to the other proofs; thus, some of them are omitted. Consider the following notation: let $u$ be the $T \times n$ matrix defined as $[u_1, ..., u_n]$ with $u_i = [u_{i1}, ..., u_{iT}]'$; define $D = diag\left\{TI_k, \sqrt{T}I_h\right\}$; let $\Upsilon = E\left(\frac{uu'}{n}\right)$ and $\Xi = \frac{uu'}{n} - \Upsilon$. It holds that (see Maciejowska, 2010)

$$\hat{K} - KH = \left(\frac{1}{n}u\Lambda K' + \Xi + \Upsilon\right)\tilde{K}D^{-2}, \tag{56}$$

$$\|\Upsilon\| = O_p\left(\sqrt{T}\right), \tag{57}$$

$$\|\Xi\| = O_p\left(\frac{T}{\sqrt{n}}\right), \tag{58}$$

with, $\tilde{K} = \hat{K}V_{nT}^{-1}$, $\left\|\tilde{K}'\tilde{K}D^{-2}\right\| = O_p(1)$ and $\left\|K'\tilde{K}D^{-2}\right\| = O_p(1)$.

We show parts *(i)* and *(ii)* for $r = 2$; the proof for general $r$ can be adapted from the proof of Lemma A.2. Consider A.5*(i)*. Using the $C_r$-inequality, $T^{-1}\left\|\hat{K} - KH\right\|^2 \leq M$ $\left(T^{-1}\left\|\frac{1}{n}u\Lambda K'\tilde{K}D^{-2}\right\|^2 + T^{-1}\left\|\Xi\tilde{K}D^{-2}\right\|^2 + T^{-1}\left\|\Upsilon\tilde{K}D^{-2}\right\|^2\right)$. It holds that $T^{-1}\left\|\frac{1}{n}u\Lambda K'\tilde{K}D^{-2}\right\|^2$ $\leq (nT)^{-1}\left\|\frac{1}{\sqrt{n}}u\Lambda\right\|^2\left\|K'\tilde{K}D^{-2}\right\|^2 = (nT)^{-1}O_p(T)O_p(1) = O_p\left(\frac{1}{n}\right)$. Also, using (57), $T^{-1}\left\|\Xi\tilde{K}D^{-2}\right\|^2 \leq T^{-1}\|\Xi\|^2\left\|\tilde{K}D^{-1}\right\|^2\left\|D^{-1}\right\|^2 = T^{-1}\ O_p\left(\frac{T^2}{n}\right)\ O_p(1)\ O_p\left(\left\|D^{-1}\right\|^2\right) = O_p\left(\frac{T}{n}\left\|D^{-1}\right\|^2\right)$. Finally, $T^{-1}\left\|\Upsilon\tilde{K}D^{-2}\right\|^2 \leq T^{-1}\|\Upsilon\|^2\left\|\tilde{K}D^{-1}\right\|^2\left\|D^{-1}\right\|^2 = T^{-1}\ O_p(T)$ $O_p(1)\,O_p\left(\left\|D^{-1}\right\|^2\right) = O_p\left(\left\|D^{-1}\right\|^2\right)$. Putting all together, $T^{-1}\left\|\hat{K}_t - H'K_t\right\|^2 = O_p\left(\frac{1}{n}\right) + O_p\left(\frac{T}{n}\left\|D^{-1}\right\|^2\right) + O_p\left(\left\|D^{-1}\right\|^2\right)$, whence, by definition of $D$, A.5*(i)*. As far as A.5*(ii)* is concerned, similarly to the proof of Lemma A.1, $T^{-1}\left\|\Delta\hat{F} - \Delta FH_F\right\|^2 \leq T^{-1}\left\|\hat{F}_t - H_F'F_t\right\|^2$ $+T^{-1}\left\|\hat{F}_{t-1} - H_F'F_{t-1}\right\|^2$, whence the desired result.

Turning to A.5*(iii)*, using (56) it holds that $T^{-1}K'\left(\hat{K} - KH\right) = \frac{1}{nT}K'K\Lambda'u'\tilde{K}D^{-2}$ $+\frac{1}{T}K'\Xi\tilde{K}D^{-2} + \frac{1}{T}K'\Upsilon\tilde{K}D^{-2} = I + II + III$. We have $I \leq \frac{1}{nT}\|K'K\|\,\|\Lambda'u\|\left\|\tilde{K}D^{-1}\right\|$ $\left\|D^{-1}\right\| = \frac{1}{nT}\ O_p\left(\|D\|^2\right)\ O_p\left(\sqrt{nT}\right)\ O_p(1)\ O_p\left(\left\|D^{-1}\right\|^2\right) = O_p\left(\frac{1}{\sqrt{nT}}\right)$. Also, $II = T^{-1}D^{-1}K'\Xi\tilde{K}D^{-1} + T^{-1}D^{-1}K'\Xi\left(\tilde{K} - KH\right)D^{-1} = II_a + II_b$. It holds that $II_a \leq T^{-1}\left\|D^{-1}K'\right\|\,\|\Xi\|\left\|\tilde{K}D^{-1}\right\|$; this has the same order as $T^{-1}\|\Xi\| = O_p\left(\frac{1}{\sqrt{n}}\right)$. Similarly, $II_b \leq T^{-1}\left\|D^{-1}K'\right\|\,\|\Xi\|\left\|\tilde{K} - KH\right\|\left\|D^{-1}\right\| = T^{-1}O_p(1)\ O_p\left(\frac{T}{\sqrt{n}}\right)\ O_p\left(\sqrt{T}\delta_{nT}^{-1}\right)$ $O_p\left(\left\|D^{-1}\right\|\right) = O_p\left(\sqrt{\frac{T}{n}}\delta_{nT}^{-1}\left\|D^{-1}\right\|\right)$, by virtue of A.5*(i)*, which is dominated. Thus, $II = O_p\left(\frac{1}{\sqrt{n}}\right)$. Finally, $III = T^{-1}D^{-1}K'\Upsilon\tilde{K}D^{-1} + T^{-1}D^{-1}K'\Upsilon\left(\tilde{K} - KH\right)D^{-1} = III_a + III_b$. It holds that $III_a \leq T^{-1}\left\|D^{-1}K'\right\|\,\|\Upsilon\|\left\|\tilde{K}D^{-1}\right\| = T^{-1}O_p(1)\ O_p\left(\sqrt{T}\right)$ $O_p(1) = O_p\left(\frac{1}{\sqrt{T}}\right)$. Also, $III_b \leq T^{-1}\left\|D^{-1}K'\right\|\,\|\Upsilon\|\left\|\tilde{K} - KH\right\|\left\|D^{-1}\right\| = T^{-1}O_p(1)$ $O_p\left(\sqrt{T}\right)\ O_p\left(\sqrt{T}\delta_{nT}^{-1}\right)\ O_p\left(\left\|D^{-1}\right\|\right) = O_p\left(\delta_{nT}^{-1}\left\|D^{-1}\right\|\right)$, which is dominated in light of the definition of $D$. Putting all together, A.5*(iii)* follows.

Turning to A.5*(iv)*, we show it by considering $T^{-1}\left(\Delta\hat{K}-\Delta KH\right)\Delta K'$ first. Using (56) $T^{-1}\Delta K'\left(\Delta\hat{K}-\Delta KH\right)=\frac{1}{nT}\Delta K'\Delta u\Lambda K'\tilde{K}D^{-2}+T^{-1}\Delta K'\widetilde{\Xi}\tilde{K}D^{-2}+T^{-1}\Delta K'\widetilde{\Upsilon}\tilde{K}D^{-2}$ $=I+II+III$, with $\widetilde{\Upsilon}=E\left(\frac{\Delta uu'}{n}\right)$ and $\widetilde{\Xi}=\frac{\Delta uu'}{n}-\widetilde{\Upsilon}$. Hence, $I\leq\frac{1}{nT}\left\|\Delta K'\Delta u\Lambda\right\|$ $\left\|K'\tilde{K}D^{-2}\right\|=\frac{1}{nT}O_p\left(\sqrt{nT}\right)O_p\left(1\right)=O_p\left(\frac{1}{\sqrt{nT}}\right)$; also, $II\leq\frac{1}{T}\left\|\Delta K\right\|\left\|\widetilde{\Xi}\right\|\left\|\tilde{K}D^{-1}\right\|$ $\left\|D^{-1}\right\|=\frac{1}{T}O_p\left(\sqrt{T}\right)O_p\left(\sqrt{T}\right)O_p\left(1\right)O_p\left(\left\|D\right\|^{-1}\right)=O_p\left(\frac{1}{\|D\|}\right)$. Finally, $III\leq\frac{1}{T}$ $\left\|\Delta K\right\|\left\|\widetilde{\Upsilon}\right\|\left\|\tilde{K}D^{-1}\right\|\left\|D^{-1}\right\|=\frac{1}{T}O_p\left(\sqrt{T}\right)O_p\left(\frac{T}{\sqrt{n}}\right)O_p\left(1\right)O_p\left(\left\|D\right\|^{-1}\right)=O_p\left(\sqrt{\frac{T}{n}}\frac{1}{\|D\|}\right)$. Thus, $T^{-1}\left(\Delta\hat{K}-\Delta KH\right)\Delta K'=O_p\left(\frac{1}{\sqrt{nT}}\right)+O_p\left(\frac{1}{\|D\|}\right)+O_p\left(\sqrt{\frac{T}{n}}\frac{1}{\|D\|}\right)$; A.5*(iv)* follows from noting that, when considering $T^{-1}\left(\Delta\hat{F}-\Delta FH_F\right)\Delta F'$, $\|D\|=O\left(T\right)$. Also, A.5*(v)* follows from combining A.5*(ii)* and A.5*(iv)*. Similarly, A.5*(vi)* and A.5*(vii)* follow from A.5*(iv)* and A.5*(ii)* respectively, using Assumption 6.

As far as A.5*(viii)* is concerned, the proof is similar to Lemma A.2*(iii)*: it holds that $T^{-1}\sum_{t=1}^{T}\left|\hat{u}_{it}-u_{it}\right|^r\leq MT^{-1}\sum_{t=1}^{T}\left|\left(\hat{\lambda}_i-H^{-1}\lambda_i\right)'K_t\right|^r+MT^{-1}\sum_{t=1}^{T}\left|\hat{\lambda}_i'\left(K_t-H'K_t\right)\right|^r=$ $I+II$. Further, Theorem 6 in Bai (2004) states that $\hat{\lambda}_i-H^{-1}\lambda_i=O_p\left(\left\|D\right\|^{-1}\right)$ - i.e. $\hat{\lambda}_i$ is $\sqrt{T}$ or $T$ consistent according as it estimates the space spanned by $\lambda_i^G$ or $\lambda_i^F$. Consider $I$; it holds that (omitting $M$) $I\leq T^{-1}\left\|\hat{\lambda}_i-H^{-1}\lambda_i\right\|^r\sum_{t=1}^{T}\left\|K_t\right\|^r$; $\sum_{t=1}^{T}\left\|K_t\right\|^r$ is $O_p\left(1\right)$ or $O_p\left(T^r\right)$ according as $G_t$ or $F_t$ is considered. This yields $I=O_p\left(T^{-\frac{1}{2}r}\right)$. Also, omitting $M$, $II\leq\left\|\hat{\lambda}_i\right\|^rT^{-1}\sum_{t=1}^{T}\left\|K_t-H'K_t\right\|^r$; A.5*(ii)* entails that $II=O_p\left(\delta_{nT}^{-r}\right)$. This proves A.5*(viii)*. Equations A.5*(ix)* and A.5*(x)* follow from the same passages as in the proof of Lemma A.2*(iv)* and Lemma A.3. ■

# Appendix B: proofs and derivations

**Proof of Proposition 1.** Consider (6). It holds that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}\left[\begin{array}{c}\Delta\hat{F}_t\\ u_{it}\end{array}\right]=\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}\left[\begin{array}{c}H'\Delta F_t\\ u_{it}\end{array}\right]+\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}\left[\begin{array}{c}\Delta\hat{F}_t-H'\Delta F_t\\ 0\end{array}\right]=I+II.$$

The weak convergence to a Brownian motion of $I$ is a standard result; a detailed proof can be found in Phillips and Solo (1992). As far as $II$ is concerned, $T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\Delta\hat{F}_t-H'\Delta F_t\right)$ $\leq M\left[T^{-1}\sum_{t=1}^{\lfloor Ts\rfloor}\left\|\Delta\hat{F}_t-H'\Delta F_t\right\|^2\right]^{1/2}$; this is $O_p\left(C_{nT}^{-1}\right)$ in view of Lemma A.1*(i)*, and it holds uniformly in $s$. Thus, (6) holds. As far as equation (7) is concerned, $T^{-2}\sum_{t=1}^{T}\hat{F}_t\hat{F}_t'=$

$H'\left(T^{-2}\sum_{t=1}^{T}F_tF_t'\right)H+H'\left[T^{-2}\sum_{t=1}^{T}F_t\left(\hat{F}_t-H'F_t\right)'\right]+\left[T^{-2}\sum_{t=1}^{T}\left(\hat{F}_t-H'F_t\right)F_t'\right]H+$ $T^{-2}\sum_{t=1}^{T}\left(\hat{F}_t-H'F_t\right)\left(\hat{F}_t-H'F_t\right)'=I+II+III+IV$. Term $I$ converges to $H'\left(\int W_\varepsilon W_\varepsilon'\right)H$; see Phillips and Solo (1992) for details. As far as $II$ and $III$ are concerned, using Lemma B.4 in Bai (2004, p. 171), $II=O_p\left(T^{-1}C_{nT}^{-1}\right)$ and similarly $III$. Lemma in B.1 Bai (2004, p. 167) also entails that $IV=O_p\left(T^{-1}C_{nT}^{-2}\right)$. Turning to (8), $T^{-1}\sum_{t=1}^{T}\hat{F}_tu_{it}=$ $H'\left(T^{-1}\sum_{t=1}^{T}F_tu_{it}\right)+T^{-1}\sum_{t=1}^{T}\left(\hat{F}_t-H'F_t\right)u_{it}=I+II$. Convergence of $I$ to $H'\int W_\varepsilon dW_{u,i}$ is a standard result; as far as $II$ is concerned, $II\leq\left[T^{-1}\sum_{t=1}^{T}\left\|\hat{F}_t-H'F_t\right\|^2\right]^{1/2}\left[T^{-1}\sum_{t=1}^{T}u_{it}^2\right]^{1/2}$ $=O_p\left(C_{nT}^{-1}\right)O_p(1)$, which is negligible. This proves (8).

Consider (9). Let the martingale approximation of $u_{it}$ (derived from the Beveridge-Nelson decomposition) be $u_{it}^*$. This is a martingale difference sequence (MDS) with variance $\sigma_{u,i}^2$; it holds that

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it}=\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it}^*+\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}R_{u,it}=I+II,$$

where $R_{u,it}$ is defined as $u_{it}^*-u_{it}$. Standard panel asymptotic arguments (Phillips and Moon, 1999) yield that $II=O_p\left(\sqrt{\frac{n}{T}}\right)$. As far as $I$ is concerned, define $\zeta_{nt}=n^{-1/2}\sum_{i=1}^{n}u_{it}^*$. The process $\zeta_{nt}$ has mean zero and is an MDS for every $n$: that $n$ passes to infinity is merely incidental. Also, consider $E\left|\zeta_{nt}\right|^{2+\delta}$. We have $E\left|\zeta_{nt}\right|^{2+\delta}\leq n^{-(1+\delta/2)}\sum_{i=1}^{n}E\left|u_{it}^*\right|^{2+\delta}\leq$ $n^{-\delta/2}\max_i E\left|u_{it}^*\right|^{2+\delta}$. Thus, in view of Assumption 1, $E\left|\zeta_{nt}\right|^{2+\delta}<\infty$ uniformly in $n$. This entails that an IP for MDS (see e.g. Theorem 4.1 in Hall and Heyde, 1980) can be applied: $T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\zeta_{nt}$ converges uniformly to a Brownian motion with variance

$$\lim_{(n,T)\to\infty}E\left(\zeta_{nt}^2\right)=\lim_{(n,T)\to\infty}\frac{1}{nT}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{t=1}^{T}E\left(u_{it}^*u_{jt}^*\right)=\lim_{n\to\infty}\frac{i_n'\Gamma(1)\Sigma_u\Gamma'(1)i_n}{n}=\sigma_u^2,$$

where the last equality holds by definition of $\sigma_u^2$; Assumption 1*(iii)* ensures that $\sigma_u^2<\infty$.

Finally, consider (10). We have $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{F}_tu_{it}=\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}H'F_tu_{it}+\frac{1}{T}\sum_{t=1}^{T}\left(\hat{F}_t-H'F_t\right)$ $\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}u_{it}\right)=I+II$. Using the Cauchy-Schwartz inequality, $II\leq\left[\frac{1}{T}\sum_{t=1}^{T}\left\|\hat{F}_t-H'F_t\right\|^2\right]^{1/2}$ $\left[\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}u_{it}\right)^2\right]^{1/2}$, which is $O_p\left(C_{nT}^{-1}\right)$ in view of Lemma B.1 in Bai (2004). As far as $I$ is concerned, let the martingale approximations to $F_t$ (from the Beveridge-Nelson decomposition) be $F_t^*$. Then

$$\begin{aligned}
\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}F_t u_{it} &= \frac{1}{T}\sum_{t=1}^{T}F_t^*\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}u_{it}^*\right)+\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}R_{it} \\
&= \frac{1}{T}\sum_{t=1}^{T}F_t^*\zeta_{nt}+\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}R_{it}=I_a+I_b,
\end{aligned}$$

where $R_{it} = F_t^* u_{it}^* - F_t u_{it}$. As shown above, an IP holds for $T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\zeta_{nt}$ and for $T^{-1/2}F_t^*$; also, $\zeta_{nt}$ and $F_t^*$ are independent for all $n$ in light of Assumption 4*(i)*. Thus, standard arguments in the theory of convergence to stochastic integrals (see e.g. Phillips, 1988) yield $I_a \xrightarrow{d} \sigma_u \int W_\varepsilon dW_u$. Finally, from Phillips and Moon (1999), it can be proved that $I_b = O_p\left(\sqrt{\frac{n}{T}}\right)$. Putting all together, (10) follows as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$. ■

**Proof of Lemma 1.** For simplicity, we suppress the subscripts in $q_F$ and $q_{u,i}$ whenever possible. Consider (14); recall (11) and

$$\Delta\hat{F}_t = \sum_{j=1}^{q}\hat{\alpha}_{q,j}\Delta\hat{F}_{t-j}+\hat{e}_{t,q}^F, \tag{59}$$

$$\Delta F_t = \sum_{j=1}^{\infty}\alpha_j\Delta F_{t-j}+e_t^F. \tag{60}$$

Using the definition of $\left\{e_{t,b}^F\right\}_{t=1}^T$,

$$\begin{aligned}
E^b\left\|e_{t,b}^F\right\|^r &= \frac{1}{T}\sum_{t=1}^{T}\left[\hat{e}_{t,q}^F-\frac{1}{T}\sum_{t=1}^{T}\hat{e}_{t,q}^F\right]^r \\
&\leq \|H\|^r\frac{1}{T}\sum_{t=1}^{T}\left\|e_t^F\right\|^r+\|H\|^r\frac{1}{T}\sum_{t=1}^{T}\left\|e_{t,q}^F-e_t^F\right\|^r+\frac{1}{T}\sum_{t=1}^{T}\left\|\hat{e}_{t,q}^F-H'e_{t,q}^F\right\|^r+\left\|\frac{1}{T}\sum_{t=1}^{T}\hat{e}_{t,q}^F\right\|^r \\
&= I+II+III+IV.
\end{aligned}$$

Assumptions 2 and 3 entail that $\|H\| = O_p(1)$. Consider $I$; Assumption 2*(i)* and the Law of Large Numbers (LLN) ensure that $\frac{1}{T}\sum_{t=1}^{T}\left\|e_t^F\right\|^r \xrightarrow{p} E\left\|e_t^F\right\|^r < \infty$. As far as $II$ is concerned, it holds that $e_{t,q}^F - e_t^F = \sum_{j=q+1}^{\infty}\alpha_j\Delta F_{t-j}$ and therefore Minkowski's inequality and the stationarity of $\Delta F_t$ yield

$$\frac{1}{T}\sum_{t=1}^{T}\left\|e_{t,q}^F-e_t^F\right\|^r=\frac{1}{T}\sum_{t=1}^{T}\left\|\sum_{j=q+1}^{\infty}\alpha_j\Delta F_{t-j}\right\|^r\leq\frac{1}{T}\sum_{t=1}^{T}\left\|\Delta F_t\right\|^r\left(\sum_{j=q+1}^{\infty}\left\|\alpha_j\right\|\right)^r.$$

Assumption 1*(ii)* entails that $\sum_{j=q+1}^{\infty} \|\alpha_j\| = o\left(q^{-s}\right)$; Assumption 2*(i)* and the LLN yield $T^{-1}\sum_{t=1}^{T} \|\Delta F_t\|^r = O_p(1)$. Thus, $II = o_p\left(q^{-rs}\right)$. As far as $III$ is concerned, we have $\hat{e}_{t,q}^F - H' e_{t,q}^F = \sum_{j=0}^{q} H'\alpha_{q,j}\left(H'\right)^{-1}\left(\Delta\hat{F}_{t-j} - H'\Delta F_{t-j}\right) - \sum_{j=1}^{q}\left[\hat{\alpha}_{q,j} - H'\alpha_{q,j}\left(H'\right)^{-1}\right]\Delta\hat{F}_{t-j}$, where $\alpha_{q,0} = -1$. Hence

$$
\begin{aligned}
III &\leq \frac{1}{T}\sum_{t=1}^{T}\left\|\sum_{j=0}^{q} H'\alpha_{q,j}\left(H'\right)^{-1}\left(\Delta\hat{F}_{t-j} - H'\Delta F_{t-j}\right)\right\|^r + \frac{1}{T}\sum_{t=1}^{T}\left\|\sum_{j=1}^{q}\left[\hat{\alpha}_{q,j} - H'\alpha_{q,j}\left(H'\right)^{-1}\right]\Delta\hat{F}_{t-j}\right\|^r \\
&= III_a + III_b.
\end{aligned}
$$

Using Minkowski's inequality, $III_a \leq MT^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t - \Delta F_t\right\|^r\left(\sum_{j=0}^{q}\|\alpha_{q,j}\|\right)^r$, with $\sum_{j=0}^{q}\|\alpha_{q,j}\| \leq \sum_{j=0}^{\infty}\|\alpha_j\| = O(1)$. Also, $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t - H'\Delta F_t\right\|^r = O_p\left(C_{nT}^{-r}\right)$ according to Lemma A.2*(i)*. Thus, $III_a = O_p\left(C_{nT}^{-r}\right)$. As far as $III_b$ is concerned, $III_b \leq T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t\right\|^r\left(\sum_{j=0}^{q}\left\|\hat{\alpha}_{q,j} - H'\alpha_{q,j}\left(H'\right)^{-1}\right\|\right)^r$. Lemma A.2 ensures $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t\right\|^r = O_p(1)$. Also, $\sum_{j=0}^{q}\left\|\hat{\alpha}_{q,j} - H'\alpha_{q,j}\left(H'\right)^{-1}\right\| \leq q\max_{1\leq j\leq q}\left\|\hat{\alpha}_{q,j} - H'\alpha_{q,j}\left(H'\right)^{-1}\right\|$, and Lemma 2 yields $\left[q\max_{1\leq j\leq q}\left\|\hat{\alpha}_{q,j} - H'\alpha_{q,j}\left(H'\right)^{-1}\right\|\right]^r = O_p\left[q^r T^{-r/2}\left(\log T\right)^{r/2} + q^r n^{-r}\right] + o_p(1)$. Thus, $III = O_p\left(C_{nT}^{-r}\right) + O_p\left(q^r\varphi_{nT}^F\right)$. Finally, consider $IV$; we have $\hat{e}_{t,q}^F = -\sum_{j=0}^{q}\hat{\alpha}_{q,j}\Delta\hat{F}_{t-j}$ with $\hat{\alpha}_{q,0} = -1$. Thus

$$
-\frac{1}{T}\sum_{t=1}^{T}\hat{e}_{t,q}^F = \sum_{j=0}^{q} H'\alpha_{q,j}\left(H'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\Delta\hat{F}_{t-j}\right) + \sum_{j=0}^{q}\left[\hat{\alpha}_{q,j} - H'\alpha_{q,j}\left(H'\right)^{-1}\right]\left(\frac{1}{T}\sum_{t=1}^{T}\Delta\hat{F}_{t-j}\right) = IV_a + IV_b.
$$

Since $T^{-1}\sum_{t=1}^{T}\Delta\hat{F}_{t-j} = T^{-1}H'\sum_{t=1}^{T}\Delta F_{t-j} + T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{F}_{t-j} - H'\Delta F_{t-j}\right) = O_p\left(T^{-1/2}\right) + o_p\left(T^{-1/2}\right)$ for all $j$s

$$
\begin{aligned}
IV_a &\leq M\left(\sum_{j=0}^{q}\|\alpha_{q,j}\|^2\right)^{1/2}\left(\sum_{j=0}^{q}\left\|\frac{1}{T}\sum_{t=1}^{T}\Delta\hat{F}_{t-j}\right\|^2\right)^{1/2} \\
&\leq O(1)\left[q\max_{1\leq j\leq q}\left\|\frac{1}{T}\sum_{t=1}^{T}\Delta\hat{F}_{t-j}\right\|^2\right]^{1/2} = O_p\left(\sqrt{\frac{q}{T}}\right);
\end{aligned}
$$

also, $IV_b \leq \left(\sum_{j=0}^{q}\left\|\hat{\alpha}_{q,j} - H'\alpha_{q,j}\left(H'\right)^{-1}\right\|^2\right)^{1/2}\left(\sum_{j=0}^{q}\left\|\frac{1}{T}\sum_{t=1}^{T}\Delta\hat{F}_{t-j}\right\|^2\right)^{1/2} \leq (q\max_{1\leq j\leq q} \left\|\hat{\alpha}_{q,j} - H'\alpha_{q,j}\left(H'\right)^{-1}\right\|^2\Big)^{1/2} O_p\left(\sqrt{\frac{q}{T}}\right)$, and thus it is dominated. Lemma 2 yields $\max_{1\leq j\leq q}\|\hat{\alpha}_{q,j} - H'\alpha_{q,j}\left(H'\right)^{-1}\Big\|^2 = O_p\left(\varphi_{nT}^F\right) + O_p\left(n^{-3/2}T^{-1/2}\right)$. Hence, $IV = O_p\left(q^{r/2}T^{-r/2}\right)$. Combining all these results

$$E^b \left\|e^F_{t,b}\right\|^r = E\left\|e^F_t\right\|^r + O_p\left(q^{-rs}\right) + O_p\left(C_{nT}^{-r}\right) + O_p\left[\left(\frac{q}{\varphi^F_{nT}}\right)^r\right] + o_p(1) = E\left\|e^F_t\right\|^r + o_p(1),$$

thus Assumption 5 ensures $E^b\left\|e^F_{t,b}\right\|^r = E\left\|e^F_t\right\|^r + o_p(1)$. Also

$$\begin{aligned} T^{1-\frac{1}{2}r}E^b\left\|e^F_{t,b}\right\|^r &= O_p\left(T^{1-\frac{1}{2}r}\right) + O_p\left(T^{1-\frac{1}{2}r}q^{-rs}\right) + O_p\left(T^{1-\frac{1}{2}r}q^r\varphi^F_{nT}\right) + \\ & \quad O_p\left(T^{1-\frac{1}{2}r}C_{nT}^{-r}\right) + O_p\left(T^{1-r}q^{\frac{1}{2}r}\right) + o_p(1). \end{aligned}$$

Thus, $T^{1-\frac{1}{2}r}E^b\left\|e^F_{t,b}\right\|^r = o_p(1)$ for $r > 2$.

As far as (15) is concerned, recall that $u_{it} = \sum_{j=1}^q \gamma^{(i)}_{q,j}u_{it-j} + e^u_{it,q}$, and let

$$\hat{u}_{it} = \sum_{j=1}^q \hat{\gamma}^{(i)}_{q,j}\hat{u}_{it-j} + \hat{e}^u_{it,q}, \tag{61}$$

$$u_{it} = \sum_{j=1}^\infty \gamma^{(i)}_j u_{it-j} + e^u_{it}. \tag{62}$$

We have

$$\begin{aligned} E^b\left|e^u_{it,b}\right|^r &= \frac{1}{T}\sum_{t=1}^T\left[\hat{e}^u_{it,q} - \frac{1}{T}\sum_{t=1}^T\hat{e}^u_{it,q}\right]^r \\ &\leq \frac{1}{T}\sum_{t=1}^T\left|e^u_{it}\right|^r + \frac{1}{T}\sum_{t=1}^T\left|e^u_{it,q} - e^u_{it}\right|^r + \frac{1}{T}\sum_{t=1}^T\left|\hat{e}^u_{it,q} - e^u_{it,q}\right|^r + \left|\frac{1}{T}\sum_{t=1}^T\hat{e}^u_{it,q}\right|^r \\ &= I + II + III + IV. \end{aligned}$$

Assumption 1*(i)* and similar arguments as in Park (2002) yield $I = O_p(1)$ and $II = o_p\left(q^{-rs}\right)$. Note that

$$III \leq \frac{1}{T}\sum_{t=1}^T\left|\sum_{j=0}^q\gamma^{(i)}_{q,j}\left(\hat{u}_{it-j} - u_{it-j}\right)\right|^r + \frac{1}{T}\sum_{t=1}^T\left|\sum_{j=1}^q\left(\hat{\gamma}^{(i)}_{q,j} - \gamma^{(i)}_{q,j}\right)\hat{u}_{it-j}\right|^r = III_a + III_b,$$

with $\beta^u_{q,0} = 1$. It holds that $III_a \leq T^{-1}\sum_{t=1}^T\left|\hat{u}_{it} - u_{it}\right|^r\left(\sum_{j=0}^q\left|\gamma^{(i)}_{q,j}\right|\right)^r$, and Lemma A.2 entails $III_a = O_p\left(q^r\delta_{nT}^{-r}\right)$ for all $i$. Also, $III_b \leq T^{-1}\sum_{t=1}^T\left|\hat{u}_{it}\right|^r\left(\sum_{j=0}^q\left|\hat{\gamma}^{(i)}_{q,j} - \gamma^{(i)}_{q,j}\right|\right)^r$. Lemma A.2 yields $T^{-1}\sum_{t=1}^T\left|\hat{u}_{it}\right|^r = O_p(1)$. Also, $\sum_{j=0}^q\left|\hat{\gamma}^{(i)}_{q,j} - \gamma^{(i)}_{q,j}\right| \leq q\max_{1\leq j\leq q}\left|\hat{\gamma}^{(i)}_{q,j} - \gamma^{(i)}_{q,j}\right|$, and from Lemma 2 we have $\left[q\max_{1\leq j\leq q}\left|\hat{\gamma}^{(i)}_{q,j} - \gamma^{(i)}_{q,j}\right|\right]^r = O_p\left(q^r\varphi^u_{nT}\right)$. Finally, similar passages as for the proof of $T^{-1}\sum_{t=1}^{T}\hat{e}_{t,q}^{F}$ above, $IV=O_p\left[\left(\sqrt{\frac{q}{T}}\right)^r\right]$. Thus

$$E^b\left|e_{it,b}^u\right|^r=E\left|e_{it}^{u(i)}\right|^r+O_p\left(q^{-rs}\right)+O_p\left(q^r\delta_{nT}^{-r}\right)+O_p\left[\left(\frac{q}{\varphi_{nT}^u}\right)^r\right]+o_p\left(1\right),$$

whence, for every $i$, $E^b\left|e_{it,b}^u\right|^r=E\left|e_{it}^{u(i)}\right|^r+o_p\left(1\right)$. Assumption 1*(i)* then entails $\max_{i,t}E^b\left|e_{it,b}^u\right|^r<\infty$. ■

**Proof of Lemma 2.** For the sake of simplicity, the proof is reported for $k=1$, and suppressing the subscripts in $q_F$ and $q_{u,i}$ whenever possible.

Consider (16). Recall (11), (59) and (60) and let

$$\Delta F_t=\sum_{j=1}^{q}\tilde{\alpha}_{q,j}\Delta F_{t-j}+\tilde{e}_{t,q}^F,$$

which is the fitted version of (11). It holds that $\max_{1\leq j\leq q}\left|\hat{\alpha}_{q,j}-H'\alpha_j\left(H'\right)^{-1}\right|\leq\max_{1\leq j\leq q}\left|\hat{\alpha}_{q,j}-H'\tilde{\alpha}_{q,j}\left(H'\right)^{-1}\right|+\max_{1\leq j\leq q}\left|H'\left(\tilde{\alpha}_{q,j}-\alpha_{q,j}\right)\left(H'\right)^{-1}\right|+\max_{1\leq j\leq q}\left|H'\left(\alpha_{q,j}-\alpha_j\right)\left(H'\right)^{-1}\right|$ $=I+II+III$. As far as $II$ is concerned, Assumptions 1*(ii)* and 2*(ii)* yield $\max_{1\leq j\leq q}\left|\alpha_{q,j}-\alpha_j\right|\leq\sum_{j=1}^{q}\left|\alpha_{q,j}-\alpha_j\right|=o\left(q^{-s}\right)$ - see e.g. Theorem 2.1 in Hannan and Kavalieris (1986). Turning to $III$, Theorem 2.1 in Hannan and Kavalieris (1986) yields $III=O_p\left(\sqrt{\log T/T}\right)$. We now show that $I=O_p\left(T^{-1/2}C_{nT}^{-1}\right)+O_p\left(C_{nT}^{-2}\right)$. This is based on adapting the proof of Lemma A.1 in Chang, Park and Song (2006): it suffices to show that $\max_{1\leq i,j\leq q}\left|T^{-1}\sum_{t=\max\{i,j\}}^{T}\Delta\hat{F}_{t-i}\Delta\hat{F}_{t-j}'-T^{-1}\sum_{t=\max\{i,j\}}^{T}H'\Delta F_{t-i}\Delta F_{t-j}'H\right|=O_p\left(T^{-1/2}C_{nT}^{-1}\right)+O_p\left(C_{nT}^{-2}\right)$. Since

$$\begin{aligned}&\frac{1}{T}\sum_{t=\max\{i,j\}}^{T}\Delta\hat{F}_{t-i}\Delta\hat{F}_{t-j}'-\frac{1}{T}\sum_{t=\max\{i,j\}}^{T}H'\Delta F_{t-i}\Delta F_{t-j}'H\\=&\frac{1}{T}\sum_{t=\max\{i,j\}}^{T}\left(\Delta\hat{F}_{t-i}-H'\Delta F_{t-i}\right)\Delta F_{t-j}'H+\frac{1}{T}\sum_{t=\max\{i,j\}}^{T}H'\Delta F_{t-i}\left(\Delta\hat{F}_{t-j}-H'\Delta F_{t-j}\right)'\\&+\frac{1}{T}\sum_{t=\max\{i,j\}}^{T}\left(\Delta\hat{F}_{t-i}-H'\Delta F_{t-i}\right)\left(\Delta\hat{F}_{t-j}-H'\Delta F_{t-j}\right)'\\=&I_a+I_b+I_c.\end{aligned}$$

Using Lemma A.1*(ii)*, $I_a$ and $I_b$ are of magnitude $O_p\left(T^{-1/2}C_{nT}^{-1}\right)$; Lemma A.1*(iii)* entails that $I_c=O_p\left(C_{nT}^{-2}\right)$. Putting all together, $\max_{1\leq j\leq q}\left|\hat{\alpha}_{q,j}-H'\alpha_j\left(H'\right)^{-1}\right|=O_p\left(\sqrt{\log T/T}\right)$

$+O_p\left(T^{-1/2}C_{nT}^{-1}\right)+O_p\left(T^{-3/2}\right)+o\left(q^{-s}\right)$.

The proof of (17) follows similar lines. Consider (12), (61) and (62), and

$$u_{it}=\sum_{j=1}^{q}\tilde{\gamma}_{q,j}^{(i)}u_{it-j}+\tilde{e}_{it,q}^{u},$$

which is the fitted version of (12). We have $\max_{1\leq j\leq q}\left|\hat{\gamma}_{q,j}^{(i)}-\gamma_j\right|\leq\max_{1\leq j\leq q}\left|\hat{\gamma}_{q,j}^{(i)}-\tilde{\gamma}_{q,j}^{(i)}\right|$ $+\max_{1\leq j\leq q}\left|\tilde{\gamma}_{q,j}^{(i)}-\gamma_{q,j}^{(i)}\right|+\max_{1\leq j\leq q}\left|\gamma_{q,j}^{(i)}-\gamma_{j}^{(i)}\right|=I+II+III$. As above, $II$ and $III$ are $o\left(q^{-s}\right)$ and $O_p\left(\sqrt{\log T/T}\right)$ respectively. As far as $I$ is concerned, we show that $I=O_p\left(C_{nT}^{-1}\right)+O_p\left(C_{nT}^{-2}\right)$. This holds because $T^{-1}\sum_{t=\max\{j,k\}}^{T}\hat{u}_{it-j}\hat{u}_{it-k}-T^{-1}\sum_{t=\max\{j,k\}}^{T}u_{it-j}u_{it-k}=O_p\left(C_{nT}^{-1}\right)+O_p\left(C_{nT}^{-2}\right)$, by adapting Lemma A.3. Thus, $\max_{1\leq i,j\leq q}\left|T^{-1}\sum_{t=\max\{j,k\}}^{T}\hat{u}_{it-j}\hat{u}_{it-k}-T^{-1}\sum_{t=\max\{j,k\}}^{T}u_{it-j}u_{it-k}\right|=O_p\left(C_{nT}^{-1}\right)+O_p\left(C_{nT}^{-2}\right)$. Putting all together, $\max_{1\leq j\leq q}\left|\hat{\gamma}_{q,j}^{(i)}-\gamma_j\right|=O_p\left(\sqrt{\log T/T}\right)+O_p\left(C_{nT}^{-1}\right)+o\left(q^{-s}\right)$. ■

**Proof of Lemma 3.** Consider (18) and note that $\sum_{j=1}^{q}\hat{\beta}_{q,j}^{(i)}=\sum_{j=1}^{\infty}\beta_{ij}-\sum_{j=q+1}^{\infty}\beta_{ij}+\sum_{j=1}^{q}\left(\hat{\beta}_{q,j}-\beta_{q,j}\right)$. Assumption 1*(ii)* and 2*(ii)* entail $\sum_{j=q+1}^{\infty}\left\|\beta_j\right\|=o\left(q^{-s}\right)$. Using Lemma 2, $\sum_{j=1}^{q}\left(\hat{\beta}_{q,j}-\beta_{q,j}\right)\leq q\max_{1\leq j\leq q}\left|\hat{\beta}_{q,j}-\beta_{q,j}\right|=o_p\left(1\right)$, where the last equality follows from Assumption 5. Thus, $\hat{\beta}_q^{-1}\left(1\right)\xrightarrow{p}\beta^{-1}\left(1\right)$. Furthermore, from the bootstrap IP in Lemma 1 it holds that $T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}e_{it,b}\xrightarrow{d^b}W_i\left(s\right)$. Also, following the same lines as Park (2002, proof of Theorem 3.3, p. 486), we have, for all $i$ and some $\delta>0$, $P^b\left[\max_t\left\|T^{-1/2}\bar{\xi}_{it,b}\right\|>\delta\right]\leq\delta^{-r}T^{1-r/2}E^b\left\|\bar{\xi}_{it,b}\right\|^r$. Using Minkowski's inequality and the fact that $\bar{\xi}_{it,b}$ is stationary by construction, $E^b\left\|\bar{\xi}_{it,b}\right\|^r\leq\left[\sum_{j=1}^{q}j\left\|\hat{\beta}_{q,j}^{(i)}\right\|\right]^r E^b\left\|\xi_{it,b}\right\|^r$. Lemma 2 yields, for all $i$, $\sum_{j=1}^{q}j\left\|\hat{\beta}_{q,j}^{(i)}\right\|=\sum_{j=1}^{\infty}j\left\|\beta_{ij}\right\|+o_p\left(1\right)$; also, from Lemma 1, $E^b\left\|\xi_{it,b}\right\|^r<\infty$. Thus, for every $i$, $T^{-1/2}\sup_{1\leq t\leq T}\left\|\bar{\xi}_{it,b}\right\|=o_p\left(1\right)$. Therefore

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}\xi_{it,b}=\hat{\beta}_{i,q}^{-1}\left(1\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}e_{it,b}\right)+o_p\left(1\right)\xrightarrow{d^b}\beta_i^{-1}\left(1\right)W_i\left(s\right).$$

■

**Proof of Theorem 1.** The proof is similar to the proof of Lemma 3.4 in Chang, Park and Song (2006); thus, some passages are omitted. Consider (19), and assume, for simplicity, that $F_{0,b}=0$. Letting $W_{\varepsilon,nT}^{(b)}\left(s\right)=T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\Delta F_{t,b}$, Lemma 3 states that, as

$(n, T) \to \infty$, $W^{(b)}_{\varepsilon,nT}(s) \overset{d^*}{\to} W_\varepsilon(s)$. Then

$$\frac{1}{T^2}\sum_{t=1}^{T} F_{t,b}F'_{t,b} \underset{d^b}{=} \int W^{(b)}_{\varepsilon,nT}(s)\, W^{(b)}_{\varepsilon,nT}(s)' + \frac{1}{T^2}\sum_{t=1}^{T} F_{T,b}F'_{T,b},$$

and $T^{-1/2}F_{T,b} = o_p(1)$, which proves (19). As far as (20) is concerned, define the martingale approximations to $F_{t,b}$ and $u_{it,b}$ as $F^*_{t,b}$ and $u^*_{it,b}$; also, let $\overline{\Delta F}_{t,b}$ and $\bar{u}_{it,b}$ be the first $k$ and the last element of $\bar{\xi}_{it,b}$ respectively. Then

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T} F_{t,b}u_{it,b} &= \frac{1}{T}\sum_{t=1}^{T} F^*_{t,b}u^*_{it,b} + \frac{1}{T}\sum_{t=1}^{T} \Delta F_{t,b}\bar{u}_{it,b} - \frac{1}{T}F_{T,b}\bar{u}_{iT,b} \\ &+ \frac{1}{T}\overline{\Delta F}_{0,b}\sum_{t=1}^{T} u^*_{it,b} - \frac{1}{T}\sum_{t=1}^{T} \overline{\Delta F}_{t-1,b}u^*_{it,b} \\ &= I + II + III + IV + V. \end{aligned}$$

It holds straightforwardly that $III + IV + V = O_p\left(T^{-1/2}\right)$; also, $II \leq \left[T^{-1}\sum_{t=1}^{T}\|\Delta F_{t,b}\|^2\right]^{1/2}$ $\left[T^{-1}\sum_{t=1}^{T}\|\bar{u}_{it,b}\|^2\right]^{1/2} = O_p(1)\, O_p\left(T^{-1/2}\right)$. Thus, $T^{-1}\sum_{t=1}^{T} F_{t,b}u_{it,b} = T^{-1}\sum_{t=1}^{T} F^*_{t,b}u^*_{it,b} + o_p(1)$. The convergence of $T^{-1}\sum_{t=1}^{T} F^*_{t,b}u^*_{it,b}$ to $\int W_\varepsilon dW_{u,i}$ follows from Lemma 3 using the same approach as in Phillips (1988). ■

**Proof of Proposition 2.** Equation (28) follows from Proposition 1 (for $\hat{\lambda}^{OLS(1)}_{i,b}$) or Theorem 1 (for $\hat{\lambda}^{OLS(2)}_{i,b}$) and the CMT. The proof of (29) is reported for $\hat{\lambda}^{OLS(2)}_{i,b}$ - the case of $\hat{\lambda}^{OLS(1)}_{i,b}$ follows very similar passages. Note that

$$\begin{aligned} \left\| T\left[\hat{\lambda}^{OLS(2)}_{i,b} - \hat{\lambda}_i\right]\right\|^{2+\delta} &= \left\| \left(\frac{1}{T^2}\sum_{t=1}^{T} F_{t,b}F'_{t,b}\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} F_{t,b}u_{it,b}\right)\right\|^{2+\delta} \\ &\leq \sqrt{k}\left\| \left(\frac{1}{T^2}\sum_{t=1}^{T} F_{t,b}F'_{t,b}\right)^{-1}\right\|_1^{2+\delta}\left\| \left(\frac{1}{T}\sum_{t=1}^{T} F_{t,b}u_{it,b}\right)\right\|^{2+\delta}. \end{aligned}$$

By symmetry, $\left\|\left(T^{-2}\sum_{t=1}^{T} F_{t,b}F'_{t,b}\right)^{-1}\right\|_1 = \ell^{-1}_{\min}\left(T^{-2}\sum_{t=1}^{T} F_{t,b}F'_{t,b}\right)$, where $\ell_{\min}(\cdot)$ denotes the smallest eigenvalue. Theorem 1 ensures that, for sufficiently large $n$ and $T$, $T^{-2}\sum_{t=1}^{T} F_{t,b}F'_{t,b} = T^{-2}H'\sum_{t=1}^{T} F_tF'_tH + o_p(1)$; thus, in light of Assumption 2*(iii)* and the invertibility of $H$, $\left\|\left(T^{-2}\sum_{t=1}^{T} F_{t,b}F'_{t,b}\right)^{-1}\right\|_1^{2+\delta}$ is bounded with probability 1. As far as $\left\|\left(T^{-1}\sum_{t=1}^{T} F_{t,b}u_{it,b}\right)\right\|^{2+\delta}$ is concerned, it holds that $\left\|\left(\frac{1}{T}\sum_{t=1}^{T} F_{t,b}u_{it,b}\right)\right\|^{2+\delta} \leq$

$\frac{1}{T^{2+\delta}}\sum_{t=1}^{T}\|F_{t,b}\|^{2+\delta}\ |u_{it,b}|^{2+\delta}$. Lemmas 1 and 2 (and the fact that $k$ is fixed) ensure that $\left\|\left(T^{-1}\sum_{t=1}^{T}F_{t,b}u_{it,b}\right)\right\|^{2+\delta}=O_p(1)$. Thus, $\left\|T\left[\hat{\lambda}_{i,b}^{OLS(2)}-\hat{\lambda}_i\right]\right\|^{2+\delta}=O_p(1)$, thereby proving (29). ■

**Proof of Proposition 3.** The proof of equation (30) is similar to the proof of Theorem 3 in Bai (2004). We report only the main passages for the proof of $\hat{\lambda}_{i,b}^{PC(2)}$. We have

$$T\left[\hat{\lambda}_{i,b}^{PC(2)}-H_1^{-1}\hat{\lambda}_i\right]=\left[\frac{1}{T^2}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC(2)}\hat{F}_{t,b}^{PC(2)\prime}\right]^{-1}\times\left[\frac{1}{T}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC(2)}u_{it,b}+\frac{1}{T}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC(2)}\left(\hat{F}_{t,b}^{PC(2)}-H_1'F_{t,b}\right)'H_1^{-1}\hat{\lambda}_i\right].\quad(63)$$

Consider the denominator. Using Lemma A.4, is given by $T^{-2}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC(2)}\hat{F}_{t,b}^{PC(2)\prime}=H_1'\left(\frac{1}{T^2}\sum_{t=1}^{T}F_{t,b}F_{t,b}'\right)H_1+O_p\left(C_{nT}^{-1}\right)$. In view of (19):

$$\frac{1}{T^2}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC(2)}\hat{F}_{t,b}^{PC(2)\prime}\xrightarrow{d^b}H_1'\left[H'\left(\int W_\varepsilon W_\varepsilon'\right)H\right]H_1\text{ in }P.$$

Turning to the numerator, Lemma A.4*(iii)* yields $T^{-1}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC(2)}\left(\hat{F}_{t,b}^{PC(2)}-H_1'F_{t,b}\right)'$ $H_1^{-1}\hat{\lambda}_i=O_p\left(C_{nT}^{-1}\right)$. Also, $\frac{1}{T}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC(2)}u_{it,b}=H_1'\frac{1}{T}\sum_{t=1}^{T}F_{t,b}u_{it,b}+\frac{1}{T}\sum_{t=1}^{T}\left(\hat{F}_{t,b}^{PC(2)}-H_1'F_{t,b}\right)$ $u_{it,b}=I+II$. As far as $II$ is concerned, it is bounded by $\left[\frac{1}{T}\sum_{t=1}^{T}\left\|\hat{F}_{t,b}^{PC(2)}-H_1'F_{t,b}\right\|^2\right]^{1/2}$ $\left[\frac{1}{T}\sum_{t=1}^{T}u_{it,b}^2\right]^{1/2}$; using Lemma A.4*(i)*, and by virtue of Lemma 1, $II=O_p\left(C_{nT}^{-1}\right)$. Equation (20) entails

$$I\xrightarrow{d^b}H_1'H'\int W_\varepsilon dW_{u,i}\text{ in }P.$$

Equation (30) follows by applying the CMT. Turning to (31), from (63)

$$\left\|T\left[\hat{\lambda}_{i,b}^{PC(2)}-H_1^{-1}\hat{\lambda}_i\right]\right\|^{2+\delta}\leq\sqrt{k}\left\|\left[\frac{1}{T^2}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC(2)}\hat{F}_{t,b}^{PC(2)\prime}\right]^{-1}\right\|_1^{2+\delta}\times\left[\left\|\frac{1}{T}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC(2)}u_{it,b}\right\|^{2+\delta}+\left\|\frac{1}{T}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC(2)}\left(\hat{F}_{t,b}^{PC(2)}-H_1'F_{t,b}\right)'H_1^{-1}\hat{\lambda}_i\right\|^{2+\delta}\right].$$

As far as the denominator is concerned, the proof is similar to that of (29), in view of (36) and of the invertibility of $H_1$. As far as the numerator is concerned, the first term is bounded by $\left\| \frac{1}{T} \sum_{t=1}^{T} \hat{F}_{t,b}^{PC(2)} u_{it,b} \right\|^{2+\delta} \leq \|H_1'\| \left\| T^{-1} \sum_{t=1}^{T} F_{t,b} u_{it,b} \right\|^{2+\delta} + \| T^{-1} \sum_{t=1}^{T} \left( \hat{F}_{t,b}^{PC(2)} - H_1' F_{t,b} \right) u_{it,b} \|^{2+\delta} = I + II$; $I$ is $O_p(1)$ following similar arguments as in the proof of (29). Further, $II \leq \left[ T^{-1} \sum_{t=1}^{T} \left\| \hat{F}_{t,b}^{PC(2)} - H_1' F_{t,b} \right\|^2 \right]^{(2+\delta)/2} \left[ T^{-1} \sum_{t=1}^{T} u_{it,b}^2 \right]^{(2+\delta)/2}$, which is $o_p(1)$ using Lemma A.4*(i)* and Lemmas 1 and 2. The second term in the numerator is bounded by $\left\| T^{-1} \sum_{t=1}^{T} \hat{F}_{t,b}^{PC(2)} \left( \hat{F}_{t,b}^{PC(2)} - H_1' F_{t,b} \right)' \right\|^{2+\delta} \left\| H_1^{-1} \right\|^{2+\delta} \left\| \hat{\lambda}_i \right\|^{2+\delta}$, which is $o_p(1)$ in light of Lemma A.4*(ii)*, the invertibility of $H_1$ and (29). ■

**Proof of Proposition 4.** Consider (34). It holds that

$$\sqrt{n} \left[ \hat{F}_{t,b}^{OLS(2)} - F_{t,b} \right] = \left[ \frac{1}{n} \sum_{i=1}^{n} \hat{\lambda}_i \hat{\lambda}_i' \right]^{-1} \left[ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{\lambda}_i u_{it,b} \right],$$

and the same expression holds for $\hat{F}_{t,b}^{OLS(1)} - \hat{F}_t$. Recall that $\hat{\lambda}_i - H^{-1} \lambda_i = O_p\left(T^{-1}\right)$; thus, the denominator of $\hat{F}_{t,b}^{OLS(2)} - F_{t,b}$ is given by $n^{-1} \sum_{i=1}^{n} \hat{\lambda}_i \hat{\lambda}_i' = H^{-1} \left( n^{-1} \sum_{i=1}^{n} \lambda_i \lambda_i' \right) (H')^{-1}$ $+ O_p\left(T^{-1}\right)$. This and Assumption 3*(ii)* yield $n^{-1} \sum_{i=1}^{n} \hat{\lambda}_i \hat{\lambda}_i' \xrightarrow{p} H^{-1} \Sigma_\Lambda (H')^{-1}$. Turning to the numerator, we showed in the proof of Lemma A.4 that a CLT holds whereby $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{\lambda}_i u_{it,b}$. Thus, $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{\lambda}_i u_{it,b} \xrightarrow{d^b} N(0, V_\lambda)$ in $P$, where $V_\lambda = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \hat{\lambda}_i \hat{\lambda}_i' u_{it,b}^2$, in view of the cross sectional independence imposed by the "one unit at a time" scheme. Hence

$$\begin{aligned}
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \hat{\lambda}_i \hat{\lambda}_i' u_{it,b}^2 &= H^{-1} \Gamma_t (H')^{-1} + \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left( \hat{\lambda}_i - H^{-1} \lambda_i \right) \lambda_i' (H')^{-1} u_{it,b}^2 \\
&\quad + \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} H^{-1} \lambda_i \left( \hat{\lambda}_i - H^{-1} \lambda_i \right)' u_{it,b}^2 \\
&\quad + \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left( \hat{\lambda}_i - H^{-1} \lambda_i \right) \left( \hat{\lambda}_i - H^{-1} \lambda_i \right)' u_{it,b}^2 \\
&= H^{-1} \Gamma_t (H')^{-1} + I + II + III,
\end{aligned}$$

where the first term, $H^{-1} \Gamma_t (H')^{-1}$ follows by definition. Term $I$ is bounded by $\left\| H^{-1} \right\|$ $\max_i \|\lambda_i\| \max_i \left\| \hat{\lambda}_i - H^{-1} \lambda_i \right\| \left( \frac{1}{n} \sum_{i=1}^{n} u_{it,b}^2 \right)$, which is $O_p\left(T^{-1}\right)$ in view of Proposition 2 and Lemma 1; the same holds for $II$, and the same logic yields $III = O_p\left(T^{-2}\right)$. Putting all together

$$\sqrt{n} \left[ \hat{F}_{t,b}^{OLS(2)} - F_{t,b} \right] \xrightarrow{d^b} H' \Sigma_\Lambda^{-1} H H^{-1} \times N[0, \Gamma_t],$$

in $P$, which proves the first part of the Proposition. Turning to (35), the proof is similar, in spirit, to the proofs of (29) and (31), and therefore we report only the main passages. It holds that $\left\|\sqrt{n}\left[\hat{F}_{t,b}^{OLS(2)}-F_{t,b}\right]\right\|^{2+\delta}\leq\sqrt{k}\left\|\left[\frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}_i\hat{\lambda}_i'\right]^{-1}\right\|_1^{2+\delta}\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\lambda}_iu_{it,b}\right\|^{2+\delta}$. For sufficiently large $T$, Lemma 2 entails that $n^{-1}\sum_{i=1}^{n}\hat{\lambda}_i\hat{\lambda}_i'=n^{-1}H^{-1}\sum_{i=1}^{n}\lambda_i\lambda_i'\left(H'\right)^{-1}$ $+o_p\left(1\right)$; Assumption 3*(ii)* ensures that the smallest eigenvalue of $n^{-1}\sum_{i=1}^{n}\lambda_i\lambda_i'$ is positive, from whence $\left\|\left[\frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}_i\hat{\lambda}_i'\right]^{-1}\right\|_1^{2+\delta}$ is bounded. As far as the numerator is concerned, $\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\lambda}_iu_{it,b}\right\|^{2+\delta}\leq\left\|H^{-1}\right\|^{2+\delta}\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\lambda_iu_{it,b}\right\|^{2+\delta}+\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\hat{\lambda}_i-H^{-1}\lambda_i\right)u_{it,b}\right\|^{2+\delta}$ $=I+II$. Term $I$ is bounded in light of Assumption 3*(i)*, Lemmas 1 and 2 and the cross sectional independence of $u_{it,b}$. As far as $II$ is concerned, it is bounded by $n^{-\delta/2}\left[E\left\|\hat{\lambda}_i-H^{-1}\lambda_i\right\|^{4+\delta}\right]^{1/2}$ $\left[E\left(u_{it,b}^{4+\delta}\right)\right]^{1/2}$. This is finite in view of Lemmas 1 and 2; (29) only stipulates that $E\left\|\hat{\lambda}_i-H^{-1}\lambda_i\right\|^{2+\delta}$ is finite, but the proof can be modified to accommodate for $4+\delta$ using the same arguments. ■

**Proof of Proposition 5.** The proof is very similar to the proof of Theorem 2 in Bai (2004, p. 171) and therefore only the main passages are reported. In light of Lemma A.4*(iv)*, under $\frac{n}{T^3}\to 0$, $\sqrt{n}\left[\hat{F}_{t,b}^{PC(2)}-H_1'F_{t,b}\right]=\frac{1}{\sqrt{nT^2}}\left(V_{nT}^b\right)^{-1}\sum_{s=1}^{T}\hat{F}_{s,b}F_{s,b}'\hat{\Lambda}'u_{t,b}$ $+o_p\left(1\right)$. It holds that

$$
\begin{aligned}
&\frac{1}{\sqrt{nT^2}}\left(V_{nT}^b\right)^{-1}\sum_{s=1}^{T}\hat{F}_{s,b}F_{s,b}'\hat{\Lambda}'u_{t,b}\\
=\;&\left(V_{nT}^b\right)^{-1}\left(\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_{s,b}F_{s,b}'\right)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\lambda}_iu_{it,b}\\
=\;&\left(V_{nT}^b\right)^{-1}\left(\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_{s,b}F_{s,b}'\right)\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}_i\hat{\lambda}_i'\right)\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}_i\hat{\lambda}_i'\right)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\lambda}_iu_{it,b}\\
=\;&H_1'\left[H^{-1}\Sigma_\Lambda\left(H'\right)^{-1}\right]^{-1}H^{-1}N\left[0,\Gamma_t\right]+o_p\left(1\right),
\end{aligned}
$$

which proves (36). The proof of (37) is very similar to the proof of (35), and thus it is omitted. ■

**Proof of Proposition 6.** Consider the case $\frac{n}{T} \to 0$. We have

$$
\begin{aligned}
\sqrt{n}\left(\hat{C}_{it,b} - \hat{C}_{it}\right) &= \hat{\lambda}_i' \left(H_1'\right)^{-1} \left[\sqrt{n}\left(\hat{F}_{t,b}^{PC} - H_1' F_t\right)\right] + \hat{F}_{t,b}^{PC\prime} \left[\sqrt{n}\left(\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right)\right] \\
&= \lambda_i' \left(H'\right)^{-1} \left(H_1'\right)^{-1} \left[\sqrt{n}\left(\hat{F}_{t,b}^{PC} - H_1' F_t\right)\right] \\
&\quad + \left(\hat{\lambda}_i - H^{-1}\lambda_i\right)' \left(H_1'\right)^{-1} \left[\sqrt{n}\left(\hat{F}_{t,b}^{PC} - H_1' F_t\right)\right] + \hat{F}_{t,b}^{PC\prime} \left[\sqrt{n}\left(\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right)\right] \\
&= I + II + III.
\end{aligned}
$$

Using Propositions 3 and 5, $II = O_p\left(T^{-1}\right)$. Also, $\hat{F}_{t,b}^{PC} = O_p\left(\sqrt{T}\right)$ by construction, whence $III = O_p\left(\sqrt{T}\right)\sqrt{n}O_p\left(T^{-1}\right) = o_p(1)$. Turning to $I$, Proposition 5 entails that $I \xrightarrow{d^b} \lambda_i' \left(H'\right)^{-1}\left(H_1'\right)^{-1} H_1' H' \Sigma_\Lambda \times N\left(0, \Gamma_t\right)$ in $P$.

As $\frac{T}{n} \to 0$, it holds that

$$
\begin{aligned}
\sqrt{T}\left(\hat{C}_{it,b} - \hat{C}_{it}\right) &= \hat{\lambda}_i' \left(H_1'\right)^{-1} \left[\sqrt{T}\left(\hat{F}_{t,b}^{PC} - H_1' F_t\right)\right] + \hat{F}_{t,b}^{PC\prime} \left[\sqrt{T}\left(\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right)\right] \\
&= \hat{\lambda}_i' \left(H_1'\right)^{-1} \left[\sqrt{T}\left(\hat{F}_{t,b}^{PC} - H_1' F_t\right)\right] + F_t' H H_1 \left[\sqrt{T}\left(\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right)\right] \\
&\quad + \left(\hat{F}_{t,b}^{PC} - \hat{F}_t' H_1\right)\left[\sqrt{T}\left(\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right)\right] + \left(\hat{F}_t - F_t' H\right) H_1 \left[\sqrt{T}\left(\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right)\right] \\
&= I + II + III + IV.
\end{aligned}
$$

Proposition 5 entails that $I = o_p(1)$; also, in view of Propositions 3 and 5, $III$ and $IV$ are both $o_p(1)$. The asymptotics is driven by $II$. The IP entails that $T^{-1/2}F_t = O_p(1)$ and $T^{-1/2}F_t \xrightarrow{d} W_\varepsilon(s)$ uniformly in $s$; using Proposition 3,

$$
II \xrightarrow{d^b} W_\varepsilon'(s) H H_1 H_1^{-1} \left[H^{-1}\left(\int W_\varepsilon W_\varepsilon'\right)^{-1}\left(\int W_\varepsilon dW_{u,i}\right)\right] \text{ in } P.
$$

Equation (41) follows from combining the two results. The proof of (42) follows from combining (31) and (37). ■

**Proof of Theorem 2.** Consider (44). Similarly to the proof of (9), we may write

$$
\begin{aligned}
\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Ts\rfloor} u_{it,b} &= \frac{1}{\sqrt{nT}}\sum_{t=1}^{\lfloor Ts\rfloor} i_n' \Gamma(1) e_{t,b}^{*u} + \frac{1}{\sqrt{nT}}\sum_{t=1}^{\lfloor Ts\rfloor} i_n' \left[\Gamma^*(1) - \Gamma(1)\right] e_{t,b}^{*u} + \frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Ts\rfloor} R_{ub,it} \\
&= I + II + III,
\end{aligned}
$$

where the superscript "*" denotes the martingale approximation, $\Gamma^*(1) = [B^*(1)]^{-1}$ and $R_{ub,it}$ denotes the remainder in the Beveridge-Nelson decomposition. Consider $I$, and define $\zeta_{nt,b} = n^{-1/2}\ i_n'\Gamma(1)\, e_{t,b}^{*u}$. The sequence $\zeta_{nt,b}$ is an MDS by construction. Also, $E\left|\zeta_{nt,b}\right|^{2+\delta} \leq n^{-\delta/2} \max_i E\left|u_{it,b}^*\right|^{2+\delta}$, where $u_{it,b}^*$ is the MDS approximation to $u_{it,b}$. In view of Lemma 1, $E\left|\zeta_{nt,b}\right|^{2+\delta}$ is bounded uniformly in $n$. Thus, an IP for MDS holds whereby, uniformly in $s$, $I \xrightarrow{d^b} W_{\zeta,b}(s)$. The variance of $W_{\zeta,b}(s)$ is $n^{-1}\ i_n'\Gamma(1)\ E\left[e_{t,b}^{*u}e_{t,b}^{*u\prime}\right]$ $\Gamma'(1)\, i_n$. We now turn to showing that $II$ and $III$ are negligible. Consider $II$. We have $II = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}\left[\sum_{i=1}^{n}\left\{\Gamma^*(1)-\Gamma(1)\right\}_{ij}\right] e_{jt,b}^{*u} = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\tilde{\zeta}_{j\lfloor Ts\rfloor}$. For each $\lfloor Ts\rfloor$, $\tilde{\zeta}_{j\lfloor Ts\rfloor}$ has mean zero and is weakly dependent. Thus, a sufficient condition for $II$ to be negligible is that, as $T\to\infty$, $\sup_j E\left(\tilde{\zeta}_{j\lfloor Ts\rfloor}^2\right) = o_p(1)$. Since

$$\begin{aligned}
\sup_j \lim_{T\to\infty} E\left(\tilde{\zeta}_{j\lfloor Ts\rfloor}^2\right) &= \sup_j \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor}\left[\sum_{i=1}^{n}\left\{\Gamma^*(1)-\Gamma(1)\right\}_{ij}\right]^2 E\left(e_{jt,b}^{*u}\right)^2 \\
&\leq M\sup_j\left[\sum_{i=1}^{n}\left|\left\{\Gamma^*(1)-\Gamma(1)\right\}_{ij}\right|\right]^2,
\end{aligned}$$

where the last inequality comes from Assumption 1*(i)* - note that this holds uniformly in $s$. Thus, a sufficient condition for $\lim_{T\to\infty}\sup_j E\left(\tilde{\zeta}_{j\lfloor Ts\rfloor}^2\right) = o_p(1)$ is that $\sup_j \sum_{i=1}^{n}\left|\left\{\Gamma^*(1)\right.\right.$ $\left.\left.-\Gamma(1)\right\}_{ij}\right| = o_p(1)$, which, by definition, is equivalent to $\left\|\Gamma^*(1)-\Gamma(1)\right\|_1 = o_p(1)$. Recall $\Gamma(1) = [B(1)]^{-1}$ and $\Gamma^*(1) = [B^*(1)]^{-1}$; using Taylor's expansion

$$\begin{aligned}
\left\|[B^*(1)]^{-1}-[B(1)]^{-1}\right\|_1 &= \left\|\Gamma^{-1}(1)\left[B^*(1)-B(1)\right]\left[\Gamma'(1)\right]^{-1}\right\|_1 \\
&\leq \left\|\Gamma^{-1}(1)\right\|_1\left\|\Gamma^{-1}(1)\right\|_\infty\left\|\left[B^*(1)-B(1)\right]\right\|_1 \\
&= O_p(1)\,O_p(1)\,o_p(1) = o_p(1),
\end{aligned}$$

by Assumption 1*(iii)* and from assuming $\left\|\left[B^*(1)-B(1)\right]\right\|_1 = o_p(1)$. Thus, $II$ is negligible. Finally, standard panel asymptotics (see e.g. Phillips and Moon, 1999) yields $III = O_p\left(\sqrt{\frac{n}{T}}\right)$, which is negligible.

Consider (45); $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}F_{t,b}u_{it,b}$ can be decomposed as

$$\frac{1}{T}\sum_{t=1}^{T}F_{t,b}^*\zeta_{nt,b} + \frac{1}{\sqrt{n}T}\sum_{t=1}^{T}F_{t,b}^*i_n'\left[\Gamma^*(1)-\Gamma(1)\right]e_{t,b}^{*u} + \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}R_{it} = I + II + III,$$

where, as before, the superscript "*" denotes the martingale approximation and $R_{it}$ is the remainder in the Beveridge-Nelson decomposition. Consider $I$. Theorem 1 and (44) ensure that an IP for MDS holds for $T^{-1/2}F_{t,b}^{*}$ and$T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\zeta_{nt,b}$ respectively. Also, $F_{t,b}^{*}$ and $\zeta_{nt,b}$ are independent by construction and for each $n$. Thus, the theory of convergence to stochastic integrals (see e.g. Phillips, 1988) entails $I\xrightarrow{d_b}\sigma_u H'\int W_{\varepsilon}dW_u$ in $P$. We now turn to $II$; we have $II=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{1}{T}\sum_{t=1}^{T}F_{t,b}^{*}\left[\sum_{i=1}^{n}\left\{\Gamma^{*}(1)-\Gamma(1)\right\}_{ij}\right]e_{jt,b}^{*u}=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\tilde{\xi}_{jT}$. Conditional on $C_F^{*}$, $\tilde{\xi}_{jT}$ is a zero mean weakly dependent sequence. As before, a sufficient condition for $II$ to be negligible is therefore that, as $T\to\infty$, $\sup_j E\left(\tilde{\xi}_{jT}^{2}\middle| C_F^{*}\right)=o_p(1)$. Since

$$
\begin{aligned}
\sup_j\lim_{T\to\infty}E\left(\tilde{\xi}_{jT}^{2}\middle| C_F^{*}\right) &= \sup_j\lim_{T\to\infty}\frac{1}{T^2}\sum_{t=1}^{T}F_{t,b}^{*}F_{t,b}^{*\prime}\left[\sum_{i=1}^{n}\left\{\Gamma^{*}(1)-\Gamma(1)\right\}_{ij}\right]^2 E\left(e_{jt,b}^{*u}\right)^2 \\
&\leq M\left(\lim_{T\to\infty}\frac{1}{T^2}\sum_{t=1}^{T}F_{t,b}^{*}F_{t,b}^{*\prime}\right)\sup_j\left[\sum_{i=1}^{n}\left|\left\{\Gamma^{*}(1)-\Gamma(1)\right\}_{ij}\right|\right]^2,
\end{aligned}
$$

where the last inequality comes from Assumption 1*(i)*. Again, this is $o_p(1)$ if $\sup_j\sum_{i=1}^{n}\left|\{\Gamma^{*}(1)\right.$ $\left.-\Gamma(1)\}_{ij}\right|=o_p(1)$. Finally, as far as $III$ is concerned, similar passages as in the proof of Theorem 1 yield $n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}R_{it}\leq n^{1/2}\sup_i\left|T^{-1}\sum_{t=1}^{T}R_{it}\right|=O_p\left(\sqrt{\frac{n}{T}}\right)$, which is of order $o_p(1)$ under $\frac{n}{T}\to 0$. The proof is now the same as above. ■

**Proof of Theorem 3.** Consider equation (48); note that $\left\|\widehat{B_q(1)}-B(1)\right\|_1=\left\|\sum_{j=1}^{q}\left(\hat{B}_{q,j}-B_{q,j}\right)\right\|_1+\left\|\sum_{j=q+1}^{\infty}B_j\right\|_1=I+II$. Assumption 1*(ii)* implies $II=o\left(q^{-s}\right)$. As far as $I$ is concerned, note $\sum_{j=1}^{q}\left(\hat{B}_{q,j}-B_{q,j}\right)=\left(\hat{B}_q-B_q\right)\left(i_q\otimes I_n\right)$. Thus, $I\leq\left\|\hat{B}_q-B_q\right\|_1\left\|i_q\otimes I_n\right\|_1=q\left\|\hat{B}_q-B_q\right\|_1$. To study the magnitude of $\left\|\hat{B}_q-B_q\right\|_1$, consider (47)

$$
\hat{B}_q=\left[\frac{1}{T}\sum_{t=q+1}^{T}\hat{u}_t\hat{u}_{qt}'\right]\left[\frac{1}{T}\sum_{t=q+1}^{T}\hat{u}_{qt}\hat{u}_{qt}'\right]^{-1}.
$$

Let $d_q=T^{-1}\sum_{t=q+1}^{T}u_{qt}u_{qt}'$ and $\hat{d}_q=T^{-1}\sum_{t=q+1}^{T}\hat{u}_{qt}\hat{u}_{qt}'$ Then we can write

$$
\begin{aligned}
\hat{B}_q &= \left[B_q d_q+\frac{1}{T}\sum_{t=q+1}^{T}\hat{u}_t\hat{u}_{qt}'-\frac{1}{T}\sum_{t=q+1}^{T}u_t u_{qt}'+\frac{1}{T}\sum_{t=q+1}^{T}e_t^{(u)}u_{qt}'\right] \\
&\quad\times\left[d_q^{-1}+d_q^{-1}\left(\hat{d}_q-d_q\right)d_q^{-1}+o_p\left(\left\|\hat{d}_q-d_q\right\|\right)\right].
\end{aligned}
\tag{64}
$$

Let $k = 1$ for simplicity; in this case, $H$ is a scalar, but we employ the matrix notation for consistency. By definition $\hat{u}_{qt} = u_{qt} + \left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right] \otimes \Lambda (H')^{-1} + \hat{F}_{q,t} \otimes \left[\hat{\Lambda} - \Lambda (H')^{-1}\right]$, where $F_{q,t} = [F_{t-1}, ..., F_{t-q}]'$, and similarly $\hat{F}_{q,t}$. Assumption 1*(ii)* yields $\left\|d_q^{-1}\right\|_1 = O_p(1)$; also, note

$$
\begin{aligned}
& \frac{1}{T}\left(\sum_{t=q+1}^{T} \hat{u}_{qt}\hat{u}_{qt}' - \sum_{t=q+1}^{T} u_{qt}u_{qt}'\right) \\
= & \left\{\frac{1}{T}\sum_{t=q+1}^{T}\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]'\right\} \otimes \Lambda \left(H'H\right)^{-1} \Lambda' \\
& + \left(\frac{1}{T}\sum_{t=q+1}^{T} \hat{F}_{q,t}\hat{F}_{q,t}'\right) \otimes \left[\hat{\Lambda} - \Lambda (H')^{-1}\right]\left[\hat{\Lambda} - \Lambda (H')^{-1}\right]' \\
& + \left\{\frac{1}{T}\sum_{t=q+1}^{T}\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]\hat{F}_{q,t}'\right\} \otimes \Lambda (H')^{-1}\left[\hat{\Lambda} - \Lambda (H')^{-1}\right]' \\
& + \left\{\frac{1}{T}\sum_{t=q+1}^{T}\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right] u_{qt}'\right\} \otimes \Lambda (H')^{-1} \\
& + \left(\frac{1}{T}\sum_{t=q+1}^{T} \hat{F}_{q,t} u_{qt}'\right) \otimes \left[\hat{\Lambda} - \Lambda (H')^{-1}\right] + A + B \\
= & \ I + II + III + IV + V + A + B,
\end{aligned}
$$

where $A$ and $B$ are the transposes of $IV$ and $V$ respectively. It holds that $\|I\|_1 \leq \left\|\Lambda (H'H)^{-1} \Lambda'\right\|_1 \left\|T^{-1}\sum_{t=q+1}^{T}\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]'\right\|_1 = O_p(n) O_p\left(qC_{nT}^{-2}\right)$, using Lemma B.1 in Bai (2004). Recalling that $\hat{\Lambda} - \Lambda (H')^{-1} = O_p\left(T^{-1}\right)$ element-wise, it holds that $\|II\|_1 \leq \left\|\left[\hat{\Lambda} - \Lambda (H')^{-1}\right]\left[\hat{\Lambda} - \Lambda (H')^{-1}\right]'\right\|_1 \left\|T^{-1}\sum_{t=q+1}^{T} \hat{F}_{q,t}\hat{F}_{q,t}'\right\|_1$ $= O_p\left(nT^{-2}\right) O_p(qT) = O_p\left(qnT^{-1}\right)$; similarly, $\|III\|_1 \leq \left\|\Lambda (H')^{-1}\left[\hat{\Lambda} - \Lambda (H')^{-1}\right]'\right\|_1$ $\left\|T^{-1}\sum_{t=q+1}^{T}\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]\hat{F}_{q,t}'\right\|_1 = O_p\left(nT^{-1}\right) O_p\left(qC_{nT}^{-1}\right)$. Considering $IV$, $\|IV\|_1 \leq \left\|\Lambda (H')^{-1}\right\|_1 \left\|T^{-1}\sum_{t=q+1}^{T}\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right] u_{q,t}'\right\|_1 = O_p(n) O_p\left(qC_{nT}^{-1}\right)$; similar calculations yield that $\|A\|_1$ has the same order. Finally, $\|V\|_1 \leq \left\|\hat{\Lambda} - \Lambda (H')^{-1}\right\|_1 \left\|T^{-1}\sum_{t=q+1}^{T} \hat{F}_{q,t} u_{qt}'\right\|_1$ $= O_p\left(nT^{-1}\right) O_p(q)$. Thus, the terms that dominate are of magnitude $O_p\left(nqC_{nT}^{-1}\right)$. Considering the numerator of (64), recall $\left\| T^{-1} \sum_{t=q+1}^{T} e_t^{(u)} u_{qt}' \right\|_1 = O_p\left(\sqrt{\frac{\log T}{T}}\right)$. Also

$$
\begin{aligned}
& \frac{1}{T}\left(\sum_{t=q+1}^{T} \hat{u}_t \hat{u}_{qt}' - \sum_{t=q+1}^{T} u_t u_{qt}'\right) \\
= & \left\{\frac{1}{T}\sum_{t=q+1}^{T}\left(\hat{F}_t - H'F_t\right)\left[\hat{F}_{q,t} - \left(H \otimes I_q\right) F_{q,t}\right]'\right\} \otimes \Lambda \left(H'H\right)^{-1} \Lambda' \\
& + \left(\frac{1}{T}\sum_{t=q+1}^{T} \hat{F}_t \hat{F}_{q,t}'\right) \otimes \left[\hat{\Lambda} - \Lambda \left(H'\right)^{-1}\right]\left[\hat{\Lambda} - \Lambda \left(H'\right)^{-1}\right]' \\
& + \left[\frac{1}{T}\sum_{t=q+1}^{T}\left(\hat{F}_t - H'F_t\right)\hat{F}_{q,t}'\right] \otimes \Lambda \left(H'\right)^{-1}\left[\hat{\Lambda} - \Lambda \left(H'\right)^{-1}\right]' \\
& + \left[\frac{1}{T}\sum_{t=q+1}^{T}\left(\hat{F}_t - H'F_t\right) u_{qt}'\right] \otimes \Lambda \left(H'\right)^{-1} \\
& + \left(\frac{1}{T}\sum_{t=q+1}^{T} \hat{F}_t u_{qt}'\right) \otimes \left[\hat{\Lambda} - \Lambda \left(H'\right)^{-1}\right] + C + D \\
= & \ I + II + III + IV + V + C + D,
\end{aligned}
$$

with $C$ and $D$ being the transposes of $IV$ and $V$. Similar results as for the denominator hold. We have $\|I\|_1 = O_p\left(nC_{nT}^{-2}\right)$, $\|II\|_1 = O_p\left(nqT^{-1}\right)$, $\|III\|_1 = O_p\left(nqT^{-1}C_{nT}^{-1}\right)$, $\|IV\|_1 = O_p\left(nqC_{nT}^{-1}\right)$ and $\|V\|_1 = O_p\left(nqT^{-1}\right)$. Putting all together, it holds that $\left\|\widehat{B_q(1)} - B(1)\right\|_1 = o\left(q^{-s}\right) + qO_p\left(nqC_{nT}^{-1}\right) + qO_p\left(\sqrt{\frac{\log T}{T}}\right)$.

Consider now (49), and let $B_q \equiv B_q^d + B_q^{od}$, where $B_q^d = \left[B_{q,1}^d|...|B_{q,q}^d\right]$ with $B_{q,j}^d = diag\left\{\gamma_{q,j}^{(i)}\right\}$ and $B_q^{od} = \left[B_{q,1}^{od}|...|B_{q,q}^{od}\right]$ defined so that $B_{q,j}^{od}$ contains the off-diagonal elements of $B_{q,j}$. As before, $\widetilde{B_q(1)} - B(1) = \left(\widetilde{B_q} - B_q\right)\left(i_q \otimes I_n\right) + \sum_{j=q+1}^{\infty} B_j$. Since $\widetilde{B_q} - B_q = \widetilde{B_q} - B_q^d - B_q^{od}$, $\left\|\widetilde{B_q} - B_q\right\|_1 \leq \left\|\widetilde{B_q} - B_q^d\right\|_1 + \left\|B_q^{od}\right\|_1$. By construction, $\left\|B_q^{od}\right\|_1 = \sup_j \sum_{i \neq j}|\tau_{ij}| = O_p\left(n^{-\phi}\right)$ where the last equality holds by assumption. Also, $\left\|\widetilde{B_q} - B_q^d\right\|_1 = \sup_{i,j}\left|\hat{\gamma}_{q,j}^{(i)} - \gamma_{q,j}^{(i)}\right| = O_p\left(\varphi_{nT}^u\right)$ in light of Lemma 2. Thus, putting everything together $\left\|\left(\widetilde{B_q} - B_q\right)\left(i_q \otimes I_n\right)\right\|_1 \leq q\left\|\widetilde{B_q} - B_q\right\|_1 \leq qO_p\left(\varphi_{nT}^u\right) + qO_p\left(n^{-\phi}\right)$; this proves (49). ■

**Proof of Proposition 7.** The proof follows the same passages as in the proofs of Lemmas 1 and 2, based on the results in Lemma A.5. ■