On Bootstrapping Panel Factor Series

Lorenzo Trapani*

Cass Business School, City University London

August 22, 2012

Abstract

This paper studies the asymptotic validity of sieve bootstrap for nonstationary panel factor series. 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 $(n, T) \to \infty$. Secondly, joint bootstrap asymptotics is also studied, investigating the conditions under which the bootstrap is valid. In particular, the extent of cross sectional dependence which can be allowed for is investigated. Whilst we show that, for general forms of cross dependence, consistent estimation of the long run variance (and therefore validity of the bootstrap) is fraught with difficulties, however we show that "one-cross-sectional-unit-at-a-time" resampling schemes yield valid bootstrap based inference under weak forms of cross-sectional dependence.

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.

1 Introduction

In recent years, factor models have achieved great popularity in applied econometrics and statistics - see e.g. Lee and Carter (1992), Forni and Reichlin (1998), 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. The literature has 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 study 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 - see Bai (2004). 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. 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 (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. We propose a "one cross sectional unit at a time" resampling algorithm, 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. In Section 4, we report validity results for bootstrap estimates of loadings, common factors and common components based on applying the PC estimator to the bootstrap sample. In the second part of the paper (Section 5), we discuss how to deal with the issue of cross dependence. In Section 5, we develop joint bootstrap asymptotics as $(n, T) \to \infty$, to accommodate for the possible presence of cross dependence in the u_{it}s by fitting an n-dimensional VAR to the vector containing the residuals $\hat{u}_{it}$. 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. Section 5 contains an inconsistency result, highlighting that consistent estimation of long run covariance matrices is not possible in this context, unless there is very little cross dependence. Theoretical findings are evaluated through a Monte Carlo simulation (Section 6). Section 7 concludes. Proofs are in Appendix.

For the sake of a concise discussion, this version of the paper only reports the main results and proofs. In an extended version (henceforth referred to as Trapani, 2012), the full set of results is reported. This includes: validity results for results for bootstrap estimates of loadings, common factors and common components based on applying the OLS estimator to the bootstrap sample; some initial results concerning the extension of the bootstrap theory to the case of (1) containing also $I(0)$ common factors and drift terms in the $I(1)$ factors; and preliminary, technical Lemmas as well as all the proofs omitted from here.[1]

NOTATION Throughout the paper, $\|A\|_p$ denotes the L_p-norm of a matrix A, i.e.

[1] The extended version is available for download at SSRN: http://ssrn.com/abstract=2062183

$\max_x \|Ax\|_p / \|x\|_p$ (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 $\int WdW$. The integer part of a number x is denoted as $\lfloor x \rfloor$. 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, assumptions and preliminary asymptotics

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.

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 the 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, $\max_{i,t} E|e_{it}^u|^{8+\delta} < \infty$ for some $\delta > 0$; *(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 Γ_j, $\max_i \sum_{j=0}^{\infty} j^s \|\Gamma_{i,j}\| < \infty$ for some $s \geq 1$; *(iii) (cross sectional dependence)* (a) $\|\Gamma(1)\|_1 \leq M$, $\left\|\Gamma^{-1}(1)\right\|_1 \leq M$, $\left\|\Gamma^{-1}(1)\right\|_\infty \leq M$ and $\|\Sigma_u\|_1 \leq M$; (b) $E\left|n^{-1/2} \sum_{i=1}^n [u_{is}u_{it} - E(u_{is}u_{it})]\right|^4 \leq M$ for every (t,s); (c) $E\left|\sum_{i=1}^n u_{it}\right|^{2+\delta} \leq ME\left|\sum_{i=1}^n u_{it}^2\right|^{\frac{2+\delta}{2}}$ for all t and $\delta > 0$; *(iv) (initial conditions)* $E|u_{i0}|^4 \leq M$ for all i.

Assumption 2: *(time series properties of ε_t)* ε_t is a k-dimensional vector random process (with finite k) and it admits an invertible $MA(\infty)$ representation where $\varepsilon_t = \alpha(L) e_t^F = \sum_{j=0}^{\infty} \alpha_j e_{t-j}^F$ 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 $E\left\|e_t^F\right\|^{8+\delta} < \infty$ for $\delta > 0$; *(ii)* $\sum_{j=0}^{\infty} \alpha_j L^j \neq 0$ for all $|L| \leq 1$ and $\sum_{j=0}^{\infty} j^s \|\alpha_j\| < \infty$

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)* $E \|F_0\|^4 \leq M$.

Assumption 3: *(identifiability)* the loadings λ_i are *(i)* either nonrandom quantities such that $\|\lambda_i\| \leq M$, or random quantities such that $E \|\lambda_i\|^4 < \infty$; *(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 Σ_Λ positive definite; *(iii)* the eigenvalues of $\Sigma_\Lambda^{1/2} \Sigma_{\Delta F} \Sigma_\Lambda^{1/2}$ 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 e.g. Assumption C in Bai (2004). Part *(ii)* is needed in order to approximate the $AR(\infty)$ polynomial with a finite autoregressive representation - see e.g. Hannan and Kavalieris (1986). Letting $E(u_{it} u_{jt}) = \tau_{ij}$, part *(iii)* entails that $\sum_{i=1}^{n} |\tau_{ij}| \leq M$ for all j, since $E(u_t u_t') = \Gamma(1) \Sigma_u \Gamma'(1)$ and $\|E(u_t u_t')\|_1 \leq \|\Gamma(1)\|_1^2 \|\Sigma_u\|_1$. Finiteness of $\left\|\Gamma^{-1}(1)\right\|_1$ could be derived 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 part (4) of Assumption C in Bai (2004). 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 $\gamma_{i,s-t} = E(u_{it} u_{is})$, which is part (2) of Assumption C in Bai (2004, p. 141). Finally, part *(iii)*(c) is a Burkholder-type inequality. This could be proved under more primitive conditions, e.g. if the u_{it}s were independent across i, and it is useful to derive joint asymptotics; see, in particular, Proposition 1 below.

Assumption 2 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). It is required that the 8-th moment of e_t^F should exist. This is in order for the bootstrap sample to satisfy the equivalent of Assumption 1*(iii)*, which in turn is needed when applying PC to the bootstrap sample (see Lemma A.4 in Appendix A). Part *(ii)* plays the same role as Assumption 1*(ii)*. Note that part *(iii)* rules out cointegration among the F_ts, which is the same as part (2) of Assumption A in Bai (2004). Also, Assumption 2 entails a Law of the Iterated Logarithm for F_t (see Phillips and Solo, 1992, Theorem 3.3) to hold, whence $\liminf_{T\to\infty} (\log\log T)\, T^{-2}\sum_{t=1}^{T} F_t F_t' = D$ with D a nonrandom positive definite matrix. This corresponds to part (3) of Assumption 2 in Bai (2004).

2.1 Inferential theory

Inference is based on standard PC. The common factors F_t are estimated by $\hat{F}_t$ under the restrictions that $T^{-2}\sum_{t=1}^{T} \hat{F}_t\hat{F}_t' = I_k$ and $n^{-1}\sum_{i=1}^{n} \hat{\lambda}_i\hat{\lambda}_i'$ is diagonal. The estimated common factor $\hat{F}_t$ is T times the eigenvectors corresponding to the k largest eigenvalues of XX' where $X = [x_1, ..., x_n]'$ with $x_i = [x_{i1}, ..., x_{iT}]'$. Then λ_i can be estimated applying OLS to

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

whence $\hat{\lambda}_i = \left[\sum_{t=1}^{T} \hat{F}_t\hat{F}_t'\right]^{-1}\left[\sum_{t=1}^{T} \hat{F}_t x_{it}\right]$. It is well known that λ_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 λ_i), finding $H'F_t$ instead of F_t and $H^{-1}\lambda_i$ instead of λ_i. The $k\times k$ matrix H is invertible and 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):

$$v_{it} = u_{it} + \lambda_i'\left(H'\right)^{-1}\left(H'F_t - \hat{F}_t\right). \tag{5}$$

Consider the following notation, which is henceforth used throughout the paper. We let W_ε be a k-dimensional Brownian motion with covariance matrix $\Sigma_{\Delta F}$; $W_{u,i}$ is a scalar Brownian motion independent of W_ε with variance $\sigma_{u,i}^2 = \Gamma_i(1)\Sigma_u\Gamma_i'(1)$, and $\Gamma_i(1) = \sum_{j=0}^{\infty}\Gamma_i$ is the i-th row of $\Gamma(1)$. Also, we define W_u and W_v as standard Brownian motions independent of W_ε; $\sigma_u^2 = \lim_{n\to\infty} n^{-1}\ i_n'\Gamma(1)\Sigma_u\Gamma'(1)\, i_n$ and $\sigma_v^2 = \lim_{n,T\to\infty} E\left[n^{-1/2}T^{-1}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(1+\bar{\lambda}'\Sigma_\Lambda^{-1}\lambda_i\right)u_{it}\right]^2$, with $\bar{\lambda} = \lim_{n\to\infty}\lambda_i$.

Proposition 1 *Let Assumptions 1-4 hold. As $(n,T)\to\infty$, it holds that for every i*

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

also, $T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor} v_{it}\xrightarrow{d} W_{u,i}(s)$, and

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

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

also, $T^{-1}\sum_{t=1}^{T}\hat{F}_t v_{it}\xrightarrow{d} H'\int W_\varepsilon dW_{u,i}$. Further, as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$:

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

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

also, $n^{-1/2}T^{-1}\sum_{t=1}^{\lfloor Ts\rfloor}\sum_{i=1}^{n}v_{it}\xrightarrow{d}\sigma_v W_v(s)$ and $n^{-1/2}T^{-1}\sum_{t=1}^{T}\sum_{i=1}^{n}\hat{F}_t v_{it}\xrightarrow{d}\sigma_v H'\int W_\varepsilon dW_v$.

Proposition 1 contains two types of results: equations (6)-(8), which are univariate, pointwise in i; and equations (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. Equations (9)-(10) highlight an important feature of joint limit theory applied to panel factor series: when deriving results that are pointwise in i, the same results hold whether u_{it} or v_{it} is present, since the extra term $\lambda_i'(H')^{-1}\left(H'F_t-\hat{F}_t\right)$ in (5) is negligible. Conversely, when studying joint asymptotics, the extra term is not negligible.

In the remainder of the paper, we show bootstrap analogues to (6)-(10) - Sections 3 and 4 - and to (9)-(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 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 imposes 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 $\xi_{it} = \sum_{j=1}^{\infty} \beta_{ij} \xi_{it-j} + e_{it}$, and $\beta_i(1) = 1 - \sum_{j=1}^{\infty} \beta_{ij}$.

3.1 The generation of the bootstrap sample

In order to preserve the autocorrelation structure of ΔF_t and u_{it}, we propose to approximate 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 $q_{u,i}$; both pass to infinity as $(n, T) \to \infty$, with no restrictions needed on the relative speed of divergence of n and T. No lower bounds are required for q_F and $q_{u,i}$. 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.

The bootstrapping algorithm is as follows:

Step 1. (PC estimation)

(1.1) Estimate λ_i and F_t in (1) using PC.

(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 $\hat{e}_{it,q} = \left[\hat{e}_{t,q}^{F\prime}, \hat{e}_{it,q}^u\right]'$.

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 $\{e_{it,b}\}_{t=1}^T$, with $e_{it,b} = \left[e_{t,b}^{F\prime}, e_{it,b}^u\right]'$.

(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 $\{x_{it,b}\}_{t=1}^T$ as $x_{it,b} = \hat{\lambda}_i' F_{t,b} + u_{it,b}$.

The algorithm is similar to Chang, Park and Song (2006) for the case of a cointegrating regression. The output of the algorithm above is 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 $e_{it}=\left[e_t^{F\prime},e_{it}^u\right]^\prime$ 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 $W_{iT}(s)\overset{d}{\to}W_i(s)$ 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 $W_{iT}^\prime(s)$, which has the same distribution as $W_{iT}(s)$ and can be chosen such that

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

where $\delta>0$, $r>2$ and M depends only on r. Such results are known as "strong" or "weak" approximations, according as $W_{iT}^\prime(s)$ is shown to converge to $W_i(s)$ a.s. or in probability (see e.g. Sakhanenko, 1980). 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). In our context, $r>8$ in view of Assumptions 1 and 2, so (13) holds.

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 "$\overset{p^b}{\to}$" and "$\overset{d^b}{\to}$" respectively.

We now report three results needed to prove the bootstrap IP. Firstly, we show the existence of moments for $\{e_{it,b}\}_{t=1}^T$. Secondly, we show the consistency of $\hat{\beta}_{q,j}^{(i)}$, estimated in Step 3.2*(a)*. Finally, we show the bootstrap IP for $T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\xi_{it,b}$. It holds that:

Lemma 1 *Let Assumptions 1-5 hold. As $(n,T)\to\infty$, for all (i,t) and $r>8$*

$$E^b\left\|e_{t,b}^F\right\|^r = E\left\|e_t^F\right\|^r+O_p\left(q_F^{-rs}\right)+O_p\left(C_{nT}^{-r}\right)+O_p\left[\left(\frac{q_F}{\varphi_{nT}^F}\right)^r\right] \tag{14}$$

$$\max_{i,t}E^b\left|e_{it,b}^u\right|^r = \max_{i,t}E\left|e_{it}^u\right|^r+O_p\left(q_{u,i}^{-rs}\right)+O_p\left(\delta_{nT}^{-r}\right)+O_p\left[\left(\frac{q_{u,i}}{\varphi_{nT}^u}\right)^r\right]. \tag{15}$$

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. Having q_F and $q_{u,i}\to\infty$ with upper bounds given by φ_{nT}^F and φ_{nT}^u respectively is necessary for the moments of the bootstrap sample to converge to the population values. Lemma 1 and equation (13) yield $T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}e_{it,b}\overset{d^b}{\to}W_i(s)$ in P.

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

$$\max_{1 \le j \le q_F} \left\| \hat{\alpha}_{q,j} - H'\alpha_j \left(H'\right)^{-1} \right\| = O_p\left(\sqrt{\frac{\log T}{T}}\right) + O_p\left(\frac{1}{n}\right) + o_p\left(\frac{1}{q_F^s}\right), \tag{16}$$

$$\max_{1 \le j \le q_{u,i}} \left| \hat{\gamma}_{q,j}^{(i)} - \gamma_j^{(i)} \right| = O_p\left(\sqrt{\frac{\log T}{T}}\right) + O_p\left(\frac{1}{\sqrt{n}}\right) + o_p\left(\frac{1}{q_{u,i}^s}\right). \tag{17}$$

Lemma 2 states that $\hat{\beta}_{q,j}^{(i)}$ is a consistent estimator of the space spanned by β_{ij}, which is a consequence of rotational indeterminacy. Equation (16) suffices for our purposes. 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 $O_p\left(1/n\right)$ and $O_p\left(\sqrt{1/n}\right)$ are due to the use of generated regressors, $\hat{F}_t$ and $\Delta\hat{F}_t$.

Finally, we show the bootstrap IP for $T^{-1/2}\sum_{t=1}^{\lfloor Ts \rfloor} \xi_{it,b}$. Define the partial sums of ξ_{it} as $W_{\xi iT}(s) = T^{-1/2}\sum_{t=1}^{\lfloor Ts \rfloor} \xi_{it}$; by Assumptions 1 and 2, $W_{\xi iT}(s) \xrightarrow{d} W_{\xi i}(s) = \beta_i^{-1}(1) W_i(s)$. Let $W_{\xi iT,b}(s) = T^{-1/2}\sum_{t=1}^{\lfloor Ts \rfloor} \xi_{it,b}$. Combining Lemmas 1 and 2:

Lemma 3 *Let Assumptions 1-5 hold. As* $(n, T) \to \infty$*, it holds that* $\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts \rfloor} \xi_{it,b} \xrightarrow{d^b} W_{\xi i}(s)$ *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 $\{\xi_{it}\}_{t=1}^{T}$. Lemma 3 requires an IP for $\{e_{it,b}\}_{t=1}^{T}$; this follows from Lemma 1. The Lemma also requires that that $\hat{\beta}_{i,q}^{-1}(1) \xrightarrow{p} \beta_i^{-1}(1)$; this follows from Lemma 2. Lemma 3 is the bootstrap counterpart to (6).

Combining Lemmas 1-3, it holds that:

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

$$\frac{1}{T^2}\sum_{t=1}^{T} F_{t,b}F_{t,b}' \xrightarrow{d^b} H'\left(\int W_\varepsilon W_\varepsilon'\right) H, \tag{18}$$

$$\frac{1}{T}\sum_{t=1}^{T} F_{t,b}u_{it,b} \xrightarrow{d^b} H'\int W_\varepsilon dW_{u,i}, \tag{19}$$

in P*, with* W_ε *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 "one unit at a time" algorithm is valid when cross dependence 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 no, or little, cross dependence. In this section we present validity results for the bootstrap estimates of loadings (Section 4.1), common factors (Section 4.2), and common components $C_{it} = \lambda_i' F_t$ (Section 4.3). We also derive square integrability results, which are required when bootstrap standard errors need to be computed.

We consider the following DGP for $x_{it,b}$, defined in Step (3.2.c) of the algorithm:

$$x_{it,b} = \hat{\lambda}_i' F_{t,b} + u_{it,b}; \tag{20}$$

alternatively, a "fixed regressors" approach could be used, viz. $x_{it,b} = \hat{\lambda}_i' \hat{F}_t + u_{it,b}$. We study the application of PC to (20): loadings and factors are extracted from $x_{it,b}$, without treating $\hat{\lambda}_i$ or $\hat{F}_t$ as observed.[2] This approach should be less dependent than OLS on the quality of the first step estimates $\left(\hat{\lambda}_i, \hat{F}_t\right)$. The same restrictions as for the computation of $\left(\hat{\lambda}_i, \hat{F}_t\right)$ can be used at each bootstrap iteration.

The issue of identification 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 knowing (λ_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

[2] In Trapani (2012) we also report results for the case of OLS estimation, whereby loadings are estimated through a time series regression using $\hat{F}_t$ as observable regressors, and factors are estimated through a cross sectional regression with $\hat{\lambda}_i$ treated as observable.

with estimated latent factors (Bai and Ng, 2006b). In these contexts, the bootstrap can be useful. Conversely, when using Factor-Augmented regression, rotational indeterminacy makes it impossible to give a structural interpretation to the slopes of factors (see also Goncalves and Perron, 2011); the bootstrap is of course unable to ameliorate this. Secondly, rotational indeterminacy also affects the bootstrap when PC is applied. Considering the estimation of the loadings as a leading example, these are estimated up to a rotation matrix H_1, given by

$$H_1 = \left[\frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}_i\hat{\lambda}_i'\right]\left[\frac{1}{T^2}\sum_{t=1}^{T}F_{t,b}\hat{F}_{t,b}^{PC\prime}\right]\left[V_{nT}^{b}\right]^{-1}, \tag{21}$$

where $\hat{F}_{t,b}^{PC}$ is the PC estimate of the common factors and V_{nT}^{b} contains the first k eigenvalues of X_bX_b' in descending order with $X_b = [x_{1,b}, ..., x_{n,b}]'$ and $x_{i,b} = [x_{i1,b}, ..., x_{iT,b}]'$. Estimation is carried out under the restrictions $\frac{1}{T^2}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}\hat{F}_{t,b}^{PC\prime} = I_k$ and $\frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}_{i,b}^{PC}\hat{\lambda}_{i,b}^{PC\prime}$ diagonal. The matrix H_1 can be computed, since it is based only on estimated quantities. Indeed, as shown by Goncalves and Perron (2011) based on Bai and Ng (2011) in the context of Factor-Augmented regression, H_1 is asymptotically diagonal, with elements equal to ± 1. This result, however, is only asymptotic.

4.1 Loadings

Consider $\hat{\lambda}_i$. Lemmas 4 and B.4 in Bai (2004) entail that $\hat{\lambda}_i - H^{-1}\lambda_i = \left[\sum_{t=1}^{T}\hat{F}_t\hat{F}_t'\right]^{-1}\left[H'\sum_{t=1}^{T}F_tu_{it}\right] + o_p(1)$. Using (7) and (8), it holds that

$$T\left(\hat{\lambda}_i - H^{-1}\lambda_i\right) \xrightarrow{d} H^{-1}\left(\int W_\varepsilon W_\varepsilon'\right)^{-1}\left(\int W_\varepsilon dW_{u,i}\right), \tag{22}$$

which is the same as Theorem 3 in Bai (2004). When applying PC, $\hat{\lambda}_{i,b}^{PC}$ estimates $H_1^{-1}\hat{\lambda}_i$, with H_1 defined in (21).

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

$$T \times H_1\left[\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right] \xrightarrow{d^b} \left[H^{-1}\left(\int W_\varepsilon W_\varepsilon'\right)^{-1}\left(\int W_\varepsilon dW_{u,i}\right)\right] \text{ in } P, \tag{23}$$

$$E^b\left\|T \times H_1\left[\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right]\right\|^{2+\delta} = O_p(1). \tag{24}$$

Equation (23) is the bootstrap counterpart to Theorem 3 in Bai (2004): the limiting distribution of $\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i$ is the same as the limiting distribution of $\hat{\lambda}_i - H^{-1}\lambda_i$, except for the presence of H_1, which anyway can be computed using (21). This is a consequence of rotational indeterminacy. The limiting distribution of $\hat{\lambda}_i - H^{-1}\lambda_i$ is approximated by $H_1\left[\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right]$. This should be used e.g. when bootstrapping t-statistics. Given (23), equation (24) ensures that the sequence $T\left\|H_1\left(\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right)\right\|^2$ is integrable, which is useful for the bootstrap computation of the standard error of $T \times H_1\left[\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right]$.

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

$$\sqrt{n}\left[\hat{F}_t - H'F_t\right] \xrightarrow{d} H'\Sigma_\Lambda^{-1} \times N\left(0, \Gamma_t\right), \tag{25}$$

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 it can be expected that the bootstrap provides valid inference on factors only when $E\left(u_{it}u_{jt}\right) = 0$ 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 $n \to \infty$, viz.

$$\frac{1}{n}\sum_{i=1}^n \sum_{j\neq i}^n E\left[\lambda_i\lambda_j' u_{it}u_{jt}\right] = o\left(1\right). \tag{26}$$

In general, consistent estimation of Γ_t is fraught with difficulties; as Bai (2003) points out in a stationary context, HAC-type estimators are not feasible since, in general, the order of cross correlation is unknown. Bai and Ng (2006b) provide a solution for the case $\Gamma_t = \Gamma$ for all t.

Similarly to the case of the loadings, PC estimates $H_1'\hat{F}_t$.

Proposition 3 *Let Assumptions 1-5 and (26) hold and let $\delta > 0$. As $(n,T) \to \infty$ with $\frac{n}{T^3} \to 0$*

$$\sqrt{n}\left(H_1'\right)^{-1}\left[\hat{F}_{t,b}^{PC} - H_1'F_{t,b}\right] \xrightarrow{d^b} \left[H'\Sigma_\Lambda^{-1} \times N\left(0, \Gamma_t\right)\right] \text{ in } P. \tag{27}$$

As $(n, T) \to \infty$ *with* $\frac{n}{T^2} \to 0$

$$E^b \left\| \sqrt{n} \left(H_1'\right)^{-1} \left[\hat{F}_{t,b}^{PC} - H_1' F_{t,b}\right] \right\|^{2+\delta} = O_p(1). \tag{28}$$

Proposition 3 ensures that $\sqrt{n} \left(H_1'\right)^{-1} \left[\hat{F}_{t,b}^{PC} - H_1' F_{t,b}\right]$ is valid in order 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). Equation (28) provides a uniform integrability result that is similar to (24); note the stronger restriction $\frac{n}{T^2} \to 0$.

4.3 Common components

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

$$\hat{C}_{it} - C_{it} = \left(\hat{F}_t - H'F_t\right)' H^{-1}\lambda_i + \hat{F}_t' \left(\hat{\lambda}_i - H^{-1}\lambda_i\right) = I + II. \tag{29}$$

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

$$\sqrt{n}\left(\hat{C}_{it} - C_{it}\right) \xrightarrow{d} \lambda_i' \Sigma_\Lambda^{-1} N(0, \Gamma_t) + \sqrt{\pi} W_\varepsilon(s) \left(\int W_\varepsilon W_\varepsilon'\right)^{-1} \int W_\varepsilon dW_{u,i}, \tag{30}$$

where the first term on the right hand side comes from I in (29) and the second one from II. When considering the bootstrap estimate, we have

$$\hat{C}_{it,b}^{PC} - \hat{C}_{it} = \left(\hat{F}_{t,b}^{PC} - H_1' F_t\right)' H_1^{-1} \hat{\lambda}_i + \hat{F}_{t,b}^{PC\prime} \left(\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right). \tag{31}$$

Proposition 4 *Let Assumptions 1-5 and (26) hold. Then, for all* (i, t) *such that* $t = \lfloor Ts \rfloor$, *as* $(n, T) \to \infty$ *with* $\frac{n}{T} \to \pi$

$$\sqrt{n}\left(\hat{C}_{it,b}^{PC} - \hat{C}_{it}\right) \xrightarrow{d^b} \lambda_i' \Sigma_\Lambda^{-1} N(0, \Gamma_t) + \sqrt{\pi} W_\varepsilon(s) \left(\int W_\varepsilon W_\varepsilon'\right)^{-1} \int W_\varepsilon dW_{u,i} \text{ in } P. \tag{32}$$

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

$$E^b \left| \sqrt{n}\left(\hat{C}_{it,b}^{PC} - \hat{C}_{it}\right) \right|^{2+\delta} = O_p(1). \tag{33}$$

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 π 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 e.g. for the bootstrap approximation of loadings, or under certain restrictions on the degree of cross dependence such as (26). However, in other cases cross sectional correlation needs to be taken into account. In this section, cross dependence is taken into account by fitting an n-dimensional VAR to the vector of the residuals $\hat{u}_{it}$. We report the bootstrap counterpart to equations (9)-(10), showing that the moment existence conditions granted by Lemma 1 are sufficient also for joint bootstrap asymptotics. However, difficulties arise when estimating the long run covariance matrix of u_t.

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 $\hat{u}_t = [\hat{u}_{1t}, ..., \hat{u}_{nt}]'$. Consider 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

$$u_t = \sum_{j=1}^{q} B_j u_{t-j} + e^u_t, \tag{34}$$

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 $B(1)$. The bootstrap sample $u_{t,b}$ can be generated using $u_{t,b} = \sum_{j=1}^{q} B^*_j u_{t-j,b} + e^u_{t,b}$. No modifications are required to the algorithm in Section 3.1 when generating $F_{t,b}$. It holds that:

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* $qn^2 < T$

$$\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{\lfloor Ts \rfloor} u_{it,b} \stackrel{d^b}{\to} \sigma_v W_v(s), \tag{35}$$

$$\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{F}^{PC}_{t,b} u_{it,b} \stackrel{d^b}{\to} \sigma_v H_1' H' \int W_\varepsilon dW_v, \tag{36}$$

in P, where W_v, W_ε and σ_v are defined in Proposition 1.

Theorem 2 contains joint asymptotics results. The distribution of $n^{-1/2}T^{-1/2}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it,b}$ is, asymptotically, the same as the distribution of $n^{-1/2}T^{-1/2}\sum_{i=1}^{n}\sum_{t=1}^{T}v_{it}$. However, when using the VAR approach, the number of parameters to be estimated is qn^2, whence the requirement that $qn^2 < T$. This constraint on the relative speed of divergence of n and T is stronger than the typical requirement that $\frac{n}{T} \to 0$. Indeed, in this context $\varphi_{nT}^{u} = \sqrt{n}$; choosing (according to Assumption 5) q as $n^{1/\omega}$, for some $\omega > 2$, entails that it must hold $\frac{n^{2+1/\omega}}{T} \to 0$, which restricts the applicability of the VAR based approach.

Equation (35) could be generalised to study multiparameter partial sum processes such as $(nT)^{-1/2}\sum_{i=1}^{\lfloor np\rfloor}\sum_{t=1}^{\lfloor Ts\rfloor}u_{it,b}$, with $(p,s) \in [0,1]\times[0,1]$. Having $\max_{i,t} E\left|u_{it,b}\right|^{2+\delta} < \infty$ yields $(nT)^{-1/2}\sum_{i=1}^{\lfloor np\rfloor}\sum_{t=1}^{\lfloor Ts\rfloor}u_{it,b} \xrightarrow{d} \sigma_v W(p,s)$, 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 (2011).

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 $(nT)^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\sum_{i=1}^{n}v_{it}$. Writing this in matrix form, a requirement for this is that $(nT)^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}i_n'\left[B^*(1)\right]^{-1}e_{t,b}^{u}$ and $(nT)^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}i_n'\left[B(1)\right]^{-1}e_t^{v}$ should have the same distribution, where e_t^v is the idiosyncratic innovation of the vector $[v_{1t},...,v_{nt}]'$. 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 $e^u_{jt,b}$ has finite variance, it is sufficient that $\sup_j \left|\sum_{i=1}^n \left\{[B^*(1)]^{-1} - [B(1)]^{-1}\right\}_{ij}\right|$ $= o_p(1)$, for which it is sufficient that $\left\|[B^*(1)]^{-1} - [B(1)]^{-1}\right\|_1 = o_p(1)$. This holds if $\|B^*(1) - B(1)\|_1 = o_p(1)$, since $\left\|[B^*(1)]^{-1} - [B(1)]^{-1}\right\|_1 \leq \left\|\Gamma^{-1}(1)\right\|_1 \left\|\Gamma^{-1}(1)\right\|_\infty$ $\|B^*(1) - B(1)\|_1$ 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 (34). Defining $u_{qt} = \left[u'_{t-1}, ..., u'_{t-q}\right]'$ and $B_q = [B_{q,1}|...|B_{q,q}]$, we have

$$u_t = B_q u_{qt} + e^u_{qt}; \tag{37}$$

B_q is estimated by

$$\hat{B}_q = \left[\sum_{t=q+1}^T \hat{u}_t \hat{u}'_{qt}\right] \left[\sum_{t=q+1}^T \hat{u}_{qt} \hat{u}'_{qt}\right]^{-1}. \tag{38}$$

Thus, $B(1)$ can be estimated by $\widehat{B_q(1)} = 1 - \sum_{j=1}^q \hat{B}_{q,j}$. Equation (38) requires the inversion of an $nq \times nq$ matrix. Also, as pointed out above, it is also required that $qn^2 < T$.

We also consider an alternative estimator of $B(1)$ which does not take into account the cross sectional correlation among the u_{it}s. Consider the $\hat{\gamma}_{q,j}$s estimated from (12), and let $\tilde{B}_{q,j}$ an $n \times n$ diagonal matrix with the $\hat{\gamma}^{(i)}_{q,j}$s on the main diagonal. We define $\widetilde{B_q(1)} = 1 - \sum_{j=1}^q \tilde{B}_{q,j}$, as an alternative estimator of $B(1)$. 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*

$$\left\|\widehat{B_q(1)} - B(1)\right\|_1 = O_p\left(q\sqrt{\frac{\log T}{T}}\right) + O_p\left(nq^2 C_{nT}^{-1}\right) + o\left(q^{-s}\right) + o_p(1). \tag{39}$$

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

$$\left\|\widetilde{B_q(1)} - B(1)\right\|_1 = O_p\left(\frac{q}{\varphi^u_{nT}}\right) + O_p\left(qn^{-\phi}\right) + o\left(q^{-s}\right). \tag{40}$$

Theorem 3 states that $\widehat{B_q(1)}$ is inconsistent in L_1-norm. The term that determines such inconsistency is the one of magnitude $O_p\left(nq^2 C_{nT}^{-1}\right)$: by the definition of $C_{nT} =$

$\min\{\sqrt{n}, T\}$, as $(n, T) \to \infty$ it holds that $nC_{nT}^{-1} \to \infty$, whence the inconsistency. This term arises due to the presence of $\left[\hat{\Lambda} - \Lambda (H')^{-1}\right] F_t$ in the $\hat{u}_{it}$s. 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 (Trapani, 2012). This result holds in spite of Assumption 1*(iii)*(a), which limits the amount of cross dependence among the u_{it}s, and it 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 $\left[\hat{\Lambda} - \Lambda (H')^{-1}\right] F_t$. 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 when $\phi > 0$, as long as Assumption 5 is modified as $q \to \infty$ 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 (40) 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 a typical approximate factor model context (see e.g. Assumption C(1) in Bai, 2004), where it suffices to have $n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n} |\tau_{ij}| = O(1)$, which is entailed by $\sup_i \sum_{j=1}^{n} |\tau_{ij}| = O(1)$. Conversely, the assumption is more general than in classical Principal Component Analysis, where $\tau_{ij} = 0$ for all $i \neq j$. Thus, in essence (40) 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. 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 $B(1)$.

Finally, we point out that, in order to take into account cross dependence, an alternative approach is to consider the case in which both $I(1)$ and $I(0)$ common factors are present in the DGP of x_{it}, viz.

$$x_{it} = \lambda_i^{F\prime} F_t + \lambda_i^{G\prime} G_t + u_{it}, \tag{41}$$

with F_t a k-dimensional nonstationary process, and G_t h-dimensional and stationary. This approach helps to avoid the issues discussed above, and of course it is valid as long as cross dependence among the u_{it}s genuinely arises from a common factor structure. In Trapani (2012), some initial routes to the extension of sieve bootstrap to (41) are presented; in particular, analogous results to Lemmas 1 and 2 are derived, from which results such as Lemma 3 and Theorem 1 can be extended to this context. Whilst this goes beyond the scope of this paper, we point out that the theory developed in this paper holds in presence of stationary common factors such as in (41), and also in presence of deterministics such as drifts in the DGP of F_t, with only minor modifications to the rates of convergence and to the choice of the truncation lags in (11)-(12).

6 Simulation results

This section contains an investigation of the finite sample performance of the bootstrap using synthetic data. The experiments reported here are based on calculating confidence intervals for the common factors; in particular, we verify whether bootstrap confidence intervals are more accurate than the asymptotic ones.

Data are generated according to the following scheme, largely based on the setup in Bai (2004, Section 6):

$$x_{it} = \lambda_i F_t + u_{it}, \tag{42}$$

where λ_i is generated as i.i.d. $N(1,1)$. Equation (42) is based on a single, nonstationary factor $F_t = F_{t-1} + \varepsilon_t$ with

$$\varepsilon_t = \rho^{\varepsilon} \varepsilon_{t-1} + w_t^F + \vartheta^{\varepsilon} w_{t-1}^F,$$

and $\rho^{\varepsilon} = \{0, 0.5\}$ and $\vartheta^{\varepsilon} = \{0, 0.5\}$. Results obtained using $\rho^{\varepsilon} = \{0, 0.5\}$ and $\vartheta^{\varepsilon} = \{0, 0.5, -0.5\}$ look very similar for all combinations $(\rho^{\varepsilon}, \vartheta^{\varepsilon})$; thus, to save space we only report empirical sizes calculated for $(\rho^{\varepsilon}, \vartheta^{\varepsilon}) = (0, 0)$; the full set of results is in Trapani

(2012). The error term is generated as

$$u_{it} = \rho^u u_{it-1} + w_t^u + \vartheta^u w_{t-1}^u,$$

with $\rho^u = \{0, 0.5\}$ and $\vartheta^u = \{0, 0.5, -0.5\}$; as above, results look similar across all combinations of (ρ^u, ϑ^u) and therefore we only report results for $(\rho^u, \vartheta^u) \in \{(0,0); (0.5, 0.5); (0, -0.5)\}$ to save space.[3] All idiosyncratic innovations are generated as i.i.d. $N(0,1)$; when generating series, the first 1,000 observations were discarded to avoid dependence on initial conditions. The bandwidths q_F and $q_{u,i}$ are been selected as $\min\left\{n^{1/3}, T^{1/3}\right\}$.

We evaluate the bootstrap approximation of confidence intervals for F_t, using a "one unit at a time" resampling scheme. In view of (25), we expect that, 95% of the times across the Monte Carlo replications, $\left|\hat{F}_t - H'F_t\right| \leq \frac{1.96}{\sqrt{n}} \hat{S}_t$, where $\hat{S}_t^2 = V_{nT}^{-1} \hat{\Gamma}_t V_{nT}^{-1}$. Indeed, by calculating how many times $\left|\hat{F}_t - H'F_t\right| > \frac{1.96}{\sqrt{n}} \hat{S}_t$, we get the empirical rejection frequency of a test for the null that F_t is the true factor. We define such empirical size as ERF_t.

The bootstrap empirical size is computed as follows. For $b = 1, ..., \beth$ iterations, we generate $F_{t,b}$ and compute $\hat{F}_{t,b}^{PC}$ as defined is Section 4.2. We define the vector $\left[\left|(H_1')^{-1}\left(\hat{F}_{t,1}^{PC} - H_1'F_{t,1}\right)\right|, ..., \left|(H_1')^{-1}\left(\hat{F}_{t,\beth}^{PC} - H_1'F_{t,\beth}\right)\right|\right]'$, with H_1 defined in (21), and calculate the rank of $\left|\hat{F}_t - H'F_t\right|$, say $r\left(\hat{F}_t\right)$. If $1 - \frac{r(\hat{F}_t)}{\beth+1} < 0.05$, this corresponds to a rejection of the null that F_t is the true factor. We define the bootstrap empirical rejection frequency as ERF_t^*. Ideally, both ERF_t and ERF_t^* should be close to 5%, for each t.

We report the averages $ERF = T^{-1}\sum_{t=1}^{T} ERF_t$ and $ERF^* = T^{-1}\sum_{t=1}^{T} ERF_t^*$. All experiments have been carried out with 1,000 replications, and $\beth = 199$ bootstrap iterations for each replication.

[Insert Table 1 somewhere here]

Table 1 shows that the asymptotic test is oversized, although this improves as (n,T) - in particular T - increase. The bootstrap based test attains, with few exceptions, the correct size: results are good even for very small sample sizes, as it can be noted from the figures in the panel corresponding to the case $(n,T) = (20,20)$: bootstrap tests strongly

[3] The full set of result for all combinations of $(\rho^\varepsilon, \vartheta^\varepsilon)$ and (ρ^u, ϑ^u) is reported in Trapani (2012).

improve the performance of asymptotic tests. Similarly, the bootstrap is robust against serial correlation in u_{it}: results differ marginally across different combinations of (ρ^u, ϑ^u). This also includes the case of negative MA roots, which is understood to be usually problematic. The asymptotic test is more variable across combinations of (ρ^u, ϑ^u), although this is attenuated as n and T increase. The impact of serial dependence in ε_t is stronger, particularly for small values of n and T: the panel corresponding to $(n, T) = (20, 20)$ shows significant differences for all three tests across different values of $(\rho^\varepsilon, \vartheta^\varepsilon)$. Finally, we point out that other, unreported experiments were carried out using different rules to choose the bandwidths; e.g. when using $\min\left\{n^{2/5}, T^{2/5}\right\}$, the results were only marginally affected by such choice. Similarly, the number of bootstrap repetitions $\beth$ does not seem to have an impact on the results either. These considerations provide some guidelines as to how to implement bootstrap based tests.

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. A possible solution is to allow for stationary factors, which may capture cross dependence; some initial results are in the extended version of this paper (Trapani, 2012).

In addition to the extension of the results to a more general framework, we note that the focus of this paper is on applying ordinary PC estimation to the bootstrap sample. More efficient estimation techniques could be employed; see e.g. the GLS-type estimators proposed by Choi (2011; see also Choi, 2012). Extensions to this context can be directly based on the theory developed in this paper: for example, the asymptotics of the GLS-type estimators of F_t in Choi (2011) is based on very similar derivations as in the context of ordinary PC. In that case, it can be expected that the bootstrap theory developed here can be readily generalised. This issue is currently under investigation by the author.

Appendix: proofs and derivations

This Appendix contains the proofs of the main results in the paper (proofs or parts thereof that are omitted can be found in Trapani, 2012). We start by reporting four preliminary Lemmas:

Lemma A.1 *Let Assumptions 1-4 hold. Then: (i)* $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t - H'\Delta F_t\right\|^2 = O_p\left(C_{nT}^{-2}\right)$*; (ii)* $T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{F}_t - H'\Delta F_t\right)'\Delta F_t = O_p\left(T^{-1/2}C_{nT}^{-1}\right)$*; (iii)* $T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{F}_t - H'\Delta F_t\right)'\Delta\hat{F}_t = O_p\left(n^{-1}\right) + O_p\left(T^{-3/2}\right)$*; (iv)* $T^{-1/2}\sum_{t=1}^{T}\left(\Delta\hat{F}_t - H'\Delta F_t\right) = O_p\left(C_{nT}^{-1}\right)$*; (v)* $T^{-1/2}\sum_{t=1}^{T}\left(\hat{F}_t - H'F_t\right) = O_p\left(C_{nT}^{-1}\right)$.

Lemma A.2 *Let Assumptions 1-4 hold. Then, for some* $r \geq 2$ *such that* $E\left|u_{it}\right|^r < \infty$ *and* $E\left\|e_t^F\right\|^r < \infty$*: (i)* $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t - H'\Delta F_t\right\|^r = O_p\left(C_{nT}^{-r}\right)$*; (ii)* $T^{-1}\sum_{t=1}^{T}\left\|\Delta\hat{F}_t\right\|^r = O_p(1)$*; (iii)* $T^{-1}\sum_{t=1}^{T}\left|\hat{u}_{it} - u_{it}\right|^r = O_p\left(\delta_{nT}^{-r}\right)$*; (iv)* $T^{-1}\sum_{t=1}^{T}\left|\hat{u}_{it}\right|^r = O_p(1)$.

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

Lemma A.4 *Let Assumptions 1-5 hold. Then, for some* $r \geq 2$ *such that* $E\left|u_{it}\right|^r < \infty$ *and* $E\left\|e_t^F\right\|^r < \infty$*: (i)* $T^{-1}\sum_{t=1}^{T}\left\|\hat{F}_{t,b} - H_1'F_{t,b}\right\|^r = O_p\left(C_{nT}^{-r}\right)$*; (ii)* $T^{-1}\sum_{t=1}^{T}\left(\hat{F}_{t,b} - H_1'F_{t,b}\right)F_{t,b}' = O_p\left(C_{nT}^{-1}\right)$*; (iii)* $T^{-1}\sum_{t=1}^{T}\left(\hat{F}_{t,b} - H_1'F_{t,b}\right)\hat{F}_{t,b}' = O_p\left(C_{nT}^{-1}\right)$*; (iv)* $\hat{F}_{t,b} - H_1'F_{t,b} = O_p\left(n^{-1/2}\right) + O_p\left(T^{-3/2}\right)$*; (v)* $T^{-1}\sum_{t=1}^{T}\left\|\left(\hat{F}_{t,b} - H_1'F_{t,b}\right)\hat{F}_{t,b}'\right\|^r = o_p(1)$.

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

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

Term I converges weakly to a Brownian motion - see Phillips and Solo (1992). As regards II, it is $o_p(1)$ by Lemma A.4*(iv)*, which proves (6). The limit of $T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}v_{it}$ follows from $T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}v_{it}=T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}u_{it}+\lambda_i'(H')^{-1}T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\left(H'F_t-\hat{F}_t\right)$; applying Lemma A.1*(iv)* yields the desired result.

We now turn to proving equation (7). We have $T^{-2}\sum_{t=1}^{T}\hat{F}_t\hat{F}_t'=H'\left(T^{-2}\sum_{t=1}^{T}F_tF_t'\right)H$ $+H'\left[T^{-2}\sum_{t=1}^{T}F_t\left(\hat{F}_t-H'F_t\right)'\right]+\left[T^{-2}\sum_{t=1}^{T}\left(\hat{F}_t-H'F_t\right)F_t'\right]H+T^{-2}\sum_{t=1}^{T}\left(\hat{F}_t-H'F_t\right)$ $\left(\hat{F}_t-H'F_t\right)'=I+II+III+IV$. Term I converges to $H'\left(\int W_\varepsilon W_\varepsilon'\right)H$ by Assumption 2. As regards II and III, 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 the proof of equation (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 in the theory of convergence to stochastic integrals (see e.g. Phillips, 1988). As regards II, $II\leq\left[T^{-1}\sum_{t=1}^{T}\left\|\hat{F}_t-H'F_t\right\|^2\right]^{1/2}\left[T^{-1}\sum_{t=1}^{T}u_{it}^2\right]^{1/2}=O_p\left(C_{nT}^{-1}\right)O_p(1)$, which is negligible. This proves (8). The limit of $T^{-1}\sum_{t=1}^{T}\hat{F}_tv_{it}$ follows from noting that $T^{-1}\sum_{t=1}^{T}\hat{F}_tv_{it}$ $=T^{-1}\sum_{t=1}^{T}\hat{F}_tu_{it}+\lambda_i'(H')^{-1}T^{-1}\sum_{t=1}^{T}\hat{F}_t\left(H'F_t-\hat{F}_t\right)$, where the last term is $O_p\left(C_{nT}^{-1}\right)$ from Lemma B.4 in Bai (2004).

We now prove (9). Let the martingale approximation of u_{it} (derived from the Beveridge-Nelson decomposition, BN henceforth) 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 asymptotics (Phillips and Moon, 1999) yields $II=O_p\left(\sqrt{\frac{n}{T}}\right)$. As regards I, define $\zeta_{nt}=n^{-1/2}\sum_{i=1}^{n}u_{it}^*$; this is an MDS for all n. Also, consider $E|\zeta_{nt}|^{2+\delta}$. By virtue of Assumption 1*(iii)*(c), $E|\zeta_{nt}|^{2+\delta}\leq$

$Mn^{-\frac{2+\delta}{2}} E \left| \sum_{i=1}^{n} |u_{it}^*|^2 \right|^{\frac{2+\delta}{2}}$. Thus

$$E |\zeta_{nt}|^{2+\delta} \leq M' n^{-\frac{2+\delta}{2}} E \left| \left(\sum_{i=1}^{n} |u_{it}^*|^{2+\delta} \right)^{\frac{2}{2+\delta}} n^{1-\frac{2}{2+\delta}} \right|^{\frac{2+\delta}{2}} \leq M'' \frac{1}{n} \sum_{i=1}^{n} E |u_{it}^*|^{2+\delta}, \quad (43)$$

where the first passage is based on Holder's inequality and the second one follows by the C_r-inequality and convexity. Given that $\max_i E |u_{it}^*|^{2+\delta} < \infty$ in view of Lemma 1, $E |\zeta_{nt}|^{2+\delta} < \infty$ 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 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 σ_u^2; Assumption 1*(iii)*(a) ensures that $\sigma_u^2 < \infty$. The limit of $\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} v_{it}$ follows from similar passages - see Trapani (2012).

Finally, we report the proof of (10). We have $\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{F}_t u_{it} = \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} H' F_t u_{it}$ $+ \frac{1}{T} \sum_{t=1}^{T} \left(\hat{F}_t - H' F_t \right) \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} u_{it} \right) = I + II$. 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 regards I, let the martingale approximation to F_t be F_t^*. Then

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

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, ζ_{nt} and F_t^* are independent for all n in light of Assumption 4*(i)*. Thus $I_a \xrightarrow{d} \sigma_u \int W_\varepsilon dW_u$ - see Phillips (1988). 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 $\frac{n}{T} \to 0$. The proof of the limit of $\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \hat{F}_t v_{it}$ follows from similar arguments. ■

Proof of Lemma 1. We suppress the subscripts in q_F and $q_{u,i}$ when no ambiguity arises, and we let "r" denote "$8+\delta$" for short. Consider (14); recall (11) and

$$\Delta \hat{F}_t = \sum_{j=1}^{q} \hat{\alpha}_{q,j} \Delta \hat{F}_{t-j} + \hat{e}_{t,q}^F, \tag{44}$$

$$\Delta F_t = \sum_{j=1}^{\infty} \alpha_j \Delta F_{t-j} + e_t^F. \tag{45}$$

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 $\|H\|^r = O_p(1)$. 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 regards$II$, it holds that $e_{t,q}^F - e_t^F = \sum_{j=q+1}^{\infty} \alpha_j \Delta F_{t-j}$; thus, Minkowski's inequality and the stationarity of ΔF_t yield $T^{-1}\sum_{t=1}^T \left\| e_{t,q}^F - e_t^F \right\|^r = T^{-1}\sum_{t=1}^T \left\| \sum_{j=q+1}^{\infty} \alpha_j \Delta F_{t-j} \right\|^r \leq T^{-1}\sum_{t=1}^T \|\Delta F_t\|^r \left(\sum_{j=q+1}^{\infty} \|\alpha_j\| \right)^r$. Assumption 1*(ii)* entails that $\sum_{j=q+1}^{\infty} \|\alpha_j\| = o(q^{-s})$; by Assumption 2*(i)*, $T^{-1}\sum_{t=1}^T \|\Delta F_t\|^r = O_p(1)$. Thus, $II = o_p(q^{-rs})$. As regards III, $\hat{e}_{t,q}^F - H' e_{t,q}^F = \sum_{j=0}^q H' \alpha_{q,j} (H')^{-1} \left(\Delta \hat{F}_{t-j} - H' \Delta F_{t-j} \right) - \sum_{j=1}^q \left[\hat{\alpha}_{q,j} - H' \alpha_{q,j} (H')^{-1} \right] \Delta \hat{F}_{t-j}$, where $\alpha_{q,0} = -1$. Hence

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

Using Minkowski's inequality, the former term is bounded by $MT^{-1}\sum_{t=1}^T \left\| \Delta \hat{F}_t - \Delta F_t \right\|^r \left(\sum_{j=0}^q \|\alpha_{q,j}\| \right)^r$, 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, the former term is of magnitude $O_p\left(C_{nT}^{-r} \right)$. As regards the latter, it is bounded by $T^{-1}\sum_{t=1}^T \left\| \Delta \hat{F}_t \right\|^r \left(\sum_{j=0}^q \left\| \hat{\alpha}_{q,j} - H' \alpha_{q,j} (H')^{-1} \right\| \right)^r$. Lemma A.2 ensures $T^{-1}\sum_{t=1}^T \left\| \Delta \hat{F}_t \right\|^r = O_p(1)$. Also, $\sum_{j=0}^q \left\| \hat{\alpha}_{q,j} - H' \alpha_{q,j} (H')^{-1} \right\| \leq q \max_{1\leq j\leq q} \left\| \hat{\alpha}_{q,j} - H' \alpha_{q,j} (H')^{-1} \right\|$, and Lemma 2 yields $\left[q \max_{1\leq j\leq q} \left\| \hat{\alpha}_{q,j} - H' \alpha_{q,j} (H')^{-1} \right\| \right]^r$ $= O_p\left[q^r T^{-r/2} (\log T)^{r/2} + q^r n^{-r} \right] + o_p(1)$. Thus, $III = O_p\left(C_{nT}^{-r} \right) + O_p\left[\left(\frac{q}{\varphi_{nT}^F} \right)^r \right]$. 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 js

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

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. By Lemma 2, $\max_{1\leq j\leq q}\|\hat{\alpha}_{q,j}$ $-\ H'\alpha_{q,j}\left(H'\right)^{-1}\Big\|^2 = O_p\left(\frac{1}{\varphi_{nT}^F}\right)+O_p\left(n^{-3/2}T^{-1/2}\right)$. Hence, $IV = O_p\left(q^{r/2}T^{-r/2}\right)$. Combining all these results

$$E^b\left\|e_{t,b}^F\right\|^r = E\left\|e_t^F\right\|^r+O_p\left(q^{-rs}\right)+O_p\left(C_{nT}^{-r}\right)+O_p\left[\left(\frac{q}{\varphi_{nT}^F}\right)^r\right]+o_p\left(1\right) = E\left\|e_t^F\right\|^r+o_p\left(1\right),$$

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

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

This proves (14). The proof of (15) is in Trapani (2012). ■

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. Passages are similar to the proof of Lemma A.1 in Chang, Park and Song (2006). Consider (16). Recall (11), (44) and (45), 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 regards II, Assumptions 1*(ii)* and 2*(ii)* yield $\max_{1\leq j\leq q}|\alpha_{q,j}-\alpha_j| \leq \sum_{j=1}^{q}|\alpha_{q,j}-\alpha_j| = o\left(q^{-s}\right)$ - 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, and it is in Trapani (2012). ■

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 (18), and assume, for simplicity, that $F_{0,b} = 0$. Letting $W_{\varepsilon,nT}^{(b)}(s) = T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\Delta F_{t,b}$, Lemma 3 states that, as $(n,T)\to\infty$, $W_{\varepsilon,nT}^{(b)}(s) \stackrel{d^*}{\to} W_{\varepsilon}(s)$. Thence, $T^{-2}\sum_{t=1}^{T}F_{t,b}F_{t,b}' \underset{d^b}{=} \int W_{\varepsilon,nT}^{(b)}(s)\,W_{\varepsilon,nT}^{(b)}(s)'$ $+\ T^{-2}\sum_{t=1}^{T}F_{T,b}F_{T,b}'$, and $T^{-1/2}F_{T,b} = o_p(1)$, which proves (18). As regards (19), 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

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}F_{t,b}u_{it,b} &= \frac{1}{T}\sum_{t=1}^{T}F_{t,b}^*u_{it,b}^* + \frac{1}{T}\sum_{t=1}^{T}\Delta F_{t,b}\bar{u}_{it,b} - \frac{1}{T}F_{T,b}\bar{u}_{iT,b} \qquad (46) \\
&+ \frac{1}{T}\overline{\Delta F}_{0,b}\sum_{t=1}^{T}u_{it,b}^* - \frac{1}{T}\sum_{t=1}^{T}\overline{\Delta F}_{t-1,b}u_{it,b}^* = I + II + III + IV + V.
\end{aligned}
$$

It holds straightforwardly that $III+IV+V = O_p\left(T^{-1/2}\right)$; also, $II \leq \left[T^{-1}\sum_{t=1}^{T}\left\|\Delta F_{t,b}\right\|^2\right]^{1/2}$

$\left[T^{-1}\sum_{t=1}^{T}\|\bar{u}_{it,b}\|^{2}\right]^{1/2} = O_p(1)\, o_p\left(T^{-1/2}\right)$, using the strong approximation for the bootstrap sample in Lemma 3 for $r > 3$. 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 (see e.g. Phillips, 1988). ■

Proof of Proposition 2. The proof of equation (23) is similar to the proof of Theorem 3 in Bai (2004); thus, we report only the main passages. We have

$$T\left[\hat{\lambda}_{i,b}^{PC} - H_1^{-1}\hat{\lambda}_i\right] = \left[\frac{1}{T^2}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}\hat{F}_{t,b}^{PC\prime}\right]^{-1} \times \left[\frac{1}{T}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}u_{it,b} + \frac{1}{T}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}\left(\hat{F}_{t,b}^{PC} - H_1'F_{t,b}\right)' H_1^{-1}\hat{\lambda}_i\right]. \quad (47)$$

Consider the denominator. Using Lemma A.4, $T^{-2}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}\hat{F}_{t,b}^{PC\prime} = H_1'\frac{1}{T^2}\sum_{t=1}^{T}F_{t,b}F_{t,b}' H_1 + O_p\left(C_{nT}^{-1}\right)$. In view of (18), we have $T^{-2}\sum_{t=1}^{T}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}\hat{F}_{t,b}^{PC\prime} \stackrel{d^b}{\to} H_1'\left[H'\left(\int W_{\varepsilon}W_{\varepsilon}'\right)H\right]H_1$ in P. Turning to the numerator, Lemma A.4*(iii)* yields $T^{-1}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}\left(\hat{F}_{t,b}^{PC} - H_1'F_{t,b}\right)' H_1^{-1}\hat{\lambda}_i = O_p\left(C_{nT}^{-1}\right)$. Also, $T^{-1}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}u_{it,b} = H_1'T^{-1}\sum_{t=1}^{T}F_{t,b}u_{it,b} + T^{-1}\sum_{t=1}^{T}\left(\hat{F}_{t,b}^{PC} - H_1'F_{t,b}\right) u_{it,b} = I + II$. As regards II, it is bounded by $\left[T^{-1}\sum_{t=1}^{T}\left\|\hat{F}_{t,b}^{PC} - H_1'F_{t,b}\right\|^2\right]^{1/2}\left[T^{-1}\sum_{t=1}^{T}u_{it,b}^2\right]^{1/2}$; using Lemma A.4*(i)*, and by virtue of Lemma 1, $II = O_p\left(C_{nT}^{-1}\right)$. Equation (19) entails $I \stackrel{d^b}{\to} H_1'H'\int W_{\varepsilon}dW_{u,i}$ in P. Equation (23) follows by the CMT.

Turning to (24), from (47)

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

Consider the denominator. By symmetry, $\left\|\left(T^{-2}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}\hat{F}_{t,b}^{PC\prime}\right)^{-1}\right\|_1 = \ell_{\min}^{-1}\left(T^{-2}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}\hat{F}_{t,b}^{PC\prime}\right)$, where $\ell_{\min}(\cdot)$ denotes the smallest eigenvalue. Equation (27) and Theorem 1 ensure that, for sufficiently large n and T, $T^{-2}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}\hat{F}_{t,b}^{PC\prime} = H_1'\left(T^{-2}\sum_{t=1}^{T}\hat{F}_t\hat{F}_t'\right)H_1 + o_p(1)$. The matrix H_1 is invertible; also, $\left(T^{-2}\sum_{t=1}^{T}\hat{F}_t\hat{F}_t'\right)$ is invertible and (conditionally on the sample) non stochastic. Hence, $\left\|\left(T^{-2}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}\hat{F}_{t,b}^{PC\prime}\right)^{-1}\right\|_1^{2+\delta}$ is bounded with probability 1. Thus

$$\left\| T\left[\hat{\lambda}_{i,b}^{PC}-\hat{\lambda}_i\right]\right\|^{2+\delta} \leq M\left\|\left(T^{-1}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}u_{it,b}\right)\right\|^{2+\delta} \tag{48}$$

$$+M\left\|T^{-1}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}\left(\hat{F}_{t,b}^{PC}-H_1'F_{t,b}\right)'H_1^{-1}\hat{\lambda}_i\right\|^{2+\delta} = I+II.$$

Consider I. We have $\left\|T^{-1}\sum_{t=1}^{T}\hat{F}_{t,b}^{PC}u_{it,b}\right\|^{2+\delta} \leq \|H_1'\| \left\|T^{-1}\sum_{t=1}^{T}F_{t,b}u_{it,b}\right\|^{2+\delta} + \left\|T^{-1}\sum_{t=1}^{T}\right.$ $\left.\left(\hat{F}_{t,b}^{PC}-H_1'F_{t,b}\right)\ u_{it,b}\right\|^{2+\delta} = I_a+I_b$. We now show that the expectation of I_a is $O_p(1)$; denote the martingale approximations of $F_{t,b}$ and $u_{it,b}$ as $F_{t,b}^{*}$ and $u_{it,b}^{*}$. The integrability of $F_{t,b}^{*}u_{it,b}^{*}-F_{t,b}u_{it,b}$ follows from Lemma 1. Thus, it suffices to show $T^{-(2+\delta)}E^b\left\|\sum_{t=1}^{T}F_{t,b}^{*}u_{it,b}^{*}\right\|^{2+\delta} <$ ∞. Since $F_{t,b}^{*}u_{it,b}^{*}$ is an MDS, Burkholder's inequality yields $E^b\left\|\sum_{t=1}^{T}F_{t,b}^{*}u_{it,b}^{*}\right\|^{2+\delta} \leq$ $E^b\left|\sum_{t=1}^{T}\left\|F_{t,b}^{*}u_{it,b}^{*}\right\|^2\right|^{\frac{2+\delta}{2}}$. By Holder's inequality, $E^b\left|\sum_{t=1}^{T}\left\|F_{t,b}^{*}u_{it,b}^{*}\right\|^2\right|^{\frac{2+\delta}{2}} \leq E^b\left|\left(\sum_{t=1}^{T}\right.\right.$ $\left.\left.\left\|F_{t,b}^{*}u_{it,b}^{*}\right\|^{2+\delta}\right)^{\frac{2}{2+\delta}}T^{1-\frac{2}{2+\delta}}\right|^{\frac{2+\delta}{2}}$. Finally, by the C_r and Jensen's inequalities, $T^{-(2+\delta)}\ E^b$ $\left\|\sum_{t=1}^{T}F_{t,b}^{*}u_{it,b}^{*}\right\|^{2+\delta} \leq T^{-1-\frac{2+\delta}{2}}\ \sum_{t=1}^{T}\ E^b\ \left\|F_{t,b}^{*}u_{it,b}^{*}\right\|^{2+\delta}$. Using the Cauchy-Schwartz inequality, this is bounded by $T^{-\frac{2+\delta}{2}}\ \sum_{t=1}^{T}\left(E^b\left\|F_{t,b}^{*}\right\|^{4+2\delta}\right)^{1/2}\left(E^b\left|u_{it,b}^{*}\right|^{4+2\delta}\right)^{1/2}$, which is bounded by $T^{-1-\frac{2+\delta}{2}}\ \sum_{t=1}^{T}\left(E^b\left\|F_{t,b}^{*}\right\|^{4+2\delta}\right)^{1/2}$ in view of (15) in Lemma 1. Consider now I_b; it holds that $I_b \leq \left[T^{-1}\sum_{t=1}^{T}\ \left\|\hat{F}_{t,b}^{PC}-H_1'F_{t,b}\right\|^2\right]^{(2+\delta)/2}\ \left[T^{-1}\sum_{t=1}^{T}u_{it,b}^2\right]^{(2+\delta)/2}$. By the Cauchy-Schwartz inequality, the expected value of I_b is bounded by the square root of $E^b\left[T^{-1}\sum_{t=1}^{T}\ \left\|\hat{F}_{t,b}^{PC}-H_1'F_{t,b}\right\|^2\right]^{(2+\delta)}\ E^b\ \left[T^{-1}\ \sum_{t=1}^{T}u_{it,b}^2\right]^{(2+\delta)}$. By convexity, this is bounded by $T^{-1}\sum_{t=1}^{T}E^b\ \left\|\hat{F}_{t,b}^{PC}-H_1'F_{t,b}\right\|^{4+2\delta}\times T^{-1}\sum_{t=1}^{T}E^b\left|u_{it,b}\right|^{4+2\delta}$. Lemmas A.4*(i)* and 1 entail that this is $o_p(1)$. Finally, consider II in (48); this is bounded by $\left\|H_1^{-1}\right\|^{2+\delta}$ $\left\|\hat{\lambda}_i\right\|^{2+\delta}\ T^{-1}\sum_{t=1}^{T}\left\|\hat{F}_{t,b}^{PC}\left(\hat{F}_{t,b}^{PC}-H_1'F_{t,b}\right)\right\|^{2+\delta}$. Both H_1 and $\hat{\lambda}_i$ are non stochastic conditional on the sample; further, H_1 is invertible. Thus, the expectation of II is bounded by $E^b\left\|\hat{F}_{t,b}^{PC}\left(\hat{F}_{t,b}^{PC}-H_1'F_{t,b}\right)\right\|^{2+\delta}$; this is $o_p(1)$ by Lemma A.4*(v)*. Putting all together, (24) follows. ■

Proof of Proposition 3. The proof of (27) 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 $\frac{n}{T^3}\to 0$, $\sqrt{n}\left[\hat{F}_{t,b}^{PC}-H_1'F_{t,b}\right] = \frac{1}{\sqrt{n}T^2}\left(V_{nT}^{b}\right)^{-1}\sum_{s=1}^{T}\hat{F}_{s,b}F_{s,b}'\hat{\Lambda}'u_{t,b}$

$+o_p(1)$. It holds that

$$\frac{1}{\sqrt{n}T^2}\left(V_{nT}^b\right)^{-1}\sum_{s=1}^{T}\hat{F}_{s,b}F_{s,b}'\hat{\Lambda}'u_{t,b}=\left(V_{nT}^b\right)^{-1}\left(\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_{s,b}F_{s,b}'\right)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\lambda}_iu_{it,b}$$

$$=\left(V_{nT}^b\right)^{-1}\left(\frac{1}{T^2}\sum_{s=1}^{T}\hat{F}_{s,b}F_{s,b}'\right)\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}_i\hat{\lambda}_i'\right)\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}_i\hat{\lambda}_i'\right)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\lambda}_iu_{it,b}$$

$$=H_1'\left[H^{-1}\Sigma_\Lambda\left(H'\right)^{-1}\right]^{-1}H^{-1}N\left[0,\Gamma_t\right]+o_p(1),$$

Turning to (28), based on Bai (2004, p. 167), we have

$$\begin{aligned}\left\|\sqrt{n}\left[\hat{F}_{t,b}^{PC}-H_1'F_{t,b}\right]\right\|^{2+\delta} &\leq \left\|\frac{1}{\sqrt{n}T^2}\left(V_{nT}^b\right)^{-1}\sum_{s=1}^{T}\hat{F}_{s,b}F_{s,b}'\hat{\Lambda}'u_{t,b}\right\|^{2+\delta}+\left\|\frac{\sqrt{n}}{T^2}\sum_{s=1}^{T}\hat{F}_{s,b}\gamma_{s-t,b}\right\|^{2+\delta} \quad (49)\\ &\quad+\left\|\frac{\sqrt{n}}{T^2}\sum_{s=1}^{T}\hat{F}_{s,b}\left(\frac{u_{s,b}'u_{t,b}}{n}-\gamma_{s-t,b}\right)\right\|^{2+\delta}+\left\|\frac{\sqrt{n}}{T^2}\sum_{s=1}^{T}\hat{F}_{s,b}F_{t,b}'\hat{\Lambda}'u_{s,b}\right\|^{2+\delta}\\ &= I+II+III+IV,\end{aligned}$$

where $\gamma_{s-t,b}=E\left(n^{-1}u_{s,b}'u_{t,b}\right)$. Consider I; by definition of H_1, $\frac{1}{\sqrt{n}T^2}\left(V_{nT}^b\right)^{-1}\sum_{s=1}^{T}\hat{F}_{s,b}F_{s,b}'\hat{\Lambda}'u_{t,b}$ $=H_1'\left(n^{-1}\sum_{i=1}^{n}\hat{\lambda}_i\hat{\lambda}_i'\right)\left(n^{-1/2}\sum_{i=1}^{n}\hat{\lambda}_iu_{it,b}\right)$. By construction, $n^{-1}\sum_{i=1}^{n}\hat{\lambda}_i\hat{\lambda}_i'$ is diagonal, invertible, and non stochastic conditional on the sample. Thus, $\left\|\left[\frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}_i\hat{\lambda}_i'\right]^{-1}\right\|_1^{2+\delta}$ is bounded and it suffices to study $E^b\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\lambda}_iu_{it,b}\right\|^{2+\delta}$. By applying: Burkholder's, Holder's, C_r, and Jensen's inequalities, $E^b\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\lambda}_iu_{it,b}\right\|^{2+\delta}\leq\frac{1}{n}\sum_{i=1}^{n}E^b\left\|\hat{\lambda}_iu_{it,b}\right\|^{2+\delta}$ $\leq\max_i\left\|\hat{\lambda}_i\right\|^{2+\delta}\max_iE^b\left|u_{it,b}\right|^{2+\delta}$, since $\hat{\lambda}_i$ is non stochastic with respect to the bootstrap sample. Thus, by Lemma 1, $E^b\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\lambda}_iu_{it,b}\right\|^{2+\delta}=O_p(1)$. Turning to II in (49), $II\leq\left\|\frac{\sqrt{n}}{T^2}\sum_{s=1}^{T}H_1'F_{s,b}\gamma_{s-t,b}\right\|^{2+\delta}+\left\|\frac{\sqrt{n}}{T^2}\sum_{s=1}^{T}\left(\hat{F}_{s,b}-H_1'F_{s,b}\right)\gamma_{s-t,b}\right\|^{2+\delta}=II_a+II_b$. Consider II_a (we omit H_1). We have $E^b(II_a)\leq\frac{n^{1+\delta/2}}{T^{3+\delta}}\sum_{s=1}^{T}E^b\left(\left\|F_{s,b}\right\|^{2+\delta}\left|\gamma_{s-t,b}\right|^{2+\delta}\right)$ $\leq\frac{n^{1+\delta/2}}{T^{3+\delta}}\max_sE^b\left\|F_{s,b}\right\|^{2+\delta}\sum_{s=1}^{T}\left|\gamma_{s-t,b}\right|^{2+\delta}$. Since the $\gamma_{s-t,b}$s are summable, and since $\max_sE^b\left\|F_{s,b}\right\|^{2+\delta}=O_p\left(T^{1+\delta/2}\right)$, we have $E^b(II_a)=O_p\left(\frac{n^{1+\delta/2}}{T^{2+\delta/2}}\right)$. Turning to II_b, after some passages we have $E^b(II_b)\leq\frac{n^{1+\delta/2}}{T^{5/2+\delta}}\left[\frac{1}{T}\sum_{s=1}^{T}E^b\left\|\hat{F}_{s,b}-H_1'F_{s,b}\right\|^{4+2\delta}\right]^{1/2}\left[\sum_{s=1}^{T}\left|\gamma_{s-t,b}\right|^{4+2\delta}\right]^{1/2}$, which is $o_p(1)$ by Lemma A.4*(i)* and the summability of the $\gamma_{s-t,b}$s. Consider III in (49); letting $\zeta_{st,b}=\frac{u_{s,b}'u_{t,b}}{n}-\gamma_{s-t,b}$, $III\leq\left\|T^{-2}\sum_{s=1}^{T}H_1'F_{s,b}\left(\sqrt{n}\zeta_{st,b}\right)\right\|^{2+\delta}+\left\|T^{-2}\sum_{s=1}^{T}\left(\hat{F}_{s,b}-H_1'F_{s,b}\right)\left(\sqrt{n}\zeta_{st,b}\right)\right\|^{2+\delta}=III_a+III_b$. As regards III_a, note that (omitting H_1) $III_a\leq T^{-(3+\delta)}\sum_{s=1}^{T}\left\|F_{s,b}\right\|^{2+\delta}\left|\sqrt{n}\zeta_{st,b}\right|^{2+\delta}$. Its expectation is bounded by $T^{-(3+\delta)}$

$\max_s E^b \|F_{s,b}\|^{2+\delta} \sum_{s=1}^{T} E^b \left|\sqrt{n}\zeta_{st,b}\right|^{2+\delta}$. We have $\max_s E^b \|F_{s,b}\|^{2+\delta} = O_p\left(T^{1+\delta/2}\right)$; also, similar passages as above, yield $E^b \left|\sqrt{n}\zeta_{st,b}\right|^{2+\delta} = O_p(1)$. Thus, $E^b(III_a) = o_p(1)$. As regards III_b, by the Cauchy-Schwartz inequality $III_b \leq T^{-(3+\delta)} \sum_{s=1}^{T} \left[E^b \left\|\hat{F}_{s,b} - H_1' F_{s,b}\right\|^{4+2\delta}\right]^{1/2}$ $\left[E^b \left|\sqrt{n}\zeta_{st,b}\right|^{4+2\delta}\right]^{1/2}$. Again, similar passages as above yield $E^b \left|\sqrt{n}\zeta_{st,b}\right|^{4+2\delta} = O_p(1)$; by Lemma A.4*(i)*, $E^b(III_a) = o_p(1)$. Finally, we turn to IV in (49). We have $IV \leq \left\|T^{-2} \sum_{s=1}^{T} H_1' F_{s,b} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{\lambda}_i' u_{is,b} F_{t,b}\right\|^{2+\delta} + \left\|T^{-2} \sum_{s=1}^{T} \left(\hat{F}_{s,b} - H_1' F_{s,b}\right) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{\lambda}_i' u_{is,b} F_{t,b}\right\|^{2+\delta}$ $= IV_a + IV_b$. It holds that $E^b(IV_a) \leq \left[\max_s E^b \|F_{s,b}\|^{2+\delta}\right] E^b \left\|n^{-1/2} T^{-2} \sum_{s=1}^{T} F_{s,b} \sum_{i=1}^{n} \hat{\lambda}_i' u_{is,b}\right\|^{2+\delta}$, with $\max_s E^b \|F_{s,b}\|^{2+\delta} = O_p\left(T^{1+\delta/2}\right)$. Also, $E^b \left\|T^{-2} \sum_{s=1}^{T} F_{s,b} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{\lambda}_i' u_{is,b}\right\|^{2+\delta} \leq$ $T^{-(3+\delta)} \sum_{s=1}^{T} \left[E^b \|F_{s,b}\|^{4+2\delta}\right]^{1/2} \left[E^b \left|\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{\lambda}_i' u_{is,b}\right|^{4+2\delta}\right]^{1/2}$. This is $O_p\left(T^{-1-\delta/2}\right)$ because $E^b \left|\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{\lambda}_i' u_{is,b}\right|^{4+2\delta}$ is bounded as shown above, and $\left[E^b \|F_{s,b}\|^{4+2\delta}\right]^{1/2} = O_p\left(s^{1+\delta/2}\right)$. Thus, $E^b(IV_a) = O_p(1)$. That $E^b(IV_b) = o_p(1)$ can be shown from the same passages as above and Lemma A.4*(i)*. Putting all together, (28) follows. ■

Proof of Proposition 4. Consider the case $\frac{n}{T} \to 0$. We have

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

Using Propositions 2 and 3, $II = O_p\left(T^{-1}\right)$. 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 3 entails that $I \stackrel{d^b}{\to} \lambda_i' \left(H'\right)^{-1} \left(H_1'\right)^{-1} H_1' H' \Sigma_\Lambda \times N\left(0, \Gamma_t\right)$ 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 3 entails that $I = o_p(1)$; also, in view of Propositions 2 and 3, III and IV are both $o_p(1)$. The asymptotics is driven by II; the IP entails $T^{-1/2} F_t = O_p(1)$ and

$T^{-1/2}F_t \xrightarrow{d} W_\varepsilon(s)$. By Proposition 2, $II \xrightarrow{d^b} W_\varepsilon'(s) H H_1 H_1^{-1}\left[H^{-1}\left(\int W_\varepsilon W_\varepsilon'\right)^{-1}\left(\int W_\varepsilon dW_{u,i}\right)\right]$ in P. Hence, equation (32) follows; (33) follows from (24) and (28). ■

Proof of Theorem 2. Consider (35). The term $(nT)^{-1/2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Ts\rfloor} u_{it,b}$ can be decomposed as

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

where "*" denotes the martingale approximation, $\Gamma^*(1) = [B^*(1)]^{-1}$ and $R_{ub,it}$ is the remainder in the BN decomposition. Consider I, and define $\zeta_{nt,b} = n^{-1/2}\, i_n'\Gamma(1)e_{t,b}^{*u}$. By construction, $\zeta_{nt,b}$ is an MDS. We now show that a Lyapunov condition holds; unlike in the proof of (9), where it follows from high-level assumptions, we show this directly. We have $E^b\left|\zeta_{nt,b}\right|^{2+\delta} \le \|\Gamma(1)\|_1^{2+\delta}\; E^b\left|n^{-1/2} i_n' e_{t,b}^{*u}\right|^{2+\delta}$, and in view of Assumption 1*(iii)*(a), $E^b\left|\zeta_{nt,b}\right|^{2+\delta} \le M\; E^b\left|n^{-1/2} i_n' e_{t,b}^{*u}\right|^{2+\delta}$. Similarly to the proof of (15)

$$\begin{aligned}
E^b\left|n^{-1/2} i_n' e_{t,b}^{*u}\right|^{2+\delta} &\le \frac{1}{T}\sum_{t=1}^{T} E\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n} e_{it}^{*u}\right|^{2+\delta} + \frac{1}{T}\sum_{t=1}^{T} E\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(e_{it,q}^{*u} - e_{it}^{*u}\right)\right|^{2+\delta} \\
&\quad + \frac{1}{T}\sum_{t=1}^{T} E\left|\frac{1}{\sqrt{n}} i_n'\left(\hat{e}_{t,q}^{*u} - e_{t,q}^{*u}\right)\right|^{2+\delta} + \frac{1}{T}\sum_{t=1}^{T} E\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{e}_{it,q}^{*u}\right|^{2+\delta} \\
&= I_a + I_b + I_c + I_d.
\end{aligned}$$

In Trapani (2012), it is shown that $I_a = O_p(1)$ and $I_b = o_p(q^{-rs})$. Consider I_c

$$\begin{aligned}
I_c &\le \frac{1}{T}\sum_{t=1}^{T} E\left|\frac{1}{\sqrt{n}} i_n'\sum_{j=0}^{q}\Gamma_{q,j}\left(\hat{u}_{t-j}^* - u_{t-j}^*\right)\right|^{2+\delta} + \frac{1}{T}\sum_{t=1}^{T} E\left|\frac{1}{\sqrt{n}} i_n'\sum_{j=0}^{q}\left(\Gamma_{q,j}^* - \Gamma_{q,j}\right)\left(\hat{u}_{t-j}^* - u_{t-j}^*\right)\right|^{2+\delta} \\
&\quad + \frac{1}{T}\sum_{t=1}^{T} E\left|\frac{1}{\sqrt{n}} i_n'\sum_{j=0}^{q}\left(\Gamma_{q,j}^* - \Gamma_{q,j}\right)u_{t-j}^*\right|^{2+\delta} = I_{c,1} + I_{c,2} + I_{c,3}.
\end{aligned}$$

It holds that $I_{c,1} \le T^{-1}\sum_{t=1}^{T} E\left|\frac{1}{\sqrt{n}} i_n'\left(\hat{u}_{t-j}^* - u_{t-j}^*\right)\right|^{2+\delta}\|\Gamma(1)\|_1^{2+\delta}$, which has the same order as $n^{\frac{2+\delta}{2}}\max_i T^{-1}\sum_{t=1}^{T} E\left|\hat{u}_t^* - u_t^*\right|^{2+\delta}$; by Lemma A.2*(iii)*, $I_{c,1} = n^{\frac{2+\delta}{2}} O_p\left(\delta_{nT}^{-\frac{2+\delta}{2}}\right)$ $= O_p(1)$. Similarly, $I_{c,2}$ is bounded by $n^{\frac{2+\delta}{2}}\max_i T^{-1}\sum_{t=1}^{T} E\left|\hat{u}_t^* - u_t^*\right|^{2+\delta}\|\Gamma^*(1) - \Gamma(1)\|_1^{2+\delta}$, which is $o_p(1)$ since $\|\Gamma^*(1) - \Gamma(1)\|_1^{2+\delta} = o_p(1)$. Finally, $I_{c,3}$ is bounded by $T^{-1}\sum_{t=1}^{T}$ $E\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n} u_{it}^*\right|^{2+\delta}\|\Gamma^*(1) - \Gamma(1)\|_1^{2+\delta}$; $E\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n} u_{it}^*\right|^{2+\delta}$ is finite as shown in (43), and the passages thereafter. Thus, $I_{c,3} = o_p(1)$. Putting all together, I_c is bounded. The proof for I_d follows from similar passages as above, and is therefore omitted. Thus, $E\left|\zeta_{nt,b}\right|^{2+\delta}$ is bounded uniformly in n. Therefore, an IP for MDS holds: $I \xrightarrow{d^b} W_{\zeta,b}(s)$. Prior to calculating the variance of $W_{\zeta,b}(s)$, we show 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; thus, a sufficient condition for II to be negligible is that $\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(\tilde{\zeta}_{i\lfloor Ts\rfloor}\tilde{\zeta}_{j\lfloor Ts\rfloor}\right) = o_p(1)$ as $(n,T)\to\infty$. Since

$$
\begin{aligned}
& \lim_{n,T\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(\tilde{\zeta}_{i\lfloor Ts\rfloor}\tilde{\zeta}_{j\lfloor Ts\rfloor}\right) && (50)\\
\leq\ & \sup_j\left[\sum_{i=1}^{n}\left\{\Gamma^*(1)-\Gamma(1)\right\}_{ij}\right]^2\left[\lim_{n,T\to\infty}\frac{1}{nT}\sum_{t=1}^{\lfloor Ts\rfloor}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(e_{it,b}^{*u}e_{jt,b}^{*u}\right)\right]\\
\leq\ & M\sup_j\left[\sum_{i=1}^{n}\left|\left\{\Gamma^*(1)-\Gamma(1)\right\}_{ij}\right|\right]^2,
\end{aligned}
$$

where the last inequality comes from Assumption 1 - note that this holds uniformly in s. Thus, a sufficient condition for $\lim_{n,T\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left(\tilde{\zeta}_{i\lfloor Ts\rfloor}\tilde{\zeta}_{j\lfloor Ts\rfloor}\right) = o_p(1)$ is that $\sup_j\sum_{i=1}^{n}\left|\left\{\Gamma^*(1)\quad -\Gamma(1)\right\}_{ij}\right| = o_p(1)$, 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, $III = O_p\left(\sqrt{\frac{n}{T}}\right)$, which is negligible (Phillips and Moon, 1999). Finally, we calculate the variance of $W_{\zeta,b}(s)$, showing that it is equal to σ_v^2. The variance is given by $E^b\left[n^{-1/2}i_n'u_{it,b}^*\right]^2 = T^{-1}\sum_{t=1}^{T}E\left[n^{-1/2}i_n'\hat{u}_{it}^*\right]^2$. Writing $\hat{u}_{it}^* = u_{it}^* + \lambda_i'F_t^* - \hat{\lambda}_i'\hat{F}_t^* = v_{it}^* - \left(\hat{\lambda}_i - H^{-1}\lambda_i\right)'\hat{F}_t^* + O_p\left(\sqrt{\frac{n}{T^3}}\right)$, where the superscript "$*$" denotes the martingale approximation and the term $O_p\left(\sqrt{\frac{n}{T^3}}\right)$ comes from Theorem 2 in Bai (2004), we have

$$
\begin{aligned}
& E^b\left[n^{-1/2} i_n' u_{it,b}^*\right]^2 \\
= \; & \frac{1}{nT}\sum_{i=1}^n\sum_{j=1}^n\sum_{t=1}^T E\left(v_{it}^* v_{jt}^*\right) + E\left[\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n\left(\hat{\lambda}_i - H^{-1}\lambda_i\right)'\left(\frac{1}{T}\sum_{t=1}^T \hat{F}_t^*\hat{F}_t^{*\prime}\right)\left(\hat{\lambda}_j - H^{-1}\lambda_j\right)\right] \\
& -2\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n E\left[\left(\hat{\lambda}_i - H^{-1}\lambda_i\right)'\left(\frac{1}{T}\sum_{t=1}^T \hat{F}_t^* v_{jt}^*\right)\right] + o_p(1) = \frac{1}{nT}\sum_{i=1}^n\sum_{j=1}^n\sum_{t=1}^T E\left(v_{it}^* v_{jt}^*\right) + I + II + o_p(1)
\end{aligned}
$$

where the $o_p(1)$ term represents the contribution from the $O_p\left(\sqrt{\frac{n}{T^3}}\right)$ in the expansion of $\hat{u}_{it}^*$. As regards I, it is bounded by $nT\max_i\left[E\left\|\hat{\lambda}_i - H^{-1}\lambda_i\right\|^4\right]^{1/2}\left[E\left\|\frac{1}{T^2}\sum_{t=1}^T F_t^* F_t^{*\prime}\right\|^2\right]^{1/2}$; recalling that $\hat{\lambda}_i - H^{-1}\lambda_i = O_p\left(T^{-1}\right)$ and that $\sum_{t=1}^T F_t^* F_t^{*\prime} = O_p\left(T^2\right)$, this is of order $O\left(\frac{n}{T}\right)$. Turning to II, $\frac{1}{T}\sum_{t=1}^T \hat{F}_t^* v_{jt}^*$ is $O_p(1)$ for each j due to standard arguments in the theory of convergence to stochastic integrals. Given that $II \leq n\max_i \left[E\left\|\hat{\lambda}_i - H^{-1}\lambda_i\right\|^2\right]^{1/2}\left[E\left\|\frac{1}{T}\sum_{t=1}^T \hat{F}_t^* v_{jt}^*\right\|^2\right]^{1/2}$, which is again $O\left(\frac{n}{T}\right)$; this is not the sharpest bound, but it suffices for our purposes. Putting all together, $E^b\left[n^{-1/2} i_n' u_{it,b}^*\right]^2 = \sigma_v^2$, whence (35) follows. The proof of (36) is in Trapani (2012). ■

Proof of Theorem 3. Consider equation (39); 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 regards I, 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 $\left\|\hat{B}_q - B_q\right\|_1$, recall $\hat{B}_q = \left[\sum_{t=q+1}^T \hat{u}_t\hat{u}_{qt}'\right]\left[\sum_{t=q+1}^T \hat{u}_{qt}\hat{u}_{qt}'\right]^{-1}$. 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

$$
\begin{aligned}
\hat{B}_q \;=\; & \left[B_q d_q + \frac{1}{T}\sum_{t=q+1}^T \hat{u}_t\hat{u}_{qt}' - \frac{1}{T}\sum_{t=q+1}^T u_t u_{qt}' + \frac{1}{T}\sum_{t=q+1}^T e_t^{(u)} u_{qt}'\right] \\
& \times\left[d_q^{-1} + d_q^{-1}\left(\hat{d}_q - d_q\right)d_q^{-1} + o_p\left(\left\|\hat{d}_q - d_q\right\|\right)\right].
\end{aligned} \tag{51}
$$

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

$\left\|d_q^{-1}\right\|_1 = O_p(1)$; also, note

$$\frac{1}{T}\left(\sum_{t=q+1}^{T}\hat{u}_{qt}\hat{u}_{qt}' - \sum_{t=q+1}^{T}u_{qt}u_{qt}'\right)$$

$$= \left\{\frac{1}{T}\sum_{t=q+1}^{T}\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]'\right\} \otimes \Lambda\left(H'H\right)^{-1}\Lambda'$$

$$+ \left(\frac{1}{T}\sum_{t=q+1}^{T}\hat{F}_{q,t}\hat{F}_{q,t}'\right) \otimes \left[\hat{\Lambda} - \Lambda\left(H'\right)^{-1}\right]\left[\hat{\Lambda} - \Lambda\left(H'\right)^{-1}\right]'$$

$$+ \left\{\frac{1}{T}\sum_{t=q+1}^{T}\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]\hat{F}_{q,t}'\right\} \otimes \Lambda\left(H'\right)^{-1}\left[\hat{\Lambda} - \Lambda\left(H'\right)^{-1}\right]'$$

$$+ \left\{\frac{1}{T}\sum_{t=q+1}^{T}\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]u_{qt}'\right\} \otimes \Lambda\left(H'\right)^{-1}$$

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

where A and B are the transposes of IV and V respectively. It holds that $\|I\|_1 \leq \left\|\Lambda\left(H'H\right)^{-1}\Lambda'\right\|_1 \left\|T^{-1}\sum_{t=q+1}^{T}\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]'\right\|_1 = O_p(n)\, O_p\left(qC_{nT}^{-2}\right)$, using Lemma B.1 in Bai (2004). Recalling that $\hat{\Lambda} - \Lambda\left(H'\right)^{-1} = O_p\left(T^{-1}\right)$ element-wise, it holds that $\|II\|_1 \leq \left\|\left[\hat{\Lambda} - \Lambda\left(H'\right)^{-1}\right]\left[\hat{\Lambda} - \Lambda\left(H'\right)^{-1}\right]'\right\|_1 \left\|T^{-1}\sum_{t=q+1}^{T}\hat{F}_{q,t}\hat{F}_{q,t}'\right\|_1 = O_p\left(nT^{-2}\right) O_p(qT) = O_p\left(qnT^{-1}\right)$; similarly, $\|III\|_1 \leq \left\|\Lambda\left(H'\right)^{-1}\left[\hat{\Lambda} - \Lambda\left(H'\right)^{-1}\right]'\right\|_1 \left\|T^{-1}\sum_{t=q+1}^{T}\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]\hat{F}_{q,t}'\right\|_1 = O_p\left(nT^{-1}\right) O_p\left(qC_{nT}^{-1}\right)$. Considering IV, $\|IV\|_1 \leq \left\|\Lambda\left(H'\right)^{-1}\right\|_1 \left\|T^{-1}\sum_{t=q+1}^{T}\left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t}\right]u_{q,t}'\right\|_1 = O_p(n)\, O_p\left(qC_{nT}^{-1}\right)$; similar calculations yield that $\|A\|_1$ has the same order. Finally, $\|V\|_1 \leq \left\|\hat{\Lambda} - \Lambda\left(H'\right)^{-1}\right\|_1 \left\|T^{-1}\sum_{t=q+1}^{T}\hat{F}_{q,t}u_{qt}'\right\|_1 = O_p\left(nT^{-1}\right) O_p(q)$. Thus, the terms that dominate are of magnitude $O_p\left(nqC_{nT}^{-1}\right)$. Considering the numerator of (51), 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

$$\frac{1}{T}\left(\sum_{t=q+1}^{T}\hat{u}_t\hat{u}_{qt}' - \sum_{t=q+1}^{T}u_tu_{qt}'\right)$$

$$= \left\{ \frac{1}{T} \sum_{t=q+1}^{T} \left(\hat{F}_t - H' F_t \right) \left[\hat{F}_{q,t} - (H \otimes I_q) F_{q,t} \right]' \right\} \otimes \Lambda \left(H' H \right)^{-1} \Lambda'$$

$$+ \left(\frac{1}{T} \sum_{t=q+1}^{T} \hat{F}_t \hat{F}'_{q,t} \right) \otimes \left[\hat{\Lambda} - \Lambda \left(H' \right)^{-1} \right] \left[\hat{\Lambda} - \Lambda \left(H' \right)^{-1} \right]'$$

$$+ \left[\frac{1}{T} \sum_{t=q+1}^{T} \left(\hat{F}_t - H' F_t \right) \hat{F}'_{q,t} \right] \otimes \Lambda \left(H' \right)^{-1} \left[\hat{\Lambda} - \Lambda \left(H' \right)^{-1} \right]'$$

$$+ \left[\frac{1}{T} \sum_{t=q+1}^{T} \left(\hat{F}_t - H' F_t \right) u'_{qt} \right] \otimes \Lambda \left(H' \right)^{-1} + \left(\frac{1}{T} \sum_{t=q+1}^{T} \hat{F}_t u'_{qt} \right) \otimes \left[\hat{\Lambda} - \Lambda \left(H' \right)^{-1} \right] + C + D$$

$$= I + II + III + IV + V + C + D,$$

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 $\left\|\widehat{B_q(1)} - B(1)\right\|_1 = o\left(q^{-s}\right) + qO_p\left(nqC_{nT}^{-1}\right) + qO_p\left(\sqrt{\frac{\log T}{T}}\right)$. Turning to (40), 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, $\widetilde{B_q(1)} - B(1) = \left(\widetilde{B_q} - B_q\right)(i_q \otimes I_n) + \sum_{j=q+1}^{\infty} B_j$. 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 $\left\|\left(\widetilde{B_q} - B_q\right)(i_q \otimes I_n)\right\|_1 \leq q\left\|\widetilde{B_q} - B_q\right\|_1 \leq qO_p\left(\varphi_{nT}^u\right) + qO_p\left(n^{-\phi}\right)$; this proves (40). ■

Acknowledgements

This paper is based on a working paper circulated under the title "Sieve Bootstrap for Large Nonstationary Panel Factor Models". Special thanks go to Giovanni Urga for reading numerous versions of this paper. I am greatly indebted to the Editor (Cheng Hsiao), an Associate Editor, and three anonymous Referees for very constructive comments. 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.

References

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

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

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

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

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

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

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

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

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

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

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

Choi, I., 2011. Efficient estimation of nonstationary factor models. Research Institute for Market Economy, Sogang University, 2011-03

Choi, I., 2012. Efficient estimation of factor models. Econometric Theory 28, 274-308

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

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

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

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

Goncalves, S., Perron, B., 2011. Bootstrapping factor-augmented regression models. Mimeo.

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

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

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

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

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

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

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

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

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

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

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

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

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

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.

Trapani, L., 2012. On bootstrapping panel factor series - Extended version. Available at SSRN: http://ssrn.com/abstract=2062183

n \ T	(ρ^u, ϑ^u)	20	50	100
20	$(0, 0)$	0.187 0.045	0.139 0.048	0.125 0.047
	$(0.5, 0.5)$	0.354 0.048	0.195 0.053	0.155 0.052
	$(0, -0.5)$	0.140 0.037	0.109 0.040	0.102 0.041
50	$(0, 0)$	0.148 0.054	0.094 0.046	0.083 0.046
	$(0.5, 0.5)$	0.354 0.055	0.160 0.054	0.108 0.053
	$(0, -0.5)$	0.107 0.038	0.077 0.042	0.072 0.043
100	$(0, 0)$	0.129 0.057	0.084 0.048	0.070 0.046
	$(0.5, 0.5)$	0.391 0.068	0.155 0.062	0.096 0.056
	$(0, -0.5)$	0.099 0.038	0.067 0.042	0.061 0.043

Table 1. Empirical rejection frequencies under data generating scheme (42), for the null that estimated factors track the true ones. The nominal significance level is, for all experiments, 5%. Each pair of numbers in the Table - under any combinations of (ρ^u, ϑ^u), (ρ^F, ϑ^F) and (n, T) - represents the empirical rejection frequency of: the asymptotic test (i.e. ERF; these are the entries at the top position of each pair); and of the bootstrap test (i.e. ERF^*; these are the entries at the bottom position of each pair). All the relevant parameters are defined in Section 6.