SAPIENZA
UNIVERSITÀ DI ROMA

Dipartimento di Scienze Statistiche
Sezione di Statistica Economica ed Econometria

Francesca Di Iorio **Stefano Fachin**

Evaluating Restricted Common Factor models for non-stationary data

DSS Empirical Economics and Econometrics Working Papers Series

DSS-E3 WP 2017/2

Dipartimento di Scienze Statistiche
Sezione di Statistica Economica ed Econometria
“Sapienza” Università di Roma
P.le A. Moro 5 – 00185 Roma - Italia

http://www.dss.uniroma1.it

Evaluating restricted Common Factor models for non-stationary data

Francesca Di Iorio
University of Naples Federico II

Stefano Fachin*
"Sapienza" University of Rome

Abstract

Approximate factor models with restrictions on the loadings may be interesting both for structural analysis (simpler structures are easier to interpret) and forecasting (parsimonious models typically deliver superior forecasting performances). However, the issue is largely unexplored. In particular, no currently available test is entirely suitable for the empirically important case of non-stationary data. Building on the intuition that de-factoring the data under a correct set of restrictions will lower the number of factors, we propose a procedure based on the comparison of the number of factors selected for the raw and de-factored data. To control and reduce the risk of rejecting valid constraints we develop a bootstrap procedure, shown analytically to be asymptotically valid and by simulation to have good small sample properties.

JEL: C12, C33, C55

Keywords: Non-stationary factor model, restricted factor models, stationary bootstrap.

* Corresponding author: Stefano Fachin, Dip. di Scienze Statistiche, "Sapienza" University, P.le A. Moro 5, 00185 Roma Italy.

E-mail: stefano.fachin@uniroma1.it. Tel: +-39-6-49910834.

Revised June 2019

1 Introduction

In recent years factor models have become increasingly popular, both in the literature and in practical usage. As is well known, the general idea of factor models is that a small number of unobserved variables, the factors, may be able to explain the dynamics of a large number of series, which can be different variables or the same variable measured over different units (countries, industries, etc.). A general survey of the more popular methods for stationary variables is provided by Stock and Watson (2011). Methods suitable for the empirically important case of non-stationary data, still scarce, are reviewed in Barigozzi, Lippi and Luciani (2016) and Banerjee, Marcellino and Masten (2016). Examples of applications for forecasting and the construction of cyclical indicators are respectively given by, *e.g.*, Giannone, Reichlin and Small (2008) and Altissimo, Cristadoro, Forni, Lippi and Veronese (2010).

Exactly as in traditional econometric modelling, economic theory or ex-post evaluation of an estimated model may suggest restrictions on the coefficients linking factors and variables (loadings). The most likely to occur are exclusion and homogeneity restrictions (respectively, a given unit is not affected by a given factor, and all units in a given group are equally affected by a given factor). Homogeneity constraints may arise in a particularly natural way in factor models applied to, *e.g.*, asset returns (for the loadings of companies active in the same industry or market) and regional datasets (for the loadings of neighbouring regions). Constrained factor models are interesting for two different reasons. First, for structural analysis, as factor loadings may be parameters of interest in their own sake, and simpler structures are easier to interpret. Second, for forecasting, as parsimonious models typically deliver superior performances.

Acknowledging this point, a stream of literature of Bayesian inspiration exploits data-driven shrinkage methods inducing sparsity in the loadings matrix (*e.g.*, Hacioglu Hoke and Tuzcuoglu, 2016, and the references therein). However, the problem of testing hypothesis on the loadings remains largely unexplored. In fact, to the best of our knowledge, the only exceptions are Reis and Watson (2010) and Amengual and Repetto (2014). The former tested an homogeneity assumption by means of separate t-tests from the regressions of each of the series of the dataset on the estimated factor.

This approach presents two major problems. First, it ignores the consequences of cross-section correlation. Second, the family wise error rate is not controlled. To avoid both shortcomings Amengual and Repetto (2014) proposed a joint $F-$type test, deriving the asymptotic distribution under the fully general assumption that the number of restrictions grows with the sample size but the restrictive one of stationary variables and factors. For the more general case of non stationary data and factors the procedure developed by Bai (2004), henceforth simply Bai, does include asymptotically

conditionally Gaussian estimators of the loadings. However, the empirical use of asymptotic tests will encounter two problems. First, the number of constraints will grow with the sample size, so that conflicts among the tests are very likely to appear (Evans and Savin, 1982). Second, the distribution will depend on some asymptotic covariance matrices, troublesome to estimate. In fact, Bai does not provide an empirical illustration, and we are not aware of any empirical applications of these results.

To circumvent these difficulties we propose a much simpler alternative, based upon the comparison of the number of factors estimated for the unconstrained and constrained models: a difference between the two values clearly provides evidence against the data-compatibility of the constraints. Since such evidence would hinge upon the model selection procedure, consistent but possibly biased in small samples, we include in the testing procedure a bootstrap evalutation of the probability of rejecting correct restrictions ("false positive"). The algorithm is shown by simulation to have overall good small sample properties, and analytically to be asymptotically valid for the special case of Bai's estimators and model selection criteria. The extension to other estimators and criteria is straightforward.

As an illustration we consider an example of a study in which allowing for non-stationarity is essential, namely the factor model estimated by Ciccarelli and Fachin (2017) for a set of 19th century value added series for the Italian economy disaggregated by industries and regions.

We now (section 2) proceed to outline the proposed procedure, then report the empirical illustration (section 3) and the results of some simulations (section 4). Section 5 concludes, while the proof of the asymptotic validity of the bootstrap procedure is reported in the Appendix.

2 Evaluating restrictions on factor loadings through model selection

Consider a panel of N non-stationary time series of length T and assume that long-run growth of this panel is driven by a small number, say r, of latent non-stationary factors. Let $\mathbf{F}$ be the $T \times r$ matrix collecting these factors, $\mathbf{u}$ a $T \times r$ matrix of independent zero mean stationary noises, and assume for the sake of presentation only that the factors follow independent random walks with no drift, so that $\mathbf{\Delta F} = \mathbf{u}$. Denoting further by $\mathbf{Y}$ the $T \times N$ data matrix, by $\mathbf{\Lambda}$ a $r \times N$ matrix of deterministic factor loadings, by ε the $T \times N$ matrix of errors which can be weakly dependent over both time and units, the approximate static factor model of $\mathbf{Y}$ is

$$\mathbf{Y} = \mathbf{F\Lambda} + \varepsilon \tag{1a}$$

$$= \textstyle\sum_{j=1}^{r} \mathbf{F}_j \mathbf{\Lambda}_j + \varepsilon \tag{1b}$$

where the $j-$th factor $\mathbf{F}_j$ is a $T \times 1$ vector (the $j-th$ column of the matrix $\mathbf{F}$) and its loadings $\mathbf{\Lambda}_j$ are collected in a $1 \times N$ vector (the $j-th$ row of the matrix $\mathbf{\Lambda}$). This model, which is a special case of Bai's under deterministic loadings, can be easily generalised to allow for dynamic or cointegrated non stationary common factors and stationary factors. However, in the interest of clarity we will first develop the analysis for the simplest case of a static model including non stationary, non cointegrated factors only, and later briefly discuss these more general cases.

Bai showed that the space spanned by the factors and their number can in all cases be consistently estimated using Principal Components (PC) and derived the asymptotic properties of the estimators under a set of assumptions which we adapt to our case as follows:

Assumption 1: $\{u_t\}$ and $\{\epsilon_t\}$ are mutually independent stochastic variables.

Assumption 2: The following assumptions in Bai hold (for details, Bai, p. 140-141): Assumption A (stochastic trends of data), C (time and cross-section dependence and heteroskedasticity), E (weak dependence of idiosyncratic errors), B (heterogenous factor loadings) for the deterministic case.

Our goal is to verify if a set of restrictions on model (1a) is sustained by the data. More precisely, we suppose that the loadings $\mathbf{\Lambda}_j$ of each factor j, $j = 1, \ldots, r$, can be obtained multiplying a $1 \times m_j$ row vector of free parameters, θ_j, by a $m_j \times N$ matrix of constraints $\mathbf{H}_j$:

$$\mathbf{\Lambda}_j = \theta_j \mathbf{H}_j$$

In this paper we shall confine our attention with no loss of generality to the case when the constraints are applied only to the first factor (*i.e.*, that associated to the highest eigenvalue), and the remaining $(r-1)$ are left unconstrained. Then, under the null hypotheses $H_0 : \mathbf{\Lambda}_1^0 = \boldsymbol{\theta_1}\mathbf{H_1}$ we can write the constrained factor model as:

$$\mathbf{Y} = \mathbf{F}_1 \theta_1 \mathbf{H}_1 + \textstyle\sum_{j=2}^{r} \mathbf{F}_j \mathbf{\Lambda}_j + \varepsilon^0 \tag{2}$$

$$= \mathbf{F}_1^0 \mathbf{\Lambda}_1^0 + \mathbf{F_2} \mathbf{\Lambda_2} + \varepsilon^0 \tag{3}$$

where $\mathbf{F}_1^0$ is the $T \times 1$ vector of the first factor assuming the constrained loadings structure, $\mathbf{F}_2$ the $T \times (r-1)$ matrix of the other factors, $\mathbf{\Lambda}_1^0 = \theta_1 \mathbf{H}_1$ the $1 \times N$ vector of loadings of the first factor obeying the desired constraints, $\mathbf{\Lambda_2}$ the $(r-1) \times N$ matrix of loadings for the remaining factors, and ε^0 a matrix of stationary errors.

Obviously, the set of constraints $\mathbf{\Lambda}_1^0 = \theta_1 \mathbf{H}_1$ applies also to the stationary model obtained taking the first differences of both sides of (3):

$$\Delta \mathbf{Y} = \Delta \mathbf{F}_1^0 \mathbf{\Lambda}_1^0 + \Delta \mathbf{F_2} \mathbf{\Lambda_2} + \Delta \varepsilon^0 \tag{4}$$

Applying to model (4) the $F-$test developed by Amengual and Repetto (2014) for the stationary factor model may appear as a simple solution to the problem of testing $H_0 : \Lambda_1^0 = \theta_1 H_1$ for the non-stationary model (1a). Unfortunately, this is not the case, as Amengual and Repetto's Assumption B1 on the moments of the model errors, which is crucial for the derivation of the asymptotic distribution of their test, is not satisfied by the errors $\mathbf{\Delta}\varepsilon^0$. Simulations, not reported here for reasons of space but available on request, showed that as a consequence the distribution of test is strongly skewed, with extremely high Type I errors (over 70%). In order to apply an $F-$test we would thus need to derive its asymptotic behaviour for the non-stationary factor model, a complex task.

Given these difficulties, we might consider taking a totally different route. Exclusively to fix ideas[1] suppose that the approximate factor model (1a) has been estimated for the data $\mathbf{Y}$ following Bai's procedure, and let $\widehat{k}$ be the number of factors selected using one of the three available consistent information criteria (IPC_1, IPC_2, IPC_3, defined in Bai's eq. (12), p. 145). Provided a maximum number of factors $kmax > r$ was allowed for, from Bai's Theorem 1 under Assumptions 1-2 it holds that $lim_{N,T\to\infty} \Pr(\widehat{k} = r) = 1$. Accordingly, we write model (1b) for unit i at time t as

$$y_{ti} = \lambda_{1i} F_{t1} + \textstyle\sum_{j=2}^{\widehat{k}} \lambda_{ji} F_{tj} + \varepsilon_{ti} \tag{5}$$

Assume, without much loss of generality, that we are interested in testing homogeneity of the loadings of $\mathbf{F}_1$ within M groups of units:

$$H_0 : \lambda_{1i} = \lambda_{1G_j},\ \forall i \in G_j,\ j = 1, \ldots, M \tag{6}$$

Then the restricted factor model for unit i at time t is

$$y_{ti} = \lambda_{1G_j} F_{t1}^0 + \textstyle\sum_{j=2}^{\widehat{k}} \lambda_{ji} F_{tj} + \varepsilon_{ti}^0, i \in G_j,\ j = 1, \ldots, M$$

where $\mathbf{F}_1^0$ is the first factor under the constrained loadings structure, and the $1 \times N$ vector of loadings of the first factor obeying the null of homogeneity is $\mathbf{\Lambda}_1^0 = [\lambda_{1G_1} \ldots \lambda_{1G_M}]$.

For our purpose it is convenient to introduce the partially de-factored data $\mathbf{Z}$, obtained subtracting the restricted common component $\mathbf{F}_1^0 \mathbf{\Lambda}_1^0$ (whose estimation will be discussed below) from $\mathbf{Y}$. Under the assumption of static, orthogonal factors $\mathbf{Z}$ has the following structure:

$$\mathbf{Z} = \mathbf{Y} - \mathbf{F}_1^0 \mathbf{\Lambda}_1^0 \tag{7}$$

Substituting for $\mathbf{Y}$ from (1b):

$$\mathbf{Z} = \mathbf{F}_1 \mathbf{\Lambda}_1 + \mathbf{F_2 \Lambda_2} + \varepsilon - \mathbf{F}_1^0 \mathbf{\Lambda}_1^0. \tag{8}$$

[1] The following arguments can be easily adapted to other factor estimators and criteria for the choice of the number of factors, such as Onatski's (2010) ED algorithm.

When H_0 holds, by Bai's Theorem 4 the estimate of the restricted common component $\mathbf{F}_1^0\mathbf{\Lambda}_1^0$ will converge to the true common component $\mathbf{F}_1\mathbf{\Lambda}_1$. Then, for N and T large enough $\mathbf{Z}$ will satisfy

$$\mathbf{Z}= \mathbf{F_2}\mathbf{\Lambda_2} + \varepsilon$$

and thus depend only upon the $k_Z =(\widehat{k}-1)$ unrestricted factors collected in $\mathbf{F_2}$.

Hence, when H_0 holds performing Bai's analysis on $\mathbf{Z}$ allowing for a maximum of $\widehat{k}$ factors with probability 1 we will asymptotically select $\widehat{k}_Z = (\widehat{k}-1)$ factors.

On the other hand, if H_0 is false $\mathbf{F}_1^0\mathbf{\Lambda}_1^0$ will not converge to $\mathbf{F}_1\mathbf{\Lambda}_1$ and the two terms will not cancel out. In this case the number of factors of$\mathbf{Z}$ will depend upon the relationship between $\mathbf{F}_1^0$ and the latent factors of $\mathbf{Y}$. If $\mathbf{F}_1^0$ is orthogonal to $\mathbf{F}_1$ and $\mathbf{F_2}$ by defactoring under a wrong null hypothesis we are actually adding a new (spurious) factor to the data, so that asymptotically we will select $\widehat{k}_Z = (\widehat{k}+1)$ factors with probability 1. In the more likely case that $\mathbf{F}_1^0$ is not orthogonal to the original factors, $\mathbf{Z}$ will instead simply depend upon the latter, so that asymptotically we will choose $\widehat{k}_Z = \widehat{k}$ with probability 1. In either of these cases, if H_0 is false asymptotically we will choose with probability 1 some $\widehat{k}_Z \geq \widehat{k}$. . To summarise, we can conclude that the difference between the number of factors estimated for the unconstrained and constrained models, $\left(\widehat{k}-\widehat{k}_Z\right)$, may be used to assess the data compatibility of the null hypothesis: $\widehat{k}_Z = (\widehat{k}-1)$ supports H_0, while $\left(\widehat{k}-\widehat{k}_Z\right) \leq 0$ does not.

Thus, an hypothesis on the loadings of the factor model can be alternatively formulated as one on the number of factors of the restricted factor model, and a test on the number of factors can be used to test restrictions on the loadings. However, only few tests for the number of factors have been developed so far, and to the best of our knowledge none of them is adequate for our needs (non-stationary data, weakly dependent errors). Connor and Korajczyk (1993) consider stationary data and a very specific two-step estimation method, with factors estimated by principal components and loadings by OLS. The version of the test by Onatski (2009) for the approximate factor model in the time domain requires stationary factors and non-autocorrelated errors. Finally, Kapetanios (2010) relies on an assumption on the eigenvalues that has been argued to require errors not cross-sectionally nor temporally correlated (see, e.g., Onatski, 2010). We thus need to devise a different way to proceed.

A natural option is to define a "Naive test" based on a direct application of the considerations made above**:**

Naive test

If $\widehat{k}_Z = (\widehat{k}-1)$: do not reject H_0.

Else, if $\widehat{k}_Z \geq \widehat{k}$: reject H_0.

The properties of this Naive test depend on the probability distribution of the estimator of the number of factors of the restricted model, $\widehat{k}_Z$. Let us consider the probability of rejecting a true null hypotheses, the analogue of the Type I error of classical tests. Define this event as a "false positive", and its probability as p_f. Since $\widehat{k}_Z$ is chosen on the basis of a consistent criterion, under the null hypothesis we have

$$\lim_{N,T\to\infty} \Pr(\widehat{k}_Z = (\widehat{k}-1)|H_0) = 1$$

so that

$$\lim_{N,T\to\infty} p_f = \Pr(\widehat{k}_Z \geq \widehat{k}|H_0) = 0$$

Hence, in large samples true null hypotheses will never be rejected. However, this is not guaranteed to happen in finite samples, when the number of factors may be overestimated and thus $p_f > 0$. The Naive test as outlined above does not allow any control of this probability, and thus does not satisfy an essential requirement of any testing procedure. We need to extend it to include an estimate of p_f. This is not a simple task: because of its degenerate shape, the asymptotic probability distribution of $\widehat{k}_Z$ offers no guidance in small samples, and we need to find a different solution. Our proposal is to construct a bootstrap test, modifying the Naive test as follows:

Boostrap test

If $\widehat{k}_Z = (\widehat{k}-1)$: do not reject H_0.

Else, if $\widehat{k}_Z \geq \widehat{k}$: compute the bootstrap estimate p_f^* of p_f. Reject H_0 if, and only if, $p_f^* \leq \overline{p_f}$, where $\overline{p_f}$ is the maximum acceptable value of the probability of a "false positive".

To recapitulate, the procedure we propose involves three successive steps:

(A) Estimation of the constrained part of model (3), $F_1^0\Lambda_1^0$. This task can be carried out through an iterative algorithm based on that proposed by Amengual and Repetto (2014). Broadly speaking, the idea is to alternate steps in which the factor is taken as given and the loadings estimated under the desired constraints, and steps in which the factor is estimated taking as given the loadings as estimated in the previous step.

(B) Estimation of the unconstrained part of model (3), $\mathbf{F_2\Lambda_2}$. This can be done applying Bai's procedure to the partially de-factored data $\mathbf{Z}$ obtained subtracting the constrained common components, as estimated in (A), from the data.

(C) Assessment of the null hypothesis on the basis of the comparison between the number of factors estimated for the unconstrained model (1a) and for the partially de-factored data obtained in Step B, and of the bootstrap estimate of the probability of a "false positive".

The details are discussed below.

2.1 A three-stages hypothesis evaluation procedure

Step A *Iterative estimation of the constrained common component* $\mathbf{F}_1^0\mathbf{\Lambda}_1^0$.

Algorithm 1 *Constrained estimation*

1. Estimate using Bai's procedure the unconstrained factor model $\mathbf{Y} = \mathbf{F}\mathbf{\Lambda} + \varepsilon$. Label the unrestricted estimates of the first factor and its loadings as $\mathbf{F}_1^{(1)}, \mathbf{\Lambda}_1^{(1)}$.

2. Given $\mathbf{F}_1^{(1)}$, obtain estimates of the loadings under the homogeneity constraint. In this set-up the easiest way is by M separate pooled time series regressions of Y on the first factor, one for each group: $Y_{ti} = \lambda_{1G_j} F_{1t}^{(1)}$, $t = 1, \ldots, T$, $i \in G_j$, $j = 1, \ldots, M$. In this way all units of a groups are forced to have the same loading. Collect this new set of estimates in the $1 \times N$ vector $\mathbf{\Lambda}_1^{(2)} = \left[\lambda_{11}^{(2)} \ldots \lambda_{1N}^{(2)}\right]$.

3. Given $\mathbf{\Lambda}_1^{(2)}$, obtain a new estimate of the first factor by running the T cross-section regressions $Y_{ti} = \lambda_{1i}^{(2)} F_{t1}$, $i = 1, \ldots, N$, with Y_{ti}, $\lambda_{1i}^{(2)}$ as observables and F_{t1} as unknown parameter. Collect this new set of estimates in the $T \times 1$ vector $\mathbf{F}_1^{(2)} = \left[F_{11}^{(2)} \ldots F_{T1}^{(2)}\right]'$.

4. Iterate steps 2 and 3 until the maximum difference over all units and periods between the estimates of the common components in two successive iterations, $\max_{t,i}(\mathbf{F}_1^{(n)}\mathbf{\Lambda}_1^{(n)} - \mathbf{F}_1^{(n-1)}\mathbf{\Lambda}_1^{(n-1)})$, meets the chosen convergence criterion. Define the restricted estimates of factor and loadings as $\widehat{\mathbf{F}}_1^0 = \mathbf{F}_1^{(n)}$, $\widehat{\mathbf{\Lambda}}_1^0 = \mathbf{\Lambda}_1^{(n)}$.

This iterative algorithm can be seen as an implementation of the EM principle based on the calibration principle[2]. Given the constraints on the loadings, defining in Step 4 of Algorithm 1 the convergence criterion in terms of the common components is fully equivalent to defining it in terms of

[2] A very general and efficient example of this principle can be found in Calzolari (2017), where an indirect inference procedure is used to simulate data guaranteed to yield parameters of statistical models respecting a given set of constraints.

loadings and factors (in other terms, convergence in terms of common components implies convergence in terms of both loadings and factors). In a set of simulation trials convergence was always reached very rapidly, less than 50 iterations (details available on request).

As anticipated above here we treated only homogeneity constraints involving all units, but other type of constraints are easily handled in this framework. Homogeneity constraints involving a subset of the units may be imposed running step 2 only for those cases, while leaving the other loadings at the values estimated unrestrictedly in step 1. Exclusion restrictions ($\lambda_{ij} = 0$ for some factor j and unit i) are trivially imposed setting these loadings at zero in step 2, and the other loadings at their restricted or unrestricted estimates as required. For constraints on the second factor, proceed as above on the data partially de-factored from the rst factor. Finally, for constraints on other factors, proceed sequentially in the same manner.

Step B *Estimation of the uncostrained common components* $\mathbf{F}_2\mathbf{\Lambda}_2$

Estimate the unconstrained common components applying Bai's procedure to the partially de-factored data $\mathbf{Z} = \mathbf{Y} - \widehat{\mathbf{F}}_1$ Note that since consistent selection requires $kmax > r$ in this step we must choose $kmax \geq (\widehat{k} + 2)$, as when H_0 is false $\mathbf{Z}$ may depend upon $(r + 1)$ factors.

Step C *Evaluation of the null hypothesis* $H_0 : \mathbf{\Lambda}_1 = \mathbf{\Lambda}_1^0$

1. Compare the number of factors estimated for the data $(\widehat{k})$ and for the data partially de-factored under the constraint of interest $(\widehat{k}_Z)$. If $\widehat{k}_Z = (\widehat{k} - 1)$ do not reject $H_0 : \mathbf{\Lambda}_1 = \mathbf{\Lambda}_1^0$; the evaluation procedure is completed. Else, if $\widehat{k}_Z \geq \widehat{k}$, proceed to steps 2-3.

2. compute the bootstrap estimate p_f^* of the probability of a "false positive";

3. for a given maximum acceptable value of the probability of a "false positive" $\overline{p_f}$, if $p_f^* \leq \overline{p_f}$: reject H_0; if $p_f^* > \overline{p_f}$: do not reject H_0.

The bootstrap estimate p_f^* of the probability of a "false positive" can be computed as the frequency of rejections of the Naive test on a large number of pseudo-datasets $\mathbf{Y}^*$ with the same time series properties of $\mathbf{Y}$ and obeying the null hypothesis H_0. Such pseudo-data can be constructed adding three elements: (i) the common components of the first factor under the desired constraints, $\widehat{\mathbf{F}}_1^0\widehat{\mathbf{\Lambda}}_1^0$; (ii) the pseudo-common components of the unconstrained factors, ($\mathbf{F}$; (iii) weakly dependent pseudo-residuals, ε^*:

$$\mathbf{Y}^* = \widehat{\mathbf{F}}_1^0\widehat{\mathbf{\Lambda}} \tag{9}$$

Let us now see in detail which resampling schemes should be applied to construct $(\mathbf{F}_2\mathbf{\Lambda}_2)^*$ and ε^*. The latter may be easily obtained resampling the residuals $\widehat{\varepsilon}$ of the unconstrained model. Since this model fully captures non-stationarity and strong cross-section dependence of the data, its residuals are only weakly dependent over time and cross-sectionally. Their dependence structure over time may be reproduced applying a resampling scheme suitable for weakly dependent series, such as the Stationary Bootstrap (SB, Politis and Romano, 1994), while that over units simply by resampling blocks of entire rows. In this way observations are swapped over time but column-wise (cross-sectionally) fixed, to obtain pseudo-data reproducing weak dependence in both dimensions. Constructing the pseudo-common components of the unconstrained factors, ($\mathbf{F}$, from the estimated common components $\widehat{\mathbf{F}}_\mathbf{2}\widehat{\mathbf{\Lambda}}_2$ known to be $I(1)$, is a more delicate problem. Parker, Paparoditis and Politis (2006) showed that asymptotically valid unit root tests can be carried out resampling by the SB the first differences of an I(1) series, and then cumulating the pseudo-differences to obtain non-stationary pseudo-levels. Di Iorio and Fachin (2014) extended this result to cointegration and panel cointegration tests. Applying the same principle we can devise the bootstrap algorithm described below. Note that since the SB can be applied to any weakly dependent series, the factors included in $\mathbf{F}_\mathbf{2}$ do not need to be simple random walks, as it was assumed for illustratory purposes in the Introduction. We do need however the following technical assumptions to hold for the factor model (Parker *et al.*, 2006):

Assumption 3
For $j = 1, \ldots, r$ and $i = 1, \ldots, N$:

(i) $E\left|u_{j1}\right|^{6+\delta} < \infty$;

(ii) $E\left|\varepsilon_{i1}\right|^{6+\delta} < \infty$;

(iii) $E\left|F_{j1}\right|^{6+\delta} < \infty$;

(iv) $\sum_k k^2[\alpha_{u_j}(k)]^{\frac{\delta}{6+\delta}} < \infty$;

(v) $\sum_k k^2[\alpha_{\varepsilon_i^y}(k)]^{\frac{\delta}{6+\delta}} < \infty$;

(vi) f_{u_j} satisfies $f_{u_j}(0) > 0$.

(vii) f_{ε_i} satisfies $f_{\varepsilon_i}(0) > 0$.

where $\alpha(k)$ denotes the strong mixing coefficient, δ is an arbitrary strictly positive constant and f_z the spectral density of z. We can now define the following resampling algorithm.

Algorithm 2 *Bootstrap estimation of the probability of a "false positive"*

1. resample applying the Stationary Bootstrap the first differences of the estimates of uncostrained common components from the unconstrained factor model (1a), $\Delta(\widehat{\mathbf{F}}_{\mathbf{2}}\widehat{\mathbf{\Lambda}}_{\mathbf{2}})$, obtaining the pseudo-differences $(\Delta\mathbf{F}_{\mathbf{2}}\mathbf{\Lambda}_{\mathbf{2}})^*$; cumulate them using as starting values the level estimates for $t = 1$, $\sum_{j=2}^{k}\widehat{\lambda}_{ji}\widehat{F}_{1j}$, $i = 1, \ldots, N$, and obtain $(\mathbf{F}_{\mathbf{2}}\mathbf{\Lambda}_{\mathbf{2}})^*$.

2. resample applying the Stationary Bootstrap the residuals $\widehat{\varepsilon}$ of the unconstrained factor model (1a) obtaining the pseudo-residuals ε^*.

3. append $[(\mathbf{F}_{\mathbf{2}}\mathbf{\Lambda}_{\mathbf{2}})^* + \varepsilon^*]$ to $\widehat{\mathbf{F}}_1^0\widehat{\mathbf{\Lambda}}_1^0$, obtaining the pseudo-data $\mathbf{Y}^* = \widehat{\mathbf{F}}_1^0\widehat{\mathbf{\Lambda}}_1^0 + (\mathbf{F}_{\mathbf{2}}\mathbf{\Lambda}_{\mathbf{2}})^* + \varepsilon^*$ which have $\widehat{k}$ common factors and obey $H_0 : \mathbf{\Lambda}_1 = \mathbf{\Lambda}_1^0$.

4. estimate the unrestricted factor model on the pseudodata $\mathbf{Y}^*$ allowing for a maximum of $(\widehat{k} + 1)$ factors. Let $\widehat{k}^*$ be the selected number of factors.

5. using the iterative algorithm described in Step A obtain an estimate of the factor $\widehat{\mathbf{F}}_1^{0*}$ and the associated loadings $\widehat{\mathbf{\Lambda}}_1^{0*}$ respecting the homogeneity constraint.

6. construct the partially de-factored pseudo-data $\widehat{\mathbf{Z}}^* = \mathbf{Y}^* - \widehat{\mathbf{F}}_1^{0*}\widehat{\mathbf{\Lambda}}_1^{0*}$.

7. estimate the unrestricted factor model on $\widehat{\mathbf{Z}}^*$ allowing for a maximum of $(\widehat{k}^* + 2)$ factors**.** Let $\widehat{k}_Z^*$ be the selected number of factors.

8. if $\widehat{k}_Z^* = (\widehat{k}^* - 1)$ do not reject $H_0 : \mathbf{\Lambda}_1 = \mathbf{\Lambda}_1^0$; else, if $\widehat{k}_Z^* \geq \widehat{k}^*$, reject it.

9. repeat 1-8 a large number of times and compute the proportion p_f^* of cases in which H_0 is rejected. This is the bootstrap estimate of the probability under H_0 of a "false positive".

The asymptotic validity of p_f^* as an estimator of p_f is proved by Lemma 1 below, in which we show that pseudodata $\mathbf{Y}^*$ respect all assumptions required by Bai's theorem 1 as well as the null hypothesis. Thus, the bootstrap probability distribution $Pr_{N,T}^*(\widehat{k}_Z^*|H_0)$ will have the same degenerate asymptotic behaviour of the empirical probability distribution $Pr_{N,T}(\widehat{k}$.

Lemma 1 *Let r^* be the number of factors of the pseudodata constructed according to Algorithm 2 and $\widehat{k}^*$ the number of factors selected using any of Bai's information criteria (IPC_1, IPC_2, or IPC_3). Then* $\lim_{N,T\to\infty} \Pr_{N,T}^*(\widehat{k}^* = r^*) = 1$.

Proof. See Appendix. ■

Lemma 1 ensures that asymptotically $p_f^* = p_f = 0$. We shall assess the accuracy of p_f^* as an estimator of p_f in small samples through a Monte Carlo experiment described in section 4.

2.2 Generalisations: dynamic, cointegrated and stationary factors

Assuming the restrictions are placed upon the loadings of a non stationary factor, the extension of the procedure describe above to the more general cases of models including dynamically loaded or cointegrated non stationary factors and stationary ones is much simpler than it might appear at first sight. Let us see the various cases in turn.

(i) *models with static non stationary and stationary factors.* The number of the former, of the latter and all loadings and factors can be consistently estimated as described in Bai's section 5. We then essentially proceed as above, with the only difference that the partially de-factored data are modelled using that generalised procedure. The bootstrap pseudodata will now be the sum of four terms, instead of three as in (**??**), as the resampled stationary common components will appear as well. These are easily obtained applying the Stationary Bootstrap to the estimated stationary common components.

(ii) *models with dynamic non stationary factors.* These models can be reparametrised as static in the non stationary factors and dynamic in their first differences, which are in fact stationary common factors (Bai, p. 152). For a lag length p, denoting by $\lambda_{ji}^{(l)}$ the loading linking the i-th unit to the j-th factor with lag l:

$$\begin{aligned} y_{ti} &= \textstyle\sum_{j=1}^{r}\sum_{l=0}^{p}\lambda_{ji}^{(l)}F_{t-l,j} \\ &= \textstyle\sum_{j=1}^{r}\left[\left(\sum_{l=0}^{p}\lambda_{ji}^{(l)}\right)F_{tj} - \left(\sum_{l=1}^{p}\lambda_{ji}^{(l)}\right)\Delta F_{tj} + \ldots - \lambda_{ji}^{(p)}\Delta F_{t-p+1,j}\right] + \varepsilon_{ti} \end{aligned}$$

In this reparametrisation the loadings of the static non stationary factors are the sum of the loadings at all lags in the dynamic model. Since in dynamic models structural hypothesis, such as homogeneity or exclusion restrictions, are always formulated as constraints on the sum of the coefficients at all lags we are exactly as in case (i).

(iii) *models with cointegrated non stationary factors.* Precisely the same applies to this case, as these models can be reparametrised to include non-stationary, non-cointegrating factors and the stationary linear combinations of the cointegrating ones (Bai, p. 153).

3 An empirical illustration

As an empirical illustration we applied our procedure to a model estimated by Ciccarelli and Fachin (2017) on Value Added (VA) data for 10 manufacturing in- dustries in the 16 Italian regions (but one industry was totally absent in one region, so that the cross-section di- mension was $N = 159$)

over 1861-1913 (thus $T = 53$). Routinary tests showed all the VA series (normalised by regional population, to avoid scale problems) to be clearly $I(1)$, with non-stationarity fully explained by a rather simple common factor structure: using Bai's procedure the selected model has only two non-stationary factors, a trend and a Kondratie cycle with a period of about 25 years. The spatial distribution of the loadings of the trend is particularly interesting: the leading role of the North-Western regions stands out clearly, and so does the well-known (see, e.g, Barro and Sala-i-Martin, 1991, p. 150-151) tendency of the South to lag behind. Setting the average of all the loadings in the NW regions at 100, the averages of the loadings in the NE and in the Centre are both about 70, and that of the loadings of the Southern industries about 60. These large regional differentials suggest testing the hypothesis that the loadings are constant for all industries and regions within each of these four macroareas. This would imply that (i) regional differences mattered more than industrial ones, and, (ii), the North-South divide has deep roots in Italian economic history. Formally we can state the null hypothesis as:

$$H_0 : \lambda_{1i} = \lambda_{1G_j}, \ \forall i \in G_j, \ j = NW, NE, Centre, South \tag{10}$$

As argued in the Introduction, in practice no classical testing procedure is available to evaluate if this restriction is supported by the data. We thus apply the procedure described above. Using ach of Bai's three selection criteria (IPC_1, IPC_2, IPC_3) we obtain slightly different results, summarised in Table 1. According to IPC_1 the number of factors for the data defactored under H_0, $\widehat{k}_Z$, turns out to be equal to that selected for the raw data, so that $\widehat{k} - \widehat{k}_Z = 2 - 2 = 0$ and the Naive test rejects H_0. However, using 5000 redrawings we obtain a bootstrap estimate p_f^* of the probability that this is a false positive equal to 85.1%. We thus definitely overturn the conclusion of the Naive test and do *not* reject H_0. This conclusion is confirmed by model selection on the basis of the second criterion, IPC_2: this also leads to $\widehat{k} - \widehat{k}_Z = 0$, thus rejection on the basis of the Naive test. The bootstrap estimate p_f^* is however also very large, 78.1%, and we similarly conclude for no rejection. Finally, if we rely on the third criterion, IPC_3, we select $\widehat{k}_Z = 1$. Thus $\widehat{k} - \widehat{k}_Z = 2 - 1 = 1$, and, contrary to the first two cases, according to the Naive test H_0 is not rejected and the evaluation procedure ends here. Note that according to the simulations reported by Bai[3] this criterion has some tendency to underestimate the number of factors, so that the results for the criteria are entirely consistent with a priori expectations.

To summarise, the homogeneity hypothesis for the loadings of all industries within the four macroareas appears largely compatible with the data. The next interesting step would be to test the hypothesis that homogeneity holds for the loadings of some industries only, a partial hypothesis which

[3]Confirmed by simulations of ours which do not report here for reasons of space.

may be tested as discussed in the remarks following Algorithm 1. This further investigation is beyond the scope of this illustration and will not be considered it here.

Table 1
Evaluation of the homogeneity hypothesis for
the factor model by Ciccarelli and Fachin (2017)

Step	*Criterion:* IPC_1
C-1	$\widehat{k}_Z = 2$: compute p_f^*
C-2	$p_f^* = 85.1$
C-3	H_0 not rejected for $\overline{p_f} = 10\%$
	Criterion: IPC_2
C-1	$\widehat{k}_Z = 2$: compute p_f^*
C-2	$p_f^* = 78.1$
C-3	H_0 not rejected for $\overline{p_f} = 10\%$
	Criterion: IPC_3
C-1	$\widehat{k}_Z = 1$: H_0 not rejected

$\widehat{k}_Z$: number of factors estimated for the data de-factored under H_0;
p_f^* : bootstrap estimate of Pr("false positive")$\times$100
$\overline{p_f}$: maximum acceptable value of Pr("false positive").

4 Monte Carlo experiment

The small sample performances of the proposed procedure have been evaluated by simulation. To facilitate the comparison with the literature we use a Monte Carlo data generating process (DGP) derived from that used by Bai, with some minor changes introduced in order to mimick the structure of the regional data used in the empirical illustration. Two statically loaded independent $I(1)$ common factors ($F_j, j = 1, 2$), the first with a drift $\delta = 0.01$, shape long-term growth over time (with $T = 50, 100, 150$) and pervasive dependence across N spatial units, hereafter regions (with $N = 50, 100, 150$). Short-term and local fluctuations of the error at time t for unit i, $\varepsilon_{t,i}$, are governed by a STARMA(1,1,1) process (that is, an ARMA(1,1) augmented with a spatial lag term at time $(t-1)$; see, *e.g*, Dai and Billard, 1998). Thus, the value of the variable Y at time t for region i is generated as follows:

$$y_{t,i} = \lambda_{1i} F_{t,1} + \lambda_{2i} F_{t,2} + \varepsilon_{t,i} \tag{11a}$$

$$\Delta F_{1,t} = \delta + u_{t,1} \tag{11b}$$

$$\Delta F_{2,t} = u_{t,2} \tag{11c}$$

$$\varepsilon_{t,i} = \phi_i \varepsilon_{t-1,i} + \rho \textstyle\sum_{j=1}^{N} w_{ij} \varepsilon_{t,j} + \xi_{t,i} + \theta_i \xi_{t-1,i} \tag{11d}$$

The spatial weight w_{ij} is binary, taking the value 1 if units (i, j) share a border and zero else (the so-called rook case, see *e.g.*, Fingleton, 1999). The choice of the spatial structure is obviously inevitably arbitrary. In view of the empirical illustration on Italian data we chose a highly stylised representation of this country, with the units distributed over a rather narrow rectangular grid of fixed width 5 units: the length of the grid is thus 10 units for $N = 50$, 20 for $N = 100$, and 30 for $N = 150$. The first-order neighbours of a given unit are defined to be, on the same row, those on its right and its left, and, in the rows above and below, those in its same column. Units at the first and last column of a row do not respectively have the left and right neighbour. Hence, for a pair of units i, j placed in the same row $w_{ij} = 1$ if $|i - j| = 1$, and for any pair located in different rows $w_{ij} = 1$ only if $|i - j| = 5$, the width of the grid. In all other cases $w_{ij} = 0$. The spatial autoregressive coefficient is fixed at $\rho = 0.25$; setting $\rho = 0$ equation (11d) collapses to an usual ARMA process with no spatial structure. In order to respect the spatial links the datasets for $N = 50$ and 100 are always obtained as subsets of the the full dataset of 150 units. Time samples will also range from 50 to 150, values quite representative of empirical applications.

The noises u_1, u_2 are independent standard gaussians as in Bai, while the $\xi_i' s$ are heteroskedastic independent normal variates with variance $\sigma^2_{\xi_i} \sim Uniform(0.5, 1.5)$. Balancing the opposite requirements of generality and simplicity we assume homogenous MA coefficients $\theta_i = 0.5$ as in Bai, but we allow AR parameters ϕ_i to be heterogenous over units generating them as $\phi_i \sim Uniform(0.4, 0.6)$.With no loss of generality, the loadings of the second factor, λ_{2i}, are generated as independent gaussians with unit mean and standard deviation.

The first aim of our experiment is to evaluate the performances of the model selection procedure when valid constraints on the loadings of the first factor are imposed. To this end, we assume that these are homogenous within two groups, taken for simplicity of equal size ($N_1 = N_2 = 0.5N$). The loadings of the first group are fixed at $\lambda_{1G_1} = 1$ and those of the second at $\lambda_{1G_2} = 2$:

$$\lambda_{1i} = \begin{cases} 1, \, i \leq 0.5N \\ 2, \, i > 0.5N \end{cases} \tag{12}$$

The true null hypothesis is thus expressed as:

$$H_0 : \lambda_{1i} = \begin{cases} \lambda_{1G_1}, \, i \leq 0.5N \\ \lambda_{1G_2}, \, i > 0.5N \end{cases} \tag{13}$$

The loadings of the second factor, λ_{2i}, are generated as $Uniform(1,3)$ and fixed across simulations as all other random parameters.Since computational convenience will be important for the experiments involving the bootstrap, we set the maximum number of factors at the smallest value required to ensure consistency, $kmax = (r+1) = 3$. Exploratory simulations confirmed that, as expected, setting a much higher value such as $kmax = 8$ as in Bai's experiment, does not change the results. Since if a single factor is selected for $\mathbf{Y}$ the factor model does not make any sense for $\mathbf{Z}$ the full two-stages procedure has been carried out (with the convergence criterion fixed at $1e^{-5}$) only if the selected number of factors was greater than 1. Finally, we used 1000 Monte Carlo replications and bootstrap redrawings. The block lengths for the Stationary Bootstrap have been generated with a geometric distribution with mean parameter computed as $1.75\sqrt[3]{T}$, as in Palm, Smeekes and Urbain (2011).

Since in our simulations, consistently with Bai's, IPC_1 and IPC_2 delivered essentially equivalent estimates of the number of factors while IPC_3 tended to somehow underestimate it, we report here in Table 2 only the results obtained using IPC_1. First of all, we notice that using the Naive test the risk of rejecting the true restriction is very high, approximately between 25% and 30%: this procedure is thus not advisable. Using the bootstrap this risk is instead always very small: for instance, with sample sizes close to those of the empirical illustration ($N = 150, T = 50$) and maximum acceptable risk$\overline{p_f} = 5\%$ the frequency of false positives is only 3.2%. Of course, in order to be recommended a testing procedure should also have a good power to detect false restrictions. Unfortunately, given the dimension of the parameter space estimating this power is an essentially unfeasible task. In some exploratory analysis with the loadings of the trending factor in the Monte Carlo DGP either $N(1,1)$ or $Uniform(0,1)$ across units and the same restriction considered above we obtained, even with the smallest sample sizes, nearly 100% rejections. This evidence is clearly extremely limited, but it nevertheless suggests that the procedure may be empirically useful.

Table 2
Naive and Bootstrap test: frequency (×100) of "false positives"

		N=50			N=100			N=150		
					Naive test					
T	50		25.0			25.9			25.0	
	100		22.7			26.4			30.6	
	150		25.1			29.2			31.3	
					Bootstrap test					
	$\overline{p_f}$	1	5	10	1	5	10	1	5	10
T	50	1.2	4.8	14.8	0.3	3.7	12.0	1.6	3.2	9.6
	100	0.9	3.5	11.0	0.8	3.0	6.8	0.0	1.6	5.6
	150	0.4	3.2	7.2	1.7	3.8	6.5	0.3	1.3	2.9

$\overline{p_f}$: maximum acceptable value of Pr("false positive")
1000 bootstrap redrawings and Monte Carlo replications
selection criterion: IPC_1.

5 Conclusions

Research in factor modeling is very active to say the least, but the problem of evaluating constraints on the loadings has not received much attention yet. We examined this issue for the non-stationary factor model by Bai (2004), developing a bootstrap procedure that can be easily adapted to other estimators and consistent model selection procedures. Our proposal is to evaluate if a set of constraints is compatible with the data by comparing the estimated number of factors of the unconstrained and constrained models: a difference between the two values clearly provides evidence against the data-compatibility of the constraints. The probability of rejecting correct constraints ("false positives") can be controlled using the bootstrap. In our simulation experiments this probability has been always found to be very low, and its bootstrap estimates accurate: the proposed evaluation procedure may thus be a useful tool for empirical factor modelling of non stationary data.

Acknowledgments

The second authors acknowledges financial support from Sapienza University. Many thanks to Luca Repetto for personal communications. All computations have been carried out using Hansl, the free programming language of the project Gretl.

6 References

References

[1] Altissimo, F., R. Cristadoro, M. Forni, M. Lippi and G. Veronese, 2010, New eurocoin: Tracking economic growth in real time. The Review of Economics and Statistics, 92, 1024–1034.

[2] Amengual, D. and L. Repetto, 2014, Testing a large number of hypotheses in approximate factor models. CEMFI Working Paper No. 1410.

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

[4] Banerjee, A., M. Marcellino and I. Masten, 2016, An Overview of the Factor-augmented Error-Correction Model, in: E. Hillebrand and S.J. Koopman, (Eds.) Dynamic Factor Models (Advances in Econometrics, 35) Emerald, Bingley, pp. 3-41.

[5] Barigozzi, M., M. Lippi and M. Luciani, 2016, Non-Stationary Dynamic Factor Models for Large Datasets. Finance and Economics Discussion Series 2016-024. Board of Governors of the Federal Reserve System, Washington. DOI: 10.17016/FEDS.2016.024.

[6] Barro, R.J. and X. Sala-i-Martin, 1991, Convergence across states and regions. Brookings Papers on Economic Activity, 1, 107-182.

[7] Calzolari, G., 2017 Econometrics exams and round numbers: Use or misuse of indirect estimation methods?, Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2017.1348514

[8] Ciccarelli, C. and S. Fachin, 2017, Common factors, spatial dependence, and regional growth in the Italian manufacturing industry. DSS-E3 WP 2017/1, Dept. of Statistics, "Sapienza" University of Rome.

[9] Ciccarelli, C. and S. Fenoaltea, 2009, La produzione industriale delle regioni d'Italia, 1861-1913: una ricostruzione statistica. 1 Le industrie non manifatturiere. Rome, Bank of Italy.

[10] Ciccarelli, C. and S. Fenoaltea, 2014, La produzione industriale delle regioni d'Italia, 1861-1913: una ricostruzione statistica. 2 Le industrie estrattivo-manifatturiere. Rome, Bank of Italy.

[11] Connor, G., and R. Korajczyk, 1993, A test for the number of factors in an approximate factor model, Journal of Finance, 1263-1291.

[12] Dai, Y. and L. Billard, 1998, A Space-Time bilinear model and its identification. Journal of Time Series Analysis, 19, 657-679.

[13] Di Iorio, F. and S. Fachin, 2014, A panel cointegration study of the long-run relationship between savings and investments in the OECD economies, 1970-2007. Empirical Economics, 46, 1271-1300.

[14] Evans, G.B.A., and N.E. Savin, 1982, Conflict among the Criteria Revisited: The W, LR and LM Tests. Econometrica, 50, 737–48.

[15] Fingleton, B., 1999, Spurious spatial regression: some Monte Carlo results with a spatial unit root and spatial cointegration. Journal of Regional Science, 39, 1-9.

[16] Giannone, D., L. Reichlin and D. Small, 2008, Nowcasting: The real-time informational content of macroeconomic data. Journal of Monetary Economics, 55, 665–676.

[17] Hacioglu Hoke, S. and K. Tuzcuoglu, 2016, Interpreting the latent dynamic factors by threshold FAVAR model. Bank of England Staff Working Paper No. 622.

[18] Kapetanios, G., A Testing procedure for determining the number of factors in approximate factor models with large datasets, Journal of Business and Economic Statistics, 28, 397-409.

[19] Onatski, A. (2009), Testing hypotheses about the number of factors in large factors models, Econometrica, 77, 1147-1479.

[20] Onatski, A. (2010), Determining the number of factors from empirical distribution of eigenvalues, The Review of Economics and Statistics, 92, 1004-1016.

[21] Palm, F.C., S. Smeekes and J.P. Urbain, 2011, Cross-sectional dependence robust block bootstrap panel unit root tests. Journal of Econometrics, 163, 85-104.

[22] Parker, C., E. Paparoditis, and D.N. Politis, 2006, Unit root testing via the stationary bootstrap, Journal of Econometrics, 133, 601-638.

[23] Politis, D.N. and J.P. Romano, 1994, The Stationary Bootstrap. Journal of the American Statistical Association, 89, 1303-1313.

[24] Reis, R. and M.W. Watson, 2010, Relative goods' prices, pure inflation, and the Phillips correlation. American Economic Journal: Macroeconomics, 2, 128-157.

[25] Stock, J.H. and M. W. Watson, 2011, Dynamic Factor Models, in: M.P. Clements and D.F. Hendry (Eds.) The Oxford Handbook of Economic Forecasting. Oxford University Press, Oxford, 35-60.

7 Appendix Proof of Lemma 1

Before moving to the proof we briefly recall the essential lines of the Stationary Bootstrap by Politis and Romano (1994). Let $L_1, \ldots, L_T$, be a sequence of values randomly chosen from a geometric distribution of parameter θ. The first L_1 observations of the pseudoseries are given by L_1 observations from the originary series starting at a location randomly chosen from a uniform distribution on $\{2, ..., T\}$, followed by L_2 observations starting at a different randomly chosen location, and so on. The process ends when the pseudoseries, which will will reproduce the weak dependence links between the data, reaches the desired length.

In order to prove Lemma 1 we need to show that the pseudodata $\mathbf{Y}^* = \widehat{\mathbf{F}}_1^0 \widehat{\mathbf{\Lambda}}_1^0 + (\mathbf{F_2 \Lambda_2})^* + \varepsilon^*$ are a function of the errors $\mathbf{u}$ and ε, exactly as the observed data $\mathbf{Y}$ object of the modelling selection procedure whose properties are stated in Bai's Theorem 1.

Under Assumption 4 the Stationary Bootstrap can be applied to $\Delta(\widehat{\mathbf{F}}_\mathbf{2} \widehat{\mathbf{\Lambda}}_\mathbf{2})$ and $\widehat{\varepsilon}$.

We now examine in turn each of the three components of the pseudodata $\mathbf{Y}^*$, that is, $\widehat{\mathbf{F}}_1^0 \widehat{\mathbf{\Lambda}}_1^0$, $(\mathbf{F_2 \Lambda_2})^*$ and ε^*.

(a) $\widehat{\mathbf{F}}_1^0 \widehat{\mathbf{\Lambda}}_1^0$ is the common component due to the first factor as estimated under the homogeneity constraint by regressions with the observed data as dependent variable, and respectively the estimated factor and loadings as the independent variable (see Algorithm 1, steps 2 and 3). With standard notation, define $\mathbf{P}_F$ and $\mathbf{P}_\Lambda$ the OLS projection matrices used in the two regressions. Then,

$$\begin{aligned} \widehat{\mathbf{\Lambda}}_1^0 &= \mathbf{P}_F \mathbf{Y} \\ &= \mathbf{P}_F (\mathbf{F\Lambda} + \varepsilon) \end{aligned}$$

since $\mathbf{F}_{t\cdot} = \sum_{s=1}^{t} \mathbf{u}_{s\cdot}$, where $\mathbf{F}_{t\cdot}$ and $\mathbf{u}_{t\cdot}$ denote respectively the t-th rows of the factor matrix $\mathbf{F}$ and of the error matrix $\mathbf{u}$, it is immediately seen that $\widehat{\mathbf{\Lambda}}_1^0$ depends on $\mathbf{u}$ and ε. Analogously,

$$\begin{aligned} \widehat{\mathbf{F}}_1^0 &= \mathbf{P}_\Lambda \mathbf{Y} \\ &= \mathbf{P}_\Lambda (\mathbf{F\Lambda} + \varepsilon) \end{aligned}$$

so that the same holds for the factor estimate $\widehat{\mathbf{F}}_1^0$.

(b) Assume for convenience and without loss of generality that there are only two factors, so that the matrix $\mathbf{F}_2$ collapses to a column vector. For each unit i and time t $(\mathbf{F}_2 \mathbf{\Lambda_2})^*$ can then be written as

$$(F_{2t} \lambda_{2i})^* = \sum_{q=1}^{t} (\Delta F_{2q} \lambda_{2i})^* \tag{14}$$

where, again, the first difference of the common component for time q and unit i in the bootstrap sample, $(\Delta F_{2q}\lambda_{2i})^*$, is simply the difference of the estimated common component for the same unit and some period q defined according to the SB rules. We now need to make explicit the dependence of q from the block structure of the SB algorithm, and write the cumulative sum in (14) as a double sum over and within blocks. To this end, for $t \in [1, T-1]$, let:

(i) L_n be the length of block n;

(ii) $K_t = inf\{k : L_1 + \ldots + L_k \geq t\}$ the number of blocks to be cumulated to have a string of length t;

(iii) $M_t = L_1 + \ldots + L_k$ their total length;

(iv) ς_{K_t} the starting observation for block K_t, drawn from a uniform distribution over $[1, T-1]$;

(v) finally, define $\tau_m = (\varsigma_{K_m} + (m - M_m))mod(T-1) + 2$.

Then, analogously to the notation used by Parker et al. (2006) in the proof of their Lemma 4, p. 626, we can write

$$\begin{aligned}(F_{2t}\lambda_{2i})^* &= \sum_{m=1}^{K_t}\sum_{s=1}^{L_m}\Delta\widehat{F}_{2\tau_m+p}\widehat{\lambda}_{2i} \\ &= \widehat{\lambda}_{2i}\left[\sum_{m=1}^{K_m}\sum_{s=1}^{L_m}(\Delta\widehat{F}_{2\tau_m+s})\right].\end{aligned} \tag{15}$$

The next step is writing the factor estimate as (Bai, p. 142):

$$\widehat{F}_{2t} = \frac{\mathbf{Y}_{t\cdot}\hat{\mathbf{\Lambda}}_2}{N}$$

where $\mathbf{Y}_{t\cdot}$ is row t of the data matrix $\mathbf{Y}$. Since

$$\mathbf{Y}_{t\cdot} = \mathbf{F}_{t\cdot}\mathbf{\Lambda} + \varepsilon_{t\cdot}$$

where $\varepsilon_{t\cdot}$ is row t of the error matrix ε, we can also write

$$\widehat{F}_{2t} = \frac{(\mathbf{F}_{t\cdot}\mathbf{\Lambda} + \varepsilon_{t\cdot})\hat{\mathbf{\Lambda}}_2}{N}$$

so that:

$$\begin{aligned}\Delta\hat{F}_{2t} &= \frac{1}{N}(\Delta\mathbf{F}_{t\cdot}\mathbf{\Lambda} + \Delta\varepsilon_{t\cdot})\hat{\mathbf{\Lambda}}_2 \\ &= \frac{1}{N}(\mathbf{u}_{t\cdot}\mathbf{\Lambda} + \Delta\varepsilon_{t\cdot})\hat{\mathbf{\Lambda}}_2.\end{aligned} \tag{16}$$

Inserting (16) into (15) we eventually have

$$
\begin{aligned}
(\lambda_{i2}F_{2t})^* =& \hat{\lambda}_{i2} \sum_{m=1}^{K_m} \sum_{s=1}^{L_m} (\Delta \hat{F}_{2\tau_m+s}) \\
=& \hat{\lambda}_{i2} \left[\sum_{m=1}^{K_m} \sum_{s=1}^{L_m} \frac{1}{N} (\mathbf{u}_{\tau_m+s} \mathbf{\Lambda} + \Delta \varepsilon_{\tau_m+s}) \hat{\mathbf{\Lambda}}_2 \right] \qquad (17)
\end{aligned}
$$

which shows that the resampled common component $(\lambda_{i2}F_{2t})^*$ also depends upon $\mathbf{u}$ e ε only.

(c) The pseudo-error for time t and unit i, ε^*_{ti}, is simply the estimated error for the same unit and a different period, say τ_t, chosen according to the SB rules. Then, denoting by $\mathbf{F}_{\tau_t \cdot}$ and $\Lambda_{\cdot i}$ respectively the row τ_t of the matrix of factors and the column i of that of the loadings,

$$
\begin{aligned}
\varepsilon^*_{ti} &= \hat{\varepsilon}_{\tau_t i} \\
&= y_{\tau_t i} - \widehat{\mathbf{F}}_{\tau_m \cdot} \widehat{\mathbf{\Lambda}}_{\cdot i} \\
&= \mathbf{F}_{\tau_t \cdot} \Lambda_{\cdot i} + \varepsilon_{\tau_t i} - \widehat{\mathbf{F}}_{\tau_t \cdot} \widehat{\mathbf{\Lambda}}_{\cdot i} \\
&= \varepsilon_{\tau_t i} - (\widehat{\mathbf{F}}_{\tau_t \cdot} \widehat{\mathbf{\Lambda}}_{\cdot i} - \mathbf{F}_{\tau_t \cdot} \mathbf{\Lambda}_{\cdot i})
\end{aligned}
$$

Assuming $N/T \to 0$ and exploiting the asymptotic convergence of the estimated common component $\widehat{\mathbf{F}}_{\tau_t \cdot} \Lambda_{\cdot i}$ to their true values (Bai, theorem 4, p. 174), we have:

$$
\varepsilon^*_{ti} = \varepsilon_{\tau_t i} - \left[\mathbf{H}' \left(\frac{\mathbf{\Lambda}' \mathbf{\Lambda}}{N} \right)^{-1} \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \lambda_j \varepsilon_{\tau_t j} + o_p(1) \right] \mathbf{H}^{-1} \mathbf{\Lambda} + O_p(1/T^{1/2})
$$

where $\mathbf{H}$ depends upon $\mathbf{F}$, hence $\mathbf{u}$ (see Bai, p. 164). The pseudo-errors ε^* are thus a function of ε and $\mathbf{u}$.

Summing up, from (a), (b) and (c) it follows that all the three components of $\mathbf{Y}^*$, that is, $\widehat{\mathbf{F}}^0_1 \widehat{\mathbf{\Lambda}}^0_1$, $(\mathbf{F_2 \Lambda_2})^*$ and ε^*, are a function of ε and $\mathbf{u}$, so that Bai theorem 1 applies. This completes the proof. ■