Journal of Econometrics 172 (2013) 127–141

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

On bootstrapping panel factor series

Lorenzo Trapani *

Cass Business School, City University London, United Kingdom

ARTICLE INFO

Article history:
Received 23 April 2010
Received in revised form
29 May 2012
Accepted 5 September 2012
Available online 12 September 2012

JEL classification:
C23

Keywords:
Bootstrap
Invariance principle
Factor series
Vector AutoRegression
Joint asymptotics

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.

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,b, 2010), and the references therein. Nonstationary panel factor series have also been paid noticeable attention to 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 $i = 1, \ldots, n$ and $t = 1, \ldots, T$ 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 et al. (2012) 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 et al. (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 et al. (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,

* Correspondence to: Cass Business School, Faculty of Finance, 106 Bunhill Row, London EC1Y 8TZ, United Kingdom. Tel.: +44 0 207 0405260.
E-mail address: L.Trapani@city.ac.uk.

0304-4076/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2012.09.001

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 $(n, T) \rightarrow \infty$, 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. $\max_x \|Ax\|_p/\|x\|_p$ (the Euclidean norm being defined simply as $\|A\|$), "i_m" indicates a unit column vector of dimension m, "$\rightarrow$" the ordinary limit, "$\stackrel{d}{\rightarrow}$" weak convergence, "$\stackrel{p}{\rightarrow}$" 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: $\delta_{nT} = \min\{\sqrt{n}, \sqrt{T}\}$, $C_{nT} = \min\{\sqrt{n}, T\}$, $\varphi^F_{nT} = \min\{n, \sqrt{T/\log T}\}$ and $\varphi^u_{nT} = \min\{\sqrt{n}, \sqrt{T/\log T}\}$.

2. Model, assumptions and preliminary asymptotics

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

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

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 $u_t = [u_{1t}, \ldots, u_{nt}]'$; then u_t admits the invertible $MA(\infty)$ representation $u_t = \Gamma(L)e^u_t = \sum_{j=0}^{\infty} \Gamma_j e^u_{t-j}$, where (i) e^u_t is i.i.d. across t with $E[e^u_t] = 0$, $E[e^u_t e^{u\prime}_t] = \Sigma_u$; also, letting e^u_{it} be the ith element of e^u_t, $\max_{i,t} E|e^u_{it}|^{8+\delta} < \infty$ 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 ith 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) $\|\Gamma(1)\|_1 \leq M$, $\|\Gamma^{-1}(1)\|_1 \leq M$, $\|\Gamma^{-1}(1)\|_\infty \leq M$ and $\|\Sigma_u\|_1 \leq M$; (b) $E|n^{-1/2}\sum_{i=1}^n [u_{is}u_{it} - E(u_{is}u_{it})]|^4 \leq M$ for every (t, s); (c) $E|\sum_{i=1}^n u_{it}|^{2+\delta} \leq ME|\sum_{i=1}^n u^2_{it}|^{\frac{2+\delta}{2}}$ 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 $\varepsilon_t = \alpha(L)e^F_t = \sum_{j=0}^{\infty} \alpha_j e^F_{t-j}$ with (i) e^F_t is i.i.d. with $E[e^F_t] = 0$, $E[e^F_t e^{F\prime}_t] = \Sigma_e$ and $E\|e^F_t\|^{8+\delta} < \infty$ 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 $\lim_{n\rightarrow\infty} n^{-1}\sum_{i=1}^n \lambda_i\lambda_i' = \Sigma_\Lambda$, if $n \rightarrow \infty$ 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 $\Sigma_\Lambda^{1/2}(T^{-2}\sum_{t=1}^T F_t F_t')\Sigma_\Lambda^{1/2}$ are a.s. distinct as $T \rightarrow \infty$.

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 $\|\Gamma^{-1}(1)\|_1$ 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 $T^{-1}\sum_{s=1}^T\sum_{t=1}^T |\gamma_{s-t}| \leq M$, 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 8th moment of e^F_t 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 $\liminf_{T\rightarrow\infty}(\log\log T)T^{-2}\sum_{t=1}^T F_t F_t' = D$ 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

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