On the applicability of the sieve bootstrap in time series panels

Citation for published version (APA):

Smeekes, S., & Urbain, J. R. Y. J. (2014). On the applicability of the sieve bootstrap in time series panels. *Oxford Bulletin of Economics and Statistics*, *76*(1), 139-151. https://doi.org/10.1111/obes.12005

Document status and date:
Published: 01/02/2014

DOI:
10.1111/obes.12005

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 76, 1 (2014) 0305-9049
doi: 10.1111/obes.12005

On the Applicability of the Sieve Bootstrap in Time Series Panels*

STEPHAN SMEEKES and JEAN-PIERRE URBAIN

Department of Quantitative Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, the Netherlands (e-mails: S.Smeekes@maastrichtuniversity.nl; J.Urbain@maastrichtuniversity.nl)

Abstract

In this article, we investigate the validity of the univariate autoregressive sieve bootstrap applied to time series panels characterized by general forms of cross-sectional dependence, including but not restricted to cointegration. Using the final equations approach we show that while it is possible to write such a panel as a collection of infinite order autoregressive equations, the innovations of these equations are not vector white noise. This causes the univariate autoregressive sieve bootstrap to be invalid in such panels. We illustrate this result with a small numerical example using a simple DGP for which the sieve bootstrap is invalid, and show that the extent of the invalidity depends on the value of specific parameters. We also show that Monte Carlo simulations in small samples can be misleading about the validity of the univariate autoregressive sieve bootstrap. The results in this article serve as a warning about the practical use of the autoregressive sieve bootstrap in panels where cross-sectional dependence of general form may be present.

I. Introduction

In this article, we investigate the validity of the autoregressive (AR) sieve bootstrap in a panel data context. The sieve bootstrap is very popular in applied work as it is one of the better performing time series bootstrap methods, and moreover easy to implement. In particular for the analysis of unit roots and cointegration, the sieve bootstrap appears to perform very well, see for example Chang, Park and Song (2006) and Palm, Smeekes and Urbain (2008).

Recently, people have also started to use the autoregressive sieve bootstrap in panel data with finite cross-sectional dimension. While a multivariate version of the sieve bootstrap exists based on the estimation of a vector autoregressive (VAR) model, this VAR sieve bootstrap becomes infeasible if the dimension of the system grows, and cannot be used in typical panel data applications where the number of cross-sectional units is too large, even in the finite N case where it remains asymptotically valid. The VAR sieve bootstrap is

*This research was supported financially by the Netherlands Organisation for Scientific Research (NWO). We are grateful to two anonymous referees, Alain Hecq, Marco Lippi and Franz Palm for helpful comments and discussions. Special thanks go to Peter Boswijk for comments that fundamentally improved the article.

JEL Classification numbers: C15, C23.

therefore only of practical use in panels with a very small cross-sectional dimension where the estimation of a VAR model is justified. In applied work people therefore often use the univariate AR sieve bootstrap, with the modification that the residuals are resampled jointly across units to preserve the cross-sectional dependence.[1] Examples for the analysis of unit roots and cointegration in panel data include Chang (2004), Smith *et al.* (2004), Cerrato and Sarantis (2007), Westerlund and Edgerton (2007), Fuertes (2008), Hanck (2009) and Di Iorio and Fachin (2011).

With the exception of Chang (2004), these articles do not provide theoretical results on the validity of the AR sieve bootstrap. Chang (2004) assumes that there is only contemporaneous dependence between units and shows that the AR sieve bootstrap is valid under those conditions. However, these conditions are likely to be violated in many empirical applications. It remains unknown if the AR sieve bootstrap is valid under more general conditions on the cross-sectional dependence.[2]

In this article, we therefore explore the properties of the AR sieve bootstrap in time series panels with more general forms of cross-sectional dependencies. Our main tool in this analysis is the final equations approach developed by Zellner and Palm (1974) and Palm (1977) among others, which shows that any VAR model can be written in final form as a system of autoregressive moving average (ARMA) equations for each individual unit. Starting from a general model that allows for various forms of cross-sectional dependencies, we use this approach as well as results of Kreiss, Paparoditis and Politis (2011) to derive univariate AR representations for each unit that are needed in order to be able to apply the AR sieve bootstrap. However, as we will see, the innovations of these AR models are not vector white noise thus invalidating the use of the AR sieve bootstrap in more general models. Our focus is on unit root testing in panels, but our results carry through to other settings. Note that our analysis assumes that the cross-sectional dimension N, although potentially large, is finite. The same assumption is for example also made in Chang (2004).

We also illustrate our analysis with a numerical example of a model in which the AR sieve bootstrap is invalid but where observed size distortions are not necessarily large. This example serves as a possible explanation of the finding that the AR sieve bootstrap performs well in some simulation studies even though it is not valid.

The structure of the article is as follows. In section II, we introduce a general sieve bootstrap algorithm used in panel data. The theoretical analysis is contained in section III. A small simulation study is considered in section IV. Section V offers some conclusions. Proofs are contained in the Appendix.

Finally, a word on notation. $|z|$ applied to a (possibly complex) number denotes its absolute value, $\|A\|$ applied to any matrix A denotes its Euclidean norm, and $\lfloor x \rfloor$ is the largest integer smaller than or equal to x. $W(r)$ denotes a standard N-dimensional Brownian

[1]To distinguish between the multivariate sieve bootstrap where a VAR model is estimated and the univariate sieve bootstrap where individual autoregressive models are estimated, we refer to the former as VAR sieve bootstrap and to the latter as AR sieve bootstrap.

[2]The sieve bootstrap has been used in combination with some kind of estimation of a factor model to filter out dependence by Pesaran *et al.* (2006), Fuertes (2008), Trapani and Urga (2010), and Trapani (2011). While this is a perfectly sensible and valid approach, in this article, we only focus on the application of the sieve bootstrap directly to the data, hence where the sieve bootstrap is the only tool that takes into account the cross-sectional dependence. Trapani (2011) provides an extensive theoretical analysis of the sieve bootstrap (both of AR and VAR type) in a model where common factors are estimated.

motion. Weak convergence (convergence in distribution) is denoted by $\xrightarrow{d}$, and bootstrap weak convergence in probability is denoted by $\xrightarrow{d^*}_p$.

II. Sieve bootstrap procedure

Here we will describe a typical setup for an AR sieve bootstrap panel unit root test on a panel of data $y_{i,t}$, $i=1,\ldots,N$, $t=1,\ldots,T$.

(i) For each $i=1,\ldots,N$, obtain $y^d_{i,t}=y_{i,t}-\hat{\beta}_i' z_t$, where $z_t=1$ or $z_t=(1,t)'$ and $\hat{\beta}_i$ is obtained for example through an OLS or GLS regression of $y_{i,t}$ on z_t.

(ii) For each $i=1,\ldots,N$, run an ADF regression with q_i lags on $\Delta y_{i,t}$ to obtain residuals

$$\hat{e}_{q_i,i,t}=\Delta y^d_{i,t}-\hat{\rho} y^d_{i,t-1}-\sum_{j=1}^{q_i}\hat{d}_{i,j}\Delta y^d_{i,t-j}.$$

Recentre the residuals to obtain

$$\tilde{e}_{q_i,i,t}=\hat{e}_{q_i,i,t}-(T-q_i-1)^{-1}\sum_{t=q_i+2}^{T}\hat{e}_{q_i,i,t}.$$

(iii) Resample with replacement from $\tilde{e}_{q,t}=(\tilde{e}_{q_1,1,t},\ldots,\tilde{e}_{q_N,N,t})'$ to obtain bootstrap errors $e^*_t=(e^*_{1,t},\ldots,e^*_{N,t})'$.

(iv) For each $i=1,\ldots,N$, construct $u^*_{i,t}$ recursively as

$$u^*_{i,t}=\sum_{j=1}^{q_i}\hat{d}_{i,j}u^*_{i,t-j}+e^*_{i,t},$$

using the estimated parameters $\hat{d}_{i,j}$ from step (ii), and construct $y^*_{i,t}$ as $y^*_{i,t}=y^*_{i,t-1}+u^*_{i,t}$.

(v) Use the bootstrap sample $y^*_t=(y^*_{1,t},\ldots,y^*_{N,t})'$ to calculate the desired panel unit root test statistic.

Lag lengths $q=(q_1,\ldots,q_N)'$ can be selected for each equation individually, for example by information criteria. We need to allow these lag lengths q in the sieve bootstrap to go to infinity at a slow enough rate.

Assumption 1. Let $q_i\to\infty$ and $q_i=o((T/\ln T)^{1/3})$ as $T\to\infty$ for all $i=1,\ldots,N$.

Under the assumptions employed in this paper the convergence rate of the $\hat{d}_{i,j}$ estimators is $O_p((\ln T/T)^{1/2})$, see for example Hannan and Kavalieris (1986, theorem 2.1) and Smeekes (2012, lemma 1).

III. Invalidity of the sieve bootstrap in panel data

In this section we will show theoretically that the AR sieve bootstrap is not valid in panel data with complex cross-dependencies. Let $y_t=(y_{1,t},\ldots,y_{N,t})'$ and

$$P(L)y_t=[(1-L)I_N-\alpha\beta' L]y_t=\Psi(L)\varepsilon_t, \tag{1}$$

where α and β are $N\times r$ matrices (with $r<N$). Furthermore we assume that

Assumption 2.

(i) ε_t are i.i.d. with $E(\varepsilon_t)=0$, $E(\varepsilon_t\varepsilon_t')=\Sigma$ and $E\|\varepsilon_t\|^4<\infty$.
(ii) $\det(\Psi(z))\neq 0$ for all $z\in\mathbb{C}$ such that $|z|\leq 1$, and $\sum_{j=0}^{\infty} j\|\Psi_j\|<\infty$.
(iii) $\det(\alpha_\perp'\beta_\perp)\neq 0$.

This DGP describes a multivariate $I(1)$ process with possible cointegration. If $\alpha=0$, there is no cointegration, otherwise there are r cointegrating relations. In the following we will not treat these two cases distinctively. We assume that $y_0=0$ and that there are no deterministic components present in the DGP. However, these assumptions are made purely for expositional simplicity and can be dispensed with without any difficulty.

By letting $\Phi(z)=\Psi(z)^{-1}P(z)$, we can derive the VAR representation

$$\Phi(L)y_t=\varepsilon_t, \tag{2}$$

where $\Phi(1)=-\Psi(1)^{-1}\alpha\beta'$. We can further rewrite this to the vector error correction model (VECM) representation

$$\Delta y_t=\Psi(1)^{-1}\alpha\beta' y_{t-1}+\Phi^*(L)\Delta y_{t-1}+\varepsilon_t, \tag{3}$$

where the specific form for $\Phi^*(z)$ is given below equation (4) in Palm, Smeekes and Urbain (2010).

From the expression in equation (3), we can derive the common trends representation (see e.g. Johansen, 1995, theorem 4.2)

$$y_t=C\sum_{s=1}^{t}\varepsilon_t+C^*(L)\varepsilon_t, \tag{4}$$

where $C=\beta_\perp'(\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp\Psi(1)$,[3] $C^*(z)=\sum_{j=0}^{\infty}C_jz^j$ has all roots outside the unit circle and $\sum_{j=0}^{\infty}j\|C_j\|<\infty$ as a consequence of Assumption 2. It then follows that

$$T^{-1/2}y_{\lfloor Tr\rfloor}\xrightarrow{d}B(r), \tag{5}$$

where $B(r)=C\Sigma^{1/2}W(r)$ is an N-dimensional Brownian motion with covariance matrix $\Omega=C\Sigma C'$, which is equal to the long-run covariance matrix of Δy_t.

In order to be valid the AR sieve bootstrap should be able to replicate the invariance principle (5). Consequently, in order to be able to apply the AR sieve bootstrap, we need first to be able to write the DGP as a diagonal VAR process for Δy_t. To derive these univariate AR representations for $\Delta y_{i,t}$, $i=1,\ldots,N$, we use the final equations approach (see e.g. Zellner and Palm, 1974, Palm, 1977, Cubadda, Hecq and Palm, 2009) and write the VAR form (2) as

$$\det(\Phi(L))y_t=\tilde{\Phi}(L)\varepsilon_t,$$

where $\tilde{\Phi}(z)$ is the adjoint matrix of $\Phi(z)$. Note that $\det(\Phi(z))$ contains $N-r$ unit roots ($r=0$ if $\alpha=0$). In order for y_t to be $I(1)$, $\tilde{\Phi}(z)$ must have $N-r-1$ unit roots. Hence, $N-r-1$ unit roots are common and cancel out. Hence, we may write $\det(\Phi(z))=a(z)(1-z)(1-z)^{N-r-1}$ and $\tilde{\Phi}(z)=B(z)(1-z)^{N-r-1}$, where $a(z)$ and $B(z)$ do not contain any factors $(1-z)$ (see Cubadda *et al.*, 2009, for more details).

[3]Note that $\alpha_\perp=I$ if $\alpha=0$ and hence $C=\Psi(1)$.

Let us further define $B_i(z)$ as the ith row of $B(z)$, and $B_{i,j}(z)$ as the (i,j)th element of the matrix $B(z)$. Then we may write for unit i that

$$a(L)\Delta y_{i,t} = B_i(L)\varepsilon_t = \sum_{j=1}^{N} B_{i,j}(L)\varepsilon_{j,t}. \tag{6}$$

We still need to show that the right hand side of equation (6) can be written as a univariate MA(∞) process that can be inverted to obtain an autoregressive representation. Theorem 1 shows this is indeed possible.

Theorem 1. Let the DGP be given by (1) and let Assumption 2 hold. For each $i=1,\ldots,N$ we may write

$$d_i(L)\Delta y_{i,t} = e_{i,t},$$

where $d_i(z)\neq 0$ for $z\in\mathbb{C}$ such that $|z|\leq 1$, $\sum_{j=0}^{\infty} j|d_{i,j}|<\infty$ and $e_{i,t}$ is white noise with $\mathrm{E}(e_{i,t})=0$, $\mathrm{E}(e_{i,t}^2)=\tilde{\sigma}_{i,i}$ and $\mathrm{E}(e_{i,t}^4)<\infty$.

Note that it would be tempting to conclude that the results in Theorem 1 provide a justification for the use of an AR sieve bootstrap procedure in panel data. Indeed, all assumptions on the DGP needed in order to apply the AR sieve bootstrap (see e.g. Chang, 2004, Assumptions 1 and 2) are satisfied, with the exception that $e_t=(e_{1,t},\ldots,e_{N,t})'$ is not a vector i.i.d. process. Unfortunately, this violation is serious enough to cause the AR sieve bootstrap to be theoretically invalid. There are two aspects of this violation that should be taken into account.

First, for each $i=1,\ldots,N$, $e_{i,t}$ is not i.i.d. but white noise. As explained in detail by Kreiss *et al.* (2011), this will only affect validity of the sieve bootstrap if the limiting distribution of the resulting test statistic depends on moments of $e_{i,t}$ higher than the second moment. The limit distributions of the unit root test statistics considered here are driven by the invariance principle (as in equation (5)) and only depend on the first two moments of the innovations. Therefore this does not lead to invalidity in this particular application, although it could do in others.

Second, while for each individual $i=1,\ldots,N$, $e_{i,t}$ is univariate white noise, e_t is not vector white noise. This is a more serious problem that causes general invalidity of the AR sieve bootstrap, as the off-diagonal elements of the long-run covariance matrix of e_t are not equal to those of the contemporaneous covariance matrix. To see why this is the case, let $D(z)=\mathrm{diag}(d_1(z),\ldots,d_N(z))$, such that $e_t=D(L)y_t$. Note that we can write, using equation (4), that

$$e_t = D(L)\Delta y_t = D(L)C\varepsilon_t + D(L)C^*(L)\Delta\varepsilon_t = \tilde{C}(L)\varepsilon_t.$$

While $E(e_{i,t}e_{i,t-s})=0$ by construction for $s>0$, this will in general not be true for $E(e_{i,t}e_{j,t-s})$ if $i\neq j$. Define $\tilde{\Sigma}=(\tilde{\sigma}_{i,j})_{i,j=1}^{N}=E(e_te_t')$ and $\tilde{\Omega}=(\tilde{\omega}_{i,j})_{i,j=1}^{N}=\lim_{T\to\infty}E[(T^{-1})\left(\sum_{t=1}^{T}e_t\right)\left(\sum_{t=1}^{T}e_t\right)']$. Then, while $\tilde{\omega}_{i,i}=\tilde{\sigma}_{i,i}$, in general we have that $\tilde{\omega}_{i,j}\neq\tilde{\sigma}_{i,j}$ for $i\neq j$.

Note that this result is irreducible, as we have found valid AR representations for each unit (making the univariate AR sieve bootstrap possible) that are derived from the fundamental ARMA models obtained through the final form. That is, there is no further transformation possible that will allow us to make the innovations vector white noise. Therefore, the AR sieve bootstrap is not valid when applied to a process generated by equation (1).

The invalidity of the univariate AR sieve bootstrap is due to the fact that the bootstrap does not estimate, i.e. reproduce, the long-run covariance matrix of the data correctly. Using established results on the sieve bootstrap, see for example Park (2002), Chang (2004) and Palm *et al.* (2010, theorem 2), we can show that

$$T^{-1/2} y^{*}_{\lfloor Tr \rfloor} \xrightarrow{d^*}_{p} B^{*}(r), \tag{7}$$

where $B^*(r)$ is an N-dimensional Brownian motion with covariance matrix $\Omega^* = D(1)^{-1} \tilde{\Sigma} D(1)^{-1\prime}$, which will in general not be equal to $\Omega = D(1)^{-1} \tilde{\Omega} D(1)^{-1\prime}$, the covariance matrix of $B(r)$ in equation (5).

Remark 1. Unfortunately there does not appear to be an intuitive way to link the DGP parameters to the parameters appearing in Ω and Ω^* (even for a bivariate VAR(1) model a complicated cumbersome expression arises that is difficult to interpret). It is therefore hard, if not impossible, to give simple conditions on the DGP considered here under which the AR sieve bootstrap is asymptotically valid. The only easily interpretable condition for sieve bootstrap validity is that $\alpha = 0$ and $\Psi(z)$ diagonal, in which case the DGP reduces to the one used by Chang (2004).

Remark 2. Our analysis assumes a finite N, yet it is not hard to see that the conclusions carry through to the case where N increases to infinity with T. The model with increasing N is analyzed in detail by Trapani (2011), who shows that, if the sieve bootstrap is applied in a panel model with a common factor structure that is estimated through principal components, the AR sieve bootstrap is valid for the long-run covariance matrix as long as the cross-sectional dependence between idiosyncratic components is 'small', i.e. vanishing with N (see his theorem 3 for details). It follows from our results that if this is not the case, for example if there are common factors in the data that are not taken care of, the sieve bootstrap will not be valid for large N, as in such a case the problem of having residuals that are not vector white noise persists even with large N.

IV. Simulations

To investigate the effect of the failure of the AR sieve bootstrap to capture the correct covariance matrix, we will now analyze a simple model by Monte Carlo simulation. We consider the following DGP:

$$\Delta y_{i,t} = \varepsilon_{i,t} + \sum_{j=1}^{N-i} \theta \varepsilon_{i+j,t-j}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T, \tag{8}$$

where $\varepsilon_t = (\varepsilon_{1,t}, \ldots, \varepsilon_{N,t})'$ is i.i.d. $N(0, I_N)$. We will consider $N = 2$ and $N = 10$. Note that this process is in final form as $e_{i,t} = \Delta y_{i,t}$ is white noise individually for each $i = 1, \ldots, N$. Of course, if $\theta \neq 0$, it is not vector white noise. Although this is a very specific and restricted DGP, this final form can capture relevant features of final forms arising from complex and rich dependence including cointegration between units, because the diagonal

AR dynamics are not of interest for AR sieve bootstrap validity, and the only important aspect is that the innovations are not vector white noise. This specific process was chosen as it is both in final form and there is no contemporaneous correlation, such that the AR sieve bootstrap cannot model any cross-sectional dependence. We will perform a group-mean (demeaned) Dickey–Fuller t-test, denoted t_{gm}, on these series using the sieve bootstrap.

Asymptotic simulations

We start with an asymptotic analysis of the AR sieve bootstrap. It follows directly (cf. Chang and Park, 2002, Chang, 2004) that

$$t_{gm} \xrightarrow{d} \frac{1}{N}\sum_{i=1}^{N} \frac{B_i^{\mu}(1)^2 - B_i^{\mu}(0)^2 - \omega_{i,i}}{2\sqrt{\omega_{i,i}\int_0^1 B_i^{\mu}(r)^2 \mathrm{d}r}},$$

where $B_i^{\mu}(r) = B_i(r) - \int_0^1 B_i(r)$ and $B(r) = (B_1(r), \ldots, B_N(r))'$ is a multivariate Brownian motion with covariance matrix $\Omega = (\omega_{i,j})_{i,j=1}^{N}$ with elements given by

$$\omega_{i,j} = \theta^{1-\delta_{ij}} + \theta^2(N - \max\{i,j\}),$$

where δ_{ij} is the Kronecker delta. This can be derived by noting that $\Omega = \Psi(1)\Psi(1)'$, where

$$\Psi(1) = \sum_{j=1}^{N-1} \Psi_j = \begin{bmatrix} 1 & \theta & \cdots & \cdots & \theta \\ 0 & 1 & \theta & \cdots & \theta \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & \theta \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}.$$

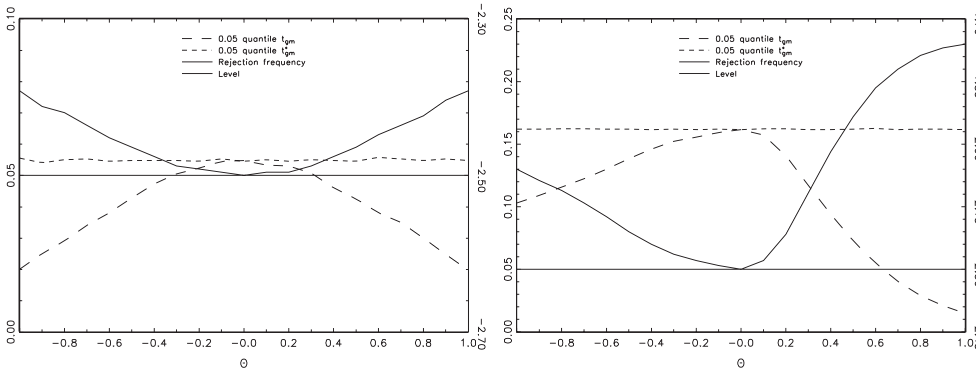
On the other hand, for the sieve bootstrap statistic we have that

$$t_{gm}^{*} \xrightarrow{d^*}_{p} \frac{1}{N}\sum_{i=1}^{N} \frac{B_i^{*\mu}(1)^2 - B_i^{*\mu}(0)^2 - \omega_{i,i}^{*}}{2\sqrt{\omega_{i,i}^{*}\int_0^1 B_i^{*\mu}(r)^2 \mathrm{d}r}},$$

where $B_i^{*\mu}(r) = B_i^{*}(r) - \int_0^1 B_i^{*}(r)$ and $B^{*}(r) = (B_1^{*}(r), \ldots, B_N^{*}(r))'$ is a multivariate Brownian motion with covariance matrix $\Omega^{*} = (\omega_{i,j}^{*})_{i,j=1}^{N}$ with elements given by

$$\omega_{i,j}^{*} = \delta_{ij}\left[1 + \theta^2(N - i)\right].$$

It may seem as if the bootstrap replicates the limiting distribution correctly, as the limiting distributions only depend explicitly on the parameters $\omega_{i,i}$ and $\omega_{i,i}^{*}$, which are equal. However, as an average of individual Dickey–Fuller distributions, the distribution of t_{gm} and t_{gm}^{*} also depends implicity on the covariance between these individual distributions and hence on the covariance between the elements of the vector Brownian motion. As $\omega_{i,j}^{*} = 0$ if $i \neq j$, the bootstrap does not correctly replicate these covariances, and therefore the bootstrap limit distribution is not the same as the original limit distribution.



(a) $N=2$ (b) $N=10$

Figure 1. Asymptotic size and quantiles. (a) $N=2$; (b) $N=10$

To investigate the asymptotic effect of the invalidity of the AR sieve bootstrap, we simulate the 0.05-quantiles of the distributions of t_{gm} and t^*_{gm} and the asymptotic rejection frequencies using a level of 0.05. The asymptotic distributions were obtained by direct simulation of the relevant limiting representations, approximating multivariate standard Brownian motion using i.i.d. $N(0, I_N)$ random variables, and the integrals by normalized sums of 2,000 steps. All simulations were performed using 500,000 Monte Carlo replications.

In Figure 1, we report the results for θ varying from -1 to 1 for $N=2$ and $N=10$. Values for the quantiles are given on the right axis, while size is given on the left axis. The invalidity of the sieve bootstrap is confirmed by the differing quantiles and corresponding asymptotic rejection frequencies. While for $N=2$ the rejection frequencies differ from the nominal level, the asymptotic size distortion is not very large. If such size distortions would be found in a small sample Monte Carlo simulation, they might easily be (incorrectly) dismissed as small sample size distortions. For $N=10$ the situation is different. There are considerably larger size distortions, in particular for large positive θ. It appears that the effect of the parameter θ, governing the cross-sectional dependence, increases with N in this particular model. Note though that if θ is close to zero size distortions are still small, although the AR sieve bootstrap remains invalid unless θ is exactly equal to zero.

Finite sample simulations

We next consider the same DGP for a finite sample Monte Carlo simulation. Sample sizes $T=50, 100, 250, 500,$ 1,000, 2,000 are considered. Results are based on 5,000 simulations and 499 bootstrap replications. Next to the AR sieve bootstrap (SB), we also consider the i.i.d. bootstrap (IID) and the moving blocks bootstrap (MBB). As equation (8) is already in final form, the augmentation with lags in the AR sieve bootstrap is unnecessary. The i.i.d. bootstrap therefore plays the same role here as the AR sieve bootstrap; the only difference with the sieve bootstrap is that there will be no finite sample effect from the selection of the lag length. As such it basically provides a 'clean' version of the sieve bootstrap here. The moving blocks bootstrap was shown to be valid in this context by Palm, Smeekes and Urbain (2011), and therefore provides a benchmark. The block length in the MBB is selected

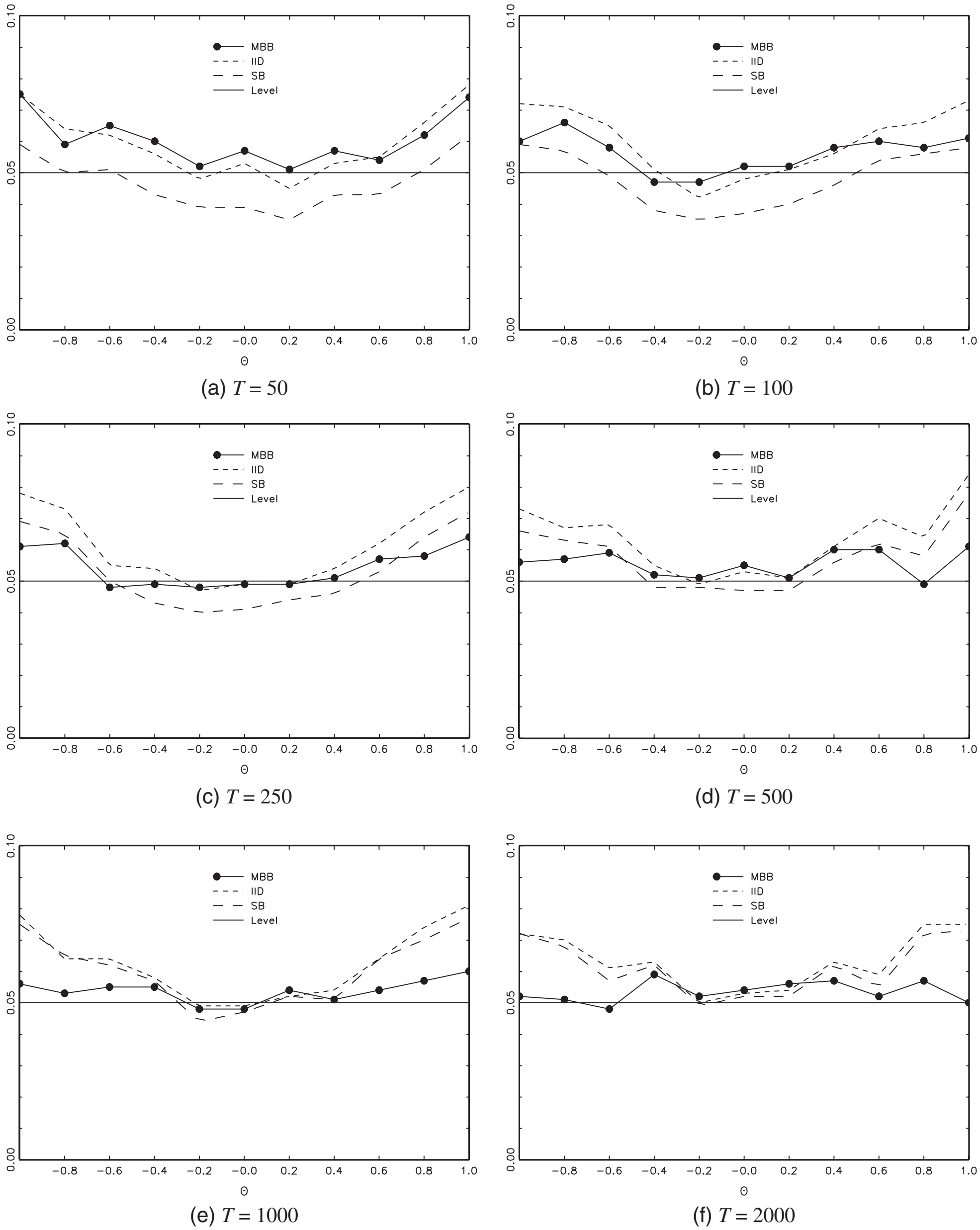


Figure 2. Size in finite samples, $N = 2$. (a) $T = 50$; (b) $T = 100$; (c) $T = 250$; (d) $T = 500$; (e) $T = 1,000$; (f) $T = 2,000$

as $1.75T^{1/3}$ as in Palm *et al.* (2011), while the lag length in the sieve bootstrap is selected by the modified Akaike information criterion (MAIC) with an upper bound of $12(T/100)^{1/4}$.[4]

[4]Clearly for the DGP with $N = 2$ we could have also used the VAR sieve bootstrap, given the small cross-sectional dimension. We are however not interested in finding the best bootstrap method for this specific DGP. As mentioned before, the DGP is not of interest in itself, but serves as illustration for more general models, capturing the most important features. Moreover, for the model with $N = 10$, implementing the VAR sieve bootstrap for the smaller time dimensions would already become impossible with the lag length selection method considered here.

Figures 2 and 3 present the results. For small to moderate T it is difficult to see any difference between the valid moving blocks bootstrap and the invalid i.i.d. and AR sieve bootstrap. This is true for $N=2$ in particular. Here T needs to increase to 1,000 before the bowl-shaped pattern found in Figure 1 really becomes visible for the invalid bootstrap methods. Surprisingly, for $N=10$ we also need a larger T (though not as extreme as for

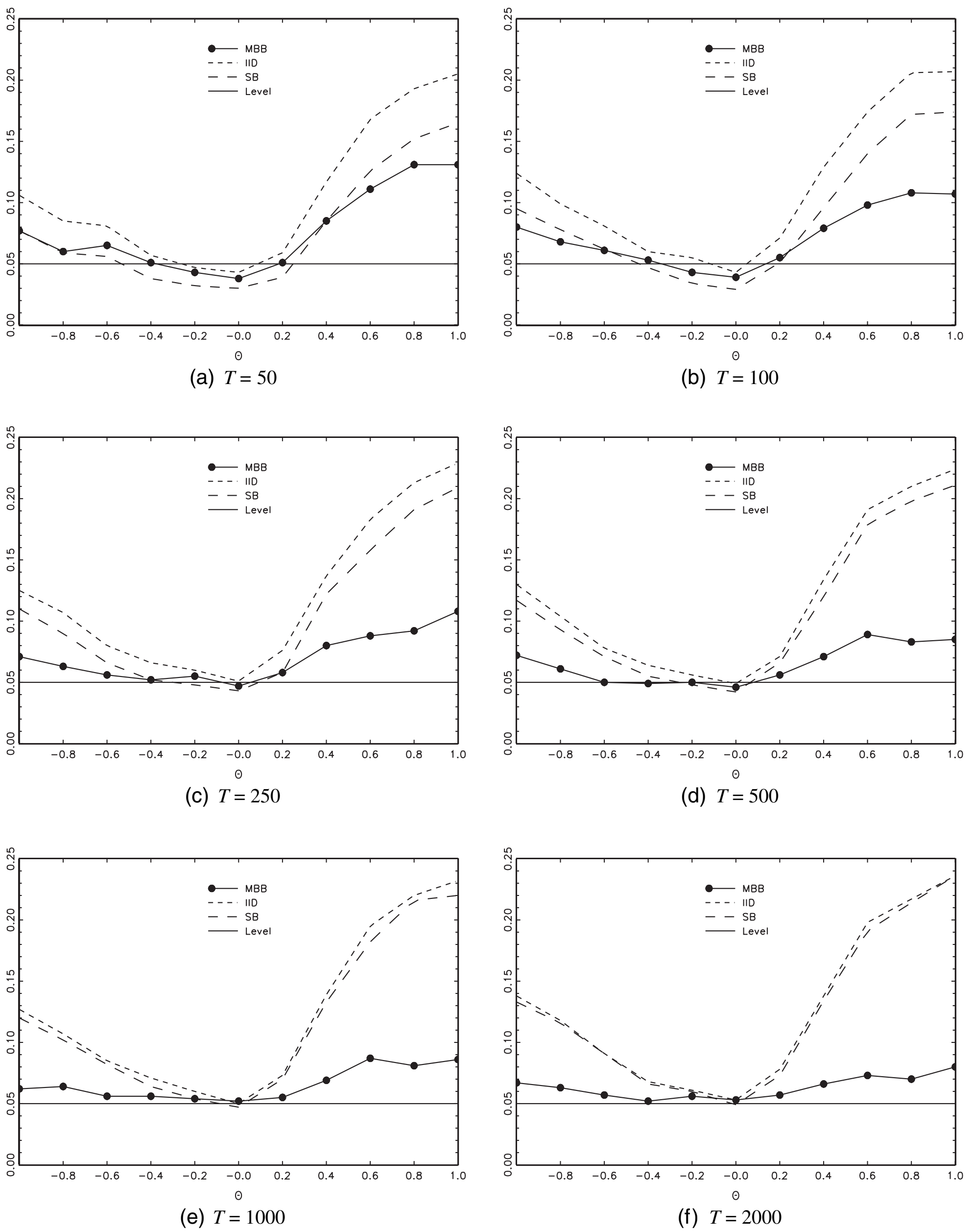


Figure 3. Size in finite samples, $N=10$. (a) $T=50$; (b) $T=100$; (c) $T=250$; (d) $T=500$; (e) $T=1,000$; (f) $T=2,000$

$N = 2$) to be able to distinguish the valid from the invalid methods. While considerable size distortions are clearly present for small T here, they are not really different for the sieve bootstrap and the block bootstrap (which uses a too small block length). One needs to increase the sample size to $T = 250$ (which is not often considered in panel data simulation studies) to really be able to distinguish the invalid sieve bootstrap from the valid block bootstrap. Therefore, if one has no knowledge of Figure 1 and only considers sample sizes commonly considered in Monte Carlo studies for panel data, one might easily draw wrong conclusions about bootstrap validity from these results.

V. Conclusion

In this article, we have investigated the validity of the univariate autoregressive sieve bootstrap applied to a non-stationary multivariate system that allowed for general forms of cross-sectional dependence, including but not restricted to cointegration. While it was shown to be possible to write this system as a collection of infinite order univariate AR models, the innovations of these equations were shown not to be vector white noise, which causes the AR sieve bootstrap to be invalid in such a general system.

This result was illustrated with a numerical example using a simple multivariate system in which the AR sieve bootstrap was invalid. It was shown that the extent of invalidity depends on the value of the parameters and the cross-sectional dimension. It was also found that in small or moderate samples the invalid AR sieve bootstrap is hard to distinguish from the valid moving blocks bootstrap, and only in larger samples the asymptotic patterns could be recovered. Also, the extent of the invalidity depends on specific parameter combinations. This can explain the observation that the AR sieve bootstrap performs well in simulation studies performed in the literature.

The results of this paper have important implications for practitioners applying the autoregressive sieve bootstrap in a panel data setup, as it will typically be invalid for many settings it is applied to, unless cross-sectional dependence in the true DGP is only of a contemporaneous nature. Moreover, the numerical example serves as a warning that relying upon Monte Carlo simulations to assess the validity of the AR sieve bootstrap can be very misleading. If one is not certain about the type of cross-sectional dependence present in the DGP, it is safer to use a different bootstrap method that is valid for a wider set of DGPs, such as the block bootstrap.

Final Manuscript Received: October 2012

References

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*, 2nd edn, Springer-Verlag, New York.

Cerrato, M. and Sarantis, N. (2007). 'A bootstrap panel unit root test under cross-sectional dependence, with an application to PPP', *Computational Statistics & Data Analysis*, Vol. 51, pp. 4028–4037.

Chang, Y. (2004). 'Bootstrap unit root tests in panels with cross-sectional dependency', *Journal of Econometrics*, Vol. 120, pp. 263–293.

Chang, Y., and Park, J. Y. (2002). 'On the asymptotics of ADF tests for unit roots', *Econometric Reviews*, Vol. 21, pp. 431–447.

Chang, Y., Park, J. Y. and Song, K. (2006). 'Bootstrapping cointegrating regressions', *Journal of Econometrics*, Vol. 133, pp. 703–739.
Cubadda, G., Hecq, A. and Palm, F. C. (2009). 'Studying co-movements in large multivariate data prior to multivariate modelling', *Journal of Econometrics*, Vol. 148, pp. 25–35.
Di Iorio, F. and Fachin, S. (2011). A Sieve Bootstrap Range Test for Poolability in Dependent Cointegrated Panels, DSS Empirical Economics and Econometrics Working Papers Series 2011/2, Sapienza University of Rome.
Fuertes, A.-M. (2008). 'Sieve bootstrap t-tests on long-run average parameters', *Computational Statistics & Data Analysis*, Vol. 52, pp. 3354–3370.
Fuller, W. A.(1996). *Introduction to Statistical Time Series*, 2nd edn, Wiley, New York.
Hanck, C. (2009). 'For which countries did PPP hold? A multiple testing approach', *Empirical Economics*, Vol. 37, pp. 93–103.
Hannan, E. J. and Kavalieris, L. (1986). 'Regression, autoregression models', *Journal of Time Series Analysis*, Vol. 7, pp. 27–49.
Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.
Kreiss, J.-P., Paparoditis, E. and Politis, D. N. (2011). 'On the range of validity of the autoregressive sieve bootstrap', *Annals of Statistics*, Vol. 39, pp. 2103–2130.
Palm, F. C. (1977). 'On univariate time series methods and simultaneous equation econometric models', *Journal of Econometrics*, Vol. 5, pp. 379–388.
Palm, F. C., Smeekes, S. and Urbain, J.-P. (2008). 'Bootstrap unit root tests: comparison and extensions', *Journal of Time Series Analysis*, Vol. 29, pp. 371–401.
Palm, F. C., Smeekes, S. and Urbain, J.-P. (2010). 'A sieve bootstrap test for cointegration in a conditional error correction model', *Econometric Theory*, Vol. 26, pp. 647–681.
Palm, F. C., Smeekes, S. and Urbain, J.-P. (2011). 'Cross-sectional dependence robust block bootstrap panel unit root tests', *Journal of Econometrics*, Vol. 163, pp. 85–104.
Park, J. Y. (2002). 'An invariance principle for sieve bootstrap in time series', *Econometric Theory*, Vol. 18, pp. 469–490.
Pesaran, M. H., Smith, R. P., Yamagata, T. and Hvozdyk, L. (2006). *Pairwise Tests of Purchasing Power Parity Using Aggregate and Disaggregate Price Measures*, Cambridge Working Papers in Economics 0634, University of Cambridge.
Smeekes, S. (2012). 'Detrending bootstrap unit root tests', *Econometric Reviews*, forthcoming.
Smith, L. V., Leybourne, S., Kim, T.-H. and Newbold, P. (2004). 'More powerful panel data unit root tests with an application to mean reversion in real exchange rates', *Journal of Applied Econometrics*, Vol. 19, pp. 147–170.
Teräsvirta, T. (1977). 'The invertibility of sums of discrete MA and ARMA processes', *Scandinavian Journal of Statistics*, Vol. 4, pp. 165–170.
Trapani, L. (2011). 'On Bootstrapping Panel Factor Series' CEA@Cass Working Paper Series WP-CEA-04–2011, Cass Business School.
Trapani, L. and Urga, G. (2010). 'Micro versus macro cointegration in heterogeneous panels', *Journal of Econometrics*, Vol. 155, pp. 1–18.
Westerlund, J. and Edgerton, D. L. (2007). 'A panel bootstrap cointegration test', *Economics Letters*, Vol. 97, pp. 185–190.
Zellner, A. and Palm, F. (1974). 'Time series analysis and simultaneous equation econometric models', *Journal of Econometrics*, Vol. 2, pp. 17–54.

Appendix A

Proof of Theorem 1

Proof. Let $u_{i,t} = a(L)\Delta y_{i,t}$ as defined in equation (6). Then the first step is to show that

$$u_{i,t} = c_i(L)e_{i,t}, \tag{9}$$

where $c_i(z) \neq 0$ for all $z \in \mathbb{C}$ such that $|z| \leq 1$ and $e_{i,t}$ is white noise. This is an extension of theorem 1 of Teräsvirta (1977). As $u_{i,t}$ is a covariance stationary process, the Wold representation applies (see e.g. Brockwell and Davis, 1991, section 5.7). Furthermore, as the Wold representation is fundamental, $c_i(z) \neq 0$ for all $z \in \mathbb{C}$ such that $|z| < 1$. We still need to ensure that $c_i(z) \neq 0$ for all $z \in \mathbb{C}$ such that $|z| = 1$, which is true if the spectral density of $u_{i,t}$ is strictly positive on $[-\pi, \pi]$. As shown in Teräsvirta (1977), the spectral density of $u_{i,t}$ is equal to

$$f_i(\lambda) = \frac{1}{2\pi} B_i(e^{-\mathrm{i}\lambda}) B_i(e^{\mathrm{i}\lambda})',$$

which is zero at some point $\lambda_{i,0} \in [-\pi, \pi]$ if and only if all polynomials $B_{i,j}(z)$ for $j = 1, \ldots, N$ have a *common* root on the unit circle at $e^{-\mathrm{i}\lambda_{i,0}}$.[5]

As $B(z)$ does not contain any factors $(1-z)$, a common root cannot arise at frequency $\lambda = 0$. Suppose that for unit i there is a common unit root at frequency $\lambda_{i,0} \neq 0$. In that case, we can factor out $(e^{-\mathrm{i}\lambda_{i,0}} - z)$ from $B_i(z)$ and consequently we have $\det(B(z)) = (e^{-\mathrm{i}\lambda_{i,0}} - z)\tilde{b}(z)$ for some polynomial $\tilde{b}(z)$. However,

$$\det(B(z)) = \frac{\det(\tilde{\Phi}(z))}{(1-z)^{N(N-r-1)}} = \frac{[\det(\Phi(z))]^{N-1}}{(1-z)^{N(N-r-1)}}$$
$$= \frac{\left[(1-z)^{N-r} a(z)\right]^{N-1}}{(1-z)^{N(N-r-1)}} = (1-z)^r a(z)^{N-1},$$

where $a(z)$ does not contain any unit roots as

$$a(z) = \frac{\det(\Phi(z))}{(1-z)^{N-r}} = \frac{\det(P(z))}{\left(\det(\Psi(z))(1-z)^{N-r}\right)} = \frac{\det(P^*(z))}{\det(\Psi(z))}, \tag{10}$$

which cannot contain any unit roots by the assumptions on $\Psi(z)$ and the fact that there are only $N-r$ unit roots in $\det(P(z))$ and consequently $\det(P^*(z)) = \det(P(z))/(1-z)^{N-r}$ does not contain any unit roots.

As furthermore the unit roots $(1-z)^r$ cannot be the common roots as they occur at frequency zero, we can conclude that for each $i = 1, \ldots, N$ there exist no common roots on the unit circle. Therefore the spectral density of $u_{i,t}$ is strictly positive for any $i = 1, \ldots, N$, which proves equation (9).

We may then define $d_i(z) = \sum_{j=0}^{\infty} d_{i,j} z^j = a(z) c_i(z)^{-1}$, and write $d_i(L)\Delta y_{i,t} = e_{i,t}$ for $i = 1, \ldots, N$. Invertibility of $d_i(z)$ then immediately follows from equation (10) and the fact that $P(z)$ does not contain roots within the unit circle.

We next show that the summability condition holds. As $\Delta y_t = C\varepsilon_t + C^*(L)\Delta\varepsilon_t$ and $\sum_{j=0}^{\infty} j\|C_j\| < \infty$ from equation (4), it follows that $\sum_{h=-\infty}^{\infty} h|\mathrm{E}\Delta y_{i,t}\Delta y_{i,t+h}| < \infty$ for any $i = 1, \ldots, N$ (see e.g. Fuller, 1996, p. 367). It then follows from lemma 2.1 of Kreiss *et al.* (2011) that $\sum_{j=0}^{\infty} j|d_{i,j}| < \infty$.

Finally, it follows from the Wold representation theorem that $e_{i,t}$ is white noise. From Assumption 2 and equation (4) one can conclude that $\mathrm{E}(\Delta y_{i,t}^4) < \infty$, and consequently that $\mathrm{E}(e_{i,t}^4) < \infty$. This concludes the proof.

[5]Teräsvirta (1977) considers MA models of finite order but his arguments are easily extended to infinite order MA models.