Downloaded from UvA-DARE, the institutional repository of the University of Amsterdam (UvA)
http://hdl.handle.net/11245/2.140

File ID uvapub:140
Filename 0215.pdf
Version unknown

SOURCE (OR PART OF THE FOLLOWING SOURCE):
Type report
Title Bootstrapping and Bartlett corrections in the cointegrated VAR model
Author(s) Pieter Omtzigt, Stefano Fachin
Faculty UvA: Universiteitsbibliotheek
Year 2003

FULL BIBLIOGRAPHIC DETAILS:
http://hdl.handle.net/11245/1.418491

Copyright

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content licence (like Creative Commons).

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
(pagedate: 2014-11-26)

UvA ECONOMETRICS

Discussion Paper: 2002/15

Bootstrapping and Bartlett corrections in the cointegrated VAR model

Pieter Omtzigt and Stefano Fachin

www.fee.uva.nl/ke/UvA-Econometrics

Department of Quantitative Economics
Faculty of Economics and Econometrics
Universiteit van Amsterdam
Roetersstraat 11
1018 WB AMSTERDAM
The Netherlands

UvA UNIVERSITEIT VAN AMSTERDAM

Bootstrapping and Bartlett corrections in the cointegrated VAR model*

Pieter Omtzigt †and Stefano Fachin ‡

January 2, 2003

Abstract

The small sample properties of tests on long-run coefficients in cointegrated systems are still a matter of concern to applied econometricians. We compare the performance of the Bartlett correction, the bootstrap and the fast double bootstrap for tests on ccointegration parameters in the maximum likelihood framework. We show by means of a theoretical result and simulations that all three procedures should be based on the unrestricted estimate of the cointegration vectors. The fast double bootstrap delivers superior size correction, whereas the Bartlett correction leads to the least loss of power. However all three perform much better than the asymptotic tests and difference between them are small.

1 Introduction

The small sample properties of tests on long-run coefficients in cointegrated systems are still a matter of concern to applied econometricians. Since the asymptotic procedures proposed by Johansen (1991) have been shown to suffer from severe size distortion (among others, see Gonzalo, 1994; Bewley et al., 1994; Li and Maddala, 1997) two natural and complementary solutions have been proposed: (i) applying Bartlett corrections to the test statistics, in the hope that the corrected statistic will follow a small sample distribution closer to the asymptotic one, and thus bring actual sizes closer to the nominal sizes (Johansen, 2000); (ii), trying to estimate the actual small sample distribution by the bootstrap,

*We are grateful to Søren Johansen for his encouragement and suggestions. We also benefited from discussions with Tor Jacobson, Bent Nielsen and Paolo Paruolo. All simulations have been run on the UNIX cluster of the CASPUR computing laboratory; we would like to thank the staff of laboratory for their efficient technical support. The usual disclaimers apply. Financial support of MURST is gratefully acknowledged.

†Universiteit van Amsterdam, Roetersstraat 11, 1018WB Amsterdam Email: P.H.Omtzigt@uva.nl

‡Università la Sapienza, DCNAPS - Facoltà di Scienze Statistiche, P.le A. Moro 5, 00185 Roma, Italy. Email: s.fachin@caspur.it

a computer-intensive technique strictly linked with the Edgeworth expansion and indeed defined by Cribari-Neto and Cordeiro (1996) 'a simulation based alternative to Bartlett and Bartlett-type corrections' (Li and Maddala, 1996, 1997; Fachin, 2000; Gredenhoff and Jacobson, 2001).

For the time being, no definite solution has however appeared. Although the only aim of both the Bootstrap and the Bartlett correction is to get the actual size closer to the nominal size, the final aim of any testing procedure must be that of distinguishing between valid and invalid hypotheses: the proportion of Type II errors of corrected tests is therefore crucial. To the best of our knowledge no evidence on the power properties of Bartlett corrected tests in the cointegrated VAR model has appeared in the literature; the only available evidence on power for bootstrapped test statistics is in Fachin (2000) and shows that the type of bootstrap test examined may have a rather high Type II error. The aim of this paper is thus examining both the size and power properties of Bartlett-corrected and bootstrap tests. With respect to the latter, we also evaluate the feasible double bootstrap, recently proposed by Davidson and MacKinnon (2000). In either cases, a key result of the paper is that the Bartlett correction and the bootstrap tests should both be based on the *unrestricted* estimate of the cointegration vectors.

The chapter is organised as follows: in section 2 we shall briefly review the model, the structure of Bartlett-corrected and bootstrap tests, as well as a theoretical result, motivating us to base both procedures on unrestricted estimates. In section 3 we shall discuss the design of the Monte Carlo experiment and in section 4 present the results of the simulations.

Some conclusions, as well as tentative recommendations for applied work, are finally drawn in section 5.

2 Bartlett-corrected and Bootstrap Tests on Cointegrating Coefficients

2.1 The model

The cointegrated p-dimensional VAR model with k lags in its autoregressive form is defined as:

$$\Delta X_t = \alpha\beta' \begin{pmatrix} X_{t-1} \\ D_t \end{pmatrix} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Psi d_t + \varepsilon_t \tag{1}$$

In this paper a linear trend is constrained to lie in the cointegration space and an unrestricted constant is included outside that space: $D_t = t$ and $d_t = 1$. We define γ and ρ by $\beta' = (\gamma', \rho')$, where γ includes the coefficients linking the stochastic variables of the system and ρ are the coefficients of the deterministic part.

Three assumptions are made to make sure this is a stable I(1) model:

Assumption (Rank) α and γ are two full rank matrices of dimension $p \times r$, $p > r$;

Assumption (No I(2)) The matrix $\alpha'_\perp \left(I - \sum_{i=1}^{k-1} \Gamma_i\right) \gamma_\perp$ is of full rank;

Assumption (No other roots) The roots z of the characteristic polynomial are either 1: $z = 1$ ($p - r$ roots are equal to unity) or larger than 1 in absolute value: $|z| > 1$.

The first two assumptions assure that the process is an I(1) process and not integrated of lower or higher order, while the third assumption excludes explosive behaviour and seasonal unit roots

The stationary, stochastic part of (1) can be written in a companion form[1]:

$$\begin{bmatrix} \gamma' X_t \\ \Delta X_t \\ \Delta X_{t-1} \\ \vdots \\ \Delta X_{t-k+2} \end{bmatrix} = \begin{bmatrix} I_r + \gamma'\alpha & \gamma'\Gamma_1 & \cdots & \gamma'\Gamma_{k-2} & \gamma'\Gamma_{k-1} \\ \alpha & \Gamma_1 & \cdots & \Gamma_{k-2} & \Gamma_{k-1} \\ 0 & I & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & I & 0 \end{bmatrix} \begin{bmatrix} \gamma' X_{t-1} \\ \Delta X_{t-1} \\ \Delta X_{t-2} \\ \vdots \\ \Delta X_{t-k+1} \end{bmatrix} + \begin{bmatrix} \gamma' \\ I \\ 0 \\ \vdots \\ 0 \end{bmatrix} \varepsilon_t$$

or

$$Y_t = PY_{t-1} + F\varepsilon_t \tag{2}$$

The Bartlett correction, which shall be discussed in the next section, depends crucially on the matrix P.

2.2 The Bartlett correction

The idea behind the Bartlett correction Bartlett (1937) is both simple and appealing. Suppose the aim is testing the following null hypothesis on the parameters Θ, $\mathcal{H}_0 : \Theta_0 \subset \Theta$. In regular cases, the LR test statistic s has an expected value of

$$\begin{aligned} E_{\hat{\theta}_0}\left[-2\ln(LR\left(\Theta_0|\Theta\right))\right] &= E_{\hat{\theta}_0}\left[l_{\hat{\theta}} - l_{\hat{\theta}_0}\right] = \\ & h\left(1 + \frac{1}{T} g\left(\theta_0\right)\right) + O\left(\frac{1}{T^2}\right) \end{aligned} \tag{3}$$

where h denotes the number of restrictions tested. Then dividing the test statistic S by $\left(1 + \frac{1}{T} g\left(\theta_0\right)\right)$ we may obtain the modified test statistic S_B and expect the resulting distribution to be closer to a χ^2 distribution. This division is called a Bartlett correction and $\frac{1}{T} g\left(\theta_0\right)$ will be referred to as the Bartlett factor.

We obviously do not know θ_0, the true value of the parameters, θ, and thus we substitute a consistent estimate of θ, $\tilde{\theta}$, in expression (3) and thus get the Bartlett factor $g\left(\tilde{\theta}\right)$.

The arguments in the following pages will revolve around which consistent estimate should substitute θ_0: $\hat{\theta}_0$ the maximum likelihood estimate under the null hypothesis or $\hat{\theta}$ the unconstrained maximum likelihood estimate. We shall argue that we need to substitute $\hat{\theta}$ and not $\hat{\theta}_0$ in the problem at hand. Even though the size correction works better

[1] The deterministic part can be taken account of by adding an extra term in d_t and D_t.

with $\hat{\theta}_0$ which is more efficient under the null, we find the power of the Bartlett corrected test-statistic with $\hat{\theta}$ extremely poor. We demonstrate this both by means of a theory and simulations in section 3. To see the differences in practise between using these two estimates, we refer to page 19, where in figure 3 we have plotted power curves for both estimates. (the DGP-value is 1 and the curves are drawn for the 5% significance level).

The problems of the Bartlett section and their solutions, carry over to the bootstrap section as well.

Lawley (1956) and Barndorff-Nielsen and Hall (1988) proved that under certain regularity conditions (which exclude cointegrated VAR models and thus the problem at hand) for any real number x

$$p\left(S_B \leq x\right) = p\left(\chi^2(h) \leq x\right) + O\left(\frac{1}{T^2}\right) \tag{4}$$

So the whole χ^2 distribution is better approximated after the correction.

Jensen and Wood (1997) showed that for the Dickey Fuller distribution (4) does not hold. This however does not mean that the size correction is not useful in practice. In fact Nielsen (1997) showed that a Bartlett correction in an AR(1) process with a unit root, does provide an improvement to the size of the test.

Under the assumption:

Assumption (Deterministics) there exist matrices K and M such that $d_t = Md_{t-1}$ and $\Delta D_t = Kd_t$ where all the eigenvalues of the matrix M equal 1 in absolute value

Johansen (2000) derived the Bartlett correction for three different kind of hypotheses on β in (1), namely:

1. $\beta = \beta^0$, a simple hypothesis on all the cointegration vectors;
2. $\beta_1 = \beta_1^0$ where β_1^0 are the first r_1 relations ($1 \leq r_1 < r$) and the other cointegration relations are unrestricted;
3. $\gamma = H\varphi$ where H is a $(p \times r)$ matrix of full rank and $s < r$. This hypothesis implies the same restriction on all relations in γ.

Corrections for other kinds of hypotheses, like restrictions of the kind $\beta_1 = H_1\varphi_1$ do not yet exist.

We therefore limit ourselves to confronting the corrections 1 and 2 with the bootstrap in this paper and do not put any dummy variables in our DGP.

The correction term itself, for which we refer to the aforementioned article, depends crucially on the total number of parameters, the variance of Y_t in (2) and a number of times on $\sum_{i=0}^{\infty} P^i$. We do not have the true value of the parameters, so we substitute estimates. Now under the null the matrix P only contains eigenvalues strictly smaller than unity in absolute value as Y_t in (2) is a stationary process. Yet under the alternative, the restricted estimate will contain at least one additional unit root, becasuse one of the relations $\gamma' X_t$ is no longer stationary. Consequently the Bartlett correction is no longer defined as the sum $\sum_{i=0}^{\infty} P^i$ diverges. We prove this fact in the following theorem:

Theorem 1 *Under the false null hypothesis $\beta = b = (b_1, b_2)$ where $b_1 \in sp(\beta^0), b_2 \notin sp(\beta^0)$ (β^0 is the true value of β), $\hat{\alpha}$ the restricted maximum likelihood estimate of α, will have reduced rank $s < r$ in the limit. Consequently the matrix P contains additional unit roots in the limit*

Proof. Partition the maximum likelihood estimate of α, $\hat{\alpha} = (\hat{\alpha}_1, \hat{\alpha}_2)$ conformably with b. It is found by ordinary least squares:

$$\hat{\alpha} = S_{01}(b_1, b_2) \begin{pmatrix} b_1' S_{11} b_1 & b_1' S_{11} b_2 \\ b_2' S_{11} b_1 & b_2' S_{11} b_2 \end{pmatrix}^{-1}$$

where S_{01} and S_{11} are defined in standard fashion (see Johansen 1995, page 90-91). From Chan and Wei (1988) we find that $S_{01}b_1, b_1'S_{11}b_1 \in O(1)$, $S_{01}b_2, b_2'S_{11}b_1 \in O(T)$ and $b_2'S_{11}b_2 \in O(T^2)$. Using standard inverse matrix formula, we find that
$\hat{\alpha}_2 = \left(S_{01}b_2 - S_{01}b_1 (b_1'S_{11}b_1)^{-1} b_1'S_{11}b_2\right) \times \left((b_2'S_{11}b_2)^{-1} - (b_2'S_{11}b_1)(b_1'S_{11}b_1)^{-1}(b_1'S_{11}b_2)\right)^{-1} \in O(T^{-1})$ which implies $\hat{\alpha}_2 \xrightarrow{P} 0$ ■

This means that if we use the restricted estimate $\hat{\theta}_0$ and the null hypothesis is false, the Bartlett correction is not defined. We shall see that in practice the absolute value the roots is underestimated, such that the additional root is estimated to be close to 1. This means that the estimated Bartlett factor can be calculated, but becomes extremely large and the null hypothesis is then easily accepted.

We thus seek an estimator which

- Whenever the null hypothesis is true is consistent.
- Whenever the null hypothesis is false, the matrix P should have stable roots, such that the Bartlett correction is defined. If possible these roots should in some sense be as stable as possible, for when they are very close to unity, the Bartlett factor explodes and a false hypothesis is accepted.

Solution 2 *Use the unrestricted estimates $\hat{\beta}$ of β in (1) and not the restricted estimates in the Bartlett correction factor. The Bartlett correction factor for $\beta = \beta^0$ and $\gamma = H\varphi$ only depends on $\hat{\beta}$, such that this defines the solution in these cases.*
In case 2 ($\beta_1 = \beta_1^0$, only some of the cointegration relations are restricted) we need estimates for both the restricted and the unrestricted vectors. In this case β_1^0 and the associated restricted estimate $\beta_2(\beta_1^0)$ should not be used, as this will lead to instability of the P matrix when $\mathcal{H}_0$ is false. Instead, we find a matrix b_1 for which $sp(b_1) \subset sp(\hat{\beta})$ and as close to β_1^0 as possible. This means that we find a matrix ξ such that

$$\xi = \left(\hat{\beta}'\hat{\beta}\right)^{-1} \hat{\beta}'\beta_1^0 \tag{5}$$

Then the estimators $b_1 = \hat{\beta}\xi$ and $b_2 = \hat{\beta}\xi_\perp$ are consistent when the null hypothesis is true and the companion matrix P is stable when it is false.

2.3 Bootstrap methods

In principle, the great advantage of the bootstrap[2] is that it can offer immediate solutions to new problems. However, in practice its ability to deliver good alternatives when reliable small sample parametric procedures are lacking must be accurately tested before its use may be recommended. This is especially true for the problem we are trying to solve, as the asymptotics of the bootstrap applied to integrated data is still largely unexplored: Horowitz (2002) summarises his survey stating that 'at present (...) there are no theoretical results on the ability of the bootstrap to provide asymptotic refinements for tests or confidence intervals when the data are integrated or cointegrated'. Recent developments in this direction covering specific cases are Chang et al. (2001), Davidson (2001), Paparoditis and Polititis (2001) and Inoue and Kilian (2002). At the opposite, a striking example of how blind implementations of the bootstrap can deliver entirely wrong results is given by Phillips (2001) for the case of spurious regression with integrated variables.

The general idea underlying bootstrap tests is to assess the value of the test statistic s obtained from the empirical analysis on the basis of the distribution of a large number of statistics s^* computed from suitably constructed pseudodata, with the null hypothesis of the former consistent with the data generating process (DGP) of the latter. To this end, $\mathcal{H}_0$ may be imposed when generating the pseudodata (as in some examples in Efron and Tibshirani, 1993), or, vice versa, the chosen DGP taken as the null hypothesis (as recommended by Hall, 1992). In both cases, $\mathcal{H}_0$ is true for the pseudodata, and thus, assuming for simplicity a one-sided test, the proportion of s^* more extreme than s in the relevant direction is a natural estimate of the p-value of the test.

With cointegrated VARs and some hypothesis on the long-run coefficients $\mathcal{H}_0 : \beta = \beta^0$, the two approaches entail respectively:

(*a*) estimating a VAR constrained under $H_0 : \beta = \beta^0$, generating the pseudodata on the basis of the estimated *constrained* estimates $\hat{\theta}_0$ and a set of random noises (we will discuss the choice of these below), and testing $\mathcal{H}_0 : \beta = \beta^0$ both on the original data and on the pseudodata;

(*b*) estimating an unconstrained VAR, generating the pseudodata on the basis of the estimated *unconstrained* estimatres $\hat{\theta}$ and a set of random noises, testing $\mathcal{H}_0 : \beta = \beta^0$ on the original data and $\mathcal{H}_0^* : \beta = \widehat{\beta}$ (where $\widehat{\beta}$ are the unconstrained estimates of β) on the pseudodata.

So far, approach (*a*) has been favoured with no exception in the applications of interest here. However, a point of crucial importance for testing in the maximum likelihood estimation of cointegrated VARs seems to have gone unnoticed: although both approaches are valid and asymptotically equivalent under H_0, this is not true any more when it is false. To see this, consider the case of a test $\mathcal{H}_0 : \beta = \beta^0$ in a model without lags and

[2]General introductions to the bootstrap are provided, *inter alia*, by Efron and Tibshirani (1993), Hall (1995) and Horowitz (2002), while a recent review especially addressed at time series applications is Berkowitz and Kilian (2000).

just one cointegration vector. If this vector is misspecified, then $\beta^{0\prime}X_{t-1}$ is clearly an $I(1)$ process, whereas ΔX_t is $I(0)$. The only congruent values for the loading factors α are therefore zero. Hence all the element of the matrix $\hat{\Pi} = \hat{\alpha}\beta^{0\prime}$ equal zero (asymptotically) and the rank of such a matrix is 0 not 1. If one were to use this matrix for the Bootstrap DGP, one would generate just random walks without any cointegration (this is essentially a different version of exactly the same issue already discussed in the previous subsection with respect to the computation of the Bartlett factor when $\mathcal{H}_0$ is false). Thus, we will consider bootstrap tests of type (*b*).

With respect to the noise, there are again essentially two alternatives: either generating it under some parametric hypothesis (typically, $MIIDN$) or by resampling from the set of residuals of a VAR. In the latter case the natural choice are the residuals of the unconstrained VAR, empirically $MIIDN$. Gredenhoff and Jacobson (2001) favoured the parametric option, while Fachin (2000) and Li and Maddala (1997) the non-parametric one[3]. Here we will consider both alternatives. Block-resampling methods, such as the "Continuous-Path Block Bootstrap" proposed by Paparoditis and Polititis (2001), which may be potentially powerful in dealing with the stochastic trends present in the system, will the subject of future research.

Defining Θ the entire parameter set of the VAR and assuming we are interested in the test $\mathcal{H}_0 : \beta = \beta^0$ we thus implement the following bootstrap procedure, which is graphically represented in figure 1

- *Bootstrap test*

1. Estimate VAR on data X; for given cointegrating rank obtain unrestricted estimates $\hat{\theta}$, unrestricted residuals $\widehat{\varepsilon}$, restricted estimates $\hat{\theta}_0$, restricted residuals $\widehat{\varepsilon}_0$ and test statistic s for the hypothesis $H_0 : \beta = \beta^0$;

2. Construct pseudodata: $X^* = \phi(\hat{\theta}, \varepsilon^*), \varepsilon^*$ drawn at random with replacement from $\widehat{\varepsilon}$ or NID.

3. Estimate VAR on pseudodata X^*; obtain $\hat{\theta}^*, \widehat{\varepsilon}^*$, $\hat{\theta}_0^*, \widehat{\varepsilon}_0^*$ and test statistic s^* for the hypothesis H_0^*: $\beta = \hat{\beta}$;

Repeat (2)-(3) a large number of times

(4) Compute bootstrap *p*-value: $p^* = prop(s^* > s)$.

The test statistic is the likelihood ratio test (which is the only one allowing a Bartlett correction).

[3]Note that there is a possible source of confusion here, as the terms 'parametric' and 'non-parametric' have been used in the bootstrap literature with different meanings. We define procedures based on resampling from estimated residuals as 'non parametric', and that involving drawings from a theoretical distribution as 'parametric'.

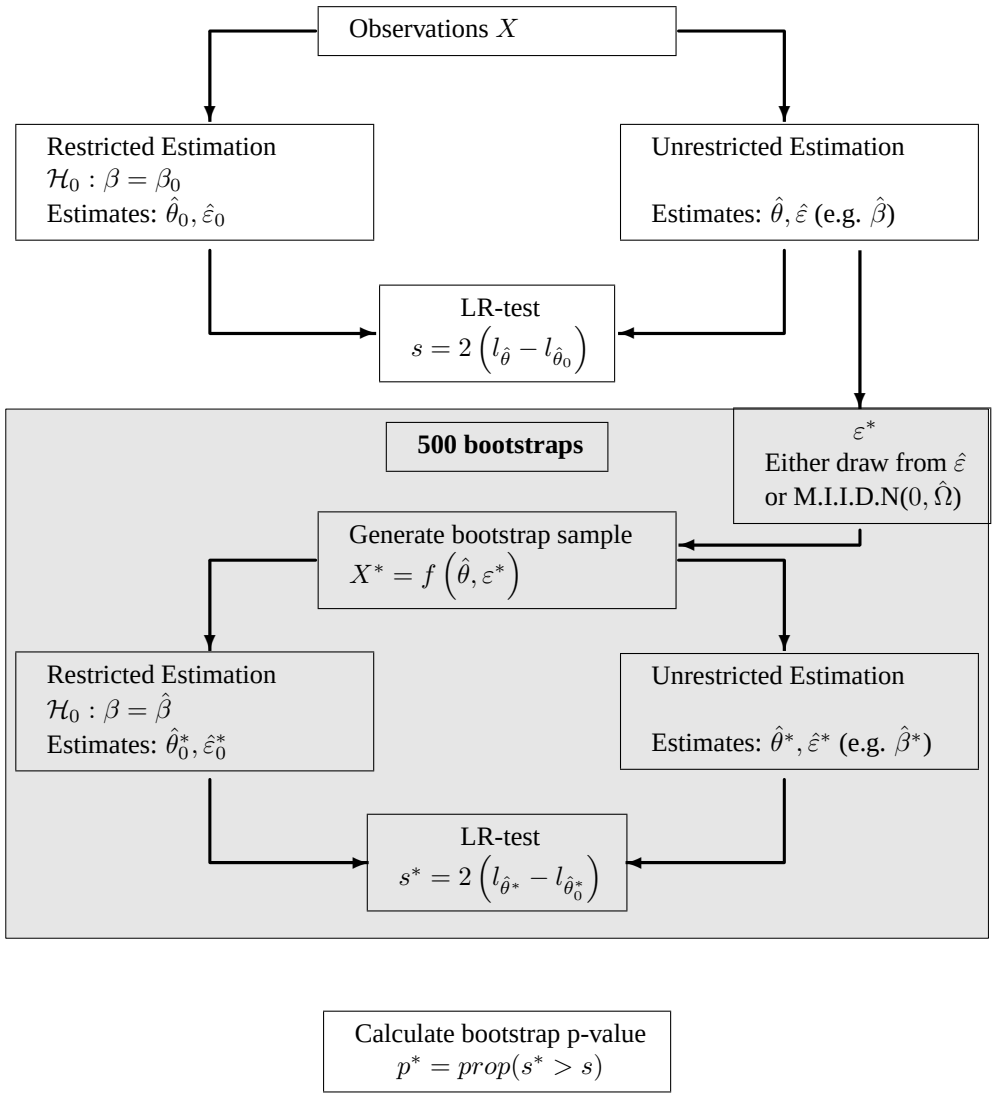


Figure 1: Bootstrap procedure for tests on the cointegration parameters

If we have a simple hypothesis on only part of the cointegration space, $\beta_1 = \beta_1^0$, we take the following null hypothesis in step 3:

$$\beta_1 = \hat{\beta}'(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'\beta_1^0 \tag{6}$$

which is easily seen to converge to β_1^0 if H_0 is true.

As mentioned in the introduction, Davidson and MacKinnon (2000) recently put forth a computationally cheap double bootstrap procedure which may deliver results superior to the standard bootstrap just outlined[4]. The idea behind the double bootstrap, proposed by Beran (1988) is that of correcting the possible bias in the bootstrap procedure implemented by a second application of the bootstrap. For instance, in the case of a test the aim of the second-level application of the bootstrap would be to estimate, and thus correct for, the bias $(p_i - i)$, where p_i is the p-value of the i-level bootstrap test. Although the principle is certainly attractive, it is also very expensive, as it involves the construction of a bootstrap pseudo-population for each bootstrap redraw. It is thus practically impossible to evaluate by means of Monte Carlo experiments with the currently available computing power. On the contrary, in Davidson and MacKinnon's method there is only one second level bootstrap redraw for each first level one, so that the computing time is of the same order of magnitude of the standard bootstrap. Monte Carlo experiments are thus feasible. Going into the details of the method is clearly beyond the scope of this paper. However, the basic intuition is very simple: if the bootstrap estimate $p^* = prop(s^* > s)$ of true p-value of the test is distorted, we may get a better estimate by replacing s with some $\widetilde{s}$ chosen so to counterbalance the distortion. Now, s is by definition the $p^* - th$ quantile of the distribution of the s^*; hence, an obvious candidate for $\widetilde{s}$ is the same quantile of the distribution of a *second-level* bootstrap distribution. If p^* is distorted downwards, such a quantile will tend to be larger than the true quantile s, and viceversa, thus delivering the desired effect.

The general structure of the fast double bootstrap test we shall implement is the following: (see figure 2 for a graphical representation)

- *Fast Double Bootstrap test*

1. Estimate VAR on data X; for given cointegrating rank obtain estimates $\hat{\theta}, \widehat{\varepsilon}, \hat{\theta}_0, \widehat{\varepsilon}_0$ and test statistic s for the hypothesis $H_0 : \beta = \beta^0$;

2. Construct pseudodata: $X^* = \phi(\hat{\theta}, \varepsilon^*), \varepsilon^*$ drawn at random with replacement from $\widehat{\varepsilon}$ or NID;

3. Estimate VAR on pseudodata X^*; obtain $\hat{\theta}^*, \widehat{\varepsilon}^*$, $\hat{\theta}_0^*, \widehat{\varepsilon}_0^*$ and test statistic s^* for the hypothesis H_0^*: $\beta = \hat{\beta}$;

[4]Although Davidson and MacKinnon's analytical results are valid only for one-sided tests with asymptotic N(0,1) distributions, some simulation evidence suggests that the properties may extend to the asymptotic χ^2 of interest here.

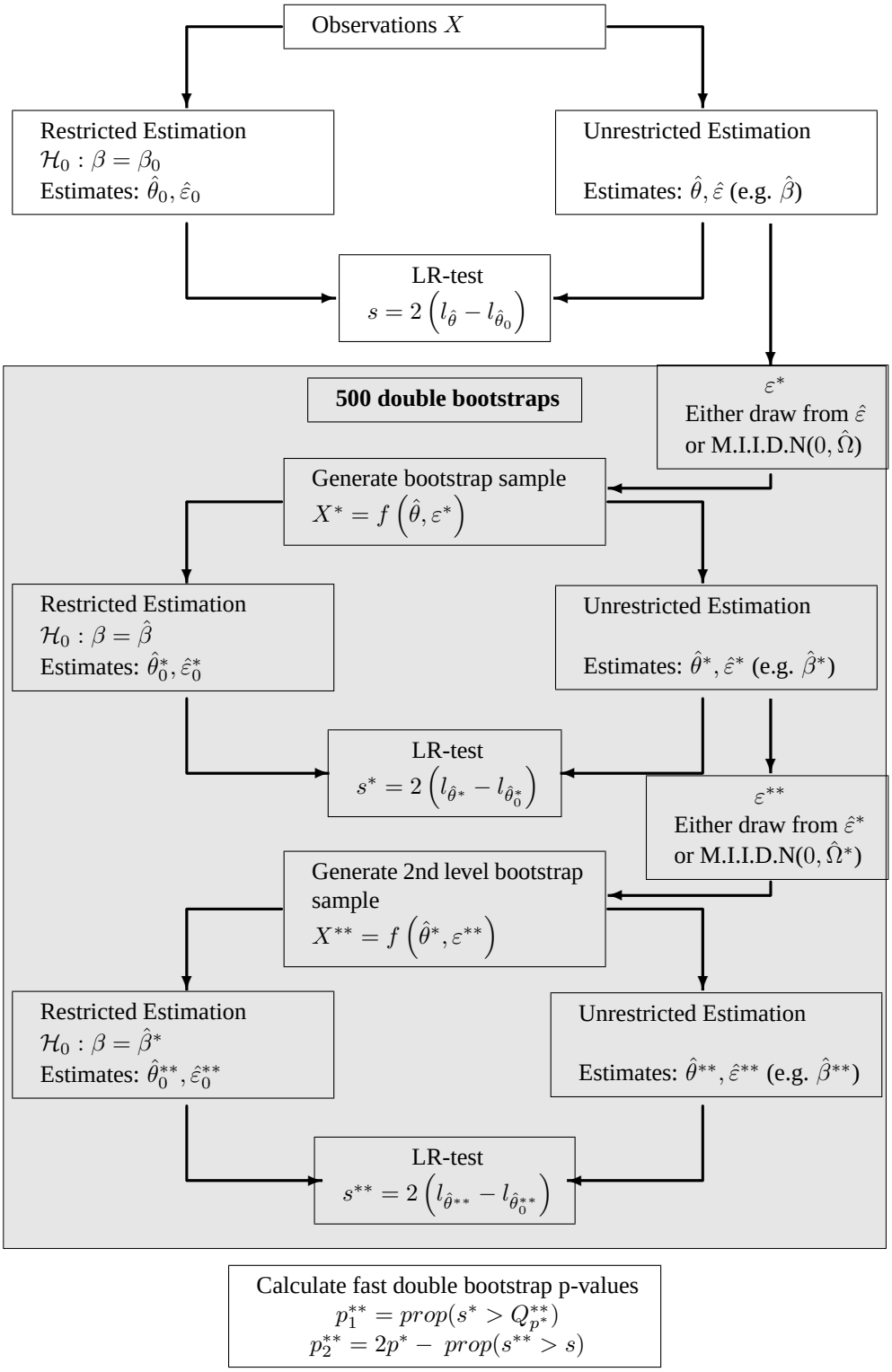


Figure 2: Fast Double Bootstrap procedure for tests on the cointegration parameters

4. Construct second-level pseudodata $X^{**} = \phi(\hat{\theta}^*, \varepsilon^{**}), \varepsilon^{**}$ drawn at random with replacement from $\widehat{\varepsilon}^*$ or NID;

5. Estimate VAR on second-level pseudodata X^{**}; obtain $\hat{\theta}^{**}, \widehat{\varepsilon}^{**}$, $\hat{\theta}_0^{**}, \widehat{\varepsilon}_0^{**}$ and test statistic s^{**} for the hypothesis H_0^{**}: $\beta = \hat{\beta}^*$;

Repeat (2)-(5) a large number of times

(6) Compute bootstrap *p*-value: $p^* = prop(s^* > s)$.

(7) Compute fast double bootstrap p-value type 1: $p_1^{**} = prop(s^* > Q_{p^*}^{**})$,where $Q_{p^*}^{**}$ is the p^* quantile of the $s^{**'}$s.

A (costless) further step is advisable:

(8) Compute fast double bootstrap p-value type 2: $p_2^{**} = 2p^* - \ prop(s^{**} > s)$.

Again, the intuition here is that if for instance $p^* > p$, we can expect $prop(s^{**} > s) > p^*$, so that p_2^{**} will be closer to p than p^*. However, p_2^{**} may not be greater than $2p^*$ and it may be negative, two undesirable features that suggest limiting its use to a reliability check: if the difference between the two p-values is sizebale neither of them should be trusted.

3 Design of the Monte Carlo Experiment

On the basis of the simulation results reported by Gredenhoff and Jacobson (2001) and Fachin (2000), the key characteristics of the DGP to be controlled in the experiments are the dimension of the system, i.e. number of variables and lags, and its long-run structure, i.e. number of the cointegrating relationships and the speed at which the system adjusts to them. Estimation of systems of higher dimension (both in terms of number of variables and lags) demand more from the data, and thus it is (ex-post) not surprising to see that both the asymptotic test and the bootstrap test proposed by Gredenhoff and Jacobson (2001) perform better in smaller systems. A crucial remark here is that the simple bivariate DGPs employed in virtually all simulation studies do suffer from loss of generality, a fact not suspected so far. The experimental design adopted here will thus generalize to a multivariate system the classical DGP used by a number of studies starting with Engle and Granger (1987), which allows an easy control of the speed of adjustment. We shall consider systems including $p = 5$ random variables and with $r = 1$ or 2 cointegrating relationships. Let $x_t = [x_{1t} \dots x_{5t}]'$ be the column vector of the realizations of the random variables of interest at time $t = 1, \dots, T$, $u_t = [u_{1t} \dots u_{5t}]'$ the errors, $\epsilon_t = [\varepsilon_{1t} \dots \varepsilon_{5t}]'$ the noise, whose stochastic structure will be discussed in detail below, and t a time trend. Our DGP is then given by

$$\begin{bmatrix} \beta_1' \\ \vdots \\ \beta_5' \end{bmatrix} \begin{bmatrix} x_t \\ t \end{bmatrix} = u_t \tag{7}$$

$$\Phi u_t = \epsilon_t \tag{8}$$

with
$\Phi = diag(\phi)$, $\phi = \begin{bmatrix} \phi_1(L) & \phi_2(L) & \phi_3(L) & \phi_4(L) & \phi_5(L) \end{bmatrix}$.

Although the Bartlett corrections do depend on the parameters of the system, in order to keep the size of the experiment within manageable dimensions in the size simulations the cointegrating coefficients will be kept fixed across trials to either zero or 1, with the vectors resembling quite closely those used by Haug (1996), while in the power simulations we shall consider a few values in the range $[0.5, 1.5]$. Given that we are using a full-information method we do not need to worry about endogeneity; we shall thus consider a very simple structure, with one stochastic trend (X_p) transmitted to the first r variables of the system, while the remaining $p - r - 1$ follow independent random walks. The details in the two cases are as follows:

(*a*) $r = 1$

$\beta_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & \beta_{15} & 0.01 \end{bmatrix}$ is the cointegration vector.
All the other relations are non-stationary:
$\beta_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$; $\beta_3 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$;
$\beta_4 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$; $\beta_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$;
$\phi_1(L) = (1, \varphi_1 L, \ldots, \varphi_k L^k)$;
$\phi_2(L) = \phi_3(L) = \phi_4(L) = \phi_5(L) = (1, -L)$.

(*b*) $r = 2$

$\beta_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0.01 \end{bmatrix}$ becomes a cointegration vector.
All the other $\beta' s$ are as in case (a).
$\phi_1(L) = \phi_2(L) = (1, \varphi_1 L, \ldots, \varphi_k L^k)$;
$\phi_3(L) = \phi_4(L) = \phi_5(L) = (1, -L)$.

Some simple considerations will allow great simplification of the design as far as the ε's are concerned. First of all, in previous work on the related topic of stationary VARs Fachin and Bravetti (1996) found that the shape of the distribution of the shocks does not appear to have a significant impact on the performances of asymptotic procedures. Further, the expectation that with a full-information method, their covariance structure should not matter either has been confirmed in the case of a simple bivariate DGP by Fachin (2000). We shall thus assume $\varepsilon = [\varepsilon_1 \ldots \varepsilon_p] \sim MIIDN(0, I)$.

Finally, the number of both Monte Carlo replications and bootstrap redrawings has been fixed to 500: on the basis of previous work and some pilot experiments we concluded that the gain in precision deliver by higher numbers of either was not worth the higher computing costs and longer calendar time required. At 0.05 the Monte Carlo standard error will thus be about 0.010.

In table 1 we give an overview of the parameter values in the various experiments. In the benchmark case we have 1 cointegration vector and test that the cointegration parameters is known. Furthermore the model has (case (a)), 2 lags, and fairly slow adjustment $(\varphi_1 = \varphi_2 = -0.35)$ We have 100 observations $(T = 100)$, and for the bootstrap algorithm, we resample with replacement from the estimated errors $\hat{\varepsilon}$ (and in the second level bootstrap from $\hat{\varepsilon}^*$).

Furthermore $\beta_{15} = 1$ and we thus test $\mathcal{H}_0 : \beta_1^0 = \beta_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0.01 \end{bmatrix}$. For the benchmark case and each of the other cases we also execute a power experiment in which we set $\beta_{15} = 0.5$ and test $\beta_1^0 = \beta_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0.01 \end{bmatrix}$.

The complexity of the DGP is such that we are unable to execute a full factorial designo over all variations, we consider relevant. We thus provide

Each time we only deviate in one respect from our benchmark DGP, which is the first experiment. In the second we test case (b), that is two cointegration vectors and test that either one or both vectors are known. Next we increase the sample size to find out whether the corrections are working with 400 observations: we do not regularly find such large samples in time series analysis, but find that even with that many observations, the asymptotic tests do not work well. In the fourth experiment we increase the lag length of the VAR to four. The sum of the adjustment coefficients is kept constant at 0.7. Subsequently we test the effect of an increase in the speed of adjustment to equilibrium. This increases the signal to noise ratio and the performance of the asymptotic test (remember that the Bartlett correction depends on $\sum_{i=0}^{\infty} P^i$). In the last experiment we try the parametric bootstrap.

Finally we compute a power curve for the benchmark case and β_{15} in $(0.5, 1.50)$, with

	Benchmark	Variation	Table	Page
Cointegration rank	$r = 1$	$r = 2$	2	20
Number of vectors tested	$s = 1$	$s = 1, 2$	2	20
Sample size	$T = 100$	$T = 400$	3	20
VAR lag length	$k = 2$	$k = 4$	4	21
	$\varphi_1 = \varphi_2 = -\frac{0.7}{2}$	$\varphi_1 = -\frac{0.7}{2}, \varphi_2 = -\frac{0.7}{3}$, $\varphi_3 = \varphi_4 = -\frac{0.7}{16}$		
Speed of adjustment	$\varphi_1 = \varphi_2 = -\frac{0.7}{2}$	$\varphi_1 = \varphi_2 = -\frac{0.4}{2}$	5	21
Resampling of errors	parametric	resample from $\widehat{\varepsilon}$ and $\widehat{\varepsilon}^*$	6	21
Power curve	based on $\hat{\theta}$		7	22
Power curve		based on $\hat{\theta}_0$	8	23

Table 1: Design of Monte Carlo Experiment for small sample corrections

$\beta_{15}^0 = 1$ as usual, both for the bootstrap and Bartlett correction we propose, that is those based on the estimates under the alternative and one based on estimates under the null.

4 Results

Although the results of the simulations amount to a considerable mass, their essence is quite simple, and summarised in Table 2; in the following tables a few details are highlighted, with baseline results repeated in different tables in order to facilitate comparisons. In all the cases reported and discussed in this section the nominal significance level of the tests is always 5%, with results for different values available on request.

First of all, in the sample sizes we typically encounter in applied econometric work (100 observations) the asymptotic tests deliver disastrous performance as far as type I errors are concerned: at times they exceed 50% at the nominal 5% level. All the alternative procedures (Bartlett correction, simple and fast double bootstrap) are able to reduce substantially the size distortion in all our experiments, but are unable to eliminate it: in the case of rank=1 the minimum rejection rate, delivered by the fast double bootstrap type 1, is 26%, while in the case of rank=2 and test on one vector the rejection rate of all bootstrap procedures and the Bartlett corrected test is 15%. The two types of fast double bootstrap p-values are always very close, confirming that the procedure is reliable in our context. The power loss from using the procedures with lower size distortion is acceptable, with the rejection rates always over 70%. This finding will be confirmed by the power curve reported in table 7. To understand the point of basing both the bootstrap and Bartlett correction on the unrestricted estimates, a glance at the power curves in table 8 suffices: whereas the size performance is indeed slightly better than in table 7, the level of type II errors is unacceptably high, such that the power curves are almost flat. In the last row we report the number of cases where the highest estimated root is explosive: when the discrepancy between DGP and model becomes large, this percentage rises rapidly and corroborates theorem 1 of this paper[5].

Another key point from Table 2, is that the size performance of all test procedures of $\mathcal{H}_0 : \beta_1 = \beta_1^0$ in a model with 2 vectors is markedly better than the hypothesis $\mathcal{H}_0 : \beta = \beta^0$ in a model with one cointegrating vector.

For $T = 400$ all corrected tests achieve correct size and 100% power while the Type I error of the asymptotic test is still higher than the nominal size (cf. Table 3).

Increasing the length of the VAR also has large adverse effects on the test (cf. Table 4): thus, contrary to somehow common wisdom and in accord with Abadir et al. (1999), parsimony in the estimation of the VAR seems to be a rather important virtue.

[5] $tr(\sum_{i=0}^{\infty} P^i)$ does not converge if P contains an explosive root. However computationally we use the standard formula $(\sum_{i=0}^{\infty} \lambda^i) = \frac{1}{1-\lambda}$ to calculate the Bartlett correction both in the convergent case (when it is valid) and the non-convergent case.

There is nothing, which prevents the Bartlett correction factor from being smaller than -1: this is a known problem in the literature. We assign a p-value of 1 to these cases. In all the published and unpublished simulations we did, this only happened in those of table 8.

How sensitive are the performances of the tests to the speed of adjustment to equilibrium? Unsurprisingly, the answer is, a lot. Cutting ϕ (the sum of the coefficients of the autoregressive polynomial describing the dynamics of the errors in the cointegrating relationships) from 0.7 to 0.4 causes generally a more than proportional fall of the Type I error (for instance, that of the fast double bootstrap type 1 falls from 26% to 10%, see table 5).

Given the good results delivered by the bootstrapped tests, it is of some interest to check if using resampled or parametrically generated errors makes any difference. The results reported in Table 6 suggest that it does not, and thus the parametric bootstrap (easier to implement) may be adopted in practice. However, some caution is needed here, as in our experiments the same parametric hypothesis (normality) is used both in the generation of the Monte Carlo and bootstrap errors. Further research with different error processes for the Monte Carlo and bootstrap DGPs (for instance a leptocurtic error distribution in the DPG and resampling for a normal distribution) is needed.

Finally, a noteworthy finding is that the power curves of the all the variants of bootstrap tests are rather steep (table 7 and figure 3(b)). Although these results are specific to a single signal/noise ratio, they do suggest that the risk of unacceptable power losses from using some type of bootstrap test rather than the asymptotic or Bartlett corrected tests is likely to be remote.

5 Conclusions

We have compared different variants of bootstrap and Bartlett-corrected tests in a DGP which is relatively unfavourable, but reproduces some features of real life empirical applications: a relatively large system (5 variables and 2 or 4 lags), and rather slow adjustment to long-run equilibrium. With such a complex DGP the caveats common to all simulation studies are even more important than usual. Our design depends on over 120 parameters, the vast majority of which had to be kept fixed across all experiments, and thus we must be extremely cautious in reaching any conclusion.

Further, the type of tests examined assumes full knowledge of the tested cointegrating vectors, a rare event in practice: however, they are the only tests for which the Bartlett correction is available. Indeed, the Bartlett correction has not been derived yet for many cases of strong empirical interest (e.g., hypotheses of the kind $\beta_i = H_i\varphi_i$ and in general models with impulse dummies) and hence the bootstrap may in fact be the only alternative to the asymptotic p-values. With all these caveats, our recommendations are the following:

(i) Asymptotic tests should be used in no circumstance;

(ii) Bartlett-corrected tests may be used provided considerable caution is exercised, as their Type I error is often much larger than the nominal size;

(iii) Bootstrap tests, with a somehow lower size distortion than the Bartlett corrected tests accompanied by limited power losses, may also be used; the fast double boot-

strap of Davidson and MacKinnon (2000) delivers the best performance, and thus it appears to be a powerful tool for applied work, especially in the many cases when the Bartlett correction is not available.

We stress that both the Bartlett correction and the bootstrap should always be based on the unrestricted estimate of β.

Among the many points that remain open, two are especially important: (*a*) the development of equivalent hypothesis, like (6) for $\mathcal{H}_0 : \beta_1 = \beta_1^0$ for more general restrictions on β, like $\beta_i = H_i\varphi_i$ with an accurate Monte Carlo study of their properties and (*b*) theoretical results on the asymptotics of the (fast double) bootstrap in cointegrated systems.

References

Abadir, K., K. Hadri, and E. Tzavalis (1999). The influence of VAR dimensions on estimator biases. *Econometrica 67*, 163–181.

Barndorff-Nielsen, O. and P. Hall (1988). On the level-error after Bartlett adjustment of the likelihood ratio statistic. *Biometrika 75*, 374–378.

Bartlett, M. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London Series A 160*, 268–282.

Beran, K. (1988). Prepivoting test statistics: A bootstrap view of asymptotic refinements. *Journal of the American Statistical Association 83*, 687–697.

Berkowitz, J. and L. Kilian (2000). Recent developments in bootstrapping time series. *Econometric Reviews 19*, 1–48.

Bewley, R., D. Orden, M. Yang, and L. Fisher (1994). Comparison of Box-Tiao and Johansen canonical estimator of cointegration vectors in VEC(1) models. *Journal of Econometrics 64*, 3–27.

Chan, N. and C. Wei (1988). Limiting distributions of least squares estimates of unstable autoregressive processes. *Annals of Mathematical Statistics 16*, 367–410.

Chang, Y., R. Sickles, and W. Song (2001). Bootstrapping unit root tests with covariates. Technical report, Rice University.

Cribari-Neto, F. and G. Cordeiro (1996). On Bartlett and Bartlett-type corrections. *Econometric Reviews 15*, 339–367.

Davidson, J. (2001). Alternative bootstrap procedures for testing cointegration in fractionally integrated processes. Technical report, Cardiff University.

Davidson, R. and J. MacKinnon (2000). Improving the reliability of bootstrap tests. Technical Report 995, Queen's University Institute for Economic Research.

Efron, B. and R. Tibshirani (1993). *An Introduction to the Bootstrap*. London: Chapman & Hall.

Engle, R. and C. Granger (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica 55*, 251–276.

Fachin, S. (2000). Bootstrap and asymptotic tests of long-run relationships in cointegrated systems. *Oxford Bulletin of Economics and Statistics 62*, 577–585.

Fachin, S. and L. Bravetti (1996). Asymptotic normal and bootstrap inference in structural VAR analysis. *Journal of Forecasting 15*, 329–341.

Gonzalo, J. (1994). Five alternative methods of estimating long-run equilibrium relationships. *Journal of Econometrics 64*, 203–233.

Gredenhoff, M. and T. Jacobson (2001). Bootstrap testing linear restrictions on cointegrating vectors. *Journal of Business and Economic Statistics 19*, 63–72.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. New York: Springer.

Hall, P. (1995). Methodology and theory for the bootstrap. In R. Engle and D. McFadden (Eds.), *Handbook of Econometrics, Vol. IV*, Amsterdam. North-Holland.

Haug, A. (1996). Tests for cointegration - a Monte Carlo comparison. *Journal of Econometrics 71*, 89–115.

Horowitz, J. (2002). The bootstrap. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics, Vol. V*, Amsterdam. North-Holland.

Inoue, A. and L. Kilian (2002). Bootstrapping autoregressive processes with possible unit roots. *Econometrica 70*, 377–391.

Jensen, J. and A. Wood (1997). On the non-existence of a Bartlett correction for unit root tests. *Statistics and Probability Letters 35*, 181–187.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive processes. *Econometrica 59*, 1551–1580.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Johansen, S. (2000). A Bartlett correction factor for tests on the cointegrating relations. *Econometric Theory 16*, 740–778.

Lawley, D. (1956). A general method for approximating the distribution for the likelihood ratio criteria. *Biometrika 43*, 293–303.

Li, H. and G. Maddala (1996). Bootstrapping time series models. *Econometric Reviews 15*, 115–158.

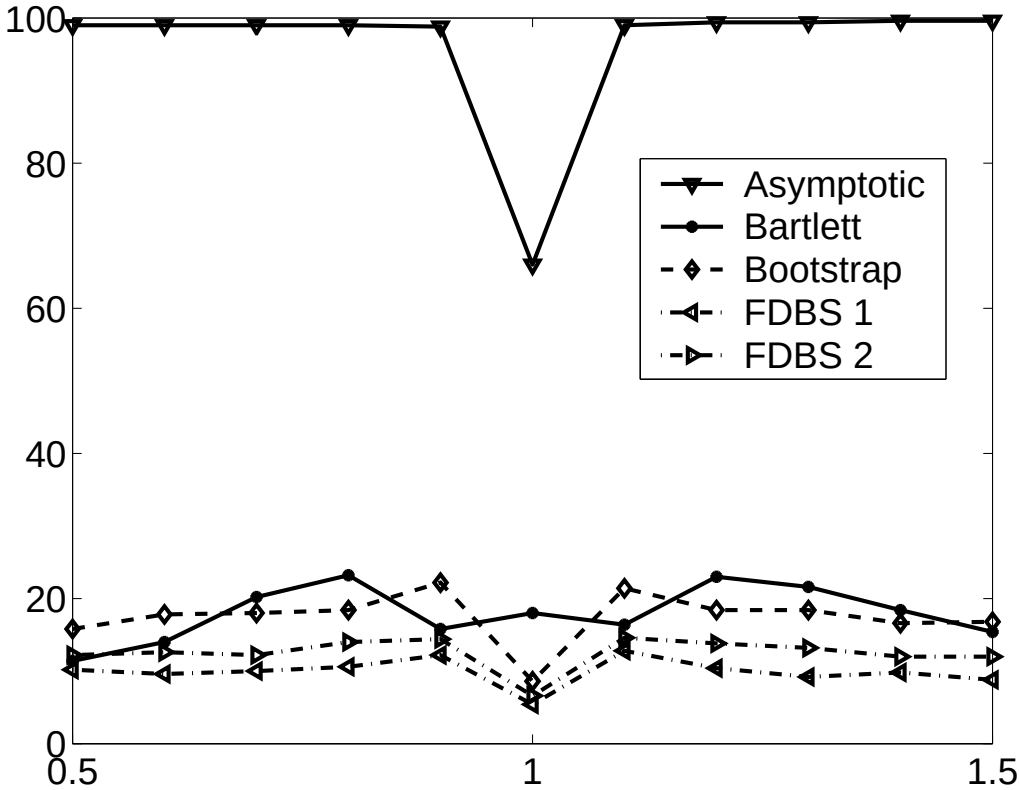
Li, H. and G. Maddala (1997). Bootstrapping cointegrating regressions. *Journal of Econometrics 80*, 297–318.

Nielsen, B. (1997). Bartlett correction of the unit root test in autoregressive models. *Biometrika 84*, 500–504.

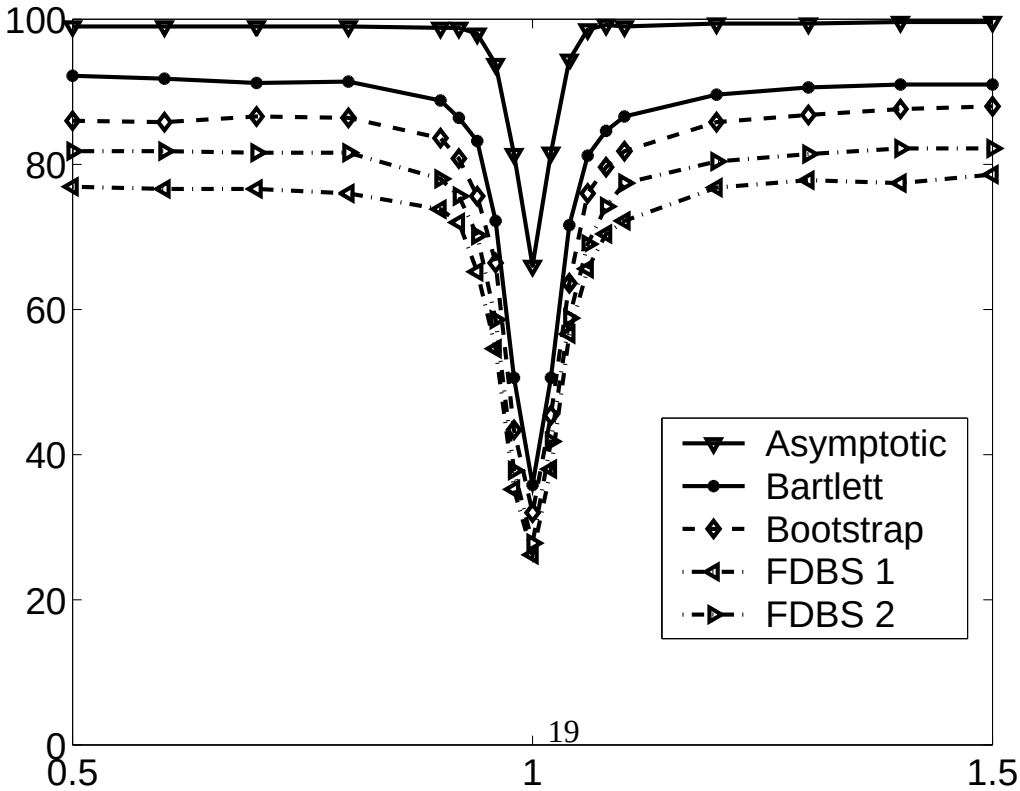
Paparoditis, E. and D. Polititis (2001). Unit-roots testing via the continuos-path block bootstrap. Technical Report 2001-06, University of California at San Diego, Department of Economics.

Phillips, P. (2001). Bootstrapping spurious regression. Technical report.

Appendix: Figures and tables



(a) power curves based on restricted estimates $\hat{\theta}_0$



(b) power curves based on unrestricted estimates $\hat{\theta}$

Figure 3: Power curves for test on cointegration coefficients

1 to 2 cointegration vectors, test on 1 to 2 vectors $\phi = 0.7$, $T = 100$, $k = 2$						
rank, tested vectors	1,1		2,1		2,2	
Test	*size*	*power*	*size*	*power*	*size*	*power*
Asymptotic	66.0	99.0	39.2	97.6	68.6	98.2
Bartlett	35.8	92.2	15.8	79.2	33.2	82.2
Bootstrap	32.0	86.0	15.2	77.2	28.2	74.6
FDB_1	26.2	76.0	13.4	68.0	20.0	62.2
FDB_2	27.8	81.8	14.2	71.4	23.6	68.2

nominal significance level: 5%; FDB_i: Fast Double Bootstrap type *i*
power simulations:
case (1,1) $H_0 : \beta_1^0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \end{bmatrix}$, DGP: $\beta_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0.5 \end{bmatrix}$
case (2,1): as case (1,1) with DGP: $\beta_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \end{bmatrix}$
case (2,2): as case (2,1) with $H_0 : \beta_2^0 = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \end{bmatrix}$

Table 2: Benchmark case small sample correction for tests on cointegration vectors

1 cointegrating vector, test on 1 vector $\phi = 0.7$, $T = 100$ and 400, $k = 2$				
T	100		400	
Test	*size*	*power*	*size*	*power*
Asymptotic	66.0	99.0	11.0	100.0
Bartlett	35.8	92.2	5.6	100.0
Bootstrap	32.0	86.0	6.2	100.0
FDB_1	26.2	76.0	5.6	100.0
FDB_2	27.8	81.8	5.8	100.0

nominal significance level: 5%
power simulations: see Table 2

Table 3: Sample size and small sample corrections

1 cointegrating vector, test on 1 vector				
$\phi = 0.7, T = 100, k = 2$ and 4				
lags	2		4	
Test	*size*	*power*	*size*	*power*
Asymptotic	66.0	99.0	85.4	99.4
Bartlett	35.8	92.2	53.2	91.0
Bootstrap	32.0	86.0	39.2	82.0
FDB_1	26.2	76.0	32.4	68.8
FDB_2	27.8	81.8	35.6	74.4
nominal significance level: 5%				
power simulations: see Table 2				

Table 4: Lag length and small sample corrections

1 cointegrating vector, test on 1 vector				
$\phi = 0.7$ and 0.4, $T = 100$, $k = 2$				
ϕ	0.7		0.4	
Test	*size*	*power*	*size*	*power*
Asymptotic	66.0	99.0	33.0	99.6
Bartlett	35.8	92.2	17.0	97.6
Bootstrap	32.0	86.0	14.2	96.8
FDB_1	26.2	76.0	10.8	94.4
FDB_2	27.8	81.8	11.8	95.6
nominal significance level: 5%				
power simulations: see Table 2				

Table 5: Speed of adjustment and small sample corrections

1 cointegrating vector, test on 1 vector				
$\phi = 0.7, T = 100, k = 2$				
Type of bootstrap	Non-Parametric		Parametric	
Test	size	power	size	power
Asymptotic	66.0	99.0	66.0	99.0
Bartlett	35.8	92.2	35.8	92.2
Bootstrap	32.0	86.0	32.0	86.4
FDB_1	26.2	76.0	25.0	76.2
FDB_2	27.8	81.8	27.0	80.2
nominal significance level: 5%				
power simulations: see Table 2				

Table 6: Non-parametric bootstrap and small sample corrections

$\phi = 0.7, T = 100, k = 2$					
Test / β_{15}	*Asymptotic*	*Bartlett*	*Bootstrap*	*FDB*$_1$	*FDB*$_2$
0.5	99.0	92.2	86.0	76.9	81.8
0.6	99.0	91.8	85.8	76.6	81.8
0.7	99.0	91.2	86.6	76.6	81.6
0.8	99.0	91.4	86.4	76.0	81.6
0.9	98.8	88.8	83.6	73.8	78.0
0.92	98.8	86.4	80.8	72.0	75.6
0.94	98.0	83.2	75.6	65.2	70.2
0.96	93.8	72.2	66.4	54.6	58.6
0.98	81.4	50.6	43.4	35.2	37.8
1.0	*66.0*	*35.8*	*32.0*	*26.2*	*27.8*
1.02	81.6	50.6	45.4	38.0	41.8
1.04	94.4	71.6	63.6	56.6	58.8
1.06	98.6	81.2	76.0	65.6	69.0
1.08	99.2	84.6	79.6	70.4	74.2
1.1	99.0	86.6	81.8	72.2	77.4
1.2	99.4	89.6	85.8	76.8	80.4
1.3	99.4	90.6	86.8	77.8	81.4
1.4	99.6	91.0	87.6	77.4	82.2
1.5	99.6	91.0	88.0	78.6	82.2
nominal significance level: 5%					

Table 7: Power curve based on unrestricted estimates

$\phi = 0.7, T = 100, k = 2$						
Test / β_{15}	*Asymptotic*	*Bartlett*	*Bootstrap*	*FDB*$_1$	*FDB*$_2$	Explosive Roots *(% of simulations)*
0.5	99.0	11.4	15.8	10.2	12.2	34.8
0.6	99.0	14.0	17.8	9.6	12.6	30.4
0.7	99.0	20.2	18.0	10.0	12.2	23.4
0.8	99.0	23.2	18.4	10.6	14.0	12.6
0.9	98.8	15.8	22.2	12.2	14.4	0.2
1.0	*66.0*	*18.0*	*8.6*	*5.4*	*6.6*	*0*
1.1	99.0	16.4	21.4	12.8	14.6	0.6
1.2	99.4	23.0	18.4	10.4	13.8	9.2
1.3	99.4	21.6	18.4	9.2	13.2	20.4
1.4	99.6	18.4	16.6	9.8	12.0	28.8
1.5	99.6	15.4	16.8	8.8	12.0	32.2
nominal significance level: 5%						

Table 8: Power curve based on restricted estimates