Department of Economics

Bootstrap Statistical Tests of Rank Determination for System Identification

Gonzalo Camba-Mendez and George Kapetanios

Working Paper No. 468 November 2002 ISSN 1473-0278

Queen Mary
University of London

Bootstrap Statistical Tests of Rank Determination for System Identification

Gonzalo Camba-Mendez[*]and George Kapetanios[†]
European Central Bank and Queen Mary University of London

October 2002

Abstract

Identification in the context of multivariate state space modelling involves the specification of the dimension of the state vector. One identification approach requires an estimate of the rank of a Hankel matrix. The most frequently used approaches of rank determination rely on information criteria methods. This paper evaluates the performance of some asymptotic tests of rank determination together with their bootstrapped versions against standard information criteria methods. This study is conducted through simulation experiments. Results show that the bootstrapped procedures significantly improve upon the performance of the corresponding asymptotic tests, and are proved better than standard Information Criterion methods.

JEL Codes:C12, C15, C32
Keywords: Rank, Bootstrap, Monte Carlo, System Identification, Hankel Operator.

[*]European Central Bank, Kaiserstrasse 29, D-60311, Frankfurt am Main, Germany, email: gonzalo.camba-mendez@ecb.int .

[†]Department of Economics, Queen Mary, University of London, Mile End Road, London E1 4NS. email: G.Kapetanios@qmul.ac.uk

1 Introduction

Identification in the context of multivariate state space modelling involves the specification of the dimension of the state vector. [1] provided one such a method based on canonical correlations. In the present paper we will focus on an alternative identification approach based on the Kronecker theorem; see [2]. This method consists of estimating the rank of the Hankel matrix. The rank of this matrix gives the minimal realisation, i.e. the minimum number of states in a state space representation that replicates the transfer function of the vector series under study. The most frequently used rank determination methods in the systems literature rely on model selection criteria techniques, see [3] and [4] and references therein.

In recent years, there has been renewed research interest in the development of tests of the rank of a matrix. *Inter alia*, [5], [6], [7] and [8] consider a situation in which a *root-T* consistent (RTC) estimator is available for the matrix of interest where T denotes the sample size. The focus of this paper will be to evaluate the performance of some of these asymptotic tests of rank determination together with their bootstrapped versions against standard Information Criterion methods in the context of System Identification.

Section 2 describes the system identification strategy based on the rank of the Hankel matrix. Section 3 describes the procedures of rank determination under study in the present paper. Section 4 describes the Monte Carlo strategy undertaken. Section 5 presents the Monte Carlo results. Finally section 6 concludes.

2 System Identification

This paper focuses on a linear state space system of the form:

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{C}\boldsymbol{s}_t + \boldsymbol{e}_t \\ \boldsymbol{s}_{t+1} &= \boldsymbol{A}\boldsymbol{s}_t + \boldsymbol{B}\boldsymbol{e}_t \end{aligned} \tag{1}$$

where $\boldsymbol{y}_t$ is a m-vector series, $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are $r \times r$, $r \times m$ and $m \times r$ parameter matrices respectively, $\boldsymbol{s}_t$ is a r-vector of unobservable state variables, and $\boldsymbol{e}_t$ is an m-vector of random variables with mean zero and positive definite covariance matrix $\boldsymbol{\Omega}$. This system can be characterised by a system transfer function matrix $\boldsymbol{G}(z) = \sum_{i=1}^{\infty} \boldsymbol{G}_i z^{-1}$, where $\boldsymbol{G}_i$ are the impulse response matrices. The order of the system, is defined as the order of the minimal state-space realization, i.e. the minimal dimension of the state vector that replicates the transfer function. Corresponding to the transfer function matrix $\boldsymbol{G}(z)$ above, the infinite dimensional Hankel matrix is defined as:

$$\boldsymbol{H} = \boldsymbol{\mathcal{O}}\boldsymbol{\mathcal{C}} = \begin{bmatrix} \boldsymbol{G}_1 & \boldsymbol{G}_2 & \boldsymbol{G}_3 & \cdots \\ \boldsymbol{G}_2 & \boldsymbol{G}_3 & \cdots & \cdots \\ \boldsymbol{G}_3 & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} \boldsymbol{CB} & \boldsymbol{CAB} & \boldsymbol{CA}^2\boldsymbol{B} & \cdots \\ \boldsymbol{CAB} & \boldsymbol{CA}^2\boldsymbol{B} & \cdots & \cdots \\ \boldsymbol{CA}^2\boldsymbol{B} & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{2}$$

where the so called observability matrix is defined as $\boldsymbol{\mathcal{O}} = \left[\boldsymbol{C}', \boldsymbol{A}'\boldsymbol{C}', (\boldsymbol{A}^2)'\boldsymbol{C}', \cdots\right]'$ and the so called controllability matrix as $\boldsymbol{\mathcal{C}} = \left[\boldsymbol{B}, \boldsymbol{AB}, (\boldsymbol{A}^2)\boldsymbol{B}, \cdots\right]$. Kronecker's theorem can be used to show that the order of the system is equal to the rank of the Hankel matrix (see [2]). The rank of $\hat{\boldsymbol{H}}$ is the focus of the investigation. The computation of the rank of the Hankel matrix is not an easy task, as it is unlikely that the impulse response matrices are given exactly, and in a majority of cases they are estimated. Furthermore, searching for the rank of the Hankel matrix is not conducted directly on (2) but rather on some *pseudo*-Hankel matrices. Two alternative strategies are described in the subsections below.

2.1 Hankel Covariance Matrix $\boldsymbol{H}^a$

An alternative characterisation of this system is in terms of a Hankel matrix of the covariances of the output vector, $\boldsymbol{y}_t$.

$$\boldsymbol{H}^a = \boldsymbol{\mathcal{O}}\overline{\boldsymbol{\mathcal{C}}} = \begin{bmatrix} \boldsymbol{\Delta}_1 & \boldsymbol{\Delta}_2 & \boldsymbol{\Delta}_3 & \cdots \\ \boldsymbol{\Delta}_2 & \boldsymbol{\Delta}_3 & \cdots & \cdots \\ \boldsymbol{\Delta}_3 & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{3}$$

where $\boldsymbol{\Delta}_i$ is the autocovariance matrix of $\boldsymbol{y}_t$ for lag i, $\overline{\boldsymbol{\mathcal{C}}} = \left[\overline{\boldsymbol{C}}, \boldsymbol{A}\overline{\boldsymbol{C}}, (\boldsymbol{A}^2)\overline{\boldsymbol{C}}, \cdots\right]$, $\overline{\boldsymbol{C}} = \boldsymbol{B\Omega} + \boldsymbol{APC}'$ and $\boldsymbol{P}$ is the covariance matrix of the state vector defined as $E\{\boldsymbol{s}_t\boldsymbol{s}_t'\}$. It follows that the rank of $\boldsymbol{H}^a$ is equal to the rank of $\boldsymbol{H}$, see [9]. Obviously one cannot use the infinite dimensional matrix above, and when working with finite data a finite truncation of the Hankel matrix will be required. Note that this Hankel Covariance matrix can be defined as the covariance matrix between the vectors $\boldsymbol{y}_+^t$ and $\boldsymbol{y}_-^t$, and defined as follows:

$$\boldsymbol{H}^a = E\left(\boldsymbol{y}_+^t \boldsymbol{y}_-^{t\,\prime}\right) = E\begin{pmatrix} \boldsymbol{y}_{t+1} \\ \cdots \\ \boldsymbol{y}_{t+k} \end{pmatrix} \begin{pmatrix} \boldsymbol{y}_t' & \cdots & \boldsymbol{y}_{t-p+1}' \end{pmatrix} = \begin{bmatrix} \boldsymbol{\Delta}_1 & \boldsymbol{\Delta}_2 & \cdots & \boldsymbol{\Delta}_p \\ \boldsymbol{\Delta}_2 & \boldsymbol{\Delta}_3 & \cdots & \boldsymbol{\Delta}_{p+1} \\ \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{\Delta}_k & \boldsymbol{\Delta}_{k+1} & \cdots & \boldsymbol{\Delta}_{k+p-1} \end{bmatrix} \tag{4}$$

The truncation parameters k and p must be fixed. For T observations, an estimator for the Hankel matrix which is computationally fast and guarantees that the upper diagonal blocks are identical to the corresponding lower diagonal blocks, is defined as $\hat{\boldsymbol{H}}^a = \frac{1}{T}\boldsymbol{\mathcal{Y}}_+'\boldsymbol{\mathcal{Y}}_-$ where,

$$\boldsymbol{\mathcal{Y}}_- = \begin{bmatrix} \boldsymbol{y}_1' & 0 & \dots & 0 \\ \boldsymbol{y}_2' & \boldsymbol{y}_1' & \dots & 0 \\ \boldsymbol{y}_3' & \boldsymbol{y}_2' & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \boldsymbol{y}_{T-1}' & \boldsymbol{y}_{T-2}' & \dots & \boldsymbol{y}_{T-p}' \end{bmatrix} \quad \boldsymbol{\mathcal{Y}}_+ = \begin{bmatrix} \boldsymbol{y}_2' & \boldsymbol{y}_3' & \dots & \boldsymbol{y}_{k+1}' \\ \dots & \dots & \dots & \dots \\ \boldsymbol{y}_{T-2}' & \boldsymbol{y}_{T-1}' & \dots & 0 \\ \boldsymbol{y}_{T-1}' & \boldsymbol{y}_T' & \dots & 0 \\ \boldsymbol{y}_T' & 0 & \dots & 0 \end{bmatrix} \tag{5}$$

As $\hat{\boldsymbol{H}}^a$ is estimated, it will always have full rank. One of the statistical procedures we will consider uses the covariance matrix of $\hat{\boldsymbol{H}}^a$ in order to assess statistically its rank. A consistent estimator of the covariance matrix of $\sqrt{T}vec(\hat{\boldsymbol{H}}^a)$, and defined as $\boldsymbol{V}^a$, is given by $\hat{\boldsymbol{V}}^a = \frac{1}{T}\sum_{t=k}^{T-k}(vec(\boldsymbol{y}_+^t\boldsymbol{y}_-^{t\,\prime}) - vec(\hat{\boldsymbol{H}}^a))(vec(\boldsymbol{y}_+^t\boldsymbol{y}_-^{t\,\prime}) - vec(\hat{\boldsymbol{H}}^a))'$. While the matrix $\boldsymbol{V}^a$ is of reduced rank, the rank of $\hat{\boldsymbol{V}}^a$ above is only of reduced rank asymptotically. This is problematic for one of the procedures consider below. An estimator of $\boldsymbol{V}^a$ with equal rank to $\boldsymbol{V}^a$ must be constructed as follows:

$$\hat{\boldsymbol{V}}^a = \frac{1}{T}\sum_{t=k}^{T-k}(vec(\boldsymbol{\mathcal{Z}}^t) - vec(\hat{\boldsymbol{H}}^a))(vec(\boldsymbol{\mathcal{Z}}^t) - vec(\hat{\boldsymbol{H}}^a))' \tag{6}$$

where $\boldsymbol{\mathcal{Z}}^t = \left((\boldsymbol{y}_t\boldsymbol{\mathcal{Y}}_-^{t-1})', (\boldsymbol{y}_t\boldsymbol{\mathcal{Y}}_-^{t-2})', \cdots, (\boldsymbol{y}_t\boldsymbol{\mathcal{Y}}_-^{t-k})'\right)'$ where $\boldsymbol{\mathcal{Y}}_-^t$ denotes the t-th row of $\boldsymbol{\mathcal{Y}}_-$ and without loss of generality we are assuming that $p < k$. Note that by construction this sample estimator of the asymptotic covariance matrix is singular and of equal rank to $\boldsymbol{V}^a$. The singularity of this estimator will present a problem as a number of procedures of rank determination require the inversion of such a matrix. This problem will be discussed in detail in the next section.

2.2 Hankel Regression Matrix $\boldsymbol{H}^b$

An alternative representation of system (1) can be accomplished by defining the vector of stacked future outputs $\boldsymbol{Y}_t^+ = \left(\boldsymbol{y}_t', \boldsymbol{y}_{t+1}', \cdots\right)'$ the vector of stacked past outputs $\boldsymbol{Y}_t^- = \left(\boldsymbol{y}_{t-1}', \boldsymbol{y}_{t-2}', \cdots\right)'$ and the vector of stacked future noise components $\boldsymbol{E}_t^+ = \left(\boldsymbol{e}_t', \boldsymbol{e}_{t+1}', \cdots\right)'$; we can then rewrite (1) as follows:

$$\boldsymbol{Y}_t^+ = \boldsymbol{\mathcal{O}}\overline{\overline{\boldsymbol{\mathcal{C}}}}\boldsymbol{Y}_t^- + \boldsymbol{M}\boldsymbol{E}_t^+ \tag{7}$$

where the observability matrix $\boldsymbol{\mathcal{O}}$ is defined as above; $\overline{\overline{\boldsymbol{\mathcal{C}}}} = [\boldsymbol{B}, (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}), (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C})^2, \cdots]$; and where

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \cdots \\ \boldsymbol{CB} & \boldsymbol{I} & \boldsymbol{0} & \ddots & \vdots \\ \boldsymbol{CAB} & \boldsymbol{CB} & \boldsymbol{I} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \cdots \end{bmatrix}$$

The rank of $\boldsymbol{H}^b = \boldsymbol{\mathcal{O}}\overline{\overline{\boldsymbol{\mathcal{C}}}}$ is equivalent to the rank of $\boldsymbol{H}$. Once more one cannot use the infinite dimensional matrix above, and when working with finite data use is made of a finite truncation of $\boldsymbol{H}^b$. Using truncation parameters k and p as above, we can rewrite (7) as:

$$\boldsymbol{y}_+^t = \boldsymbol{H}^b \boldsymbol{y}_-^t + \boldsymbol{M}_{k,p} \boldsymbol{E}_{(k,p)t}^+ \tag{8}$$

where $\boldsymbol{y}_+^t$ and $\boldsymbol{y}_-^t$ are truncated versions of the quantities defined above. $\boldsymbol{M}_{k,p}$ and $\boldsymbol{E}_{(k,p)t}^+$ are truncated correspondingly, see [10]. The important feature is that one can get an estimate of $\boldsymbol{H}^b$, denoted as $\hat{\boldsymbol{H}}^b$, by standard regression methods, i.e. $\hat{\boldsymbol{H}}^b = \boldsymbol{\mathcal{Y}}_+' \boldsymbol{\mathcal{Y}}_- (\boldsymbol{\mathcal{Y}}_-' \boldsymbol{\mathcal{Y}}_-)^{-1} = \hat{\boldsymbol{H}}^a T (\boldsymbol{\mathcal{Y}}_-' \boldsymbol{\mathcal{Y}}_-)^{-1}$. A consistent estimator of the covariance matrix of $\sqrt{T} vec(\hat{\boldsymbol{H}}^b)$, $\boldsymbol{V}^b$ is given by $\hat{\boldsymbol{V}}^b = \boldsymbol{D} \hat{\boldsymbol{V}}^a \boldsymbol{D}'$ where $\boldsymbol{D} = [T(\boldsymbol{\mathcal{Y}}_-' \boldsymbol{\mathcal{Y}}_-)^{-1} \otimes \boldsymbol{I}]$. Note that the rank of $\hat{\boldsymbol{V}}^b$ is equal to the rank of $\boldsymbol{V}^b$. As stated above this is a desirable property to bear in mind in the context of one of the rank tests described below.

3 Estimating the rank of the Hankel matrix

Two types of procedures will be used to provide an estimate of the rank: Statistical tests of rank and information criteria. In what follows whenever we refer to $\boldsymbol{H}$ and $\boldsymbol{V}$ it should be understood that results apply directly to both *pseudo*-Hankel matrices $\boldsymbol{H}^a$ and $\boldsymbol{H}^b$ with their respective covariance matrices $\boldsymbol{V}^a$ and $\boldsymbol{V}^b$. When results do not direcly apply to both matrices the particular superscript ^{a} or ^{b} is used.

3.1 Statistical Tests of Rank

The procedures which involve statistical tests consider the following null hypothesis $H_0 : r(\boldsymbol{H}) = r^*$ against the alternative $H_1 : r(\boldsymbol{H}) > r^*$. Starting with the null hypothesis of $r^* = 1$, a sequence of tests is performed. If the null is rejected r^* is augmented by one and the test is repeated. When the null cannot be rejected, r^* is adopted as the estimate of the rank of $\boldsymbol{H}$. Nevertheless, the rank estimate provided by this approach will not converge in probability to the true value of the rank of the Hankel matrix, denoted by r^0. The reason is that even if the null hypothesis tested is true, the testing procedure will reject it with probability α, where α is the significance level. The rank estimate will converge to its true value, r^0, as T goes to infinity, if α is made to depend on T and goes to zero as T goes to infinity but not faster than a given rate. We denote this α by α_T, where the subscript T now denotes dependence of the significance level on the sample size. [11] shows that if α_T goes to zero as the sample size T goes to infinity and also $\lim_{T\to\infty} \ln \alpha_T / T = 0$, then the rank estimate provided by the sequential testing procedure will converge in probability to r^0, see also [6].

3.1.1 Cragg and Donald (1996)

This method applies to both $\boldsymbol{H}^a$ and $\boldsymbol{H}^b$. The procedure proposed by [5] is based on the transformation of the matrix $\boldsymbol{H}$ using Gaussian elimination with complete pivoting[1]. r^* steps of Gaussian elimination with full pivoting on matrix $\boldsymbol{H}$ amounts to the following operations:

$$\boldsymbol{Q}_{r^*}\boldsymbol{R}_{r^*}\boldsymbol{Q}_{r^*-1}\boldsymbol{R}_{r^*-1}\dots\boldsymbol{Q}_1\boldsymbol{R}_1\boldsymbol{H}\boldsymbol{C}_1\dots\boldsymbol{C}_{r^*-1}\boldsymbol{C}_{r^*} = \begin{bmatrix} \boldsymbol{H}_{11}(r^*) & \boldsymbol{H}_{12}(r^*) \\ \boldsymbol{0} & \boldsymbol{H}_{22}(r^*) \end{bmatrix}$$

where $\boldsymbol{R}_i$ and $\boldsymbol{C}_i$ are pivoting matrices for step i and $\boldsymbol{Q}_i$ are Gauss transformation matrices. The pivoting matrices used to perform the first r^* steps of Gaussian elimininination are applied to $\boldsymbol{H}$ to obtain the following relation

$$\boldsymbol{R}_{r^*}\boldsymbol{R}_{r^*-1}\dots\boldsymbol{R}_1\boldsymbol{H}\boldsymbol{C}_1 \dots \boldsymbol{C}_{r^*-1}\boldsymbol{C}_{r^*} = \boldsymbol{R}\boldsymbol{H}\boldsymbol{C} = \boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}_{11}(r^*) & \boldsymbol{F}_{12}(r^*) \\ \boldsymbol{F}_{21}(r^*) & \boldsymbol{F}_{22}(r^*) \end{bmatrix} \tag{9}$$

where $\boldsymbol{F}$ is partitioned accordingly, i.e. $\boldsymbol{F}_{11}(r^*)$ is of dimension $r^* \times r^*$. Note that in this case $\boldsymbol{F}_{11}(r^*)$ has full rank, under the null hypothesis that $rk(\boldsymbol{H}) = r^*$. It then follows, (see [5]), that $\boldsymbol{F}_{22}(r^*) - \boldsymbol{F}_{21}(r^*)\boldsymbol{F}_{11}^{-1}(r^*)\boldsymbol{F}_{12}(r^*) = 0$. The estimated counterpart of the above relation, i.e. $\hat{\boldsymbol{F}}_{22} - \hat{\boldsymbol{F}}_{21}\hat{\boldsymbol{F}}_{11}^{-1}\hat{\boldsymbol{F}}_{12} = \hat{\boldsymbol{\Lambda}}_{22}(r^*)$, may be used as a test statistic of the hypothesis that the rank of $\boldsymbol{H}$ is r^*. Under regularity conditions, including the requirement that the covariance matrix of the asymptotically normally distributed matrix $\sqrt{T}vec(\hat{\boldsymbol{H}} - \boldsymbol{H})$ has full rank, the following result can be shown, under H_0.

$$\sqrt{T}vec(\hat{\boldsymbol{\Lambda}}_{22}(r^*)) \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{\Gamma}\boldsymbol{V}\boldsymbol{\Gamma}')$$

where $\boldsymbol{\Gamma} = \boldsymbol{\Phi}_2 \otimes \boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_1 = \begin{bmatrix} -\boldsymbol{F}_{21}\boldsymbol{F}_{11}^{-1} & \boldsymbol{I}_{mk-r^*} \end{bmatrix} \boldsymbol{R}$, $\boldsymbol{\Phi}_2 = \begin{bmatrix} -\boldsymbol{F}_{12}'\boldsymbol{F}_{11}^{-1'} & \boldsymbol{I}_{mp-r^*} \end{bmatrix} \boldsymbol{C}'$ and $\xrightarrow{d}$ denotes convergence in distribution. Then,

$$\hat{\xi} = T vec\ \hat{\boldsymbol{\Lambda}}_{22}(r^*)'(\hat{\boldsymbol{\Gamma}}\hat{\boldsymbol{V}}\hat{\boldsymbol{\Gamma}}')^{-1} vec\ \hat{\boldsymbol{\Lambda}}_{22}(r^*) \xrightarrow{d} \chi^2_{(mk-r^*)(mp-r^*)}$$

where $\hat{\boldsymbol{\Gamma}}$ and $\hat{\boldsymbol{V}}$ are the sample estimates of $\boldsymbol{\Gamma}$ and $\boldsymbol{V}$ and χ^2_l denotes the χ^2 distribution with l degrees of freedom.

While the matrix $\boldsymbol{V}$ is of reduced rank, the rank of $\hat{\boldsymbol{V}}$ above is only of reduced rank asymptotically, where we use $\hat{\boldsymbol{V}}$ to refer to either $\hat{\boldsymbol{V}}^a$ or $\hat{\boldsymbol{V}}^b$. This is problematic for the Cragg and Donald procedure because it uses the inverse of the covariance matrix of the Hankel matrix. [13] have shown that if additionally the rank of $\boldsymbol{V}$ is known and $rk[\hat{\boldsymbol{V}}] = rk[\boldsymbol{V}], \quad \forall T$, then

$$\hat{\xi} = T vec\ \hat{\boldsymbol{\Lambda}}_{22}(r^*)'(\hat{\boldsymbol{\Gamma}}\hat{\boldsymbol{V}}\hat{\boldsymbol{\Gamma}}')^{+} vec\ \hat{\boldsymbol{\Lambda}}_{22}(r^*) \xrightarrow{d} \chi^2_\beta \tag{10}$$

where $^+$ denotes the Moore-Penrose inverse of a matrix, and the number of degrees of freedom β is given by the minimum between the number of rows in $\hat{\boldsymbol{\Gamma}}$ and the rank of $\hat{\boldsymbol{V}}$; i.e.

[1]For details on Gaussian elimination with complete pivoting see [5] or [12].

$\min\{(mp - r^*) \times (mk - r^*), (k + p - 1) * m\}$. Note that by construction, both estimators for the covariance matrices of the Hankel matrix presented above have the required property. We will refer to the Cragg and Donald procedure using generalised inverses as the GE procedure when applied to the $\hat{\boldsymbol{H}}^a$, $\hat{\boldsymbol{V}}^a$ pair; and as CC procedure when applied to the $\hat{\boldsymbol{H}}^b$, $\hat{\boldsymbol{V}}^b$ pair.

3.1.2 Bartlett (1947)

An alternative method to estimate the rank of the Hankel matrix is based on the computation of canonical correlations. This method only applies directly to the $\boldsymbol{H}^a$ matrix. A well known result in canonical correlation analysis is that given two random vector series $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, each of dimensions k, the rank of the covariance matrix between those two random vectors is equal to the number of nonzero cannonical correlations, see [14] for further details. The Hankel covariance matrix $\boldsymbol{H}^a$ defined in equation (4) above is the covariance matrix between the random vectors $\boldsymbol{y}^t_+$ and $\boldsymbol{y}^t_-$. Compute the QR decomposition of the matrices $\boldsymbol{\mathcal{Y}}_+$ and $\boldsymbol{\mathcal{Y}}_-$ given in (5) above, i.e. $\boldsymbol{\mathcal{Y}}_+ = \boldsymbol{Q}_+\boldsymbol{R}_+$ and $\boldsymbol{\mathcal{Y}}_- = \boldsymbol{Q}\_\boldsymbol{R}_-$. The canonical correlations between the vectors $\boldsymbol{y}^t_+$ and $\boldsymbol{y}^t_-$, are the singular values of $\boldsymbol{Q}'_+\boldsymbol{Q}_-$. [15] provided a LR criterion for testing the null $H_k : r_{r+1} = \cdots = r_{\text{min(k,p)}m} = 0$. Under the null hypothesis

$$BA = \left[\frac{m(k+p)+1}{2} - T\right] ln \prod_{i=r^*+1}^{\text{min(k,p)}m} \left(1 - \hat{r}_i^2\right) \xrightarrow{d} \chi^2_{(mk-r^*)(mp-r^*)}$$

3.2 Information criteria

[16] and [17] showed that the number of linearly independent components of the projections of the previously defined $\boldsymbol{y}_t^+$ onto the linear space spanned by the components of $\boldsymbol{y}_t^-$ is identical to the number of nonzero canonical correlations between $\boldsymbol{y}_t^+$ and $\boldsymbol{y}_t^-$. When $\boldsymbol{y}_t$ is Gaussian, canonical correlation analysis between $\boldsymbol{y}_t^+$ and $\boldsymbol{y}_t^-$ is equivalent to maximum likelihood estimation of the linear model: $\boldsymbol{y}_t^+ = \boldsymbol{\Psi}\boldsymbol{y}_t^- + \boldsymbol{\varepsilon}_t$ See [14]. The number of free parameters for this model is: $F(r^*) = \{[s^+(s^+ + 1)]/2\} + \{[s^-(s^- + 1)]/2\} + r^*(s^+ + s^- - r^*)$ where s^+ denotes the dimension of the vector $\boldsymbol{y}_t^+$ and s^- denotes the dimension of $\boldsymbol{y}_t^-$. The first and second term are the number of free parameters of the covariance matrices of $\boldsymbol{y}_t^+$ and $\boldsymbol{y}_t^-$ respectively, and the last term gives the number of free parameters in matrix $\boldsymbol{\Psi}$. [17] defined a loss function for model fitting, and by extension rank determination, as:

$$AIC(r^*) = T \ln \prod_{i=1}^{r^*} (1 - \hat{\rho}_i^2) + 2F(r^*)$$

where $\hat{\rho}_i$ are the estimated canonical correlation coefficients previously defined. This criterion penalises models with a large number of parameters, and by extension large rank, and favours parsimonious representations. Note that, as discussed in [14, pp. 499], when $\rho_i = 0$ then

$\rho_i^2 = O_p(T^{-1})$ implying that $\ln(1-\rho_i^2) = O_p(T^{-1})$ where $O_p(.)$ denotes order in probability. This suggests that there is a positive probability that AIC will be minimised for some $r^* > r^0$ since the probability that $T\sum_{i=r^0+1}^{r^*} \ln(1-\hat{\rho}_i^2) < 2(F(r^0) - F(r^*))$ may be greater than zero. Therefore the estimated rank will not converge in probability to r^0 when AIC is used.

Alternatively, one could use an estimate of the covariance matrix of the error term $\boldsymbol{e}_t$ for a model specified for a certain rank r^*, and denoted as $\hat{\boldsymbol{\Omega}}_{r^*}$ in the expression to be minimised. In order to compute $\hat{\boldsymbol{\Omega}}_{r^*}$ use is made of the regression based method for the estimation of the system parameter matrices described in the appendix. Let's denote this loss function to be minimised as:

$$AIC_c(r^*) = T\ln\left|\hat{\boldsymbol{\Omega}}_{r^*}\right| + 2F(r^*)$$

[18] suggested an alternative penalty on increasing the number of parameters, and in the present paper we explore the performance of this criterion in searching for the rank. His loss function is:

$$BIC(r^*) = T\ln\prod_{i=1}^{r^*}(1-\hat{\rho}_i^2) + \ln(T)F(r^*)$$

The penalty used by BIC is much more severe than that used by AIC. In fact it is easy to see that the rank estimate obtained by the BIC will converge in probability to r^0. Nevertheless, BIC is likely to underestimate the rank in small samples. Similar to the AIC case we could also use as a loss function

$$BIC_c(r^*) = T\ln\left|\hat{\boldsymbol{\Omega}}_{r^*}\right| + \ln(T)F(r^*)$$

[19] suggested a further alternative penalty on increasing the number of parameters. The loss function correponding to this penalty are:

$$HQ(r^*) = T\ln\prod_{i=1}^{r^*}(1-\hat{\rho}_i^2) + 2*\ln(\ln(T))F(r^*)$$

$$HQ_c(r^*) = T\ln\left|\hat{\boldsymbol{\Omega}}_{r^*}\right| + 2*\ln(\ln(T))F(r^*)$$

3.3 Bootstrap Testing Procedures

Theoretical work on the advantages of bootstrapping pivotal statistics for time series is limited. [20] and [21] show that, for univariate AR and MA models with independent innovations, the approximation error of the bootstrap distribution of least squares estimators of the AR and MA coefficients is $o_p(T^{-1/2})$ whereas it is $O_p(T^{-1/2})$ for the first order limiting distribution. One might conjecture that an extension of the results of [20] and [22] to a multivariate framework should provide a rigorous proof for similar advantages of the bootstrap distribution for statistics obtained from state space models, in particular for bootstrap versions of the above B and GE test statistics; see also [23]. Of course, given the asymptotic nature of the above arguments, these conclusions may prove misleading for the performance of the bootstrap in finite samples; see, for example, [24]. The bootstrap estimate of the distribution of test statistic i, $i = GE$, CC and BA, is given by $F_i(x) = N^{-1}\sum_{n=1}^{N}\mathbf{1}(\tau_i^n < x)$, where τ_i^n denotes a generic bootstrap test statistic, $n = 1, ..., N$, $\mathbf{1}(.)$ is the indicator function and N is the number of replications. For each sequence of $m-1$ tests of the null hypotheses $H_{0,r} : rk(\boldsymbol{H}) = r$, $r = 1, \ldots, m-1$, we need $m-1$ bootstrap distributions. Let a generic test statistic of the hypothesis $H_{0,r}$ be denoted as $\psi_r(\hat{\boldsymbol{H}})$ where $\hat{\boldsymbol{H}}$ is the estimated Hankel matrix on which the test is carried out. Also, let the general form of the test statistic, as given in (10) for example, be denoted by $\psi_r(\hat{\boldsymbol{H}}, \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ where $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ are the true value of the coefficient matrices of the state space model on which the test is carried out. For the null hypothesis $H_{0,r^*} : rk(\boldsymbol{H}) = r^*$, we estimate the data using an state space model with order fixed to r^*. A regression based approach to estimate the coefficient matrices is reviewed in the appendix. Use of maximum likelihood estimation in connection with the bootstrap is not feasible for the purpose of the Monte Carlo exercise due to limited computer power. The estimates of the coefficient matrices denoted by $\hat{\boldsymbol{A}}$, $\hat{\boldsymbol{B}}$ and $\hat{\boldsymbol{C}}$ are then used to generate the bootstrap samples, denoted by $\boldsymbol{y}_t^{bo}$, $t = 1, \ldots, T$, $bo = 1, \ldots, B$ where B is the number of bootstrap replications. The errors used to generate these samples are obtained by random resampling from the residuals of the estimated state space model. Consistency of the estimation of the state space model guarantees that these residuals will converge in probability to the error terms of the model. The bootstrap samples are then used to generate the test statistics for the test of H_{0,r^*}, $\psi_{r^*}(\hat{\boldsymbol{H}}^{bo}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{C}})$, $bo = 1, \ldots, B$. [2] The bootstrap versions of the GE, CC and BA tests using this bootstrapping method will be referred to as GE_1, CC_1 and BA_1 respectively.

4 Monte Carlo Design

All the procedures discussed above will be considered. Firstly the statistical tests of rank, i.e. GE, GE_1, CC, CC_1, BA and BA_1, and secondly the information criteria, i.e. AIC, AIC_c, BIC, BIC_c, HQ and HQ_c.

We concentrate on a state space model like that in equation (1). The dimension of the vector series $\boldsymbol{y}_t$ is fixed to three. The rank of the Hankel matrix is equal to the dimension of the state vector $\boldsymbol{s}_t$ which is fixed to three as well. Matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{\Omega}$ are built as follows. $\boldsymbol{B}$ and $\boldsymbol{C}$ are (3×3) matrices of values drawn from a standard normal distribution; $\boldsymbol{\Omega}$ is fixed to an identity matrix of dimension (3×3). Note that $\boldsymbol{A}$ is a key matrix to explain

[2] We have also studied the use of nonparametric bootstrap methods, but for the sample sizes under study in this paper, results were not satisfactory and therefore are not reported.

the dynamics of the series; the degree of persistence of shocks will depend on the eigenvalues of $\boldsymbol{A}$. To control the experiment for this, we have chosen to build $\boldsymbol{A} = \tilde{\boldsymbol{E}}\boldsymbol{\Lambda}\tilde{\boldsymbol{E}}'$. $\boldsymbol{\Lambda}$ is a 3×3 quasi upper triangular matrix; the last element of the diagonal corresponds to the modulus assigned to that experiment, and the 2×2 block matrix in the left upper corner is computed in such a way that the modulus of the complex pair of eigenvalues of this 2×2 block is also equal to the modulus assigned to the eigenvalues of that experiment; the remaining values are fixed to a value of one. $\tilde{\boldsymbol{E}}$ is an orthonormal matrix generated from a standard normal matrix using Gram-Schmidt orthogonalisation. For the Monte Carlo experiments presented below, 3 different moduli making 3 alternative experiments will be used, namely experiment a, large eigenvalues, with moduli given by (0.8,0.8,0.8); experiment b, small eigenvalues, (0.4,0.4,0.4) and experiment c, uneven eigenvalues, (0.8,0.8,0.2), which allows to check for the robustness of the procedures when one of the eigenvalues is very small.

Using these matrices and random normal disturbances generated by the GAUSS random number generator with an identity covariance matrix a sample from a process following (1) is obtained. The sample sizes considered are: 200, 400 and 600. For each simulated sample, 200 initial observations have been discarded, to minimise the effect of starting values which are set to zero. 100 replications for 10 different random sets of matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are carried out, making a total of 1000 replications. The reason for this design is to reduce the dependence of the results on arbitrary parameter values. In each bootstrap application, for reasons of computational feasibility, only 99 replications were carried out. The experiments have been carried out for equal values for the truncation parameters k and p, i.e. $k = p$, and for values $k = 2$ and $k = 3$. The notation for the different experiments in the tables will also indicate the dimension of $k = p$, i.e. $a2$ refers to an experiment with moduli of eigenvalues a and $k = p = 2$. All computations were carried out using the GAUSS programming language.

5 Monte Carlo Results

For each of the experiments, tables 1 and 2 present the probability of rejection of the null hypothesis $H_{0,r} : rk(\boldsymbol{H}) = r$, $r = 1, \ldots, r^*$ where r^* is the true rank, using critical values associated with a nominal 0.05 significance level with * and ** indicating a significant difference between estimated actual and nominal size at the 0.05 and 0.01 levels respectively. As expected, performance worsens with respect to the size of the Hankel matrix, for experiments with smaller moduli of the eigenvalues, and when the sample size is small. But there are some interesting disparities in the relative performance of the alternative methods under study.

The size properties of GE and CC are not very good particularly for three blocks in the Hankel matrix, $k = 3$. The size of their bootstrapped versions, GE_1 and CC_1 is clearly superior. The price to pay for this improvement comes in terms of less power; this is of serious concern for the CC_1 test, but less so for the GE_1 test. The BA method appears to be less sensitive than the others to all dimensions explored in the Monte Carlo exercise. In particular BA is much more robust than the others to large number of blocks in the Hankel matrix, $k = 3$. The bootstrapped version of the Barlett test, BA_1, helps significantly to improve the size properties of its asymptotic counterpart when the number of blocks in the Hankel matrix is large, i.e. experiments $a3$, $b3$ and $c3$. Furthermore, the loss in power of BA_1 is small for all experiments compared to the asymptotic test. Another interesting feature of the results is the nonmonotonicity of the power function with respet to sample size. This is likely to be the result of varying randomly the coefficient matrices.

Table 3 presents results for the mean, standard deviation and root mean square error (RMSE) of the rank estimates over the Monte Carlo exercises for the Statistical tests. Similar results for the Information Criterion Methods are given in table 4. As for the statistical tests, the performance of CC_1 is clearly worse than the asymptotic version of this test, particularly when the number of blocks is $k = 3$ and for experiments of type b and c. The performance of GE_1 is in line with its asymptotic equivalent GE. Generally speaking the performance of BA and BA_1 is good for all experiments, coming almost always as best.

Results presented in table 4 show that information criteria of the type AIC_c, BIC_c and HQ_c display similar results to their equivalent counterparts AIC, BIC and HQ. The performance of information criteria methods don't deteriorate very much when the number of blocks in the Hankel increases. But information criteria methods appear more sensitive than statistical tests to sample size. This is particularly the case for the BIC method which displays a poor performance for a sample size of 200. Among the information criteria methods BIC and HQ display the best results for sample sizes of 400 and 600. Generally speaking, and ignoring the CC_1 results, the performance of statistical tests is best compared to information criteria methods for most experiments. Statistical tests such as GE_1, BA and BA_1 provide very good results, and are shown to be less sensitive to all the dimensions explored in the Monte Carlo exercies. The GE tests should only be used for small values for k. Tables 5 and 6 present the distribution of the estimated rank for the statistical tests and the information criteria. Results highlight the same conclusions.

6 Conclusion

This paper has studied the performance of statistical tests of rank and their bootstrap versions against information criteria methods in determining the rank of the Hankel matrix.

This study has been conducted through Monte Carlo simulations. The results presented show that the bootstrap served to improve the size properties of the tests. Further, statistical tests, in particular GE_1, BA and BA_1 were shown to be superior to information criteria methods.

Appendix: Estimation of System Matrices

The regression approach described in this appendix is that proposed by [10]. This method relies on a rank approximation of the Hankel matrix and a decomposition of this approximation into an observability and controllability matrix. The rank approximation is achieved by means of its singular value decomposition, i.e. $\hat{\boldsymbol{H}} = \hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\hat{\boldsymbol{V}}$ where $\hat{\boldsymbol{U}}$ and $\hat{\boldsymbol{V}}$ are orthonormal matrices, and $\hat{\boldsymbol{\Sigma}}$ is a rectangular matrix of zero nondiagonal elements and with values $\hat{\sigma}_1 \geq \hat{\sigma}_2 \geq \ldots \geq \hat{\sigma}_{\mathrm{min(k,p)}m}$. Once the rank has been estimated, say r, a r rank approximation to the Hankel matrix $\boldsymbol{H}$ is given by $\hat{\boldsymbol{H}}_r = \hat{\boldsymbol{U}}_r\hat{\boldsymbol{\Sigma}}_r\hat{\boldsymbol{V}}_r$ where only the largest r singular values and singular vectors are included.

A Decomposition of the Hankel covariance matrix into an observability matrix $\boldsymbol{O}$ and a controllability matrix $\boldsymbol{K}$ is defined as: $\hat{\boldsymbol{H}}_r = \hat{\boldsymbol{O}}_r\hat{\boldsymbol{K}}_r$ and it follows that $\hat{\boldsymbol{K}}_r^+ = \hat{\boldsymbol{V}}_r\hat{\boldsymbol{\Sigma}}_r^{-\frac{1}{2}}$ and $\hat{\boldsymbol{O}}_r^+ = \hat{\boldsymbol{\Sigma}}_r^{-\frac{1}{2}}\hat{\boldsymbol{U}}_r'$ where $^+$ denotes generalised inverses.

Given a r rank approximation to $\hat{\boldsymbol{H}}^b$, and estimate of the state vector is given by $\hat{\boldsymbol{s}}_t = \hat{\boldsymbol{K}}_r^b\boldsymbol{y}_-^t$. This allows to estimate $\boldsymbol{C}$ from the regression:

$$\boldsymbol{y}_t = \boldsymbol{C}\hat{\boldsymbol{s}}_t + \boldsymbol{e}_t$$

then $\hat{\boldsymbol{C}} = (\sum_{t=1}^T \boldsymbol{y}_t\hat{\boldsymbol{s}}_t')(\sum_{t=1}^T \hat{\boldsymbol{s}}_t\hat{\boldsymbol{s}}_t')^{-1}$, this regression will also provide $\hat{\boldsymbol{\Omega}} = T^{-1}(\sum_{t=1}^T \hat{\boldsymbol{e}}_t\hat{\boldsymbol{e}}_t')$, where $\hat{\boldsymbol{e}}_t$ are the estimated residuals. Finally, matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ can be estimated from the regression:

$$\hat{\boldsymbol{s}}_{t+1} = \boldsymbol{A}\hat{\boldsymbol{s}}_t + \boldsymbol{B}\hat{\boldsymbol{e}}_t + \boldsymbol{\varepsilon}_t$$

References

[1] H. Akaike, “Canonical correlation analysis of time series and the use of and information criterion,” in *System Identification: Advances and Case Studies* (R. K. Mehra and D. G. Lainiotis, eds.), New York, Academic Press, 1976.

[2] T. Kailath, *Linear Systems*. Englewood Cliffs, N. J.: Prentice-Hall, 1980.

[3] D. Bauer, “Order estimation for subspace methods.” Institute f. Econometrics, Operation Research and System Theory, TU Wien, Austria, Working Paper No 190, 2000.

[4] D. Bauer, “Some asymptotic theory for the estimation of linear systems using maximum likelihood methods or subspace algorithms.” Institute f. Econometrics, Operation Research and System Theory, TU Wien, Austria, PhD Thesis, 1998.

[5] J. G. Cragg and S. G. Donald, “On the asymptotic properties of ldu-based tests of the rank of a matrix,” *Journal of the American Statistical Association*, vol. 91, no. 435, pp. 1301–1309, 1996.

[6] J. G. Cragg and S. G. Donald, “Inferring the rank of a matrix,” *Journal of Econometrics*, vol. 76, no. 1–2, pp. 223–250, 1997.

[7] L. Gill and A. Lewbel, “Testing the rank and definiteness of estimated matrices with applications to factor, state-space and arma models,” *Journal of the American Statistical Association*, vol. 87, no. 419, pp. 766–776, 1992.

[8] J. M. Robin and R. J. Smith, “Tests of rank,” *Econometric Theory*, vol. 16, pp. 151–175, 2000.

[9] P. L. Faurre, “Stochastic realization algorithms,” in *System Identification: Advances and Case Studies* (R. K. Mehra and D. G. Lainiotis, eds.), New York, Academic Press, 1976.

[10] W. E. Larimore, “System identification, reduced order filters and modeling via canonical variate analysis,” in *Proc. 1983 Amer. Control Conference 2* (H. S. Rao and P. Dorato, eds.), Piscataway, NJ. IEEE Service Center, 1983.

[11] Y. Hosoya, “Hierarchical statistical models and a generalised likelihood ratio test,” *Journal of the Royal Statistical Society B*, vol. 51, no. 3, pp. 435–448, 1989.

[12] G. H. Golub and C. F. V. Loan, *Matrix Computation*. North Oxford Academic, 1983.

[13] G. Camba-Mendez and G. Kapetanios, "Testing the rank of the hankel covariance matrix: A statistical approach," *IEEE Transactions on Automatic Control*, vol. 46, no. 2, pp. 331–336, 2001.

[14] T. W. Anderson, *Introduction to Multivariate Statistical Analysis*. John Wiley and Sons, New York, 1984.

[15] M. Bartlett, "Multivariate analysis," *Journal of the Royal Statistical Society, Series B*, vol. 9, pp. 176–197, 1947.

[16] H. Akaike, "Stochastic theory of minimal realisations," *IEEE Transactions on Automatic Control*, vol. AC-19, no. 6, pp. 667–674, 1974.

[17] H. Akaike, "Canonical correlation analysis of time series and the use of an information criterion," in *System Identification* (R. Mehra and D. Lainiotis, eds.), Academic Press, 1976.

[18] G. Schwarz, "Estimating the dimension of a model," *Annals of Statistics*, vol. 6, pp. 461–464, 1978.

[19] E. J. Hannan and B. G. Quinn, "The determination of the order of an autoregression," *Journal of the Royal Statistical Society. Series B*, vol. 41, pp. 190–195, 1979.

[20] A. Bose, "Edgeworth correction by bootstrap in autoregressions," *Annals of Statistics*, vol. 16, pp. 1709–1722, 1988.

[21] A. Bose, "Bootstrap in moving average models," *Annals of the Institute of Statistical Mathematics*, vol. 40, pp. 752–768, 1990.

[22] F. Gotze and C. Hipp, "Asymptotic expansions for sums of weakly dependent random vectors," *Z. Wahrsch. verw. Gebiete*, vol. 64, pp. 211–239, 1983.

[23] P. Hall and J. L. Horowitz, "Bootstrap critial values for tests based on gmm estimators," *Econometrica*, vol. 64, no. 4, pp. 891–916, 1996.

[24] L. Kilian, "Pitfalls in constructing bootstrap confidence intervals for asymptotically pivotal statistics." Working Paper, University of Michigan, 1998.

Table 1: Size and Power of Rank Tests. (I).

Exp	H_0	Tests	Sample Size			Exp	H_0	Tests	Sample Size		
			200	400	600				200	400	600
		GE	1.000	1.000	1.000			GE	1.000	1.000	1.000 =
		GE_1	0.999	1.000	1.000			GE_1	0.993	0.994	0.999
		CC	1.000	1.000	1.000			CC	1.000	1.000	1.000
	$r=1$	CC_1	0.878	0.964	0.998		$r=1$	CC_1	0.327	0.184	0.383
		BA	1.000	1.000	1.000			BA	1.000	1.000	1.000
		BA_1	1.000	1.000	1.000			BA_1	1.000	1.000	1.000
		GE	0.866	0.794	0.997			GE	1.000	1.000	1.000
		GE_1	0.881	0.797	1.000			GE_1	0.555	0.525	0.664
		CC	0.954	0.795	0.995			CC	1.000	1.000	1.000
$a2$	$r=2$	CC_1	0.845	0.804	0.997	$a3$	$r=2$	CC_1	0.026	0.052	0.022
		BA	0.999	0.833	0.999			BA	1.000	1.000	1.000
		BA_1	0.999	0.827	0.998			BA_1	0.999	1.000	1.000
		GE	0.010**	0.005**	0.017**			GE	0.192**	0.094**	0.135**
		GE_1	0.066*	0.053	0.048			GE_1	0.093**	0.032**	0.045
		CC	0.018**	0.015**	0.019**			CC	0.672**	0.473**	0.556**
	$r=3$	CC_1	0.064	0.046	0.046		$r=3$	CC_1	0.080**	0.068*	0.058
		BA	0.071**	0.064	0.065			BA	0.116**	0.121**	0.110**
		BA_1	0.073**	0.063	0.060			BA_1	0.077**	0.063	0.056
		GE	1.000	1.000	1.000			GE	1.000	1.000	1.000
		GE_1	0.972	0.999	0.999			GE_1	0.909	1.000	1.000
		CC	1.000	1.000	1.000			CC	1.000	1.000	1.000
	$r=1$	CC_1	0.896	0.983	0.999		$r=1$	CC_1	0.332	0.262	0.280
		BA	0.998	1.000	1.000			BA	1.000	1.000	1.000
		BA_1	0.996	1.000	1.000			BA_1	0.999	1.000	1.000
		GE	0.773	0.794	0.848			GE	0.963	0.999	0.954
		GE_1	0.628	0.716	0.793			GE_1	0.333	0.630	0.419
		CC	0.694	0.701	0.889			CC	0.999	1.000	1.000
$b2$	$r=2$	CC_1	0.538	0.708	0.809	$b3$	$r=2$	CC_1	0.053	0.032	0.050
		BA	0.710	0.897	0.890			BA	0.808	0.975	0.846
		BA_1	0.686	0.880	0.885			BA_1	0.746	0.959	0.799
		GE	0.083**	0.155**	0.056			GE	0.473**	0.626**	0.493**
		GE_1	0.085**	0.084**	0.067*			GE_1	0.141**	0.175**	0.133**
		CC	0.086**	0.052	0.044			CC	0.690**	0.741**	0.544**
	$r=3$	CC_1	0.123**	0.089**	0.066*		$r=3$	CC_1	0.054	0.077**	0.058
		BA	0.042	0.055	0.050			BA	0.104**	0.104**	0.088**
		BA_1	0.080**	0.062	0.053			BA_1	0.107**	0.083**	0.078**

Table 2: Size and Power of Rank Tests. (II).

Exp	H_0	Tests	Sample Size 200	400	600	Exp	H_0	Tests	Sample Size 200	400	600
$c2$	$r=1$	GE	1.000	1.000	1.000	$c3$	$r=1$	GE	1.000	0.996	1.000
		GE_1	1.000	1.000	1.000			GE_1	0.975	0.977	1.000
		CC	1.000	1.000	1.000			CC	1.000	1.000	1.000
		CC_1	0.984	0.997	1.000			CC_1	0.455	0.461	0.000
		BA	1.000	1.000	1.000			BA	1.000	1.000	1.000
		BA_1	1.000	1.000	1.000			BA_1	1.000	1.000	1.000
	$r=2$	GE	0.697	0.819	0.821		$r=2$	GE	0.922	0.961	1.000
		GE_1	0.713	0.843	0.854			GE_1	0.293	0.599	0.825
		CC	0.638	0.860	0.831			CC	1.000	1.000	1.000
		CC_1	0.620	0.858	0.826			CC_1	0.040	0.044	0.011
		BA	0.696	0.839	0.930			BA	0.919	0.922	1.000
		BA_1	0.673	0.814	0.911			BA_1	0.887	0.910	1.000
	$r=3$	GE	0.030**	0.055	0.028**		$r=3$	GE	0.322**	0.372**	0.883**
		GE_1	0.054	0.070*	0.083**			GE_1	0.089**	0.149**	0.079**
		CC	0.014**	0.021**	0.066*			CC	0.499**	0.651**	0.819**
		CC_1	0.070*	0.050	0.093**			CC_1	0.064	0.086**	0.048
		BA	0.050	0.064	0.070*			BA	0.110**	0.125**	0.090**
		BA_1	0.070*	0.059	0.076**			BA_1	0.081**	0.075**	0.049

Table 3: Mean, St. Deviation and RMSE of Estimated Rank. Statistical Methods.

		Sample Size								
		200			400			600		
Exp.	Tests	$\overline{x}^a$	σ	$\Re$	$\overline{x}$	σ	$\Re$	$\overline{x}$	σ	$\Re$
	GE	2.88	0.36	0.38	2.80	0.41	0.46	3.02	0.16	0.16
	GE_1	2.95	0.46	0.46	2.86	0.53	0.55	3.06	0.29	0.29
	CC	2.97	0.25	0.25	2.81	0.43	0.47	3.01	0.15	0.15
$a2$	CC_1	2.69	0.79	0.85	2.78	0.58	0.62	3.04	0.26	0.26
	BA	3.08	0.30	0.31	2.90	0.49	0.50	3.07	0.26	0.27
	BA_1	3.10	0.40	0.41	2.91	0.56	0.57	3.08	0.34	0.35
	GE	3.20	0.41	0.46	3.10	0.30	0.31	3.14	0.35	0.38
	GE_1	2.60	0.60	0.72	2.55	0.58	0.73	2.68	0.51	0.60
	CC	3.76	0.60	0.96	3.48	0.51	0.70	3.58	0.54	0.79
$a3$	CC_1	1.33	0.48	1.74	1.19	0.41	1.86	1.40	0.53	1.69
	BA	3.13	0.39	0.41	3.15	0.43	0.45	3.12	0.38	0.39
	BA_1	3.13	0.55	0.56	3.13	0.58	0.59	3.11	0.56	0.57
	GE	2.86	0.54	0.56	2.96	0.62	0.62	2.91	0.45	0.46
	GE_1	2.68	0.67	0.74	2.79	0.58	0.62	2.85	0.51	0.53
	CC	2.78	0.60	0.63	2.75	0.54	0.60	2.94	0.41	0.41
$b2$	CC_1	2.51	0.83	0.96	2.77	0.63	0.67	2.88	0.52	0.54
	BA	2.75	0.53	0.59	2.96	0.41	0.41	2.94	0.40	0.41
	BA_1	2.78	0.66	0.69	2.96	0.49	0.50	2.96	0.48	0.48
	GE	3.46	0.62	0.78	3.70	0.60	0.92	3.49	0.65	0.81
	GE_1	2.31	0.74	1.01	2.74	0.67	0.72	2.50	0.67	0.84
	CC	3.71	0.51	0.88	3.89	0.67	1.11	3.56	0.53	0.77
$b3$	CC_1	1.36	0.53	1.73	1.27	0.46	1.79	1.29	0.47	1.78
	BA	2.93	0.59	0.59	3.09	0.40	0.41	2.94	0.52	0.52
	BA_1	2.92	0.81	0.82	3.10	0.58	0.59	2.91	0.67	0.67
	GE	2.73	0.51	0.58	2.88	0.47	0.49	2.85	0.43	0.46
	GE_1	2.77	0.54	0.59	2.92	0.49	0.49	2.95	0.51	0.52
	CC	2.65	0.51	0.61	2.88	0.38	0.40	2.90	0.48	0.49
$c2$	CC_1	2.66	0.63	0.72	2.91	0.48	0.49	2.92	0.54	0.54
	BA	2.75	0.55	0.61	2.91	0.47	0.48	3.00	0.39	0.39
	BA_1	2.76	0.64	0.68	2.89	0.54	0.55	3.01	0.48	0.48
	GE	3.26	0.62	0.68	3.34	0.60	0.69	3.95	0.43	1.05
	GE_1	2.29	0.57	0.90	2.66	0.68	0.76	2.90	0.50	0.51
	CC	3.51	0.53	0.74	3.66	0.49	0.82	3.84	0.43	0.94
$c3$	CC_1	1.48	0.54	1.62	1.49	0.56	1.61	1.00	0.00	2.00
	BA	3.05	0.51	0.51	3.06	0.50	0.50	3.10	0.35	0.37
	BA_1	3.04	0.71	0.71	3.04	0.63	0.63	3.08	0.43	0.43

[a] $\overline{x}$, σ and $\Re$ denote respectively the mean standard error and RMSE of the estimated rank over the Monte Carlo samples.

Table 4: Mean, St. Deviation and RMSE of Estimated Rank. Information Criteria Methods

		Sample Size								
		200			400			600		
Exp.	Tests	$\overline{x}$[a]	σ	$\Re$	$\overline{x}$	σ	$\Re$	$\overline{x}$	σ	$\Re$
	AIC	3.16	0.58	0.60	3.21	0.47	0.51	3.17	0.42	0.45
	AIC_c	3.01	0.43	0.43	3.08	0.28	0.29	3.26	0.46	0.53
$a2$	BIC	2.90	0.31	0.32	3.00	0.00	0.00	2.99	0.08	0.08
	BIC_c	2.90	0.34	0.35	3.01	0.08	0.08	3.08	0.31	0.32
	HQ	2.93	0.34	0.35	3.01	0.09	0.09	3.01	0.08	0.08
	HQ_c	2.95	0.37	0.37	3.02	0.14	0.14	3.16	0.38	0.41
	AIC	3.24	0.53	0.58	3.21	0.61	0.65	3.23	0.50	0.55
	AIC_c	3.05	0.62	0.63	3.19	0.73	0.76	3.46	0.71	0.85
$a3$	BIC	2.86	0.34	0.37	2.84	0.37	0.40	3.00	0.04	0.04
	BIC_c	2.68	0.59	0.67	2.89	0.52	0.54	3.12	0.58	0.59
	HQ	2.97	0.19	0.19	2.90	0.31	0.33	3.00	0.04	0.04
	HQ_c	2.84	0.57	0.59	2.97	0.62	0.62	3.29	0.61	0.67
	AIC	2.96	0.58	0.58	3.08	0.53	0.54	3.15	0.52	0.54
	AIC_c	2.80	0.69	0.72	3.17	0.58	0.61	2.99	0.51	0.51
$b2$	BIC	2.41	0.64	0.87	2.79	0.41	0.46	2.72	0.45	0.53
	BIC_c	2.40	0.75	0.96	2.96	0.49	0.49	2.74	0.50	0.56
	HQ	2.64	0.52	0.63	2.90	0.33	0.34	2.90	0.33	0.34
	HQ_c	2.59	0.66	0.78	3.03	0.53	0.53	2.83	0.50	0.53
	AIC	3.03	0.58	0.58	3.08	0.54	0.55	3.07	0.58	0.58
	AIC_c	2.69	0.96	1.01	3.27	0.83	0.87	3.10	0.63	0.63
$b3$	BIC	2.08	0.71	1.16	2.54	0.50	0.68	2.70	0.46	0.55
	BIC_c	2.15	0.85	1.20	2.42	0.76	0.96	2.69	0.51	0.60
	HQ	2.53	0.59	0.75	2.78	0.43	0.48	2.78	0.42	0.47
	HQ_c	2.41	0.90	1.08	2.94	0.76	0.77	2.88	0.51	0.52
	AIC	3.15	0.53	0.55	3.17	0.48	0.51	3.20	0.48	0.52
	AIC_c	3.08	0.50	0.50	3.05	0.53	0.53	3.15	0.57	0.59
$c2$	BIC	2.80	0.41	0.45	2.77	0.42	0.48	2.88	0.32	0.34
	BIC_c	2.82	0.46	0.49	2.70	0.50	0.58	2.78	0.57	0.61
	HQ	2.91	0.36	0.37	2.91	0.31	0.33	2.97	0.21	0.21
	HQ_c	2.93	0.47	0.47	2.87	0.49	0.51	2.94	0.55	0.55
	AIC	3.23	0.53	0.58	3.21	0.53	0.57	3.23	0.51	0.56
	AIC_c	3.32	0.67	0.75	3.51	0.87	1.01	3.28	0.55	0.62
$c3$	BIC	2.68	0.47	0.57	2.60	0.49	0.63	2.85	0.35	0.38
	BIC_c	2.89	0.51	0.52	2.92	0.74	0.74	3.06	0.48	0.48
	HQ	2.92	0.28	0.29	2.81	0.40	0.44	2.98	0.16	0.16
	HQ_c	3.12	0.54	0.56	3.21	0.78	0.80	3.16	0.48	0.50

[a] $\overline{x}$, σ and $\Re$ denote respectively the mean standard error and RMSE of the estimated rank over the Monte Carlo samples.

Table 5: Distribution of Estimated Rank (%). Statistical Tests[a]

Exp.	Tests	200: 1	200: 2	200: 3	200: 4	400: 1	400: 2	400: 3	400: 4	600: 1	600: 2	600: 3	600: 4
$a2$	GE	0.0	13.4	85.7	0.9	0.0	20.6	79.2	0.2	0.0	0.3	98.0	1.5
	GE_1	0.1	11.9	82.0	4.9	0.0	20.3	75.2	3.2	0.0	0.0	95.2	3.7
	CC	0.0	4.6	93.6	1.8	0.0	20.5	78.1	1.4	0.0	0.5	97.6	1.9
	CC_1	12.2	13.3	69.6	3.7	3.6	19.1	73.2	3.7	0.2	0.3	94.9	4.2
	BA	0.0	0.1	92.8	6.6	0.0	16.7	76.9	5.8	0.0	0.1	93.4	6.4
	BA_1	0.0	0.1	92.6	5.4	0.0	17.3	76.7	3.9	0.0	0.2	93.8	4.1
$a3$	GE	0.0	0.0	80.8	18.6	0.0	0.0	90.6	9.3	0.0	0.0	86.5	13.3
	GE_1	0.7	43.8	51.0	4.1	0.6	46.9	50.1	1.8	0.1	33.5	64.5	1.7
	CC	0.0	0.0	32.8	58.9	0.0	0.0	52.7	46.8	0.0	0.0	44.4	53.3
	CC_1	67.3	32.3	0.4	0.0	81.6	17.8	0.6	0.0	61.7	36.8	1.3	0.1
	BA	0.0	0.0	88.4	9.9	0.0	0.0	87.9	9.9	0.0	0.0	89.0	9.8
	BA_1	0.0	0.1	92.2	5.2	0.0	0.0	93.7	3.1	0.0	0.0	94.4	2.9
$b2$	GE	0.0	22.7	69.1	8.0	0.0	20.6	64.1	14.5	0.0	15.2	79.2	5.4
	GE_1	2.8	34.5	55.8	6.1	0.1	28.3	64.5	6.4	0.1	20.6	74.2	4.5
	CC	0.0	30.6	60.8	8.3	0.0	29.9	65.0	4.9	0.0	11.1	84.5	3.9
	CC_1	10.4	38.5	42.0	8.0	1.7	28.1	62.6	6.8	0.1	19.0	75.5	4.1
	BA	0.2	28.8	66.8	4.0	0.0	10.3	84.2	5.1	0.0	11.0	84.0	4.8
	BA_1	0.4	31.0	60.9	5.9	0.0	12.0	81.8	4.9	0.0	11.5	83.2	3.6
$b3$	GE	0.0	3.7	49.4	43.7	0.0	0.1	37.3	55.6	0.0	4.6	46.1	45.3
	GE_1	9.1	57.6	26.3	6.8	0.0	37.0	53.3	8.7	0.0	58.1	35.4	5.0
	CC	0.0	0.1	30.9	66.4	0.0	0.0	25.9	61.5	0.0	0.0	45.6	52.6
	CC_1	66.8	31.0	2.1	0.1	73.8	25.5	0.7	0.0	72.0	27.4	0.6	0.0
	BA	0.0	19.2	70.4	8.7	0.0	2.5	87.1	9.2	0.0	15.4	75.8	8.0
	BA_1	0.1	25.3	63.9	7.2	0.0	4.1	87.6	5.2	0.0	20.1	72.6	5.0
$c2$	GE	0.0	30.3	66.8	2.9	0.0	18.1	76.4	5.4	0.0	17.9	79.3	2.7
	GE_1	0.0	28.7	66.6	4.2	0.0	15.7	77.7	5.9	0.0	14.6	77.2	7.1
	CC	0.0	36.2	62.4	1.4	0.0	14.0	83.9	2.1	0.0	16.9	76.5	6.5
	CC_1	1.6	37.3	55.9	4.3	0.3	14.2	80.8	4.0	0.0	17.4	74.0	7.4
	BA	0.0	30.4	64.6	4.7	0.0	16.1	77.5	6.1	0.0	7.0	86.0	6.5
	BA_1	0.0	32.7	60.6	5.1	0.0	18.6	75.5	4.6	0.0	8.9	83.6	5.5
$c3$	GE	0.0	7.8	60.1	30.3	0.4	3.9	58.9	35.3	0.0	0.0	11.7	81.3
	GE_1	2.5	68.2	27.0	1.9	2.3	38.2	51.7	7.0	0.0	17.5	76.2	5.5
	CC	0.0	0.0	50.1	48.5	0.0	0.0	34.9	64.5	0.0	0.0	18.1	79.7
	CC_1	54.5	43.3	2.2	0.0	53.9	43.9	1.7	0.5	100.0	0.0	0.0	0.0
	BA	0.0	8.1	80.9	9.0	0.0	7.8	79.7	10.8	0.0	0.0	91.0	7.8
	BA_1	0.0	11.3	80.8	4.3	0.0	9.0	83.5	4.7	0.0	0.0	95.1	3.4

[a]Values are in (%). *Exp* denotes experiment according to what is explained in the text.

Table 6: Distribution of Estimated Rank (%). Information Criteria Methods.[a]

Exp.	Tests	Sample Size											
		200				400				600			
		1	2	3	4	1	2	3	4	1	2	3	4
	AIC	0.0	6.3	74.4	16.9	0.0	0.0	82.1	15.5	0.0	0.0	84.5	13.9
	AIC_c	0.0	8.6	82.0	9.3	0.0	0.0	91.9	7.9	0.0	0.3	74.1	24.8
$a2$	BIC	0.0	10.3	89.6	0.1	0.0	0.0	100.0	0.0	0.0	0.6	99.4	0.0
	BIC_c	0.0	11.1	87.5	1.4	0.0	0.0	99.3	0.7	0.0	1.0	89.7	9.3
	HQ	0.0	9.8	87.6	2.6	0.0	0.0	99.1	0.9	0.0	0.0	99.4	0.6
	HQ_c	0.0	9.6	86.3	4.0	0.0	0.0	98.3	1.6	0.0	0.6	82.9	16.5
	AIC	0.0	0.4	79.6	15.8	0.0	4.9	73.9	17.2	0.0	0.0	80.8	16.1
	AIC_c	0.0	16.3	63.7	19.2	0.0	14.8	55.8	25.6	0.0	0.0	66.6	21.6
$a3$	BIC	0.0	13.6	86.4	0.0	0.0	16.4	83.6	0.0	0.0	0.2	99.8	0.0
	BIC_c	1.1	34.6	59.3	4.9	0.0	20.0	71.5	8.4	0.0	7.5	77.6	10.6
	HQ	0.0	3.1	96.4	0.5	0.0	10.5	89.3	0.2	0.0	0.0	99.8	0.2
	HQ_c	0.0	25.5	65.3	9.0	0.0	19.4	65.3	14.1	0.0	1.2	75.2	17.0
	AIC	0.0	17.2	71.4	10.2	0.0	7.8	78.4	12.1	0.0	4.8	77.2	16.3
	AIC_c	0.4	35.0	49.0	15.6	0.0	6.6	73.0	17.3	0.0	13.0	76.0	10.2
$b2$	BIC	8.3	42.3	49.4	0.0	0.0	21.0	79.0	0.0	0.0	28.4	71.6	0.0
	BIC_c	13.5	36.2	47.4	2.9	0.0	13.2	78.1	7.9	0.0	28.8	68.5	2.7
	HQ	1.1	34.8	63.5	0.6	0.0	10.9	88.4	0.7	0.0	10.7	88.3	1.0
	HQ_c	3.6	40.0	50.0	6.4	0.0	11.1	75.9	11.6	0.0	22.2	72.6	5.0
	AIC	0.1	12.9	72.3	13.1	0.0	8.9	76.4	12.7	0.0	9.9	76.1	11.4
	AIC_c	11.9	29.0	39.0	18.2	0.0	15.7	49.5	27.7	0.0	12.8	67.0	17.8
$b3$	BIC	21.6	48.6	29.8	0.0	0.1	45.9	54.0	0.0	0.0	30.2	69.8	0.0
	BIC_c	25.5	37.6	32.9	4.0	10.8	42.4	40.7	6.1	0.0	33.8	63.9	2.3
	HQ	4.8	37.4	57.4	0.4	0.0	22.7	76.9	0.4	0.0	22.0	77.8	0.2
	HQ_c	17.7	34.0	38.1	10.0	0.6	28.0	51.2	17.5	0.0	19.6	73.2	7.0
	AIC	0.0	3.9	80.1	13.6	0.0	1.6	81.9	14.3	0.0	0.3	82.3	15.0
	AIC_c	0.0	8.7	75.3	15.8	0.0	10.7	74.1	14.5	0.0	6.5	75.5	14.6
$c2$	BIC	0.0	20.5	79.4	0.1	0.0	23.0	77.0	0.0	0.0	11.6	88.4	0.0
	BIC_c	0.0	21.4	75.6	3.0	0.0	31.9	66.4	1.7	0.0	29.1	64.6	5.7
	HQ	0.0	11.3	86.5	2.2	0.0	9.9	89.4	0.7	0.0	3.7	95.4	0.9
	HQ_c	0.0	14.8	77.6	7.6	0.0	19.2	75.1	5.4	0.0	16.7	74.0	7.7
	AIC	0.0	0.4	80.0	15.9	0.0	2.1	78.6	16.4	0.0	0.0	80.1	17.1
	AIC_c	0.0	4.1	67.0	21.7	0.0	9.5	44.6	32.3	0.0	1.1	74.1	20.8
$c3$	BIC	0.0	32.3	67.7	0.0	0.0	40.1	59.9	0.0	0.0	14.6	85.4	0.0
	BIC_c	0.0	18.6	73.9	7.1	0.0	29.3	52.1	16.2	0.0	8.0	78.8	12.6
	HQ	0.0	8.1	91.6	0.3	0.0	19.1	80.3	0.6	0.0	2.5	97.5	0.0
	HQ_c	0.0	7.0	76.0	14.7	0.0	15.9	53.6	24.7	0.0	3.3	79.0	16.3

[a]Values are in (%). *Exp* denotes experiment according to what is explained in the text. The numbers 200, 400 and 600 refer to the different sample sizes

Queen Mary
University of London

This working paper has been produced by
the Department of Economics at
Queen Mary, University of London

Department of Economics
Queen Mary, University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 5096 or Fax: +44 (0)20 8983 3580
Email: j.conner@qmul.ac.uk
Website: www.econ.qmul.ac.uk/papers/wp.htm