Econometric Theory, **21**, 2005, 593–620. Printed in the United States of America.
DOI: 10.1017/S0266466605050334

A MONTE CARLO STUDY ON THE SELECTION OF COINTEGRATING RANK USING INFORMATION CRITERIA

ZIJUN WANG AND DAVID A. BESSLER
Texas A&M University

We conduct Monte Carlo simulations to evaluate the use of information criteria (Akaike information criterion [AIC] and Schwarz information criterion [SC]) as an alternative to various probability-based tests for determining cointegrating rank in multivariate analysis. First, information criteria are used to determine cointegrating rank given the lag order in a levels vector autoregression. Second, information criteria are used to determine the lag order and cointegrating rank simultaneously. Results show that AIC has an advantage over trace tests for cointegrated or stationary processes in small samples. AIC does not perform well in large samples. The performance of SC is close to that of the trace test. SC shows better large sample results than AIC and the trace test, even if the series are close to nonstationary series or they contain large negative moving average components. We also find evidence that supports the joint estimation of lag order and cointegrating rank by the SC criterion. We conclude that information criteria can complement traditional parametric tests.

1. INTRODUCTION

Cointegration of time-ordered observational data has received considerable attention in the past decade. Various procedures have been proposed in the literature to determine cointegrating rank. They include single equation methods such as the Engle–Granger residual-based test (Engle and Granger, 1987) and the ECM test of Kremers, Ericsson, and Dolado (1992). Recently, empirical researchers have relied more on multiple-equation or system-based methods, for example, the principal components test of Stock and Watson (1988), Johansen (1988, 1991), and the likelihood ratio test of Ahn and Reinsel (1990) based on canonical correlation analysis.

Because these parametric test procedures are probability-based, they are likely to have problems of low power and size distortions when, for example, errors

We are grateful to Peter C.B. Phillips and an anonymous referee for their comments, which significantly improved the paper. Address correspondence to Zijun Wang, Private Enterprise Research Center, Academic Building West, Room 3028, Texas A&M University, College Station, TX 77843, USA; e-mail: z-wang@tamu.edu.

 0266-4666/05 $12.00

(innovations) are not independent and identically distributed (i.i.d.) or the series are close to nonstationary ones. Podivinsky (1990), Cheung and Lai (1993), Toda (1995), Haug (1996), and Gonzalo and Pitarakis (1999) provide simulation evidence that these tests may either over- or underspecify cointegrating rank, especially in small (finite) samples. An alternative to the preceding parametric procedures is to consider various information criteria (IC) in determining rank restrictions. This application of the model selection approach was first suggested and implemented in Phillips and McFarland (1997). For practical purposes, model specification ultimately involves a trade-off between model parsimony (complexity) and fit, given the fact that the true model is rarely, if ever, known. As various IC take into account both model fit and parsimony, they have become increasingly important tools for specifying models; in particular, they now have a rich history in selecting lag order in both univariate and multivariate modeling. Because the determination of cointegrating rank is essentially a model specification problem just like the lag order selection, it is quite natural to consider IC in the determination of cointegrating rank (Phillips, 1996).

An advantage of IC over traditional probability-based test procedures in determining rank order is that researchers are exempt from first having to select an "appropriate" significance level to implement a test procedure. Although a 5% (or perhaps a 10%) significance level is "traditionally" chosen as a benchmark, such a choice may generate concerns. For example, Maddala and Kim (1998, ch. 6) suggest that researchers should be more conservative in testing for a unit root, that it may be better to use the 25% level instead of the 5% level. Furthermore, to many empirical researchers, as argued by Maddala and Kim, the goal of cointegration tests is not to uncover the *true number* of cointegrating relationships per se but rather to have a useful guide in imposing restrictions on vector autoregression (VAR) models and error correction models (ECM) that may lead to more efficient estimation and improve forecasting performance (p. 233). When forecasting performance of a model is of interest, clearly both fit and complexity have to weigh in at the same time.

Another attractive feature of using IC is that it allows researchers to conduct cointegration analysis within a single step, instead of a two-step procedure. As is well known, the choice of lag order in a VAR has an important impact on the cointegration test performance (e.g., Boswijk and Frasnes, 1992). However, because choices of lag order and cointegrating rank are two separate steps in application of the trace test and other probability-based procedures, it is essentially impossible to comment on the underlying probability distribution of the final results. In contrast, it is possible to determine the lag order and cointegrating rank in one step by minimizing an IC over a domain of models with different lag orders and cointegrating ranks.

Phillips (1996) formally shows how cointegrating rank, lag length, and trend degree in a VAR can be jointly determined using the model selection method. Gonzalo and Pitarakis (1998) evaluate both the theoretical and applied properties of the model selection approach for the estimation of the cointegrating rank given lag orders in the models. Aznar and Salvador (2002) establish the consistency of a general IC that includes the Schwarz information criterion (SC) as a special case. Kapetanios (2004) recently derived the asymptotic distribution of the cointegrating rank estimator based on the Akaike information criterion (AIC). He shows that the estimator is inconsistent, a result similar to that found when AIC is used as a tool for lag order selection. For the purpose of model specification in a (partially) nonstationary framework, researchers have also proposed other IC. For example, extending the analysis of Phillips and Ploberger (1996), Chao and Phillips (1999) show the consistency of the posterior information criterion (PIC) in the joint determination of cointegrating ranks and VAR lag order. They also provide Monte Carlo evidence that shows that PIC performs at least as well and sometimes better than SC and AIC. On the empirical side, Phillips and McFarland (1997) use the SC criterion to jointly estimate the lag order and cointegrating rank in the VAR analysis of the Australian exchange market. Wang and Bessler (2002) apply a similar procedure in studying U.S. meat demand systems.

The goal of this paper is to provide more comprehensive evidence on the performance of the model selection approach (IC) in cointegration analysis. We conduct three Monte Carlo simulations. The design of the first simulation borrows from Toda (1995). Here we provide evidence on the performance of the two widely used IC procedures, SC and AIC, in testing the cointegrating rank when the lag order of the VAR is known. In the second simulation, employing a data generating process (DGP) used by Haug (1996) and others, we evaluate the performance of IC in determining the lag order and cointegrating rank simultaneously. These two DGPs allow us to investigate the test performance under a great variety of model specifications, including moving average components, closeness to a unit root, correlation between innovations, and so on. The third simulation evaluates the use of IC in a larger, five-variable system. Throughout the paper, we pay special attention to the small sample performance of the approach, as it is probably more relevant to many macroeconomic series (the sample size considered by Gonzalo and Pitarakis, 1998, and Chao and Phillips, 1999, is at least 150).

For comparison purposes, we also examine the performance of Johansen's trace test, which is chosen for its current popularity in empirical applications.[1] Recently, Johansen (2000, 2002) has proposed the use of so-called Bartlett correction to improve the small sample performance of the trace test. In this paper, we will have an opportunity to see how the correction factor fits into the simulation models.

The paper is organized as follows. Section 2 briefly discusses the basic model, the trace test statistic, and the AIC and SC formulas. Section 3 reports the first Monte Carlo simulation results. The design and results of the second experiment are summarized in the fourth section. Section 5 offers a real life example and simulation results corresponding to it. A short summary concludes the paper.

2. THE MODEL AND TEST STATISTICS

The basic model is an m-dimensional VAR model. Using conventional notation, the model can be described as

$$z_t = A_1 z_{t-1} + \cdots + A_p z_{t-p} + \mu + \varepsilon_t, \qquad t = 1, \ldots, T, \tag{1}$$

where z_t is an $m \times 1$ vector of time series, $z_{t-1}, \ldots z_{t-p}$, are 1 up to p lags of z_t, ε_t are i.i.d. random variates following multivariate $N(0, \Sigma)$ with Σ being positive definite, $A_1, \ldots, A_p$ are conformable parameter matrices, and μ is an $m \times 1$ vector of parameters. The error correction form of (1) is

$$\Delta z_t = \Gamma_1 \Delta z_{t-1} + \cdots + \Gamma_{p-1} \Delta z_{t-p+1} + \Pi z_{t-1} + \mu + \varepsilon_t, \qquad t = 1, \ldots, T, \tag{2}$$

with

$$\Gamma_i = (-A_{i+1} + A_{i+2} + \cdots + A_p), \quad \text{for } i = 1, 2, \ldots, p-1$$

and

$$\Pi = -(I_m - A_1 - A_2 - \cdots - A_p).$$

The hypothesis of cointegration in the vector process z_t can be formulated as testing the rank of the Π matrix (Johansen, 1988, 1991). When the null hypothesis is that the cointegrating rank is r, the trace test statistic (λ_{tr}) is given by

$$\lambda_{\text{tr}} = -T \sum_{i=r+1}^{m} \ln(1 - \lambda_i), \tag{3}$$

where r is the cointegrating rank order and λ_i is the ith largest eigenvalue related to the Π matrix. The sequential tests start from the null hypothesis $r = 0$ (namely, all eigenvalues are 0's). If this hypothesis is rejected, one continues to test $r \leq 1$ and stops testing the first time the hypothesis is not rejected or after $r \leq m - 1$. For $0 < r < m$, z_t is a cointegrated process; otherwise, it is nonstationary if $r = 0$ (or stationary if $r = m$). The asymptotic distributions of the trace statistic are affected by the assumption on the time trend in the process. If the constant in (2) is restricted to the cointegration space, the process contains a stochastic trend. If it is unrestricted, then the process contains both a linear time trend and a stochastic trend. Although other assumptions on the time trend have also been considered (Johansen, 1996), these two are used most often in applied studies.

The small sample correction for the preceding test statistic proposed by Johansen (2000, 2002) is of Bartlett type. The idea is to approximate the expectation of the likelihood ratio test statistic and to thereby correct it to have the same mean as the asymptotic distribution. The correction factor depends on moments of functions of the random walk and functions of the parameters. The exact formulas and coefficients necessary to compute the factor can be found in Johansen (2002).

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

We examine the performance of two widely used IC in the simulations: AIC (Akaike, 1973) and SC (Schwarz, 1978).[2] They are computed according to the following equations:

$$AIC = \ln(\det(\hat{\Sigma})) + 2K/T, \tag{4}$$

$$SC = \ln(\det(\hat{\Sigma})) + K\ln(T)/T, \tag{5}$$

where $\hat{\Sigma}$ is the maximum likelihood estimate of the variance-covariance Σ of the innovation (ε_t's) and K is the number of free parameters in the model, which, other things being equal, increases with the lag order (p) and the cointegrating rank (r) assumed in the model. The first term in equations (4) and (5) is the log determinant of $\hat{\Sigma}$, which measures lack of fit of the model. The second term penalizes overparameterization of the model. It is clear from (4) and (5) that SC punishes model overparameterization more than AIC for sample sizes equal to or larger than eight.

Clearly, the IC method and the trace test are closely related. They both condition on the feature of matrix Π. The trace test detects the rank of Π by testing the statistical significance of the eigenvalues related to Π. The IC method determines cointegrating rank by balancing the benefit and cost of adding additional restriction vectors (cointegrating vectors) to the model. Specifically, if Π has rank r, it may be written as the product of two matrices: $\Pi = \alpha\beta'$, where α and β are of dimension $m \times r$. We may regard $\beta' z_{t-1}$ as r linear restrictions/combinations on right-hand-side variables z_{t-1}. If a restriction is true, it must carry some useful information to explain the variation in the left-hand-side variables z_t. The more significant the restriction is (correspondingly, the larger the associated eigenvalues of Π), the more information it can convey. If the restriction is true, the useful information it contains should be enough to offset its cost (introducing more parameters to the model). IC would accept the restriction in this case. If, on the other hand, the restriction is insignificant or false, the information it carries cannot offset the cost; IC would reject the addition of such a restriction.

3. MONTE CARLO SIMULATION I

In this section, we investigate the sampling properties of AIC, SC, and two forms of the trace test in determining the cointegrating rank of model (1) assuming the lag order is known. To this end, Toda (1994, 1995) shows that, without loss of generality, we can study a "canonical form" model.[3] Consider the following m-dimensional process:

$$w_t = \begin{pmatrix} w_{1,t} \\ w_{2,t} \end{pmatrix} = \begin{pmatrix} 0 \\ \delta e_{m-r} \end{pmatrix} + \begin{pmatrix} \Psi & 0 \\ 0 & I_{m-r} \end{pmatrix} \begin{pmatrix} w_{1,t-1} \\ w_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}, \qquad t = 1,2,\ldots,T, \tag{6}$$

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

where the dimensions of w_t, $w_{1,t}$, and $w_{2,t}$ are m, r, and $(m-r)$, respectively, δ is a nonnegative scalar, $e_{m-r} = (0,\ldots,0,1)'$ is an $(m-r)$-dimensional vector, all eigenvalues of Ψ lie inside the unit circle, and finally,

$$\begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \sim \text{i.i.d. } N\left(0, \begin{pmatrix} I_r & \Theta \\ \Theta' & I_{m-r} \end{pmatrix}\right).$$

Clearly, the subvector process $w_{1,t}$ is stationary and $w_{2,t}$ is nonstationary. They are correlated by a matrix Θ. The process $w_{2,t}$ also contains a linear deterministic trend unless $\delta = 0$. Following Toda (1995), we consider a bivariate VAR(1). There are three possibilities in regard to the cointegration relations of DGP (6).

First, if $r = 0$, that is, (6) contains only nonstationary components, then model (6) simplifies to

$$w_t = \delta e_2 + w_{t-1} + \varepsilon_t \tag{7}$$

and $\varepsilon_t \sim$ i.i.d. $N(0, I_2)$. In (7), the only parameter is δ.

Second, if $r = 1$, (6) becomes a cointegrated process with the following explicit form:

$$\begin{pmatrix} w_{1,t} \\ w_{2,t} \end{pmatrix} = \begin{pmatrix} 0 \\ \delta \end{pmatrix} + \begin{pmatrix} \psi & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} w_{1,t-1} \\ w_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \tag{8}$$

and

$$\begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \sim \text{i.i.d. } N\left(0, \begin{pmatrix} 1 & \theta \\ \theta & 1 \end{pmatrix}\right),$$

where $|\psi| < 1$ and $|\theta| < 1$.

Third, if $r = 2$, (6) becomes a stationary process

$$w_t = \Psi w_{t-1} + \varepsilon_t, \tag{9}$$

where $\varepsilon_t \sim$ i.i.d. $N(0, I_2)$ and Ψ is further assumed to be diagonal

$$\Psi = \begin{pmatrix} \psi_a & 0 \\ 0 & \psi_b \end{pmatrix},$$

where $|\psi_a| < 1$ and $|\psi_b| < 1$.

In this and the next simulation, we examine the IC performance for four sample sizes: 30, 50, 100, and 200. The number of replications for each sample size is 5,000. Following tradition, for each replication we generate an additional 50 random observations to eliminate start-up effects. Toda (1995) explicitly considers the impact of starting values of w_t on the test performance. It is clear from his reported results that the relative performance of the trace test does not change significantly for three different sets of starting values. To save

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

space, and also to follow tradition, we only report the simulation results using $w_0 = 0$ in all cases.

The results reported in Table 1 are based on DGP (7), a nonstationary process without cointegration relations ($r = 0$). Each entry is the percentage of times the four test procedures, AIC, SC, trace test (λ_{tr}), and the small sample adjusted trace test (λ^*_{tr}) of Johansen (2002), correctly determine the cointegrating rank of the simulated data.[4] When the sample size is small (30) and the DGP has no linear trend ($\delta = 0$), the AIC correctly finds the cointegrating rank ($r = 0$) only 39.3% of the time. The performance of SC is much better (86.1%). The probability that SC chooses alternative models with $r = 1$ and $r = 2$ is 12.2% and 1.7%, respectively, whereas the numbers are 39.3% and 21.4% for AIC (not reported in the table). The result, that AIC tends to choose more complicated models (in this case, models with higher cointegrating ranks) than SC, is as expected. The trace test, λ_{tr}, selects the correct model in 93.1% of the cases, whereas the performance of the small sample adjusted trace test, λ^*_{tr}, shows even further improvement (94.3%), close to the test size (recall that we use the 5% significance level for the two trace tests throughout the paper). It is clear from the table that the performance of all four procedures deteriorates when a linear trend is present in the model ($\delta > 0$). The effect of trend on the two ICs is more noticeable than that on the trace tests. Nevertheless, the results also show that all tests perform better when the trend signal is strong ($\delta = 1$).

TABLE 1. Performance of IC and trace tests for cointegrating rank: $r = 0$, Simulation I

	$T = 30$			$T = 50$		
	$\delta = 0.0$	$\delta = 0.2$	$\delta = 1.0$	$\delta = 0.0$	$\delta = 0.2$	$\delta = 1.0$
AIC	39.3	27.2	33.3	40.5	29.3	34.4
SC	86.1	70.9	75.8	93.2	82.8	86.4
λ_{tr}	93.1	87.9	92.7	93.6	89.1	93.9
λ^*_{tr}	94.3	89.1	93.7	94.4	89.9	94.3
	$T = 100$			$T = 200$		
	$\delta = 0.0$	$\delta = 0.2$	$\delta = 1.0$	$\delta = 0.0$	$\delta = 0.2$	$\delta = 1.0$
AIC	42.2	31.9	36.7	42.4	35.5	38.0
SC	97.6	92.8	93.7	99.5	97.3	97.4
λ_{tr}	94.1	92.0	94.2	94.3	93.4	94.8
λ^*_{tr}	94.2	92.1	94.6	94.5	93.5	94.9

Note: Each entry is the percentage that the cointegrating rank of the simulated data is correctly determined based on the AIC, SC, trace test (λ_{tr}), and small sample adjusted trace test (λ^*_{tr}) procedures. The DGP is a bivariate VAR(1) without cointegration relations.

The performance of SC improves considerably when the sample size increases. For example, when $T = 50$, the frequencies with which it correctly identifies the rank are 93.2%, 82.8% and 86.4% when $\delta = 0.0$, 0.2, and 1.0, respectively. At $T = 100$, the success ratios of SC are all larger than 90% for the different values of δ. When the sample size further increases to 200, SC almost always finds the correct rank (99.5%) in the case of $\delta = 0$. In models with a linear trend, the percentages are also high, 97.3% or higher. In contrast, the change in sample size has little impact on AIC, which confirms that SC is consistent whereas AIC is not.

Using DGP (8), we examine the performance of the four procedures when the true model is a bivariate cointegrated process with $r = 1$. We use following parameter values in the simulations: $\delta = 0.0$, 0.2, and 1.0, $\psi = 0.8$ and 0.9, and $\theta = 0.0$, 0.4, and 0.8 (results based on negative θ are omitted because they are very close to their positive counterparts). Table 2 contains the simulation results for two small sample sizes, $T = 30$ and 50. Clearly, all four procedures perform

TABLE 2. Performance of IC and trace tests for cointegrating rank: $r = 1$, $T = 30$, 50, Simulation I

	$\delta = 0.0$			$\delta = 0.2$			$\delta = 1.0$		
	$\theta = 0.0$	$\theta = 0.4$	$\theta = 0.8$	$\theta = 0.0$	$\theta = 0.4$	$\theta = 0.8$	$\theta = 0.0$	$\theta = 0.4$	$\theta = 0.8$
$T = 30$, $\psi = 0.8$									
AIC	40.5	42.6	48.2	42.2	43.3	49.1	63.0	64.8	70.0
SC	14.4	15.4	36.6	23.8	26.0	42.7	29.4	32.2	54.2
λ_{tr}	7.8	9.4	27.4	9.3	10.4	22.5	9.0	10.2	24.8
λ_{tr}^*	6.6	8.3	26.5	8.4	9.6	23.7	8.4	9.4	24.8
$T = 30$, $\psi = 0.9$									
AIC	38.6	39.2	41.1	40.2	40.3	44.2	55.5	56.4	60.3
SC	11.3	11.6	18.2	20.8	21.0	25.7	23.3	24.1	30.0
λ_{tr}	6.0	6.7	13.1	7.8	7.4	10.1	7.1	7.0	10.2
λ_{tr}^*	5.5	6.2	12.6	7.4	7.3	10.7	6.3	6.3	10.0
$T = 50$, $\psi = 0.8$									
AIC	47.4	49.0	53.0	50.8	53.0	57.2	74.0	75.3	76.9
SC	11.8	15.9	56.4	26.1	31.2	62.9	30.7	36.2	73.8
λ_{tr}	12.5	16.1	56.7	14.8	18.4	49.7	14.5	18.7	56.7
λ_{tr}^*	11.3	15.2	56.9	14.3	17.9	51.2	13.7	17.5	56.6
$T = 50$, $\psi = 0.9$									
AIC	40.7	40.7	44.4	43.6	45.1	52.5	60.4	61.1	67.0
SC	7.0	7.9	20.6	15.1	16.7	29.4	16.5	18.2	34.0
λ_{tr}	6.4	7.7	22.2	8.0	8.5	17.6	7.1	7.9	19.1
λ_{tr}^*	5.9	7.1	22.3	7.8	8.6	19.3	6.7	7.7	19.7

Note: Each entry is the percentage that the cointegrating rank of the simulated data is correctly determined based on the AIC, SC, trace test (λ_{tr}), and small sample adjusted trace test (λ_{tr}^*) procedures. The DGP is a bivariate VAR(1) with cointegrating rank of 1.

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

poorly in small sample sizes. In models with $\theta = 0.0$ and $\delta = 0.0$, namely, no correlation between the stationary and nonstationary components and no linear trend, λ_{tr} makes the correct rank choice only 7.8% of the time (the small sample correction actually leads to slightly worse performance, 6.6%). The success ratio of SC is higher but still low (14.4%). AIC performs best in this case. When the two component series are more closely related (larger θ), the performance of all procedures improves. The frequencies that SC and λ_{tr} find correct rank, when $\theta = 0.8$, increase to 36.6% and 27.4%, respectively. The presence of a linear trend in the data appears to have different impacts for IC and the trace tests. Compared to the results when $\delta = 0.0$, the performance of SC when $\delta = 1.0$ improves considerably (29.4% vs. 14.4%), whereas AIC's performance also increases to 63.0%. The numbers remain small for both λ_{tr} and λ_{tr}^* (9.0% and 8.4%). As in the model with $\delta = 0.0$, the test power increases for larger correlation ($\theta = 0.8$ and $\delta = 1.0$). And SC (54.2%) still leads the trace tests (24.8%).

The second part of Table 2 repeats the preceding simulations with the autoregressive coefficient further approaching 1 ($\psi = 0.9$). Not surprisingly, all four test procedures are less powerful, as the stationary component is now closer to a nonstationary component. In the case of $\delta = 1$ and $\theta = 0.8$ (the rightmost column), SC concludes with zero or two unit roots 70.0% (100% − 30.0%) of the time, even if the true model has only one unit root, whereas the trace test is more likely to err (about 90%). Correcting for small sample bias does not help.

The results in the third and fourth sections of Table 2 are based on the slightly larger sample size of 50. In general, all procedures, especially SC, λ_{tr}, and λ_{tr}^*, perform better when $\psi = 0.8$ in the DGP. The improvement is obvious for $\theta = 0.8$. The procedures λ_{tr} and λ_{tr}^* now have similar power as SC in models without linear trends. Nevertheless, at this sample size, AIC and SC are still better choices than are the trace tests when a linear trend is present ($\delta > 0$). Comparing the results in the fourth section with the second section, we find that the increase of T from 30 to 50 does not lead to much improvement on performance for any of the four tests if $\psi = 0.9$ with the exception of $\theta = 0.8$.

Results in Table 3 are also based on DGP (8) using a moderate sample size of 100 and a larger size of 200. When $T = 100$, both λ_{tr} and λ_{tr}^* outperform SC in the model with $\delta = 0.0$ but are outperformed by the latter if $\delta \neq 0.0$. Although the performance of AIC is primarily affected by δ, the correlation coefficient θ is an important factor in determining the performance of SC, λ_{tr}, and λ_{tr}^*, given ψ. For example, when $\theta = 0.0$ or 0.4, and $\psi = 0.8$, all procedures perform poorly, even when the sample size is 100. In contrast, all perform quite well when the two innovations are strongly correlated. The low power of the test procedures against large ψ is still evident when $T = 200$. If $\psi = 0.8$, all four procedures perform well, except AIC in the models with $\delta = 0$. When $\psi = 0.9$, the performance of SC, λ_{tr}, and λ_{tr}^* deteriorates significantly (lower than 60%) in models with low correlations. These three procedures perform equally well (around 90%) when $\theta = 0.8$; in other cases, λ_{tr} and λ_{tr}^* outperform SC. Notice

TABLE 3. Performance of IC and trace tests for cointegrating rank: $r = 1$, $T = 100, 200$, Simulation I

	$\delta = 0.0$			$\delta = 0.2$			$\delta = 1.0$		
	$\theta = 0.0$	$\theta = 0.4$	$\theta = 0.8$	$\theta = 0.0$	$\theta = 0.4$	$\theta = 0.8$	$\theta = 0.0$	$\theta = 0.4$	$\theta = 0.8$
$T = 100$, $\psi = 0.8$									
AIC	57.2	57.7	56.3	67.1	66.7	65.9	83.5	83.2	80.3
SC	22.3	34.4	91.3	50.9	62.0	87.3	55.5	67.2	94.7
λ_{tr}	37.9	50.0	93.0	45.3	55.3	83.3	47.3	60.3	91.9
λ^*_{tr}	37.3	49.4	94.1	44.9	55.1	84.8	46.5	59.4	92.9
$T = 100$, $\psi = 0.9$									
AIC	45.9	48.6	54.4	59.9	61.5	65.4	74.5	76.4	77.8
SC	4.5	6.7	42.1	15.3	19.3	57.5	16.1	21.0	62.2
λ_{tr}	11.9	15.2	55.7	13.6	17.3	50.9	13.0	16.9	55.3
λ^*_{tr}	11.9	15.1	56.7	13.3	17.1	52.4	12.7	16.7	56.1
$T = 200$, $\psi = 0.8$									
AIC	58.3	57.7	58.1	75.1	75.3	74.6	84.4	84.2	82.2
SC	74.7	88.9	97.4	92.8	94.4	94.6	95.6	97.1	97.4
λ_{tr}	90.7	93.7	94.9	88.9	89.7	90.1	93.7	94.5	94.0
λ^*_{tr}	90.7	94.1	95.5	89.2	90.0	90.7	93.8	94.7	94.3
$T = 200$, $\psi = 0.9$									
AIC	56.7	57.4	56.7	75.1	75.4	74.0	84.7	84.1	80.5
SC	9.2	17.6	88.0	34.9	47.9	92.8	35.4	48.7	95.4
λ_{tr}	35.3	48.5	92.9	44.5	57.2	88.7	45.2	58.6	92.5
λ^*_{tr}	35.3	48.6	94.1	44.3	57.3	89.9	44.9	58.4	93.1

Note: Each entry is the percentage that the cointegrating rank of the simulated data is correctly determined based on the AIC, SC, trace test (λ_{tr}), and small sample adjusted trace test (λ^*_{tr}) procedures. The DGP is a bivariate VAR(1) with cointegrating rank of 1.

in the fourth section of the table ($\psi = 0.9$) that the power of SC is extremely low when $\delta = 0.0$ and θ is small, even at the sample size of 200 (9.2% and 17.6% for $\theta = 0.0$ and 0.4, respectively). In additional simulations we conducted (not reported in detail here), only when the sample size increases to 400 does the ratio increase to 52% and 76.1%. These percentages increase to 79.1% and 93.4% when $T = 500$.

Table 4 contains simulation results assuming the DGP is a stationary process (equation (9)). The parameters that affect the distribution of the test statistics in this model are the two autoregressive coefficients, for which we use $\psi_a = 0.5$, 0.7, and 0.9 and $\psi_b = 0.7$, 0.8, and 0.9. In small samples, AIC performs much better than SC, λ_{tr}, and λ^*_{tr}. SC outperforms λ_{tr} and λ^*_{tr} when $T = 30$ but does not outperform them when $T = 100$. The later three procedures perform similarly when $T = 50$. As before, the power of the procedures decreases as the series considered approach nonstationary processes (ψ_a and/or ψ_b approaches 1). For example, there is a high probability (about 70%) that SC, λ_{tr}, and λ^*_{tr} con-

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

TABLE 4. Performance of IC and trace tests for cointegrating rank: $r = 2$, Simulation I

	$\psi_a = 0.5$			$\psi_a = 0.7$			$\psi_a = 0.9$		
	$\psi_b = 0.7$	$\psi_b = 0.8$	$\psi_b = 0.9$	$\psi_b = 0.7$	$\psi_b = 0.8$	$\psi_b = 0.9$	$\psi_b = 0.7$	$\psi_b = 0.8$	$\psi_b = 0.9$
$T = 30$									
AIC	97.7	91.5	77.6	94.7	87.3	73.7	72.5	66.7	56.7
SC	82.5	66.2	46.8	63.2	49.2	34.3	33.9	26.0	18.8
λ_{tr}	72	54.6	35.8	50.5	36.6	23.7	24.0	17.2	11.5
λ^*_{tr}	69	51.8	33.6	46.8	33.4	21.2	21.6	15.2	9.7
$T = 50$									
AIC	99.9	99.0	88.0	99.8	98.6	87.7	87.2	84.1	72.6
SC	98.2	87.1	57.3	89.3	72.9	44.4	44.8	31.3	17.7
λ_{tr}	98.4	88.1	58.5	90.5	74.9	46.3	46.8	33.6	19.1
λ^*_{tr}	98.2	87.3	57.2	89.3	72.9	44.1	44.8	31.3	17.5
$T = 100$									
AIC	100.0	100.0	98.9	100.0	100.0	99.0	98.7	98.6	96.2
SC	100.0	99.7	80.8	100.0	99.3	79.3	79.4	66.3	32.0
λ_{tr}	100.0	99.9	89.4	100.0	99.9	89.3	89.5	83.7	55.8
λ^*_{tr}	100.0	99.9	89.0	100.0	99.8	88.6	88.9	82.7	53.8
$T = 200$									
AIC	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SC	100.0	100.0	99.1	100.0	100.0	99.1	99.1	99.0	83.9
λ_{tr}	100.0	100.0	99.9	100.0	100.0	99.9	99.9	99.9	99.0
λ^*_{tr}	100.0	100.0	99.9	100.0	100.0	99.9	99.9	99.9	98.9

Note: Each entry is the percentage that the cointegrating rank of the simulated data is correctly determined based on the AIC, SC, trace test (λ_{tr}), and small sample adjusted trace test (λ^*_{tr}) procedures. The DGP is a stationary bivariate VAR(1).

clude that the DGP (9) is nonstationary, without cointegration relations, when $T = 50$ and $\psi_a = \psi_b = 0.9$ (not reported in the table). Finally, if T increases to 200, all procedures work well even for $\psi_a = \psi_b = 0.9$ (the lowest score is SC's 83.9%).

We end the discussion of the first simulation by noting that the trace test tends to perform quite well for the cases where the DGP is of full rank ($r = 2$). This is not surprising, as Johansen (1992) has shown analytically that the probability of overestimation of rank does not go to zero if a fixed significance level is used in the sequential trace tests. However, in the full rank DGP ($r = m$), it is impossible to have overestimation in the trace test, which helps explain the good performance of the trace tests in this model. Similarly, as mentioned in the Introduction, AIC is also inconsistent and overestimates the cointegrating rank. This explains why AIC does quite well for models with $r = 2$ although performing poorly in other cases.[5]

4. MONTE CARLO SIMULATION II

As discussed in the Introduction, both the lag order determination and the test of cointegration relations in a multivariate model relate to model specification. In the preceding simulations, we have assumed that the lag order of the DGP is known, which is rarely true in empirical studies. When the lag order is unknown, the practice is to determine the lag order first using either IC or sequential likelihood ratio tests. In the second stage, a parametric test such as λ_{tr} is used to determine the cointegrating rank conditional on the lag order chosen in the first stage. In this section, we examine the performance of AIC and SC when they are used to determine the lag order and the cointegrating rank simultaneously. For the purpose of comparison, we also implement the two-step procedure to the new DGP. The new DGP again includes two series: y_t and x_t. Specifically, the DGP is described by the following equations:

$$
\begin{aligned}
y_t - x_t &= v_t, \\
a_1 y_t + x_t &= \psi_t, \\
v_t &= \rho v_{t-1} + w_t, \\
\psi_t &= \psi_{t-1} + \rho_t, \\
\rho_t &= \varphi_t + \theta \varphi_{t-1},
\end{aligned}
$$

and

$$
\begin{bmatrix} w_t \\ \varphi_t \end{bmatrix} = \text{i.i.d. } N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \eta\sigma \\ \eta\sigma & \sigma^2 \end{pmatrix}\right], \qquad \textbf{(10)}
$$

Many researchers have used DGPs similar to (10). A short list includes Banerjee, Dolado, Hendry, and Smith (1986), Engle and Granger (1987), Blangiewicz and Charemza (1990), Hansen and Phillips (1990), Gonzalo (1994), and Haug (1996). To examine the impact of the model parameters on the test performance, we use the following values in the experiments: $a_1 = (0, 1)$, $\rho = (0.1, 0.3, 0.5, 0.7, 0.85, 0.9, 0.95, 1)$, $\theta = (-0.8, -0.5, -0.25, -0.1, 0.0, 0.1, 0.25, 0.5, 0.8)$, and $\eta = (-0.8, -0.5, -0.25, -0.1, 0.0, 0.1, 0.25, 0.5, 0.8)$. Because, unlike other cointegrating rank tests, both the IC and trace tests do not depend on the standard deviation of the second innovation process φ_t, σ, we fix it at $\sigma = 1$. Nor, in the case of equation (10), do results depend on whether x_t is weakly exogenous ($a_1 = 1$) or endogenous ($a_1 = 0$) to the system.

As before, all basic simulations are conducted for four different sample sizes, 30, 50, 100, and 200. Tables A.1, A.2, and A.3 summarize the major results for $\rho = 0.5$, 0.85, and 1.0, respectively. In each section, we report the performance of four two-stage procedures and two one-step procedures. After using AIC or SC to determine the optimal lag order of VARs in the first stage, we use the same criterion in the second stage to determine cointegrating rank. Denote this two-step procedure by AICAIC and SCSC, respectively (they correspond to the

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

notations AIC and SC used in Simulation I). We denote the procedures λ_{tr} and λ^*_{tr} that use the trace and small sample adjusted trace test to determine the cointegrating rank with lag order chosen by SC in the first stage. Here, we simply use AIC and SC to denote, respectively, the one-step procedures that use AIC and SC to simultaneously determine the lag order and cointegrating rank.

Figures 1–3 are graphical presentations of the simulation results for two sample sizes ($T = 50$ and 200). Because the performances of λ_{tr} and λ^*_{tr} are very

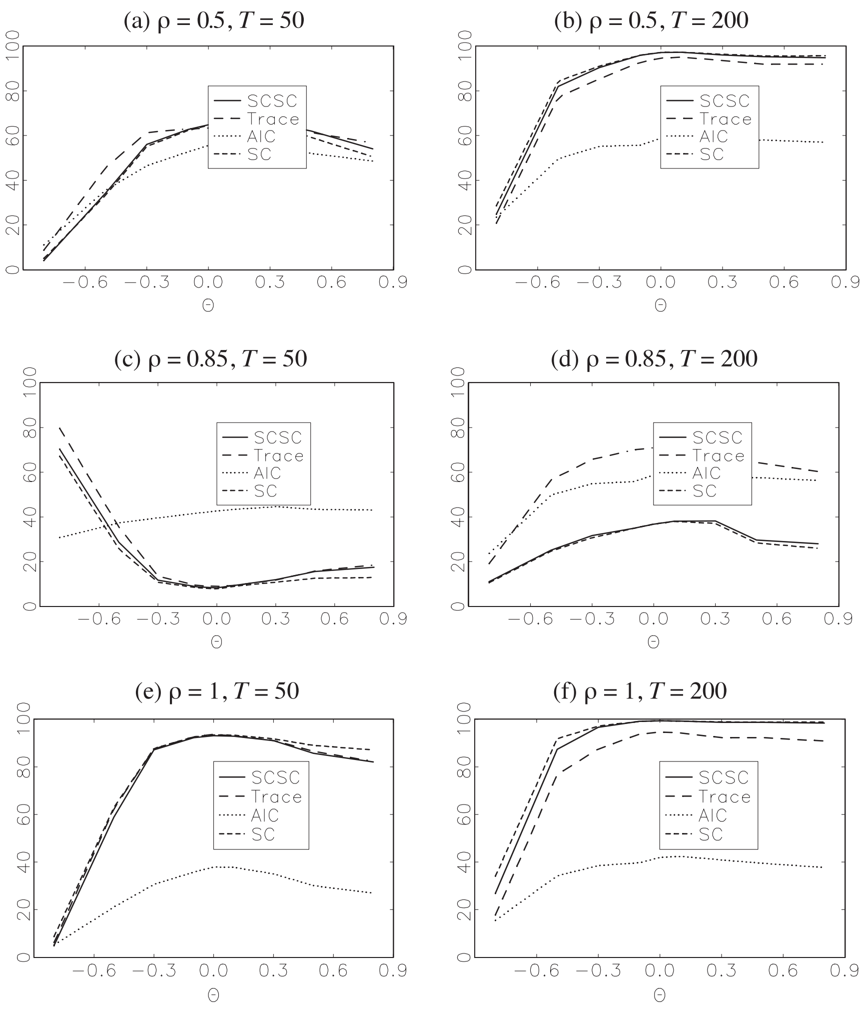


FIGURE 1. Performance of IC and trace tests for cointegrating rank with different values of θ (given $\eta = 0$), Simulation II.

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

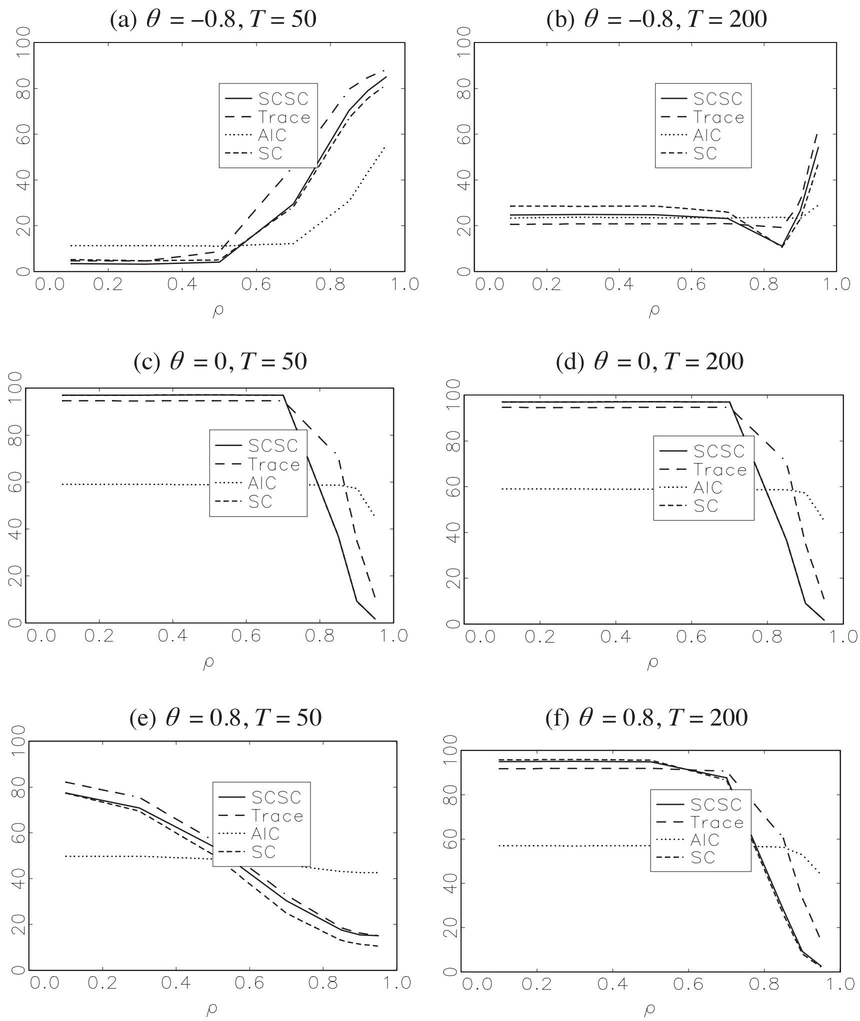


FIGURE 2. Performance of IC and trace tests for cointegrating rank with different values of ρ ($\rho < 1$) (given $\eta = 0$), Simulation II.

similar, and the one-step AIC performs slightly better than the two-step AIC-AIC for almost all parameters and sample sizes, we only compare in each figure the performances of SCSC, λ_{tr}, AIC, and SC.

In Figure 1, we examine the impact of the parameter θ on the test performance while fixing η at 0.0 and ρ at three values: 0.5, 0.85, and 1. First, we discuss Figures 1a and 1b. The DGP is a cointegrated process with a moderate value on ρ(0.5). When there is a large negative moving average component ($\theta = -0.8$) in the DGP, all four procedures perform poorly. The percentages of

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

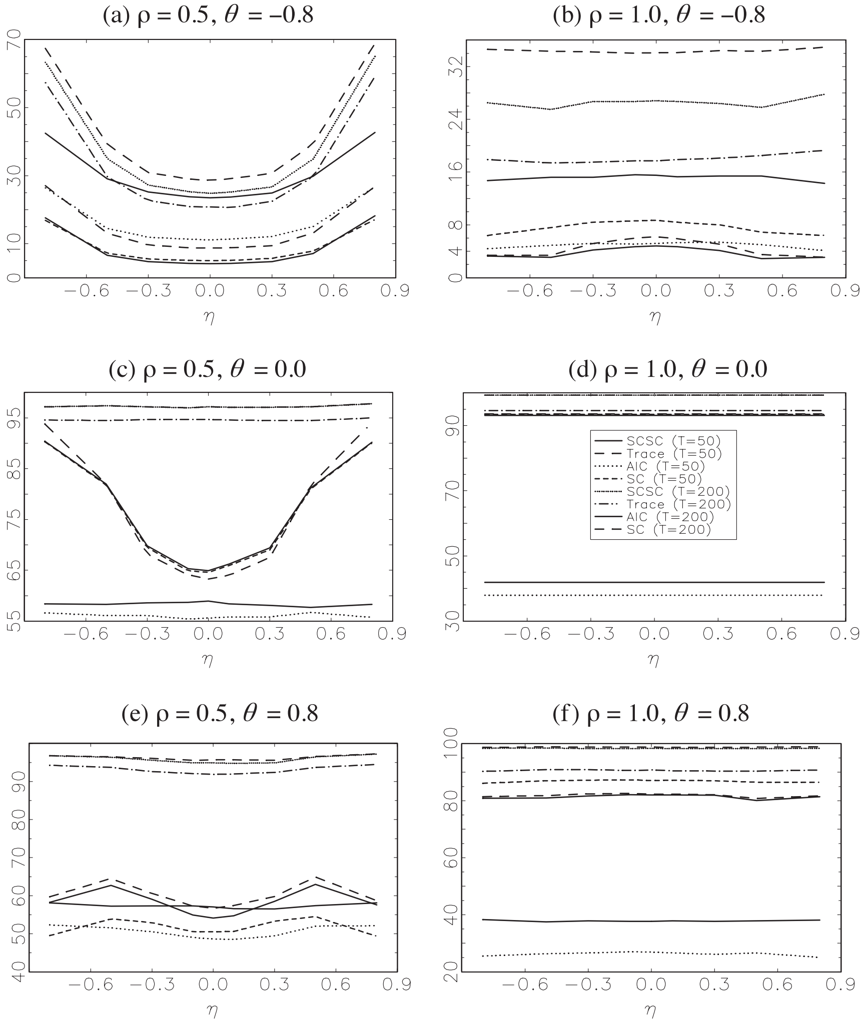


FIGURE 3. Performance of IC and trace tests for cointegrating rank with different values of η, Simulation II. The legends in graphs (a), (b), (c), (e), and (f) are the same as those in (d). For ease of reading, they are omitted.

correct choices on rank are only 4.1%, 8.7%, 11.1%, and 5.0% for SCSC, λ_{tr}, AIC, and SC, respectively, when the sample size is small (50). The performances do not improve significantly if the sample size increases to 200 (Figure 1b).

It is clear from Figure 1a that as θ gets larger, all four procedures perform better, although AIC improves relatively slowly. When the magnitude of the moving average component θ is small in the DGP, the procedures perform best

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

(around 60%). When a large and positive θ is involved, the performances also deteriorate but are still better than when θ is large and negative. The procedure λ_{tr} performs better than both SCSC and SC with θ being large in magnitude, whereas it is outperformed by the latter two for small θ (in absolute values). The two-step SCSC performs slightly better than the one-step SC for all θ values except $\theta = -0.8$. AIC is better than SC only for $\theta < 0.5$. When T increases to 200, SCSC, λ_{tr}, and SC can correctly find the rank in 90% or more of the cases, unless the DGP has a large and negative moving average component. In contrast, AIC's performance always falls below 60% for all θ. The one-step SC performs slightly better than the two-step SCSC over the entire range of θ, although the difference is smaller for θ near or at zero. The percentages for both procedures are also slightly higher than those of λ_{tr}. Nevertheless, considering that the significance level of 95% is used for the λ_{tr} test, their performances can be labeled as similar.

Results summarized in Figures 1c and 1d are also based on a cointegrated DGP with rank 1, but here the process is closer to a nonstationary one without cointegrations ($\rho = 0.85$). In this case, all procedures are less powerful than they are in the models with $\rho = 0.5$ (Figures 1a and 1b) for corresponding model parameters. The exception is that all test procedures perform much better when $\theta = -0.8$ for sample sizes smaller than 100. Also note in Figure 1c that SCSC, λ_{tr}, and SC perform worst when θ is around 0, which is the opposite of the pattern found in Figure 1a. Figure 1d shows that the performances of the model selection approaches, SCSC and SC, appear more sensitive than the trace test to the magnitude of ρ (especially when the correlation between the innovations η equals zero, as in the graphs). For example, when $\theta = 0.8$, λ_{tr}'s performance decreases from 91.9% in Figure 1b where $\rho = 0.5$ to 60.3% in Figure 1d with $\rho = 0.85$. At the same time, SC's performance decreases from 95.7% to only 26.0%. In contrast, the impact of changing ρ from 0.5 to 0.85 on AIC is quite small. AIC also performs better than SCSC and SC for most θ values.

When $\rho = 1$, DGP (10) is a nonstationary process without cointegration relations. Figures 1e and 1f summarize the cointegration test performances of the four procedures under this assumption. First, the two graphs show that SCSC, λ_{tr}, and SC are much more powerful when the true model is a nonstationary process without cointegration than for the cointegrated process summarized in Figures 1a–d if the model contains either no moving average or positive moving average components. Second, the one-step SC consistently performs better than the two-step SCSC, especially for large values of θ when $T = 50$. SC also outperforms λ_{tr} when $T = 200$.

Figure 2 offers more details on how the test performances change over parameter ρ by fixing θ at -0.8, 0, and 0.8. Here, the DGP is cointegrated with rank 1 ($0 < \rho < 1$). For $\theta = -0.8$ (Figures 1a and 1b), the power of all four procedures in finding correct rank is low for small or moderate ρ values. Performances improve when the DGP is closer to a nonstationary process (ρ approaches 1). When $T = 50$, AIC performs best, followed by SC, λ_{tr}, and

SCSC in order, for $\rho = 0.1$ and 0.3. However, for larger ρ, λ_{tr} is the best procedure. When the sample size is large (200), the one-step SC finds the rank most accurately for ρ up to 0.7. As Figures 1c and 1d, SCSC, λ_{tr}, and SC work quite well for small ρ when $\theta = -0.8$ and 0. This is true even if the sample size is only 50 (Figures 2c and 2e). The performances of SCSC, λ_{tr}, and SC quickly deteriorate as ρ gets larger. When $\rho > 0.7$, all tests perform poorly, even if $T = 200$, although λ_{tr} does better than SCSC and SC.

So far in both Figures 1 and 2, we have maintained the assumption that fundamental innovations in the DGPs are not correlated ($\eta = 0$). Figure 3 provides evidence on whether the test performances also vary with η. First, the U-shaped patterns in Figures 3a, 3c, and 3e indicate that, for the cointegrated process, all procedures perform better when the fundamental innovations are correlated than when they are not, especially when either T is 50 or $\theta = -0.8$. At the same time, the sign of the correlation η matters little; that is, the impact of η is symmetric. Second, for nonstationary process ($\rho = 1$), the effect of θ is either small ($\theta = -0.8$ in Figure 3b) or essentially zero ($\theta = 0.0$ in Figure 3d, 0.8 in Figure 3f). The bell-shaped curves in Figure 1b indicate that all procedures perform better when the fundamental innovations are not correlated than when they are correlated.

Before ending presentation of Simulation II, we turn to the Appendix tables for some interesting results that are not seen in the preceding graphs. First, we note that some of the results in Table A.2 are comparable to those of Haug (1996). In our simulations with $\theta = 0.0$ and $T = 100$, the power of λ_{tr} is 37.1%, 22.1%, and 35.3% for $\eta = -0.5$, 0.0, and 0.5, respectively, which are slightly higher than the maximum eigenvalue test (λ_{max}(SC)) results (31.1%, 17.6%, and 29.9%) reported in Haug (1996, Table 1, p. 104).[6] Second, because both IC and the trace tests perform poorly when $\theta = -0.8$ for all sample sizes considered previously, we conduct additional simulations to see how they perform under larger sample sizes. The simulation results are tabulated in Table A.4. As expected from the consistency of the SC criterion, both SCSC and SC do better than λ_{tr}. For example, when $T = 1{,}000$, the performance of SC is better than 77%, whereas the performance of λ_{tr} is always less than 67%. AIC's performance is only about 35% at best. Third, comparing the results in Table 1 for $\delta = 0.0$ with those in Table A.3 under the column $\theta = 0.0$, we find that the performance of all procedures is similar in the two Monte Carlo designs. This is because the two simulation designs are quite similar under the current assumptions: the true DGP are nonstationary processes without cointegration relations, and they include stochastic trends and no moving average components.

5. AN EXAMPLE OF THE U.S. HOG DATA AND MONTE CARLO SIMULATION III

So far, we have concentrated on bivariate models. The purpose of this section is twofold: first, we provide evidence of cointegration analysis using the SC and trace test procedures on a real life example.[7] Second, we conduct a third Monte Carlo simulation to compare the performance of the model selection approach with the parametric trace test in a five-variable VAR where the DGP uses the parameter values estimated from the real example.

The data set we analyze contains five variables related to the U.S. hog market. Specifically, we study annual observations from 1867 to 1948 on the farm wage rate, hog supply, hog price, corn price, and corn supply. These data were first edited and studied by Quenouille (1957, pp. 88–101). He logarithmically transformed each variable and linearly coded the logs. Several other authors, including Box and Tiao (1977), Tiao and Tsay (1983, 1989), Reinsel (1983), Velu, Reinsel, and Wichern (1986), Reinsel and Velu (1998, ch. 5), and Wang and Bessler (2004), have analyzed these data from various perspectives. The original data are included in Quenouille (1957).

Following Box and Tiao (1977) and Reinsel and Velu (1998) we shift backward by one period the wage rate and hog price. We first implement the two-step procedure. To this end, in the first step, both SC and a likelihood ratio test are applied. They agree on two lags (the maximum number of lags in the levels VAR used in the test is four), which is also consistent with the aforementioned literature. In the second step, we calculate SC values for $r = 0, 1, 2, \ldots, 5$ and λ_{tr} and λ^*_{tr} statistics for $r \le 0, 1, 2, \ldots, 4$, respectively, based on a VAR(2) model. The results are presented in section A of Table 5. For comparison, we also reproduce the likelihood ratio test statistics of Reinsel and Velu (1998) in Table 5. SC is minimized at $r = 2$, the same choice based on the other three parametric tests. Section B contains SC values for all combinations of lag order and cointegrating rank ($p = 0, 1, 2, 3, 4$; $r = 0, 1, 2, \ldots, 5$). Again, SC is minimized at $r = 2$ and $p = 2$. Therefore, in this example, the one-step SC agrees with the four two-step procedures in choices of both the order of autoregression and the cointegrating rank. The one-step results are also illustrated in Figure 4, where the surface of SC values against possible p and r values is displayed in Figure 1a and a cross-section of the surface at $\hat{p} = 1$ in Figure 1b (where $\hat{p}$ is the number of lags in the ECM, which, of course, is $\hat{p} = p - 1$).

Assuming $\hat{p} = 1$ and $r = 2$, the parameter estimates of the hog data, using notations in model (2), are as follows:

$$\hat{\mu} = (159.033, -421.906, -142.852, 106.623, -10.688)',$$

$$\hat{\Pi} = \begin{pmatrix} -0.133 & -0.062 \\ 0.328 & -0.177 \\ 0.059 & -0.62 \\ -0.055 & 0.277 \\ 0.004 & -0.107 \end{pmatrix} \begin{pmatrix} 1.000 & -2.824 & 1.580 & 0.370 & 1.370 \\ 2.169 & 1.000 & 0.342 & -1.142 & -1.264 \end{pmatrix},$$

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

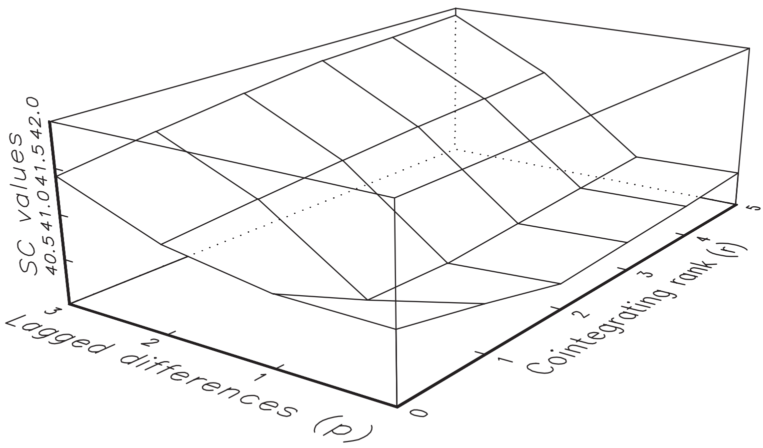
TABLE 5. Determination of the cointegrating rank for Quenouille's U.S. hog data

	SC	Trace statistics	Trace statistics (adjusted)	Reinsel and Velu likelihood ratio statistics
A: Two steps (assume $p = 2$)				
$r = 0$	40.781	138.773 (68.681)	124.938 (68.681)	142.9 (71.1)
1	40.241	58.797 (47.208)	53.909 (47.208)	61.7 (49.4)
2	**40.204**	**25.655** (29.376)	**24.067** (29.376)	**30.2** (31.7)
3	40.263	8.437 (15.340)	8.006 (15.340)	10.5 (18.0)
4	40.323	0.033 (3.841)	0.029 (3.841)	0.04 (8.16)
5	40.379			
	$p = 1$	2	3	4
B: One step (SC)				
$r = 0$	40.769	40.781	40.974	41.418
1	40.540	40.241	41.060	41.567
2	40.287	**40.204**	41.138	41.698
3	40.326	40.263	41.208	41.812
4	40.362	40.323	41.248	41.852
5	40.415	40.379	41.305	41.906

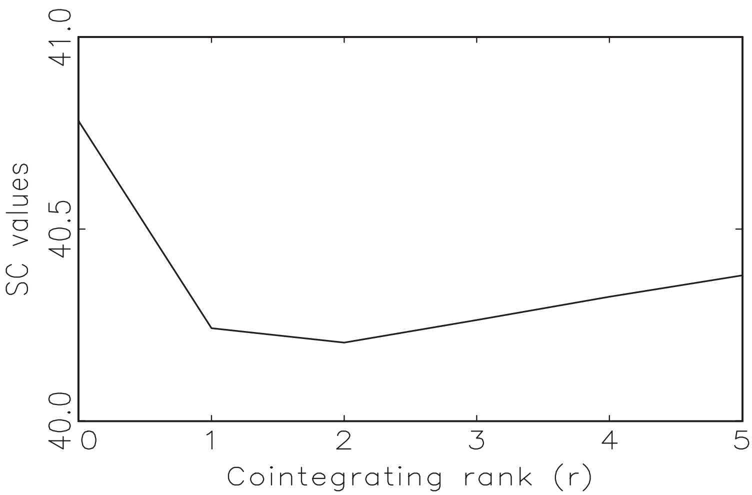
Note: Numbers within parentheses are critical values at the 5% significance level. The trace test critical values are from Hansen and Juselius (1995, Table B.3, p. 81). Reinsel and Velu's likelihood ratio test results are from Reinsel and Velu (1998, Table 5.3, p. 149). The bold number in the SC column represents the minimum SC among all possible rank orders (and all lag orders in the VAR in the one-step procedure). The bold numbers in the columns Trace statistics, Trace statistics (adjusted), and Reinsel and Velu likelihood ratio statistics are the statistics at which the null hypothesis of the cointegrating rank equals the specified value fails to be rejected.

$$\hat{\Gamma}_1 = \begin{pmatrix} -0.165 & -0.345 & 0.291 & 0.234 & 0.243 \\ -0.195 & 0.496 & -0.275 & -0.312 & 0.493 \\ 0.154 & 0.417 & -0.076 & -0.172 & 0.252 \\ -0.267 & -0.327 & 0.039 & -0.319 & 0.289 \\ 0.148 & 0.092 & 0.030 & -0.014 & 0.360 \end{pmatrix},$$

$$\hat{\Sigma} = \begin{pmatrix} 550.705 & -368.370 & 308.664 & -175.909 & 192.249 \\ -368.370 & 3{,}843.067 & 1{,}811.162 & 362.860 & 1{,}145.455 \\ 308.664 & 1{,}811.162 & 9{,}655.385 & -4{,}796.441 & 1{,}687.970 \\ -175.909 & 362.860 & -4{,}796.441 & 4{,}974.835 & 105.510 \\ 192.249 & 1{,}145.455 & 1{,}687.970 & 105.510 & 1{,}372.676 \end{pmatrix}.$$



(a)



(b)

FIGURE 4. SC values, lag order, and cointegrating rank of Quenouille's U.S. hog data. (a) Plots the SC values for different values of cointegrating rank r and autoregressive order p in the hog model. (b) Plots the SC values for different values of cointegrating rank r given $\hat{p} = 1$.

Assuming the DGP in equation (2) is described by these parameters, we proceed with the third simulation as follows: first, sequences of i.i.d. standard normal random numbers are generated; second, these pseudo numbers are multiplied by a factor of variance and covariance $\hat{\Sigma}$ to derive sequences of multivariate

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

normal errors/innovations (following Doan, 1996, p. 10-2). A random sample is formed with these generated errors and the preceding parameter matrices according to equation (2). Repeating the previous process, we obtain 5,000 random samples.

Table 6 summarizes the IC and the trace test performance based on the previously simulated samples. As in the first two bivariate model simulations, the performance of the one-step and two-step AIC procedures is poor, which reflects the fact that these procedures are known to be inconsistent. The performance of SC is similar to that of λ_{tr} for all sample sizes. In this large system, λ_{tr}^* performs significantly worse than λ_{tr} when sample size is 30 and 50. This is similar to the finding in Johansen (2002) that the power function of the trace test is actually shifted down by the correction factor in the simulation on the Danish data. SC correctly determines the cointegrating rank 30.2% of time in the one-step procedure when $T = 30$, which is lower than λ_{tr}'s 46.2%. However, when the sample size is larger than 50, the two methods perform similarly. The right half of Table 6 also includes the performance of AIC and SC in finding correct lag order and cointegrating rank in the one-step procedure. Except when the sample size is small ($T = 30$), both IC criteria appear to be able to find correct ranks if they can find correct lag orders (comparing the last two columns with the two to their left).

Table 6. Performance of IC and trace tests for cointegrating rank: Simulation based on Quenouille's U.S. hog data

Correct choices of r						Correct choices of both p and r	
Two-step				One-step		One-step	
AICAIC	SCSC	λ_{tr}	λ_{tr}^*	AIC	SC	AIC	SC
$T = 30$							
7.5	40.4	46.2	21.6	0.7	30.2	0.2	14.3
$T = 50$							
12.1	62.8	64.7	47.5	10.4	59.4	9.1	51.0
$T = 100$							
16.7	95.2	89.9	92.1	17.2	94.9	16.9	94.8
$T = 200$							
19.7	98.8	92.3	93.6	19.5	98.7	19.4	98.7

6. SUMMARY

Information criteria are widely used in selecting the lag order of time series models. In this paper, we investigate whether they are also useful in the cointegration analysis by conducting three separate Monte Carlo simulations. We summarize the major findings as follows.

First, the design of the first simulation is of "canonic form," which is invariant to the nonsingular transformation of the original series. The second simulation design allows investigation of the test performance under more detailed assumptions on model specifications. Simulation results from these two different designs agree when the underlying DGPs are similar (e.g., when $r = 0$, or when $r = 1$ in the models free of moving average correlation). This suggests that the DGPs used here are likely to be representative in other cointegration analyses.

Second, the IC approach can either be used to determine the lag order and cointegrating rank of the VAR in two steps, or it can be used to determine them in one step. The one-step AIC in general performs better than the two-step AICAIC. There is also some gain in using the one-step SC if the underlying DGP is nonstationary or the sample size is 100 or higher.

Third, AIC performs better than SC and the trace tests when the true DGP is stationary (of full rank). But in most cases, it converges to true models slowly in the first simulation. It does not converge in the second simulation. This result agrees with the theoretical result that AIC is inconsistent in selecting lag order or cointegrating rank.

Fourth, although SC's performance is close to that of the trace test in most cases, it may perform better than equally as well as, or worse than the trace tests depending on the presence of linear time trends, the strength of correlation between two series, and the absence of moving averages in the innovations. The results show that SC is consistent in the joint estimation of lag order and cointegrating rank. Furthermore, when the sample size is larger than 100, SC performs at least as well as, and many times better than, the trace test in selecting cointegrating rank for all model specifications.

The simulation based on a five-variable system shows that the IC performance obtained from the bivariate models may extend to larger VAR models. In particular, the one-step SC still performs quite well in the selection of both lag order and cointegrating rank if the sample size is at least moderate (larger than 50).

To conclude, future research could proceed by providing additional simulation and empirical evidence on the IC performance in larger systems and by considering alternative IC.

NOTES

1. Johansen's likelihood ratio test can be implemented in two forms. In this study, we will use his trace test statistic. Because the alternative maximum eigenvalue test has similar power to the trace test, it is not examined here.

2. We also investigated PIC as recently discussed in Chao and Phillips (1999). Using the approximation given in their equation (20), we find our Monte Carlo results on PIC close to those discussed here for SC. Results are available from the senior author.

3. To make our results directly comparable to the relevant parts of Toda (1995) for this simulation, and those of Haug (1996) for the second simulation, we deliberately use the same notations to specify the DGPs as their corresponding sources. Thus, the same parameters (e.g., ψ and θ) do not have the same meaning in the two simulations. Defining them in two sections separately, we hope to minimize the confusion caused by this abuse of notation.

4. The actual output contains more specific information on what exact model ($r = 0$, 1, or 2) is chosen. To save space here we only report the success ratio for each test, namely, the percentage that models with $r = 0$ are chosen in the case of Table 1.

5. We thank an anonymous referee for the suggestion to add the discussion in this paragraph.

6. Correspondingly, the empirical size of λ_{tr} is $1 - 0.878 = 0.122$ when $\theta = 0.8$, $\eta = -0.5$, and $T = 100$ (Table A.3), which is also slightly higher than the λ_{max}(SC)'s 0.094 in Haug (1996, Table 3, p. 105), implying more serious size distortion.

7. Both the one-step AIC and the two-step AICAIC conclude with the highly parameterized model ($r = 4$). To save space, the details of these results are not presented in Table 6.

REFERENCES

Ahn, S.K & C.G. Reinsel (1990) Estimation for partially nonstationary multivariate autoregressive models. *Journal of the American Statistical Association* 85, 813–823.

Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. In B.N. Petrov & F. Csaki (eds.), *Second International Symposium on Information Theory*, pp. 267–281. Akademiai Kiado.

Aznar, A. & M. Salvador (2002) Selecting the rank of cointegration space and the form of the intercept using an information criterion. *Econometric Theory* 18, 926–947.

Banerjee, A., J.J. Dolado, D.F. Hendry, & G.W. Smith (1986) Exploring equilibrium relationships in econometrics through static models: Some Monte Carlo evidence. *Oxford Bulletin of Economics and Statistics* 48, 253–277.

Blangiewicz, M. & W. Charemza (1990) Cointegration in small samples: Empirical percentiles, drifting moments and customized testing. *Oxford Bulletin of Economics and Statistics* 52, 303–315.

Boswijk, H.P. & P.H. Frasnes (1992) Dynamic specification and cointegration. *Oxford Bulletin of Economics and Statistics* 54, 369–381.

Box, G.E.P. & G.C. Tiao (1977) A canonical analysis of multiple time series. *Biometrika* 64, 355–365.

Chao, J.C. & P.C.B. Phillips (1999) Model selection in partially nonstationary vector autoregressive processes with reduced rank structure. *Journal of Econometrics* 91, 227–271.

Cheung, Y.W. & K.S. Lai (1993) Finite sample sizes of Johansen's likelihood ratio tests for cointegration. *Oxford Bulletin of Economics and Statistics* 55, 313–332.

Doan, T.A. (1996) *User's Manual: RATS 4.0*. Estima.

Engle, R.F. & C.W.J. Granger (1987) Cointegration and error correction: Representation, estimation and testing. *Econometrica* 5, 251–276.

Gonzalo, J. (1994) Comparison of five alternative methods of estimating long-run equilibrium relationships. *Journal of Econometrics* 60, 203–233.

Gonzalo, J. & J.Y. Pitarakis (1999) Dimensionality effect in cointegration analysis. In R.F. Engle & H. White (eds.), *Cointegration, Causality and Forecasting: Festschrift in Honour of Clive Granger*, pp. 212–229. Oxford: Oxford University Press.

Gonzalo, J. & J.Y. Pitarakis (1998) Specification via model selection in vector error correction models. *Economics Letters* 60, 321–328.

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

Hansen, B.E. & P.C.B. Phillips (1990) Estimation and inference in models of cointegration: A simulation study. *Advances in Econometrics* 8, 225–248.

Hansen, H. & K. Juselius (1995) *CATS in RATS*. Estima.

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

Johansen, S. (1988) Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12, 231–254.

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

Johansen, S. (1992) Determination of cointegration rank in the presence of a linear trend. *Oxford Bulletin of Economics and Statistics* 54, 383–397.

Johansen, S. (1996) *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

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

Johansen, S. (2002) A small sample correction for the test of cointegrating rank in the vector autoregressions. *Econometrica* 70, 1929–1961.

Kapetanios, G. (2004) The asymptotic distribution of the cointegration rank estimator under the Akaike information criterion. *Econometric Theory* 20, 735–742.

Kremers, J.J., N.R. Ericsson, & J.J. Dolado (1992) The power of co-integration tests. *Oxford Bulletin of Economics and Statistics* 54, 325–348.

Maddala, G.S. & I.M. Kim (1998) *Unit Roots, Cointegration, and Structural Change*. Cambridge University Press.

Phillips, P.C.B. (1996) Econometric model determination. *Econometrica* 64, 763–812.

Phillips, P.C.B. & J.W. McFarland (1997) Forward exchange market unbiasedness: The case of the Australian dollar since 1984. *Journal of International Money and Finance* 16, 885–907.

Phillips, P.C.B. & W. Ploberger (1996) An asymptotic theory of Bayesian inference for time series. *Econometrica* 64, 381–412.

Podivinsky, J.M. (1990) Testing for a unit root in time series regression. *Biometrika* 75, 335–346.

Quenouille, M.H. (1957) *The Analysis of Multiple Time Series*. Griffin.

Reinsel, G.C. (1983) Some results on multivariate autoregressive index models. *Biometrika* 70, 145–156.

Reinsel, G.C. & R.P. Velu (1998) *Multivariate Reduced-Rank Regression.* Springer.

Schwarz, G. (1978) Estimating the dimension of a model. *Annals of Statistics* 8, 461–464.

Stock, J.H. & M.W. Watson (1988) Testing for common trends. *Journal of the American Statistical Association* 83, 1097–1107.

Tiao, G.C. & R.S. Tsay (1983) Multiple time series modeling and extended sample cross-correlations. *Journal of Business and Economic Statistics* 1, 43–56.

Tiao, G.C. & R.S. Tsay (1989) Model specification in multivariate time series (with discussion). *Journal of the Royal Statistical Society, Series B* 51, 157–213.

Toda, H.Y. (1994) Finite sample properties of likelihood ratio tests for cointegrating ranks with linear trends are present. *Review of Economics and Statistics* 76, 66–79.

Toda, H.Y. (1995) Finite sample performances of likelihood ratio tests for cointegrating ranks in vector autoregressions. *Econometric Theory* 11, 1015–1032.

Velu, R.P., G.C. Reinsel, & D.W. Wichern (1986) Reduced rank models for multiple time series. *Biometrika* 73, 105–118.

Wang, Z. & D.A. Bessler (2002) The homogeneity restriction and forecasting performance of VAR-type demand systems: An empirical examination of U.S. meat consumption. *Journal of Forecasting* 21, 193–206.

Wang, Z. & D.A. Bessler (2004) Forecasting performance of multivariate time series models with full and reduced rank: An empirical examination. *International Journal of Forecasting* 20, 683–695.

APPENDIX: TABULAR RESULTS ON SIMULATION II

TABLE A.1. Performance of IC and trace tests for cointegrating rank: $\rho = 0.5$, Simulation II

	$\theta = -0.8$			$\theta = 0.0$			$\theta = 0.8$		
	$\eta = -0.5$	$\eta = 0.0$	$\eta = 0.5$	$\eta = -0.5$	$\eta = 0.0$	$\eta = 0.5$	$\eta = -0.5$	$\eta = 0.0$	$\eta = 0.5$
$T = 30$									
AICAIC	15.4	13.3	15.2	51.3	50.9	52.3	44.4	42.3	45.1
SCSC	20.6	18.2	22.4	54.8	40.9	56.3	61.2	45.4	62.3
λ_{tr}	46.1	43.6	45.8	42.0	28.6	43.0	58.1	37.8	58.5
λ^*_{tr}	46.6	44.1	45.8	39.2	25.1	39.7	52.1	29.8	52.7
AIC	18.2	16.1	18.1	53.0	52.2	54.0	45.9	44.5	46.5
SC	21.1	18.8	21.9	54.0	39.9	55.2	50.9	40.4	52.0
$T = 50$									
AICAIC	12.3	9.3	13.1	55.2	54.8	55.9	50.4	46.6	51.1
SCSC	6.6	4.1	7.1	82.0	64.9	81.2	62.7	54.1	63.0
λ_{tr}	13.1	8.7	13.1	81.9	63.3	81.7	64.5	56.7	64.9
λ^*_{tr}	13.4	9.1	13.4	80.8	61.2	80.9	59.8	50.1	60.1
AIC	14.5	11.1	15.1	56.1	55.6	56.7	51.5	48.6	52.0
SC	7.2	5.0	7.8	81.8	64.6	81.0	53.9	50.5	54.5
$T = 100$									
AICAIC	19.0	14.7	20.0	56.8	57.0	57.8	54.3	53.9	56.4
SCSC	14.6	8.9	15.4	95.1	94.2	94.8	85.2	81.0	84.4
λ_{tr}	15.7	9.9	16.4	95.1	94.0	94.6	89.1	85.5	88.7
λ^*_{tr}	15.8	10.0	16.6	95.3	94.3	95.0	88.5	84.1	87.6
AIC	21.0	16.2	22.2	57.2	57.5	58.3	55.1	55.0	57.1
SC	16.8	10.9	18.0	95.2	94.2	94.9	83.5	79.9	82.6
$T = 200$									
AICAIC	27.8	22.5	28.7	58.1	58.7	57.6	56.9	56.5	57.1
SCSC	35.0	24.8	34.9	97.4	97.2	97.2	96.4	94.9	96.5
λ_{tr}	29.7	20.8	29.7	94.5	94.7	94.5	93.7	91.9	93.7
λ^*_{tr}	29.9	21.1	29.9	94.8	94.7	94.6	94.2	92.3	94.1
AIC	29.1	23.5	29.8	58.3	58.9	57.7	57.2	57.0	57.4
SC	39.3	28.6	39.3	97.4	97.2	97.2	96.5	95.7	96.6

Note: Each entry is the percentage that the cointegrating rank of the simulated data is correctly determined based on the two-step AICAIC, SCSC, trace test (λ_{tr}), and small sample adjusted trace test (λ^*_{tr}) and the one-step AIC and SC procedures. The DGP is a bivariate VAR with cointegrating rank of 1.

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

TABLE A.2. Performance of IC and trace tests for cointegrating rank: $\rho = 0.85$, Simulation II

	$\theta = -0.8$			$\theta = 0.0$			$\theta = 0.8$		
	$\eta = -0.5$	$\eta = 0.0$	$\eta = 0.5$	$\eta = -0.5$	$\eta = 0.0$	$\eta = 0.5$	$\eta = -0.5$	$\eta = 0.0$	$\eta = 0.5$
$T = 30$									
AICAIC	44.9	44.7	44.9	42.1	42.4	41.4	41.8	42.4	41.3
SCSC	72.3	70.6	72.5	15.9	14.6	16.0	37.6	29.0	38.9
λ_{tr}	76.8	71.2	76.4	10.7	8.3	9.8	30.6	21.5	31.4
λ_{tr}^*	74.9	68.0	74.7	9.2	6.5	8.8	25.0	15.0	25.6
AIC	46.0	46.3	46.3	41.4	41.5	41.0	42.0	42.7	41.7
SC	69.5	68.3	69.3	14.7	13.3	14.8	28.3	22.3	28.6
$T = 50$									
AICAIC	31.4	29.3	31.4	45.5	43.1	45.0	45.8	43.7	44.2
SCSC	71.3	70.4	70.9	12.7	8.3	13.0	27.7	17.5	28.4
λ_{tr}	79.9	79.6	78.8	12.9	8.9	13.4	28.4	18.5	28.6
λ_{tr}^*	80.2	79.7	79.1	12.4	8.3	12.8	24.7	14.3	24.9
AIC	32.9	30.8	32.4	45.3	42.7	44.7	44.9	43.1	43.9
SC	65.0	67.1	65.7	12.4	7.8	12.5	16.7	12.9	17.8
$T = 100$									
AICAIC	19.6	16.0	20.0	54.9	52.2	55.8	49.6	47.6	51.6
SCSC	42.6	43.0	41.4	21.6	10.2	21.2	20.2	11.4	20.1
λ_{tr}	47.8	46.3	47.1	37.1	22.1	35.3	33.0	24.6	32.0
λ_{tr}^*	48.0	46.9	47.1	36.7	21.7	35.4	29.1	21.5	28.5
AIC	20.4	17.2	21.1	54.9	52.2	55.8	49.6	47.9	51.8
SC	37.2	38.4	35.6	21.5	10.1	21.1	14.5	9.7	14.3
$T = 200$									
AICAIC	27.5	22.5	28.0	57.4	58.5	56.8	57.3	55.7	56.1
SCSC	14.2	11.0	14.4	69.2	36.8	69.2	52.3	28.0	52.5
λ_{tr}	25.9	19.2	27.0	88.3	70.9	88.5	76.2	60.3	77.1
λ_{tr}^*	25.9	19.6	27.1	88.5	70.7	88.7	74.8	58.0	75.6
AIC	28.6	23.7	28.9	57.6	58.7	57.1	57.6	56.3	56.5
SC	13.1	10.6	12.5	69.2	36.8	69.1	45.0	26.0	45.6

Note: Each entry is the percentage that the cointegrating rank of the simulated data is correctly determined based on the two-step AICAIC, SCSC, trace test (λ_{tr}), and small sample adjusted trace test (λ_{tr}^*) and the one-step AIC and SC procedures. The DGP is a bivariate VAR with cointegrating rank of 1.

TABLE A.3. Performance of IC and trace tests for cointegrating rank: $\rho = 1.0$, Simulation II

	$\theta = -0.8$			$\theta = 0.0$			$\theta = 0.8$		
	$\eta = -0.5$	$\eta = 0.0$	$\eta = 0.5$	$\eta = -0.5$	$\eta = 0.0$	$\eta = 0.5$	$\eta = -0.5$	$\eta = 0.0$	$\eta = 0.5$
$T = 30$									
AICAIC	2.1	2.3	2.4	27.0	27.0	27.0	10.3	11.2	10.3
SCSC	4.4	9.5	4.3	84.3	84.3	84.3	60.8	65.1	60.3
λ_{tr}	10.5	22.1	11.2	91.3	91.3	91.3	73.0	75.7	71.7
λ_{tr}^*	12.4	25.8	13.2	92.9	92.9	92.9	82.0	83.8	80.9
AIC	3.5	3.8	3.8	32.0	32.0	32.0	14.8	16.5	14.9
SC	7.3	12.3	7.9	85.8	85.8	85.8	71.5	74.1	71.1
$T = 50$									
AICAIC	3.5	3.5	3.4	35.4	35.4	35.4	21.9	22.5	22.4
SCSC	3.1	4.8	2.9	93.1	93.1	93.1	81.0	82.0	80.1
λ_{tr}	3.4	6.2	3.5	93.6	93.6	93.6	81.8	82.3	80.8
λ_{tr}^*	3.6	6.8	3.8	94.3	94.3	94.3	86.5	86.8	85.6
AIC	4.9	5.2	5.0	37.9	37.9	37.9	26.3	26.9	26.6
SC	7.6	8.7	6.9	93.4	93.4	93.4	87.0	87.1	86.5
$T = 100$									
AICAIC	7.4	7.7	7.8	39.3	39.3	39.3	31.0	31.1	30.7
SCSC	10.1	9.0	10.4	97.7	97.7	97.7	94.5	94.1	94.1
λ_{tr}	8.5	7.5	8.6	93.9	93.9	93.9	87.8	88.3	87.2
λ_{tr}^*	8.8	8.0	9.0	94.0	94.0	94.0	90.4	90.7	89.8
AIC	8.9	9.4	9.5	40.5	40.5	40.5	33.5	33.9	33.2
SC	15.6	15.5	16.0	97.7	97.7	97.7	96.1	95.7	95.9
$T = 200$									
AICAIC	13.9	14.1	14.3	41.3	41.3	41.3	36.4	36.5	36.6
SCSC	25.5	26.8	25.8	99.3	99.3	99.3	98.5	98.4	98.3
λ_{tr}	17.4	17.7	18.5	94.6	94.6	94.6	90.9	90.8	90.4
λ_{tr}^*	17.9	18.2	18.9	94.7	94.7	94.7	92.2	91.9	91.7
AIC	15.2	15.5	15.4	41.9	41.9	41.9	37.5	37.7	37.9
SC	34.3	34.1	34.3	99.3	99.3	99.3	98.9	98.8	98.8

Note: Each entry is the percentage that the cointegrating rank of the simulated data is correctly determined based on the two-step AICAIC, SCSC, trace test (λ_{tr}), and small sample adjusted trace test (λ_{tr}^*) and the one-step AIC and SC procedures. The DGP is a bivariate VAR without cointegration relations.

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press

Table A.4. Performance of IC and trace tests for cointegrating rank: $\theta = -0.8$, Simulation II

	$\rho = 0.5$		$\rho = 0.85$		$\rho = 1.0$	
	$\eta = 0.0$	$\eta = 0.5$	$\eta = 0.0$	$\eta = 0.5$	$\eta = 0.0$	$\eta = 0.5$
$T = 300$						
AICAIC	25.8	30.9	26.1	30.9	17.6	17.1
SCSC	36.6	48.4	21.8	28.1	41.2	42.9
λ_{tr}	29.2	39.2	29.0	38.6	27.4	28.2
λ^*_{tr}	29.3	39.4	29.3	38.9	28.0	28.9
AIC	26.4	31.6	26.6	31.5	18.5	18.0
SC	40.4	52.2	22.7	28.1	50.2	50.3
$T = 500$						
AICAIC	29.3	33.9	29.0	33.1	21.0	20.1
SCSC	55.1	65.7	53.0	64.6	64.3	64.8
λ_{tr}	41.6	51.6	41.8	52.2	41.1	41.5
λ^*_{tr}	41.7	51.8	42.0	52.5	41.7	41.9
AIC	29.4	34.0	29.2	33.3	21.3	20.4
SC	57.9	68.8	55.5	67.5	70.4	70.6
$T = 1{,}000$						
AICAIC	28.7	34.6	28.5	34.1	20.7	19.4
SCSC	76.0	83.6	75.9	83.6	85.4	85.6
λ_{tr}	57.1	66.5	57.2	66.6	58.2	57.9
λ^*_{tr}	57.3	66.6	57.4	66.8	58.6	58.4
AIC	28.7	34.7	28.5	34.1	20.8	19.4
SC	77.2	84.8	77.5	85.0	88.1	88.0

Note: Each entry is the percentage that the cointegrating rank of the simulated data is correctly determined based on the two-step AICAIC, SCSC, trace test (λ_{tr}), and small sample adjusted trace test (λ^*_{tr}) and the one-step AIC and SC procedures. The DGP is a bivariate VAR (equation (10)).

https://doi.org/10.1017/S0266466605050334 Published online by Cambridge University Press