JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **14:** 563–577 (1999)

NUMERICAL DISTRIBUTION FUNCTIONS OF LIKELIHOOD RATIO TESTS FOR COINTEGRATION

JAMES G. MACKINNON,[a]* ALFRED A. HAUG[b] AND LEO MICHELIS[c]

[a]*Department of Economics, Queen's University, Kingston, Ontario, Canada K7L 3N6*
[b]*Department of Economics, University of Canterbury, Private Bag 4800, Christchurch, New Zealand*
[c]*Department of Economics, Ryerson Polytechnic University, 350 Victoria Street, Toronto, Ontario, Canada M5B 2K3*

SUMMARY

This paper employs response surface regressions based on simulation experiments to calculate asymptotic distribution functions for the Johansen-type likelihood ratio tests for cointegration. These are carried out in the context of the models recently proposed by Pesaran, Shin, and Smith (1997) that allow for the possibility of exogenous variables integrated of order one. The paper calculates critical values that are very much more accurate than those available previously. The principal contributions of the paper are a set of data files that contain estimated asymptotic quantiles obtained from response surface estimation and a computer program for utilizing them. This program, which is freely available via the Internet, can be used to calculate both asymptotic critical values and *P*-values. Copyright © 1999 John Wiley & Sons, Ltd.

1. INTRODUCTION

Since the influential work of Engle and Granger (1987), several procedures have been proposed for testing the null hypothesis that two or more non-stationary time series are not cointegrated, meaning that there exist no linear combinations of the series that are stationary. One approach is to use likelihood ratio tests based on estimating a vector autoregression. This approach was first proposed by Johansen (1988) and refined further by Johansen and Juselius (1990) and Johansen (1991, 1992, 1994). There are two different test statistics, which are called the Trace and $\lambda_{\max}$ statistics.

Johansen's approach, which has been used extensively in applied work, provides a unified framework for estimation and testing in the context of a multivariate vector autoregressive model in error correction form (VECM) with normal errors. The normality assumption allows a neat application of maximum likelihood theory, which produces both the test statistics and the maximum likelihood estimators (MLE) of the parameters of interest. Phillips (1991) noted several desirable properties of the MLE for this model and demonstrated that asymptotically optimal inferences can be based on the MLE of the cointegrating vectors. Gonzalo (1994) showed that these properties hold in finite samples even without the normality assumption. Haug (1996), among others, has provided Monte Carlo evidence that the Trace and $\lambda_{\max}$ statistics generally have reasonable size and power properties. Just how accurate they are will depend on the sample size, the number of lags in the vector autoregression, and the data-generating process; see Cheung and Lai (1993).

* Correspondence to: Prof. J. MacKinnon, Dept of Economics, Queen's University, Kingston, Ontario, Canada, K7L 3N6. E-mail: jgm@qed.econ.queensu.ca

Contract/grant sponsor: Social Sciences and Humanities Research Council of Canada.

CCC 0883-7252/99/050563–15$17·50

Received 19 September 1996
Revised 8 December 1998

In the literature, tables of critical values have been computed by simulating the expressions to which the two tests converge asymptotically for five cases (models) of interest; see, in particular, Osterwald-Lenum (1992) and Johansen (1995). Recently, Pesaran, Shin, and Smith (1997), henceforth PSS, has extended the analysis in two ways. PSS includes exogenous I(1) variables in the VECM, and it modifies some of the models so as to make their trending behaviour invariant with respect to the cointegration rank of the whole system. Section 2 provides a comparison of the PSS framework of analysis with that of Johansen (1995).

A major problem with the studies just cited is that their results are not very accurate. There are two reasons for this. First, they employ experiments with no more than 100,000 replications, and sometimes as few as 6000. These are not large numbers for the estimation of tail quantiles. Second, all of them simulate the asymptotic quantities to which the two tests converge by using a discrete approximation with either 400 or 500 steps, instead of using response surfaces.

In this paper, we obtain extremely accurate critical values and marginal significance levels, or *P*-values, for the Trace and $\lambda_{\max}$ tests in the context of the PSS framework of analysis. This is done by adopting the response surface approach of MacKinnon (1994, 1996). To facilitate comparisons with the PSS results, the five different models considered in that paper are analysed for up to 12-dimensional systems with between 0 and 8 exogenous variables. The basic idea is to estimate a large number of quantiles of the distributions of the test statistics, for a number of different sample sizes, by means of Monte Carlo experiments. Response surface regressions, in which the estimated quantiles are regressed on negative powers of the sample size, are then used to estimate the quantiles of the asymptotic distribution. Some of the estimated quantiles from the response surface regressions directly provide asymptotic critical values. The quantiles can also be used as input to a computer program which can calculate the asymptotic *P*-value for any test statistic.

Both the tables of estimated asymptotic quantiles and a computer program called `lrcdist` that uses them are available via the Internet; for details, see the Appendix. The `lrcdist` program is run interactively and prompts the user for input. For those who wish to compute large numbers of critical values or *P*-values, two sets of routines, called `johrouts.f` and `lrcdists.f`, are also provided. These users simply need to write their own main program to call the appropriate routine, which in turn reads the appropriate files and calls other routines to do the calculations.

The rest of the paper is organized as follows. Section 2 discusses the five models and the two likelihood ratio tests for cointegration. Section 3 discusses the simulation experiments and the response surface regressions. Section 4 presents a very small subset of our results in tabular form and explains why they are more accurate than previous results. Section 5 discusses how *P*-values and critical values may be calculated using the response surface estimates.

2. THE MODELS AND TEST STATISTICS

The maximum likelihood theory of systems of potentially cointegrated stochastic variables presupposes that the variables are integrated of order one, or I(1), and that the data-generating process is a Gaussian vector autoregressive model of finite order l, or VAR(l), possibly including some deterministic components. If $\mathbf{Z}_t$ denotes an m-dimensional column vector of I(1) variables, the VAR(l) model can be written as

$$\boldsymbol{\Phi}(L)\mathbf{Z}_t - \boldsymbol{\mu} - \boldsymbol{\gamma}t) = \mathbf{e}_t \tag{1}$$

Here $\mathbf{\Phi}(L)$ is an (m, m) matrix polynomial of order l in the lag operator L, $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ are m-vectors of unknown coefficients, and $\mathbf{e}_t$ is an m-vector of unknown error terms assumed to be NID($\mathbf{0}$, $\mathbf{\Omega}$), with $\mathbf{\Omega}$ positive definite. It is convenient to express equation (1) in the VECM form

$$\Delta \mathbf{Z}_t = \mathbf{\Pi} \mathbf{Z}_{t-1} + \sum_{i=1}^{l-1} \mathbf{\Gamma}_i \Delta \mathbf{Z}_{t-i} + \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t + \mathbf{e}_t \quad t = 1, \ldots, T \tag{2}$$

where $\mathbf{\Pi}$ and the $\mathbf{\Gamma}_i$ are (m, m) matrices of coefficients, and $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ are, respectively, m-vectors of constant and trend coefficients defined by

$$\boldsymbol{\mu}_0 \equiv -\mathbf{\Pi}\boldsymbol{\mu} + (\mathbf{\Gamma} + \mathbf{\Pi})\boldsymbol{\gamma}, \quad \text{and} \quad \boldsymbol{\mu}_1 \equiv -\mathbf{\Pi}\boldsymbol{\gamma} \tag{3}$$

with $\mathbf{\Gamma} = \mathbf{I}_m - \Sigma_{i=1}^{l-1} \mathbf{\Gamma}_i$.

The VECM representation (2) is convenient, because the hypothesis of cointegration can be stated in terms of the rank of the long-run impact matrix, $\mathbf{\Pi}$. In the existing literature, multivariate cointegration has been analysed in the context of five different models corresponding to specific restrictions on the vector parameters $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$; see the five models in Johansen (1995) or, equivalently, the five cases in Osterwald-Lenum (1992).

The PSS paper departs from the existing literature in two important ways. First, it allows for the effects of exogenous I(1) variables on the distribution of the likelihood ratio tests for cointegration; see also Harbo *et al.* (1998). Second, it incorporates explicitly the restrictions (3) into the analysis and accounts for their effects on the properties of the process $\mathbf{Z}_t$ and the null distributions of the Trace and $\lambda_{\max}$ tests. In particular, restricting $\boldsymbol{\mu}_1$ to lie in the range space of $\mathbf{\Pi}$ eliminates the quadratic trend in the level of the process $\mathbf{Z}_t$ that one obtains without the restriction; see, for instance, Johansen (1994). This is desirable, because it makes the trending behaviour of $\mathbf{Z}_t$ independent of the cointegration rank of the system. Otherwise, the number of independent quadratic deterministic trends in $\mathbf{Z}_t$ decreases as the cointegration rank of the system increases.

To describe briefly the models and test statistics that arise in the PSS framework, let $\mathbf{Z}_t$ be partitioned into a p-vector $\mathbf{Y}_t$ and a k-vector $\mathbf{X}_t$, where $k \equiv m - p$, and $\mathbf{X}_t$ is assumed to be weakly exogenous with respect to $\mathbf{\Pi}$. Partitioning the parameters and the error term $\mathbf{e}_t$ of model (2) conformably with $\mathbf{Z}_t = (\mathbf{Y}_t', \mathbf{X}_t')'$, so that $\mathbf{\Pi} = (\mathbf{\Pi}_y' \quad \mathbf{\Pi}_x')'$, and so on, it is easy to derive the conditional VECM for Y_t:

$$\Delta \mathbf{Y}_t = \mathbf{\Pi}_y \mathbf{Z}_{t-1} + \sum_{i=1}^{l-1} \mathbf{\Psi}_i \Delta \mathbf{Z}_{t-i} + \mathbf{\Lambda} \Delta \mathbf{X}_t + \mathbf{c}_0 + \mathbf{c}_1 t + \mathbf{U}_t \quad t = 1, \ldots, T \tag{4}$$

where

$$\mathbf{c}_0 = -\mathbf{\Pi}_y \boldsymbol{\mu} + \mathbf{\Theta}\boldsymbol{\gamma}, \quad \text{and} \quad \mathbf{c}_1 = -\mathbf{\Pi}_y \boldsymbol{\gamma} \tag{5}$$

In model (5), $\mathbf{\Theta}$ is a (p, k) matrix of unknown parameters derived, like the other parameters of (4), from the parameters of the unconditional model (2). Since $\mathbf{e}_t$ is normally distributed, so is $\mathbf{U}_t$.

Further, under the assumption of weak exogeneity with respect to $\mathbf{\Pi}$, which requires $\mathbf{\Pi}_x = 0$, the data-generating process for $\mathbf{X}_t$ takes the VECM form

$$\Delta \mathbf{X}_t = \sum_{i=1}^{l-1} \mathbf{\Gamma}_{xi} \Delta \mathbf{Z}_{t-i} + \boldsymbol{\mu}_{x0} + \mathbf{V}_t \quad t = 1, \ldots, T$$

where $\boldsymbol{\mu}_{x0} = \mathbf{\Gamma}_x \boldsymbol{\gamma}$, in which $\mathbf{\Gamma}_x = [\mathbf{O} : \mathbf{I}_k] - \Sigma_{i=1}^{l-1} \mathbf{\Gamma}_{xi}$, and $\mathbf{V}_t$ is a k-vector of error terms that is multivariate normal with mean vector zero. Therefore, in this framework, the cointegration analysis is based on the assumption that there are at most p cointegrating vectors and that $\text{rank}(\mathbf{\Pi}) \equiv \text{rank}(\mathbf{\Pi}_y)$.

The hypothesis of cointegration can be stated in terms of the conditional long-run impact matrix $\mathbf{\Pi}_y$, which can be written as

$$\mathbf{\Pi}_y = \boldsymbol{\alpha} \boldsymbol{\beta}' \tag{6}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are respectively (p, r) and (m, r) matrices of full rank. If $r = 0$, then $\mathbf{\Pi}_y = \mathbf{0}$, and there exists no linear combination of the elements of $\mathbf{Y}_t$ that is stationary. At the other extreme, if $\text{rank}(\mathbf{\Pi}_y) = p$, then $\mathbf{Y}_t$ is a stationary process if $k = 0$, but will in general be non-stationary if $\mathbf{X}_t$ is I(1). In the intermediate case, when $0 < r < p$, there exist r stationary linear combinations of the elements of $\mathbf{Y}_t$, along with $m - r$ stochastic trends.

Under the hypothesis (6), different restrictions on $\mathbf{c}_0$ and $\mathbf{c}_1$ are crucial in determining the properties of the process Y_t and the various cases of interest that can arise. Following PSS, we consider five submodels of the general model (4), which are ordered from most to least restrictive:

$$\begin{aligned}
\text{Case I: } & \mathbf{c}_0 = \mathbf{0}, \quad \mathbf{c}_1 = \mathbf{0} \\
\text{Case II: } & \mathbf{c}_0 = -\mathbf{\Pi}_y \boldsymbol{\mu}, \quad \mathbf{c}_1 = \mathbf{0} \\
\text{Case III: } & \mathbf{c}_0 \neq \mathbf{0}, \quad \mathbf{c}_1 = \mathbf{0} \\
\text{Case IV: } & \mathbf{c}_0 \neq \mathbf{0}, \quad \mathbf{c}_1 = -\mathbf{\Pi}_y \boldsymbol{\gamma} \\
\text{Case V: } & \mathbf{c}_0 \neq \mathbf{0}, \quad \mathbf{c}_1 \neq \mathbf{0}
\end{aligned}$$

In addition to the above restrictions, which are explicitly imposed on the likelihood function in each case, there are some additional restrictions that are implicitly assumed. First, for cases I, II, and III, $\boldsymbol{\gamma} = \mathbf{0}$, and hence $\boldsymbol{\mu}_{x0} = \mathbf{0}$. If this restriction is violated, then a nuisance parameter appears in the limit distributions of the two likelihood ratio tests, which makes inference difficult; see Theorem 2 in Harbo *et al.* (1998). Second, $\boldsymbol{\gamma} = \mathbf{0}$ also implies, via model (5), that $\mathbf{c}_0 = -\mathbf{\Pi}_y \boldsymbol{\mu}$. Even though case II fully incorporates this restriction, case III ignores it. Third, the original model (1), which is maintained throughout the analysis, implies the restriction $\mathbf{c}_1 = -\mathbf{\Pi}_y \boldsymbol{\gamma}$. Although case IV imposes this restriction, case V does not.

Since PSS introduces exogenous I(1) variables into the analysis, the five cases I to V above are not directly comparable to the five cases 0, 1*, 1, 2*, and 2 in Osterwald-Lenum (1992). However, in the special case in which $k = 0$, when there are no exogenous variables in the VAR, cases I, II and IV of PSS are the same as cases 0, 1, and 2* of Osterwald-Lenum. Cases III and V of PSS, however, are different from cases 1 and 2 of Osterwald-Lenum, because the former together with the definitions (3) do not allow for a linear and quadratic trend, respectively, in the level of the process $\mathbf{Y}_t$, whereas the latter do allow for them.

Because of the normality assumption, it is natural to test for the reduced rank of $\mathbf{\Pi}_y$ by using a likelihood ratio test. The procedure uses the technique of reduced rank regression first introduced by Anderson (1951) and applied to systems of I(1) variables independently by Johansen (1988) and Ahn and Reinsel (1990). This technique is appealing because it delivers at once the MLE of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and the eigenvalues needed to construct likelihood ratio tests.

PSS shows how this technique can be applied when there are exogenous I(1) variables in the VAR. Consider the problem of testing the null hypothesis that there are at most r cointegrating vectors against the unrestricted model (4). The null hypothesis is that $\text{rank}(\mathbf{\Pi}_y) = r$, and the alternative is that $\text{rank}(\mathbf{\Pi}_y) = p$. The likelihood ratio test statistic, which is called the Trace statistic, is

$$\text{Trace} = -T \sum_{i=r+1}^{p} \log(1 - \lambda_i) \tag{7}$$

where the λ_i are the eigenvalues, ordered from smallest to largest, which arise in the solution of the reduced rank regression problem. The testing is performed sequentially either for $r = p-1, \ldots, 0$ or for $r = 0, \ldots, p-1$. The testing sequence terminates when the null is rejected for the first time in the former case or when it is not rejected for the first time in the latter case. It is also possible to test the null that $\text{rank}(\mathbf{\Pi}_y) = r$ against the alternative that $\text{rank}(\mathbf{\Pi}_y) = r+1$. In this case, the likelihood ratio statistic, which is called the $\lambda_{\max}$ statistic, is

$$\lambda_{\max} = -T \log(1 - \lambda_{r+1}) \tag{8}$$

Of course, the $\lambda_{\max}$ statistic is equal to the Trace statistic when $p - r = 1$.

The asymptotic distributions of the Trace and $\lambda_{\max}$ statistics are given, respectively, by the trace and maximal eigenvalue of

$$\int_0^1 \mathrm{d}\mathbf{B}_{p-r}\mathbf{F}'_{m-r}\left(\int_0^1 \mathbf{F}_{m-r}\mathbf{F}'_{m-r}\,\mathrm{d}u\right)^{-1}\int_0^1 \mathbf{F}_{m-r}\,\mathrm{d}\mathbf{B}'_{p-r} \tag{9}$$

where $u \in [0, 1]$, $\mathbf{B}_{p-r}$ is a standard $(p-r)$-dimensional Brownian motion on the unit interval, and $\mathbf{F}_{m-r}$ contains functions of standard $(m-r)$-dimensional Brownian motions on the unit interval. The stochastic process $\mathbf{F}_{m-r}$ depends in addition on the restrictions imposed on the deterministic component of the VECM (4). Definitions of $\mathbf{F}_{m-r}$ for the five cases are provided in Table I.

3. THE SIMULATION EXPERIMENTS

The simulations used 12 different discrete-time approximations to expression (9), in which the Brownian motion was replaced by a Gaussian random walk. Let each element of the $(m-r)$-vector $\mathbf{z}_t$ follow an independent random walk with $N(\mathbf{0}, \mathbf{I})$ innovations, with $\mathbf{z}_0 = \mathbf{0}$, where $\mathbf{z}_t$ is partitioned into a $(p-r)$-vector $\mathbf{y}_t$ and a k-vector $\mathbf{x}_t$, with $k = m - p$. In expression (9), $\mathrm{d}\mathbf{B}_{p-r}$ is replaced by $\mathbf{y}_t - \mathbf{y}_{t-1}$, and $\mathbf{F}_{m-r}$ is replaced by $\mathbf{F}_t$, where the latter is obtained by replacing $\mathbf{B}_{m-r}$ by $\mathbf{z}_{t-1}$ in the various definitions of $\mathbf{F}_{m-r}$; see Table I. Thus equation (9) is replaced in the

Table I. An overview of the various cases simulated

A. Specifications of VAR and VECM

Case	Case in O-L	Deterministic component in the levels VAR			Deterministic component in the VECM: $\mathbf{c}_0 + \mathbf{c}_1 t$	
		Intercept	t	t^2		
I	0	No	No	No	$\mathbf{c}_0 = \mathbf{0}$	$\mathbf{c}_1 = \mathbf{0}$
II	1*	Rest.	No	No	$\mathbf{c}_0 = -\mathbf{\Pi}_y \boldsymbol{\mu}$	$\mathbf{c}_1 = \mathbf{0}$
III		Unrest.	No	No	$\mathbf{c}_0 \neq \mathbf{0}$	$\mathbf{c}_1 = \mathbf{0}$
IV	2*	Unrest.	Rest.	No	$\mathbf{c}_0 \neq \mathbf{0}$	$\mathbf{c}_1 = -\mathbf{\Pi}_y \boldsymbol{\gamma}$
V		Unrest.	Unrest.	No	$\mathbf{c}_0 \neq \mathbf{0}$	$\mathbf{c}_1 \neq \mathbf{0}$
	1	Unrest.	Rest.	No	$\mathbf{c}_0 \neq \mathbf{0}$	$\mathbf{c}_1 = \mathbf{0}$
	2	Unrest.	Unrest.	Rest.	$\mathbf{c}_0 \neq \mathbf{0}$	$\mathbf{c}_1 \neq \mathbf{0}$

B. Asymptotic distributions

Case	$\mathbf{F}_{m-r}$	$\mathbf{F}_t$
I	$\mathbf{B}_{m-r}(u)$	$\mathbf{z}_{t-1}$
II	$[\mathbf{B}'_{m-r}(u), 1]'$	$[\mathbf{z}'_{t-1}, 1]'$
III	$\tilde{\mathbf{B}}_{m-r}(u)$	$\tilde{\mathbf{z}}_{t-1}$
IV	$[\tilde{\mathbf{B}}'_{m-r}(u), u - 0{\cdot}5]'$	$[\tilde{\mathbf{z}}'_{t-1}, t - 0{\cdot}5T]'$
V	$\hat{\mathbf{B}}_{m-r}(u)$	$\hat{\mathbf{z}}_{t-1}$

Notes: In Panel A, 'Case in O-L' refers to Osterwald-Lenum (1992) with $k = 0$. In Panel B, $\mathbf{B}_{m-r}(u)$ is a standard $(m-r)$-dimensional Brownian motion on the unit interval, and $\mathbf{z}_t$ is an $(m-r)$-vector that follows a random walk with innovations that are $N(\mathbf{0}, \mathbf{I})$. The following functions of $\mathbf{B}_{m-r}(u)$ also appear in panel B:

$$\tilde{\mathbf{B}}_{m-r}(u) \equiv \mathbf{B}_{m-r}(u) - \int_0^1 \mathbf{B}_{m-r}(u)\, \mathrm{d}u, \quad \text{and}$$

$$\hat{\mathbf{B}}_{m-r}(u) \equiv \tilde{\mathbf{B}}_{m-r}(u) - 12(u - 0{\cdot}5) \int_0^1 \mathbf{B}_{m-r}(u)\, \mathrm{d}u$$

The discrete-time analogues of these are $\tilde{\mathbf{z}}_{t-1}$, which is $\mathbf{z}_{t-1}$ minus its sample mean, and $\hat{\mathbf{z}}_{t-1}$, which is $\mathbf{z}_{t-1}$ minus the fitted values obtained by regressing it on a constant and t.

simulations by

$$\sum_{t=1}^{T} (\mathbf{y}_t - \mathbf{y}_{t-1}) \mathbf{F}'_t \left(\sum_{t=1}^{T} \mathbf{F}_t \mathbf{F}'_t \right)^{-1} \sum_{t=1}^{T} \mathbf{F}_t (\mathbf{y}_t - \mathbf{y}_{t-1})' \tag{10}$$

where T is the number of steps in the discrete-time approximation. Realizations of the quantities to which the Trace and $\lambda_{\max}$ statistics tend asymptotically were computed as the trace and maximum eigenvalue of expression (10) for each of cases I to V.

In order to be able to estimate asymptotic distributions using response surface regressions, we used 12 different values of T. The approach is similar to the one used by MacKinnon (1994, 1996) to compute the asymptotic distributions of Dickey–Fuller unit root and cointegration tests. For every $(T, p-r)$ pair, with $p - r = 1, \ldots, 12$, we performed 50 experiments, each with 100,000 replications. We did this because it would have been impossible to keep results for all 5 million replications in memory at once, and because the observed variation among the 50 experiments

provided an easy way to measure experimental randomness. For each $(T, p-r)$ pair, results for all values of $k = 0, \ldots, 8$ were computed simultaneously.

This approach made it convenient to divide the experiments among several computers. The experiments were performed on 10 different computers, half of them IBM RS/6000 machines running AIX, and half of them 200 MHz Pentium Pro machines running Linux, over a period of several months. They would have required about two years of CPU time on a single Pentium Pro machine. Because it would have been impractical to store all the simulated test statistics, 221 quantiles were estimated and stored for each experiment. These quantiles were: 0·0001, 0·0002, 0·0005, 0·001, . . . , 0·010, 0·015, . . . , 0·985, 0·990, 0·991, . . . , 0·999, 0·9995, 0·9998, 0·9999. The 221 quantiles provide more than enough information about the shapes of the cumulative distribution functions of the various test statistics.

Because so many random numbers were used, it was important to use a pseudo-random number generator with a very long period. The generator employed was also used in MacKinnon (1994, 1996). It combines two different uniform pseudo-random number generators recommended by L'Ecuyer (1988). The two generators were started with different seeds and allowed to run independently, so that two independent uniform pseudo-random numbers were generated at once. The procedure of Marsaglia and Bray (1964) was then used to transform them into two $N(0,1)$ pseudo-random variates.

The estimated finite-sample quantiles from the simulation experiments were used to estimate response surface regressions, one for each of the 221 asymptotic quantiles used to describe each asymptotic distribution. Consider the estimation of the α quantile for some test statistic. Let $q^{\alpha}(T_i)$ denote the estimate of that quantile based on the ith experiment, for which the sample size is T_i. Then the response surface regressions have the form

$$q^{\alpha}(T_i) = \theta_{\infty}^{\alpha} + \theta_1^{\alpha} T_i^{-1} + \theta_2^{\alpha} T_i^{-2} + \theta_3^{\alpha} T_i^{-3} + \varepsilon_i \tag{11}$$

The first parameter here, θ_{∞}^{α}, is the α quantile of the asymptotic distribution, which is what we are trying to estimate. The other three parameters allow the finite-sample distributions to differ from the asymptotic ones.

Based on preliminary experiments and the experience obtained in earlier work, we used the following 12 sample sizes: 80, 90, 100, 120, 150, 200, 400, 500, 600, 800, 1000, 1200. It is obvious that it is desirable for there to be some large values of T_i. What may be less obvious is that it is also desirable for there to be several small values of T_i, because the smaller is the smallest value of T_i, the more trouble the other regressors in equation (11) have explaining the constant term, and thus the smaller is the standard error of θ_{∞}^{α}. However, none of the T_i's should be too small, because then equation (11) may not fit satisfactorily; it is, after all, just an approximation. Computation costs, of course, increase as T_i increases.

Equation (11) was estimated 221 times for each of 1035 different test statistics; there are 1035, rather than $1080 = 5 \times 2 \times 12 \times 9$, because the Trace and $\lambda_{\max}$ statistics are the same when $p - r = 1$. Each estimation normally used 600 observations. However, in a few cases when $p - r + k$ was large, it was necessary to drop the observations corresponding to one or more of the smallest values of T_i.

As in MacKinnon (1996), we employed a form of GMM estimation to allow for the fact that the error terms of equation (11) are heteroscedastic. Let $\mathbf{q}^{\alpha}$ denote the regressand, $\boldsymbol{\theta}$ the vector of coefficients, and $\mathbf{V}$ the matrix of regressors in equation (11). Further, let $\boldsymbol{\Omega}$ denote the covariance matrix of the error terms. This matrix is diagonal, because all the experiments are independent.

The estimator we used was

$$\hat{\boldsymbol{\theta}} = (\mathbf{V}'\mathbf{W}(\mathbf{W}'\hat{\boldsymbol{\Omega}}\mathbf{W})^{-1}\mathbf{W}'\mathbf{V})^{-1}\mathbf{V}'\mathbf{W}(\mathbf{W}'\hat{\boldsymbol{\Omega}}\mathbf{W})^{-1}\mathbf{W}'\mathbf{q}^{\alpha} \tag{12}$$

where $\mathbf{W}$ is a matrix of up to 12 zero-one dummy variables, the first equal to 1 when $T_i = 80$, the second equal to 1 when $T_i = 90$, and so on. The matrix $\hat{\boldsymbol{\Omega}}$ was obtained by first running an OLS regression of $\mathbf{q}^{\alpha}$ on $\mathbf{W}$ and then regressing the squared residuals on a constant, $1/T$, and $1/T^2$. The fitted values from this auxiliary regression were then used as the diagonal elements of $\hat{\boldsymbol{\Omega}}$.

This GMM estimation procedure automatically generates a statistic for testing the specification of the response surface equation (11). The test statistic is the minimized value of the criterion function,

$$(\mathbf{q}^{\alpha} - \mathbf{V}\boldsymbol{\theta})'\mathbf{W}(\mathbf{W}'\hat{\boldsymbol{\Omega}}\mathbf{W})^{-1}\mathbf{W}'(\mathbf{q}^{\alpha} - \mathbf{V}\boldsymbol{\theta}) \tag{13}$$

Standard results about GMM estimation imply that, under the null hypothesis that equation (11) is a correct specification, equation (13) is asymptotically distributed as $\chi^2(d)$, where d is equal to the number of distinct T_i's (which may be 12 or less) minus the number of parameters in equation (11).

The GMM test statistic (13) played a key role in the specification of the response surfaces. In order to avoid discontinuities caused by changes in functional form, the same response surface regression was estimated for every one of the 221 quantiles for a given distribution. The average value of the 221 test statistics was used to decide whether to set $\theta_3^{\alpha} = 0$ in equation (11) and, in a few cases, to determine how many small values of T_i to drop. Since the objective was to obtain efficient estimates of θ_{∞}^{α}, it was desirable to set $\theta_3^{\alpha} = 0$, if possible. On average, for a correctly specified response surface, reducing the number of distinct T_i's by 1, or dropping the constraint that $\theta_3^{\alpha} = 0$ in equation (11), would be expected to reduce the value of expression (13) by 1·0, because the mean of a random variable with a $\chi^2(d)$ distribution is d. In most cases, we chose to reject a model when such a change reduced the value of (13) by more than 2·5. The fit of the response surface regressions tended to deteriorate as either $p - r$ or k increased. For small values of $p - r + k$, the restriction that $\theta_3^{\alpha} = 0$ was almost always compatible with the data. Only for a few large values of $p - r + k$ was it ever necessary to drop observations corresponding to one or more of the smallest values of T_i.

The simulations described above are not the only ones we performed. In preliminary work, we also performed simulations for cases 0, 1, 1*, 2, and 2* of Osterwald-Lenum (1992). For these simulations, the actual values of the test statistics were calculated instead of asymptotic approximations, different values of the T_i were used, and the details of how equation (11) was estimated were not quite the same. The two sets of simulations yielded almost identical results for cases that are comparable, but the earlier ones provide the only results we have for cases 1 and 2.

4. NUMERICAL DISTRIBUTIONS

The principal results of this paper are 228,735 ($= 221 \times 1035$) estimates of θ_{∞}^{α}. These estimates, which are very much more accurate than any published previously, allow us to construct tables of asymptotic critical values directly. In addition, as we will discuss in the next section, they allow us to obtain asymptotic P-values for any observed test statistic.

Tables II–VI present asymptotic critical values at the 5% level for all the tests we examined. These critical values differ substantially from those previously published by Johansen and Juselius (1990), Osterwald-Lenum (1992), Johansen (1995), and PSS, especially when $p-r$ is large. There appear to be two reasons for this. First, we used more replications than did the earlier authors; Johansen and Juselius (1990) and Osterwald-Lenum (1992) each used 6000, Johansen (1995) used 100,000, and PSS used 10,000. Our critical values therefore suffer from much less experimental error. Second, the estimates of θ_{∞}^{α} from equation (11) really are estimates of the quantiles of an asymptotic distribution, while the values previously published by others are merely approximations based on a discrete random walk with a finite number of steps, either $T = 400$ or $T = 500$.

It appears that approximations based on finite T are simply not very accurate, especially when $p-r$ is large. As an example, for case II with $p-r = 12$ and $k = 0$, which is the same as case 1* with $p-r = 12$, Table III reports a 5% critical value of 348·98 for the Trace test. In contrast, Johansen (1995) reports a critical value of 338·10, and PSS reports a critical value of 341·2. The discrepancies between these numbers are mainly due to the fact that neither Johansen nor PSS actually computed asymptotic critical values. Instead, Johansen used a discrete approximation

Table II. Five per cent critical values for case I

$p-r$ / k	0	1	2	3	4	5	6	7	8
1	4·13	8·11	11·37	14·36	17·20	19·95	22·63	25·26	27·84
	4·13	8·11	11·37	14·36	17·20	19·95	22·63	25·26	27·84
2	11·23	14·96	18·33	21·49	24·53	27·47	30·35	33·16	35·92
	12·32	18·25	23·72	28·95	34·03	39·01	43·91	48·75	53·53
3	17·80	21·43	24·83	28·08	31·21	34·27	37·23	40·16	43·05
	24·28	32·20	39·79	47·16	54·39	61·53	68·57	75·54	82·47
4	24·16	27·76	31·18	34·46	37·66	40·77	43·81	46·82	49·78
	40·17	50·12	59·77	69·23	78·57	87·79	96·96	106·07	115·08
5	30·42	33·99	37·41	40·73	43·96	47·10	50·22	53·27	56·28
	60·06	71·98	83·69	95·22	106·62	117·92	129·16	140·35	151·49
6	36·61	40·16	43·59	46·93	50·18	53·39	56·54	59·63	62·67
	83·94	97·86	111·59	125·18	138·63	152·00	165·30	178·54	191·72
7	42·76	46·29	49·75	53·09	56·37	59·60	62·76	65·88	68·96
	111·79	127·73	143·49	159·12	174·65	190·08	205·43	220·74	235·98
8	48·87	52·40	55·83	59·19	62·50	65·73	68·92	72·07	75·18
	143·64	161·59	179·33	196·99	214·58	232·03	249·43	266·76	284·02
9	54·97	58·49	61·92	65·31	68·62	71·88	75·09	78·27	81·39
	179·48	199·41	219·24	238·91	258·51	278·02	297·44	316·80	336·12
10	61·04	64·54	67·98	71·36	74·68	77·95	81·18	84·38	87·54
	219·38	241·31	263·12	284·82	306·43	327·96	349·42	370·84	392·22
11	67·06	70·58	74·02	77·42	80·74	84·02	87·26	90·47	93·65
	263·25	287·20	310·97	334·69	358·33	381·86	405·37	428·79	452·16
12	73·10	76·60	80·04	83·41	86·76	90·06	93·34	96·56	99·77
	311·09	336·98	362·81	388·56	414·19	439·77	465·29	490·71	516·16

Note: The top entry in each cell is the 5% critical value for $\lambda_{\max}$, and the bottom entry is the 5% critical value for Trace.

Table III. Five per cent critical values for case II

$p-r$ / k	0	1	2	3	4	5	6	7	8
1	9·17	12·34	15·28	18·10	20·84	23·51	26·14	28·71	31·27
	9·17	12·34	15·28	18·10	20·84	23·51	26·14	28·71	31·27
2	15·88	19·21	22·35	25·38	28·31	31·18	34·00	36·76	39·48
	20·25	25·64	30·82	35·89	40·85	45·74	50·57	55·35	60·09
3	22·30	25·68	28·91	32·04	35·09	38·07	40·98	43·87	46·71
	35·19	42·70	50·03	57·23	64·33	71·38	78·33	85·26	92·14
4	28·58	31·99	35·27	38·46	41·57	44·63	47·63	50·58	53·49
	54·09	63·66	73·08	82·38	91·59	100·71	109·82	118·85	127·83
5	34·80	38·22	41·53	44·75	47·90	51·02	54·07	57·08	60·05
	76·96	88·59	100·08	111·46	122·74	133·95	145·12	156·24	167·31
6	40·95	44·37	47·70	50·97	54·17	57·32	60·42	63·46	66·46
	103·84	117·49	131·02	144·45	157·80	171·08	184·31	197·49	210·58
7	47·06	50·51	53·87	57·14	60·36	63·54	66·65	69·72	72·78
	134·70	150·40	165·98	181·47	196·89	212·21	227·51	242·73	257·90
8	53·15	56·59	59·95	63·25	66·51	69·69	72·85	75·97	79·04
	169·54	187·29	204·90	222·41	239·87	257·24	274·56	291·82	309·04
9	59·26	62·69	66·06	69·38	72·65	75·86	79·04	82·18	85·27
	208·41	228·13	247·79	267·34	286·84	306·26	325·59	344·90	364·15
10	65·30	68·73	72·12	75·43	78·71	81·95	85·15	88·31	91·44
	251·31	273·05	294·71	316·26	337·80	359·23	380·63	401·96	423·27
11	71·33	74·76	78·16	81·50	84·77	88·01	91·23	94·41	97·55
	298·16	321·92	345·56	369·16	392·69	416·14	439·57	462·93	486·27
12	77·35	80·79	84·19	87·52	90·82	94·09	97·31	100·51	103·69
	348·98	374·70	400·43	426·05	451·62	477·11	502·57	527·96	553·33

Note: The top entry in each cell is the 5% critical value for $\lambda_{\max}$, and the bottom entry is the 5% critical value for Trace.

with $T = 400$, and PSS used one with $T = 500$. The averages of our experimental results for $T = 400$ and $T = 500$ were 338·00 and 340·12, respectively, which are quite close to the numbers reported by Johansen and PSS. These results demonstrate clearly that, if one wishes to obtain accurate asymptotic critical values, one must use a response surface analysis.

There are good reasons to believe that our estimated asymptotic distributions are extremely accurate. The estimates of θ_{∞}^{α} have standard errors associated with them, and, except for the extreme tail quantiles, these are all very small. For example, for the 5% critical value, the estimated standard errors range from 0·0023 to 0·0433, with the larger standard errors being associated with the larger values of $p-r$ and k. Of course, because of the pretesting involved in choosing which version of equation (11) to estimate, the estimated standard errors may be rather too optimistic.

In at least one special case, it is possible to compare our estimates with outside benchmarks. When $p-r = 1$ and $k = 0$, the distribution of both test statistics for case I (or case 0) is that of the square of the corresponding Dickey–Fuller test. Table VII therefore compares our critical values for this special case with ones reported by Nielsen (1997) based on the analytic formulae of

Table IV. Five per cent critical values for case III

$p-r$ / k	0	1	2	3	4	5	6	7	8
1	8·19	11·42	14·39	17·23	19·97	22·64	25·27	27·85	30·39
	8·19	11·42	14·39	17·23	19·97	22·64	25·27	27·85	30·39
2	15·02	18·36	21·52	24·54	27·48	30·35	33·17	35·93	38·65
	18·11	23·62	28·88	33·98	38·95	43·87	48·72	53·50	58·24
3	21·49	24·87	28·11	31·24	34·29	37·27	40·17	43·06	45·90
	31·88	39·56	46·99	54·27	61·41	68·48	75·46	82·39	89·30
4	27·80	31·20	34·49	37·67	40·78	43·83	46·84	49·78	52·68
	49·64	59·42	68·97	78·34	87·61	96·78	105·91	114·97	123·95
5	34·03	37·44	40·75	43·98	47·13	50·24	53·29	56·30	59·27
	71·44	83·26	94·89	106·36	117·71	128·97	140·19	151·33	162·40
6	40·19	43·61	46·94	50·20	53·40	56·55	59·65	62·69	65·69
	97·26	111·11	124·77	138·30	151·72	165·06	178·33	191·53	204·67
7	46·31	49·75	53·11	56·38	59·60	62·77	65·88	68·96	72·00
	127·05	142·93	158·66	174·26	189·75	205·15	220·48	235·73	250·95
8	52·41	55·84	59·20	62·50	65·75	68·93	72·09	75·21	78·28
	160·87	178·80	196·54	214·15	231·70	249·12	266·49	283·80	301·05
9	58·51	61·94	65·31	68·63	71·89	75·10	78·28	81·42	84·51
	198·72	218·63	238·41	258·04	277·62	297·11	316·49	335·85	355·14
10	64·56	67·99	71·37	74·69	77·97	81·20	84·39	87·56	90·69
	240·58	262·48	284·26	305·92	327·53	349·05	370·49	391·89	413·22
11	70·59	74·03	77·41	80·75	84·03	87·27	90·47	93·65	96·80
	286·39	310·33	334·09	357·78	381·40	404·93	428·42	451·82	475·15
12	76·61	80·06	83·45	86·78	90·08	93·34	96·57	99·76	102·94
	336·22	362·07	387·90	413·63	439·30	464·86	490·33	515·77	541·21

Note The top entry in each cell is the 5% critical value for $\lambda_{\max}$, and the bottom entry is the 5% critical value for Trace.

Abadir (1995). It is clear from the table that our estimates are very accurate in this case, much more accurate than those of Osterwald-Lenum (1992), Johansen (1995), and PSS.

No benchmarks are available for most of our estimates. We have tried to verify that equation (11) is accurate even for values of T much larger than the ones we actually used by performing a (very expensive) set of simulations for the case $T = 5000$, $p - r = 12$, $k = 0$. When equation (11) was re-estimated for this case using these additional simulation results, the GMM test statistics (13) rose only modestly, and the estimated asymptotic distributions did not change appreciably. This is precisely what we would expect to observe if equation (11) is correctly specified.

5. LOCAL APPROXIMATIONS

The response surface coefficient estimates obtained in Section 3 may be used to obtain approximate *P*-values as well as approximate critical values. A program which does both is available via the Internet; see the Appendix.

Table V. Five per cent critical values for case IV

$p-r$ / k	0	1	2	3	4	5	6	7	8
1	12·52	15·46	18·26	20·98	23·63	26·24	28·81	31·35	33·87
	12·52	15·46	18·26	20·98	23·63	26·24	28·81	31·35	33·87
2	19·38	22·50	25·50	28·43	31·28	34·09	36·84	39·56	42·24
	25·86	31·05	36·09	41·03	45·90	50·72	55·50	60·22	64·91
3	25·83	29·04	32·15	35·19	38·15	41·08	43·94	46·77	49·56
	42·92	50·25	57·45	64·54	71·56	78·52	85·44	92·29	99·12
4	32·12	35·38	38·55	41·66	44·71	47·70	50·64	53·55	56·42
	63·87	73·31	82·62	91·81	100·96	110·03	119·03	128·00	136·94
5	38·32	41·62	44·84	47·99	51·10	54·14	57·13	60·10	63·03
	88·79	100·29	111·69	122·96	134·16	145·30	156·44	167·47	178·46
6	44·47	47·79	51·06	54·24	57·38	60·47	63·52	66·53	69·50
	117·69	131·23	144·66	158·01	171·33	184·53	197·70	210·80	223·88
7	50·58	53·94	57·21	60·44	63·61	66·72	69·79	72·82	75·85
	150·55	166·15	181·67	197·07	212·39	227·68	242·90	258·09	273·21
8	56·68	60·03	63·33	66·57	69·76	72·89	76·00	79·08	82·12
	187·44	205·08	222·61	240·05	257·43	274·77	292·03	309·24	326·43
9	62·75	66·11	69·43	72·69	75·91	79·08	82·21	85·31	88·39
	228·32	247·96	267·54	287·03	306·47	325·83	345·11	364·38	383·59
10	68·81	72·17	75·50	78·77	82·00	85·18	88·34	91·47	94·57
	273·20	294·87	316·45	337·95	359·40	380·79	402·18	423·48	444·75
11	74·83	78·21	81·54	84·82	88·07	91·27	94·43	97·61	100·73
	322·03	345·74	369·31	392·86	416·35	439·78	463·15	486·46	509·74
12	80·84	84·23	87·57	90·87	94·13	97·36	100·54	103·70	106·84
	374·84	400·54	426·17	451·78	477·28	502·72	528·13	553·48	578·84

Note: The top entry in each cell is the 5% critical value for $\lambda_{\max}$, and the bottom entry is the 5% critical value for Trace.

In order to obtain a *P*-value for any test statistic or a critical value for any desired test size, some procedure for interpolating between the 221 tabulated values is needed. Many such procedures could be devised, but the one we used, which was proposed by MacKinnon (1996), is appealing and seems to work well. First, consider the regression

$$\Phi^{-1}(\alpha) = \gamma_0 + \gamma_1 \hat{q}(\alpha) + \gamma_2 \hat{q}^2(\alpha) + \gamma_3 \hat{q}^3(\alpha) + e_\alpha \tag{14}$$

where $\Phi^{-1}(\alpha)$ is the inverse of the cumulative standard normal distribution function, evaluated at α. It may seem a bit odd that the regressors in (14) are stochastic and the regressand is not. However, because the estimated quantiles are very accurate, the errors in variables bias that this induces is trivially small; see MacKinnon (1994). If the distribution from which the estimated quantiles were obtained were in fact normal with any mean and variance, regression (14) would be correctly specified with $\gamma_2 = \gamma_3 = 0$. Since that is not the case here, this regression can only be valid as an approximation. Therefore, we estimate it using only a small number of points in the neighbourhood of the observed test statistic, say τ. After (14) has been estimated, the estimated *P*-value associated with τ is given by $\hat{P} = \Phi(\hat{\gamma}_0 + \hat{\gamma}_1\tau + \hat{\gamma}_2\tau^2 + \hat{\gamma}_3\tau^3)$.

Table VI. Five per cent critical values for case V

$p-r$ / k	0	1	2	3	4	5	6	7	8
1	11·64	14·59	17·39	20·11	22·76	25·37	27·94	30·48	33·00
	11·64	14·59	17·39	20·11	22·76	25·37	27·94	30·48	33·00
2	18·55	21·67	24·68	27·60	30·46	33·26	36·01	38·73	41·41
	23·94	29·16	34·22	39·18	44·06	48·88	53·67	58·38	63·08
3	25·03	28·24	31·35	34·38	37·35	40·28	43·14	45·95	48·75
	39·92	47·31	54·54	61·66	68·70	75·67	82·58	89·47	96·29
4	31·34	34·60	37·77	40·87	43·92	46·90	49·85	52·76	55·63
	59·79	69·31	78·66	87·90	97·08	106·17	115·19	124·17	133·11
5	37·55	40·85	44·06	47·21	50·32	53·35	56·35	59·32	62·24
	83·63	95·22	106·67	118·00	129·23	140·42	151·55	162·61	173·62
6	43·71	47·03	50·30	53·48	56·62	59·70	62·74	65·75	68·72
	111·45	125·10	138·59	151·99	165·35	178·59	191·77	204·91	218·01
7	49·83	53·18	56·46	59·68	62·84	65·95	69·02	72·06	75·07
	143·29	158·97	174·54	190·02	205·39	220·69	235·96	251·16	266·30
8	55·92	59·29	62·57	65·81	69·00	72·13	75·24	78·32	81·35
	179·08	196·84	214·44	231·94	249·37	266·76	284·06	301·28	318·52
9	62·01	65·36	68·67	71·93	75·15	78·33	81·45	84·55	87·62
	218·91	238·67	258·33	277·87	297·37	316·78	336·09	355·39	374·62
10	68·07	71·43	74·75	78·02	81·24	84·42	87·58	90·71	93·82
	262·76	284·53	306·19	327·78	349·27	370·73	392·13	413·46	434·75
11	74·10	77·48	80·80	84·08	87·32	90·52	93·69	96·86	99·97
	310·55	334·34	358·00	381·62	405·17	428·64	452·06	475·41	498·75
12	80·11	83·49	86·83	90·13	93·38	96·61	99·80	102·95	106·09
	362·35	388·15	413·85	439·51	465·09	490·59	516·06	541·43	566·80

Note: The top entry in each cell is the 5% critical value fo $\lambda_{\max}$, and the bottom entry is the 5% critical value for Trace.

If we are interested in obtaining approximate critical values, equation (14) has to be turned around. Consider the regression

$$\hat{q}(\alpha) = \delta_0 + \delta_1 \Phi^{-1}(\alpha) + \delta_2 (\Phi^{-1}(\alpha))^2 + \delta_3 (\Phi^{-1}(\alpha))^3 + e_\alpha^* \tag{15}$$

This is not actually the inverse of equation (14). However, if the distribution from which the estimated quantiles were obtained were in fact normal with any mean and variance, equation (15) would be correctly specified with $\delta_2 = \delta_3 = 0$. In that case, equation (14) would have $\gamma_2 = \gamma_3 = 0$, and (15) would be the inverse of (14).

Regressions (14) and (15) could be estimated by OLS, but this would ignore both heteroscedasticity and serial correlation. In MacKinnon (1996), it is shown how to take both of these into account. Therefore, when estimating these equations, our program actually uses the form of feasible GLS estimation proposed in that paper. As discussed above, equations (14) and (15) are to be fitted only to a small number of points near the specified test statistic or test size. Experimentation suggests that 11 points is a good number to use. Also, in many cases, it is possible to set γ_3 or δ_3 equal to zero on the basis of a t-test.

Table VII. Comparisons of critical values. Case I, $p - r = 1$, $k = 0$

	1%	5%	10%
Benchmark	6·9383	4·1293	2·9776
Our estimate	6·9321	4·1296	2·9786
O-L estimate	6·51	3·84	2·86
Johansen estimate	7·02	4·14	2·98
PSS estimate		4·16	3·04

6. CONCLUSION

In this paper, we have used computer simulation and response surface estimation to obtain excellent approximations to the asymptotic distributions of the Trace and $\lambda_{\max}$ tests for VECM systems with up to 12 variables and from 0 to 8 structurally exogenous I(1) variables. Although the paper contains tables of critical values, which are far more accurate than those previously available, the principal results are tables of estimated asymptotic quantiles, in machine-readable form, along with a computer program that uses these to calculate critical values and *P*-values. Both of these are available via the Internet.

APPENDIX

The tables of asymptotic quantiles from the response surface regressions and the associated computer programs may be obtained from the *Journal of Applied Econometrics* Data Archive:

http://www.econ.queensu.ca/jae/

In addition to zipped files that contain the estimated quantiles of the asymptotic distribution functions, the archive contains the Fortran source files `lrcdist.f`, `lrcrouts.f`, and `johrouts.f`, along with a compiled version of `lrcdist` for personal computers (486DX or later) running DOS, Windows, or OS/2. Detailed instructions are provided in the `readme` file.

ACKNOWLEDGEMENTS

This research was supported, in part, by grants from the Social Sciences and Humanities Research Council of Canada to the first two authors. The second author is on leave from York University, Toronto, Canada. We are grateful to Søren Johansen, Bent Nielsen, the Editor, and two referees for comments on earlier drafts.

REFERENCES

Abadir, K. M. (1995), 'The limiting distribution of the *t* ratio under a unit root', *Econometric Theory*, **11**, 775–793.

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

Anderson, T. W. (1951), 'Estimating linear restrictions on regression coefficients for multivariate normal distributions', *Annals of Mathematical Statistics*, **22**, 327–351.

Cheung, Y.-W. and K. S. Lai (1993), 'Finite-sample sizes of Johansen's likelihood ratio tests for cointegration', *Oxford Bulletin of Economics and Statistics*, **55**, 313–328.

Engle, R. F. and C. W. J. Granger (1987), 'Co-integration and error correction: representation, estimation and testing', *Econometrica*, **55**, 251–276.
Gonzalo, J. (1994), 'Five alternative methods of estimating long-run equilibrium relationships', *Journal of Econometrics*, **60**, 203–233.
Harbo, I., S. Johansen, B. Nielsen and A. Rahbek (1998), 'Asymptotic inference on cointegrating rank in partial systems', *Journal of Business and Economic Statistics*, **16**, 388–399.
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 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. (1994), 'The role of the constant and linear terms in cointegration analysis of nonstationary variables', *Econometric Reviews*, **13**, 205–229.
Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.
Johansen, S. and K. Juselius (1990), 'Maximum likelihood estimation and inference on cointegration — with applications to the demand for money', *Oxford Bulletin of Economics and Statistics*, **52**, 169–210.
L'Ecuyer, P. (1988), 'Efficient and portable random number generators', *Communications of the ACM*, **31**, 742–751.
MacKinnon, J. G. (1994), 'Approximate asymptotic distribution functions for unit-root and cointegration tests', *Journal of Business and Economic Statistics*, **12**, 167–176.
MacKinnon, J. G. (1996), 'Numerical distribution functions for unit root and cointegration tests', *Journal of Applied Econometrics*, **11**, 601–618.
Marsaglia, G. and T. A. Bray (1964), 'A convenient method for generating normal variables', *SIAM Review*, **6**, 260–264.
Nielsen, B. (1997), 'Bartlett correction of the unit root test in autoregressive models', *Biometrika*, **84**, 500–504.
Osterwald-Lenum, M. (1992), 'A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics', *Oxford Bulletin of Economics and Statistics*, **54**, 461–471.
Pesaran, M. H., Y. Shin, and R. J. Smith (1997), 'Structural analysis of vector error correction models with exogenous I(1) variables'. University of Cambridge DAE Working Paper No. 9706. Revised November 1998. Available from **http://www.econ.cam.ac.uk/faculty/pesaran/public.htm**
Phillips, P. C. B. (1991), 'Optimal inference in cointegrated systems', *Econometrica*, **59**, 283–306.

