Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration

by

James G. MacKinnon

Department of Economics
Queen's University
Kingston, Ontario, Canada
K7L 3N6

jgm@qed.econ.queensu.ca

Alfred A. Haug

Department of Economics
York University
North York, Ontario, Canada
M3J 1P3

haug@rsrch.econ.yorku.ca

and

Leo Michelis

Department of Economics
Ryerson Polytechnic University
350 Victoria Street
Toronto, Ontario, Canada
M5B 2K3

michelis@acs.ryerson.ca

Please address correspondence to the first author.

This research was supported, in part, by grants from the Social Sciences and Humanities Research Council of Canada to the first two authors. We are grateful to Søren Johansen and Bent Nielsen for comments on an earlier draft.

September, 1996.

Abstract

This paper employs response surface regressions based on simulation experiments to calculate asymptotic distribution functions for the likelihood ratio tests for cointegration proposed by Johansen. The paper provides tables of critical values that are very much more accurate than those available previously. However, 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 easily be used to calculate asymptotic critical values and P values. Graphs of some of the tabulated distribution functions are also provided. An empirical example, motivated by the European Economic and Monetary Union proposed in the Maastricht Treaty, suggests that not all the countries of the European Union may qualify initially for participation in the EMU.

JEL Classification Number: C22

Keywords: cointegration tests, Johansen tests, vector autoregressions, response surfaces, critical values, approximate P values, simulation, EMU, European Union

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 nonstationary 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 $\lambda_{\max}$ tests.

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 have reasonable size and power properties.

Although the asymptotic theory of the Trace and $\lambda_{\max}$ statistics is well understood, relatively little work has been done to obtain accurate critical values for the two tests, especially when the dimension of the system is large. In the literature, tables of critical values have been computed by simulating the expressions to which the two test statistics converge asymptotically. Tables of critical values using this method include those of Johansen (1988), Johansen and Juselius (1990), Osterwald-Lenum (1992), and Johansen (1995). To our knowledge, there has been no work which would allow practitioners to calculate the marginal significance level, or P value, of a test statistic.

A major problem with the studies just cited is that their results are not very accurate. There are two reasons for this. First, all but Johansen (1995) used experiments with only 6000 replications, which is not a large number for the estimation of tail quantiles. Second, all of them simulated the asymptotic quantities to which the two tests converge by using a discrete approximation with only 400 steps. As we shall show in Section 4, this number is too small.

In this paper, we obtain extremely accurate critical values for the Trace and $\lambda_{\max}$ tests by adopting the response surface approach of MacKinnon (1991, 1994, 1996). To facilitate comparisons with existing results, the five different models considered by Osterwald-Lenum (1992) are analyzed for up to 12–dimensional systems. The basic idea is to estimate a large number of quantiles of the finite-sample 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. Moreover, it is easy to use all the quantiles 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 `johdist` that uses them are available via the Internet; for details, see the Appendix. The `johdist` program is run interactively and prompts the user for input. For those who wish to compute large numbers of critical values or P values, a set of routines called `johrouts.f` is also provided. These users simply need to write their own main programs to call the routine `johval`, which in turn reads the appropriate files and calls other routines to do the calculations.

The paper does not concern itself with the important practical problem that asymptotic P values can be seriously inaccurate in finite samples. 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).

The rest of the paper is organized as follows. Section 2 discusses the five cases and the two likelihood ratio tests for cointegration. Section 3 discusses the simulation experiments and the response surface regressions. Section 4 presents tables of critical values for all the cases dealt with in the paper and also provides graphs of the cumulative distribution functions for a few cases. Section 5 discusses how P values may be calculated using the response surface estimates. Finally, Section 6 presents an empirical example. The example, which is motivated by the European Economic and Monetary Union proposed in the Maastricht Treaty, deals with the cointegration of certain macroeconomic variables among the countries of the European Union.

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 k, or VAR(k), possibly including some deterministic components. Let $\boldsymbol{X}_t$ be a p–dimensional column vector of $I(1)$ variables. The VAR(k) model can be written in VECM form as

$$\Delta \boldsymbol{X}_t = \boldsymbol{\Pi} \boldsymbol{X}_{t-1} + \sum_{i=1}^{k-1} \boldsymbol{\Gamma}_i \Delta \boldsymbol{X}_{t-i} + \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t + \boldsymbol{U}_t, \quad t = 1, \ldots, T, \tag{1}$$

where $\boldsymbol{\Pi}$ and the $\boldsymbol{\Gamma}_i$ are $p \times p$ matrices of coefficients, and $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ are $p \times 1$ vectors of constant and trend coefficients, respectively. It will be convenient to let $\boldsymbol{\mu}_t \equiv \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$ denote the deterministic part of the model. The error vector $\boldsymbol{U}_t$, which is $p \times 1$, is assumed to be multivariate normal with mean vector zero and covariance matrix $\boldsymbol{\Omega}$, and to be independent across time periods.

The VECM representation (1) is convenient because the hypothesis of cointegration can be stated in terms of the long run impact matrix, $\boldsymbol{\Pi}$. This matrix can always be written as

$$\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}', \tag{2}$$

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

Under the hypothesis (2), the relation between $\boldsymbol{\alpha}$ and the deterministic term $\boldsymbol{\mu}_t$ is crucial in determining the properties of the process $\boldsymbol{X}_t$ and the various cases of interest that can arise. Consider the decomposition of $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ in the directions $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}_\perp$, where $\boldsymbol{\alpha}_\perp$ is a $p \times (p - r)$ matrix orthogonal to $\boldsymbol{\alpha}$. We can write

$$\boldsymbol{\mu}_i = \boldsymbol{\alpha}\boldsymbol{\beta}_i + \boldsymbol{\alpha}_\perp\boldsymbol{\gamma}_i, \quad i = 0, 1, \tag{3}$$

where $\boldsymbol{\beta}_i = (\boldsymbol{\alpha}'\boldsymbol{\alpha})^{-1}\boldsymbol{\alpha}'\boldsymbol{\mu}_i$ and $\boldsymbol{\gamma}_i = (\boldsymbol{\alpha}_\perp{}'\boldsymbol{\alpha}_\perp)^{-1}\boldsymbol{\alpha}_\perp{}'\boldsymbol{\mu}_i$. Different restrictions on $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ imply different submodels of the general model (1). Following Osterwald-Lenum (1992), we consider five submodels, which are ordered from most to least restrictive:

Case 0:	$\boldsymbol{\mu}_t = \mathbf{0}$
Case 1*:	$\boldsymbol{\mu}_t = \boldsymbol{\alpha}\boldsymbol{\beta}_0$
Case 1:	$\boldsymbol{\mu}_t = \boldsymbol{\mu}_0$
Case 2*:	$\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\alpha}\boldsymbol{\beta}_1 t$
Case 2:	$\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$

The interpretation of each of these models becomes clear by considering the solution of $\boldsymbol{X}_t$ in (1) using a version of the Granger Representation Theorem; see Johansen (1991, Theorem 4.1). Let $\boldsymbol{W}_t$ denote a stationary process, $\boldsymbol{A}$ denote a vector such that $\boldsymbol{\beta}'\boldsymbol{A} = \mathbf{0}$, and $\boldsymbol{C} = \boldsymbol{\beta}_\perp(\boldsymbol{\alpha}_\perp{}'\boldsymbol{\Gamma}\boldsymbol{\beta}_\perp)^{-1}\boldsymbol{\alpha}_\perp{}'$, where $\boldsymbol{\Gamma} = \mathbf{I}_p - \sum_{i=1}^{k-1}\boldsymbol{\Gamma}_i$ and $\boldsymbol{\beta}_\perp$ is a $p \times (p - r)$ matrix of full rank orthogonal to $\boldsymbol{\beta}$. Then this solution is

$$\boldsymbol{X}_t = \boldsymbol{C}\sum_{i=1}^{t}\boldsymbol{U}_i + \tfrac{1}{2}\boldsymbol{\tau}_2 t^2 + \boldsymbol{\tau}_1 t + \boldsymbol{\tau}_0 + \boldsymbol{W}_t + \boldsymbol{A}, \tag{4}$$

where $\boldsymbol{\tau}_2 = \boldsymbol{C}\boldsymbol{\mu}_1$. The representation (4) makes it clear that, in general, the inclusion of a linear time trend in (1) gives rise to a quadratic time trend in the process $\boldsymbol{X}_t$.

The five submodels describe different behaviors of the process $\boldsymbol{X}_t$ and the cointegrating relations $\boldsymbol{\beta}'\boldsymbol{X}_t$. In Case 0, $\boldsymbol{X}_t$ has no deterministic terms and all the stationary components have zero mean. In Case 1*, $\boldsymbol{X}_t$ has neither a quadratic trend, since $\boldsymbol{\mu}_1 = \mathbf{0}$ and hence $\boldsymbol{\tau}_2 = \boldsymbol{C}\boldsymbol{\mu}_1 = \mathbf{0}$, nor a linear trend, since $\boldsymbol{\alpha}_\perp{}'\boldsymbol{\mu}_0 = \mathbf{0}$ and hence $\boldsymbol{\tau}_1$, which is equal to $\boldsymbol{C}\boldsymbol{\mu}_0$ in this case, is equal to $\mathbf{0}$; see Johansen (1991, Theorem 4.1) and Johansen (1994). However, both $\boldsymbol{X}_t$ and the cointegrating relations, $\boldsymbol{\beta}'\boldsymbol{X}_t$, are allowed a constant term. In Case 1, where $\boldsymbol{\alpha}_\perp{}'\boldsymbol{\mu}_0 \neq \mathbf{0}$, we have $\boldsymbol{\tau}_1 \neq \mathbf{0}$, and $\boldsymbol{X}_t$ therefore has a linear trend. This trend is eliminated in the cointegrating relations $\boldsymbol{\beta}'\boldsymbol{X}_t$ because $\boldsymbol{\beta}'\boldsymbol{\tau}_1 = \boldsymbol{\beta}'\boldsymbol{C}\boldsymbol{\mu}_0 = \mathbf{0}$; see the definition of $\boldsymbol{C}$ preceding (4). In Case 2*, $\boldsymbol{X}_t$ has no quadratic trend since $\boldsymbol{\alpha}_\perp{}'\boldsymbol{\mu}_1 = \mathbf{0}$ and hence $\boldsymbol{\tau}_2 = \mathbf{0}$, but $\boldsymbol{X}_t$ has a linear trend which is present even in the cointegrating relations. Finally, Case 2 allows for a quadratic trend in the process $\boldsymbol{X}_t$, because $\boldsymbol{\mu}_1 \neq \mathbf{0}$ and hence $\boldsymbol{\tau}_2 \neq \mathbf{0}$. However, the cointegrating relations have a linear trend only, because $\boldsymbol{\beta}'\boldsymbol{\tau}_2 = \boldsymbol{\beta}'\boldsymbol{C}\boldsymbol{\mu}_1 = \mathbf{0}$.

Because of the normality assumption, it is natural to test for the reduced rank of $\boldsymbol{\Pi}$ 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. Consider the problem of testing the null hypothesis that there are at most r cointegrating vectors against the unrestricted model (1). The null hypothesis is that $\text{rank}(\boldsymbol{\Pi}) = r$ and the alternative is that $\text{rank}(\boldsymbol{\Pi}) = p$. The likelihood ratio test statistic, called the Trace statistic by Johansen and Juselius (1990), is given by

$$\text{Trace} = -T \sum_{i=r+1}^{p} \log(1 - \lambda_i), \tag{5}$$

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 $\text{rank}(\boldsymbol{\Pi}) = r$ against the alternative that $\text{rank}(\boldsymbol{\Pi}) = r+1$. In that case, the likelihood ratio statistic, which is called the $\lambda_{\max}$ statistic, is

$$\lambda_{\max} = -T \log(1 - \lambda_{r+1}). \tag{6}$$

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

The asymptotic distributions of the test statistics (5) and (6) are nonstandard. They are given by the trace and maximal eigenvalue, respectively, of the expression

$$\int_0^1 d\boldsymbol{B}\boldsymbol{F}' \left[\int_0^1 \boldsymbol{F}\boldsymbol{F}' du \right]^{-1} \int_0^1 \boldsymbol{F} d\boldsymbol{B}', \tag{7}$$

where $\boldsymbol{B}$ is a standard $(p-r)$–dimensional Brownian motion on the unit interval and $\boldsymbol{F}$ depends on $\boldsymbol{B}$ and on which restrictions are imposed on the deterministic terms. For each case, expression (7) is independent of nuisance parameters and depends only on $p-r$.

In the literature, asymptotic critical values for the Trace and $\lambda_{\max}$ statistics (5) and (6) have been calculated by Monte Carlo simulations of (7), where $\boldsymbol{B}$ is approximated by a $(p-r)$–dimensional discrete random walk, generally with 400 steps. However, we show in Section 4 that using this approach with only 400 steps leads to quite inaccurate results, especially when $p-r$ is large. In the present study, we instead generate simulated data and compute (5) and (6) directly for a number of sample sizes, and then employ response surface regressions to estimate the quantiles of the asymptotic distributions. The simulation experiments are described in the next section.

3. The Simulation Experiments

The simulation experiments which are at the heart of this paper are similar to those used by MacKinnon (1994, 1996) to compute the asymptotic distributions of Dickey-Fuller unit root tests and Engle-Granger cointegration tests. These experiments were used to estimate distributions for several finite sample sizes, from which the asymptotic distributions were then estimated using response surface regressions. Hendry (1984) and Ericsson (1991) discuss response surface methods in econometrics. Each experiment had 100,000 replications, and there were 50 experiments for each $p = 1, \ldots, 12$ for each of up to 12 sample sizes. What sample sizes were used, and why, will be discussed below.

There were several reasons for doing 50 experiments of 100,000 replications for each p and each sample size instead of a single experiment with five million replications. First, because computer memory capacities are finite, it would have been quite difficult to handle five million replications at once. Second, the observed variation among the 50 experiments provides an easy way to measure experimental randomness. Third, it was sometimes convenient to divide the experiments among two or more computers. The experiments were performed on six different computers, primarily two IBM RS/6000s, a Model 3AT and a Model 590, over a period of several months.

For each replication with a given p, we generated each element of $\boldsymbol{X}_t$ either as an independent random walk with no drift (for Cases 0, 1^*, and 2^*), as an independent random walk with drift (for Case 1), or as an independent random walk with drift and trend (for Case 2). We then computed the Trace statistic for $r = 0$ against $r = p$ and the $\lambda_{\max}$ statistic for $r = p - 1$ against $r = p$ for the simplest possible VAR appropriate to each model, in which all the $\boldsymbol{\Gamma}_i$'s were zero. 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: .0001, .0002, .0005, .001, . . . , .010, .015, . . . , .985, .990, .991, . . . , .999, .9995, .9998,

.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 $N(0,1)$ 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 i^{th} experiment, for which the sample size is T_i. Then the response surface regressions have the form

$$q^\alpha(T_i) = \theta^\alpha_\infty + \theta^\alpha_1 T_i^{-1} + \theta^\alpha_2 T_i^{-2} + \theta^\alpha_3 T_i^{-3} + \varepsilon_i. \tag{8}$$

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.

Because the functional form of the response surface regressions is known to be (8), it is possible to choose the sample sizes for the experiments more or less optimally. If we write (8) as

$$\boldsymbol{q}^\alpha = \boldsymbol{Z\theta} + \boldsymbol{\varepsilon} = \theta^\alpha_\infty + \boldsymbol{Z}^*\boldsymbol{\theta}^* + \boldsymbol{\varepsilon}, \tag{9}$$

the standard error of the OLS estimate of θ^α_∞ will be proportional to $(\boldsymbol{\iota}'\boldsymbol{M}^*\boldsymbol{\iota})^{-1/2}$, where $\boldsymbol{\iota}$ is a vector of ones and $\boldsymbol{M}^* = \mathbf{I} - \boldsymbol{Z}^*(\boldsymbol{Z}^{*\prime}\boldsymbol{Z}^*)^{-1}\boldsymbol{Z}^{*\prime}$. For any set of T_i's, we can evaluate the standard error of θ^α_∞. Based on estimates of how computation cost varies with T, we can also evaluate the cost of using that set of T_i's. Some very rough calculations, based on these two quantities, suggest that it is always desirable for there to be several small values of T_i, because the smaller the smallest value of T_i, the more trouble $\boldsymbol{Z}^*$ has explaining a constant term, and thus the larger is $\boldsymbol{\iota}'\boldsymbol{M}^*\boldsymbol{\iota}$. However, none of the T_i's should be too small, because then equation (8) may not fit satisfactorily. It is also desirable for there to be some values of T_i that are quite large, but it is not cost-effective for there to be any intermediate values. This led us to choose the following set of 12 sample sizes: 30, 40, 50, 60, 75, 90, 100, 110, 500, 750, 1000, and 1250. For all but the smallest values of $p - r$, it was necessary to omit some of the smaller values of T_i in order to obtain response surfaces that fit acceptably well.

Equation (8) was estimated 221 times for each of 115 different test statistics. There were up to 600 observations, depending on how many of the smaller T_i's had to be dropped. As in MacKinnon (1996), we used a form of GMM estimation to allow for the fact that the error terms of (8) are heteroskedastic. Let $\boldsymbol{\Omega}$ denote the covariance matrix of the error vector $\boldsymbol{\varepsilon}$ in equation (9). This matrix is diagonal, because all the experiments are independent. The estimator we used was

$$\hat{\boldsymbol{\theta}} = \left(\boldsymbol{Z}'\boldsymbol{W}(\boldsymbol{W}'\hat{\boldsymbol{\Omega}}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{W}(\boldsymbol{W}'\hat{\boldsymbol{\Omega}}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{q}^p, \tag{10}$$

where $\boldsymbol{W}$ is a matrix of up to 12 zero-one dummy variables, the first equal to 1 when $T_i = 30$, the second equal to 1 when $T_i = 40$, and so on, and $\hat{\boldsymbol{\Omega}}$ estimates $\boldsymbol{\Omega}$. The matrix $\hat{\boldsymbol{\Omega}}$ is obtained by regressing the squared residuals from an OLS regression of $\boldsymbol{q}^p$ on $\boldsymbol{W}$ on a constant, $1/T$, and $1/T^2$. The fitted values from this auxiliary regression were then used as the diagonal elements of $\hat{\boldsymbol{\Omega}}$. The estimator (10) can easily be computed by a weighted least squares regression with as many observations as there are distinct values of T_i; see MacKinnon (1994) for details.

This GMM estimation procedure automatically generates a statistic for testing the specification of the response surface equation (8). The test statistic is the minimized value of the objective function:

$$(\boldsymbol{q}^p - \boldsymbol{Z\theta})'\boldsymbol{W}\left(\boldsymbol{W}'\hat{\boldsymbol{\Omega}}\boldsymbol{W}\right)^{-1}\boldsymbol{W}'(\boldsymbol{q}^p - \boldsymbol{Z\theta}). \tag{11}$$

Standard results about GMM estimation imply that, under the null hypothesis that (8) is a correct specification, (11) is asymptotically distributed as $\chi^2(l)$, where l is equal to the number of distinct T_i's (which may be 12 or less) minus the number of parameters in (8).

The GMM test statistic (11) 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 choose whether to set $\theta_3^p = 0$ in (8) and to determine how many small values of T_i to drop. Since the objective was to obtain efficient estimates of θ_∞^p, it was desirable to set $\theta_3^p = 0$, if possible, and to throw out as few small T_i's as possible. On average, for a correctly specified response surface, reducing the number of distinct T_i's by one, or dropping the constraint that $\theta_3^p = 0$ in (8), would be expected to reduce the value of (11) by 1.0, because the mean of a random variable with a $\chi^2(l)$ distribution is l. In most cases, we chose to reject a model when such a change reduced the value of (11) by more than 1.2. There were some very clear patterns in the response surface estimates. More values of T_i had to be dropped, and/or the restriction that $\theta_3^p = 0$ relaxed, as $p - r$ was increased.

4. Numerical Distributions

The principal results of this paper are 25,415 (= 221 × 115) estimates of θ^p_∞. These estimates 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 1 through 5 present asymptotic critical values at several conventional levels for all the tests we examined. These critical values differ substantially from those previously published by Johansen and Juselius (1990), Osterwald-Lenum (1992), and Johansen (1995), 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, and Johansen (1995) used 100,000. Our critical values therefore suffer from less experimental error. Second, the estimates of θ^p_∞ from equation (8) really are estimates of the quantiles of an asymptotic distribution, while the values previously published by others are simply approximations to expression (7) based on a discrete random walk with a finite number of steps. All these authors used $T = 400$ steps, and it appears that, especially when $p - r$ is large, this number is much too small.

In order to investigate the discrepancy between our estimates and those of Johansen (1995), we carried out some additional Monte Carlo experiments for the asymptotic approximations based on equation (7) for case 1*, using 100,000 replications with $p-r = 12$. For T=400, the 5% critical value was estimated to be 338.13, which is extremely close to the value of 338.10 reported in Table 15.2 of Johansen (1995). However, it is not close to the value of 348.99 reported in Table 2. We therefore tried several other values of T. Invariably, as the value of T rose, so did the estimated critical value; the largest one we obtained was 346.52 for $T = 2000$. We then estimated a response surface, in which the regressors were just a constant and $1/T$, using results from experiments for nine values of T evenly spaced between 400 and 2000. The resulting estimate of the asymptotic 5% critical values was 348.65, which is very close to the number in Table 2. These results demonstrate clearly that approximations to equation (7) based on 400, or even on 2000, steps are not adequate. If one wishes to obtain accurate asymptotic critical values, one must use a response surface analysis, whether based on approximations to equation (7) or on the test statistics themselves.

There are good reasons to believe that our estimated asymptotic distributions are extremely accurate. The estimates of θ^p_∞ have standard errors associated with them, and except for the extreme tail quantiles, these are all very small. For example, for the .05 critical value, the estimated standard errors range from .0018 to .0410, with the larger standard errors being associated with the larger values of $p - r$. Of course, because of the pretesting involved in choosing which version of equation (8) to estimate and which sample sizes to drop, the estimated standard errors may be a bit too optimistic.

For a few cases, it is possible to compare our estimates with outside benchmarks. When $p - r = 1$, the distribution of both test statistics for Case 0 is that of the square of the corresponding Dickey-Fuller test, and for Cases 1 and 2 it is $\chi^2(1)$; see Johansen and Juselius (1990). Table 6 therefore compares some of our critical values for Case 0 with ones reported by Nielsen (1996) based on the analytic formulae of Abadir (1995). It also compares some of our critical values for Cases 1 and 2 with critical values from the $\chi^2(1)$ distribution. It is clear from the table that our estimates are very accurate when $p - r = 1$. For comparison, the table also reports the estimates of Osterwald-Lenum (1992), which are very much less accurate, and (for Case 0 only) the estimates of Johansen (1995). Note that the critical values in Tables 1, 2, and 4 for $p - r = 1$ are the benchmark ones.

It is interesting to see what the distributions of the various test statistics look like. Figures 1 and 2 graph the asymptotic distribution functions of the $\lambda_{\max}$ and Trace statistics, respectively, for Case 0 for all 12 values of $p - r$, and Figures 3 and 4 graph the same functions for Case 2. The other cases are quite similar to these ones and are omitted to save space. Each curve simply joins the 221 estimated quantiles for a given test statistic, without any smoothing. One striking feature of these figures is the regular and predictable way in which all the curves move to the right as $p - r$ increases. Another feature is that the distributions are highly skewed for small values of $p - r$, but much less so for large values.

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 called `johdist` which does both is available via the Internet; see the Appendix.

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

$$\Phi^{-1}(p) = \gamma_0 + \gamma_1 \hat{q}(p) + \gamma_2 \hat{q}^2(p) + \gamma_3 \hat{q}^3(p) + e_p, \tag{12}$$

where $\Phi^{-1}(p)$ is the inverse of the cumulative standard normal distribution function, evaluated at p. If the distribution from which the estimated quantiles were obtained were in fact normal with any mean and variance, regression (12) 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 neighborhood of the test statistic that is of interest. It may seem a bit odd that the regressors in (12) 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 we are interested in obtaining approximate critical values, equation (12) has to be turned around. Consider the regression

$$\hat{q}_p = \delta_0 + \delta_1 \Phi^{-1}(p) + \delta_2 \left(\Phi^{-1}(p)\right)^2 + \delta_3 \left(\Phi^{-1}(p)\right)^3 + e_p^*. \tag{13}$$

This is not actually the inverse of equation (12). However, if the distribution from which the estimated quantiles were obtained were in fact normal with any mean and variance, equation (13) would be correctly specified with $\delta_2 = \delta_3 = 0$. In that case, equation (12) would have $\gamma_2 = \gamma_3 = 0$, and (13) would be the inverse of (12).

Regressions (12) and (13) could be estimated by OLS, but this would ignore both heteroskedasticity and serial correlation. In MacKinnon (1996), it is shown how to take both of these into account. Therefore, when estimating these equations, the `johdist` program actually uses the form of feasible GLS estimation proposed in that paper. As discussed above, equations (12) and (13) 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.

6. An Empirical Example: Is Europe Ready for the EMU?

In this section, we analyze the prospects for the formation of the European Economic and Monetary Union (EMU) using cointegration analysis. The example is intended to be of interest in its own right, although it is primarily designed to illustrate how the results of this paper, in particular the `johdist` program, may be used in applied work.

The EMU represents the third and final stage of a complete economic and monetary union among the countries participating in the European Union (EU), namely, Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, and the United Kingdom. In addition to these countries, Austria, Finland, and Sweden have recently joined the EU. The legal, institutional, and monetary aspects of European integration were laid down in the Maastricht Treaty (1992), following the recommendations of the Delors Report (1989). Although the Maastricht Treaty does not set a specific date for the start of the EMU, it is generally understood that it should begin by January, 1997, at the earliest or by January, 1999, at the latest (Eichengreen and Wyplosz, 1993).

Regarding the monetary and fiscal aspects of the Treaty, four "nominal convergence" criteria were laid down that would have to be met by a member country in order to qualify for participation in the EMU. These were:

(a) no devaluation of its currency in the two years preceding the entrance into the union;

(b) inflation rate no higher than 1.5 percent above the average of the three countries with the lowest inflation rates;

(c) long-term interest rate not in excess of 2 percent of the three countries with the lowest inflation rates, and

(d) government deficits and debts not exceeding 3 percent and 60 percent of the GDP respectively.

In view of the fact that the macroeconomic variables involved are typically nonstationary, a minimum requirement for the formation of the EMU is the existence of stable long-run relationships that tie together the variables in each of the above criteria. Otherwise, the chances for the success of the EMU would be very slim. This issue is the focus of our empirical study. If the variables for the EU countries are found to be cointegrated within each of the four criteria, deviations from certain linear combinations will be stationary, implying that the variables will be tied together in the long run. Because the nominal convergence criteria imply comovement of specific variables over time, the cointegration approach is well-suited to assess the potential of the EMU. Since the late 1970s, monetary and, to some extent, fiscal policy coordination among the EU member countries has been the focus of the European Commission's efforts. In addition, the Exchange Rate Mechanism (ERM) of the European Monetary System (EMS) was introduced in March, 1979 to provide intra-EMS exchange rate stability.

MacDonald and Taylor (1991) also use Johansen's method and present evidence of long run convergence in real and nominal exchange rates and in money supplies. However, they consider only four EMS countries over the period 1979 to 1988 and do not include other variables in their study. Based on conditional variance comparisons and persistence of real exchange shocks, von Hagen and Neumann (1994) report results supportive of the idea of a two-speed Europe, with Germany, its smaller neighbors, and France forming a viable monetary union. Bayoumi and Taylor (1995) examine the comparative behavior of real output growth and inflation rates of ERM and non-ERM participants and conclude that the ERM has contributed to macro-policy coordination among ERM members.

In our empirical example, we analyze the behavior of the monthly nominal spot exchange rates per European Currency Unit (ECU) and per German Mark (DM). Inflation rates based on the consumer price index (CPI) seem to be mostly $I(0)$ series for the countries considered, and the concept of cointegration therefore does not apply to them. Instead, we analyze real monthly DM exchange rates derived using the CPI. We also study long-term interest rates, using monthly long-term government bond yields, and deficit/GDP ratios, using quarterly data because monthly data for deficits are not available. We do not study quarterly debt/GDP ratios because the data are not available for most of the countries considered in our example.

Most of the data were taken from the IFS CD-ROM, September 1995. We collected data for all of the 12 original EU countries and did not include the three new members. The time periods considered are determined by the availability of data. When possible, we start in 1979:3, when the ERM was introduced. Monthly nominal end-of-period spot exchange rates per US dollar are from the IFS, line ae (or line ag if line ae is not available; line ag is the inverse of line ae). The period covered is 1979:3 to 1995:7. Monthly end-of-period nominal ECU rates are taken from the IFS, line ea (or line ec if line ea is not available). ECU rates for Greece and Portugal are not available prior to 1981:1 and 1985:7, respectively. We used ECU rates per US dollar and US dollar spot rates for these two countries to construct ECU rates for the months with missing data. ECU data cover the period 1979:7 to 1995:7. Real exchange rates are constructed from the IFS, line ae (or ag) and the monthly CPI from line 64 in the IFS. The period considered is 1979:3 to 1995:5. Monthly average long-term government bond yields are from the IFS, line 61, except for Denmark. This IFS data series was incomplete for Denmark and we therefore used for Denmark the corresponding data from the OECD Main Economic Indicators that reports end-of-month figures instead of averages. The period covered is 1979:3 to 1995:3.

Complete quarterly government deficit (or surplus) data are available only for a few countries: France, Germany, Italy, the Netherlands, Spain, and the U.K. The surplus is given in the IFS, line 80. To construct the deficit/GDP ratios, we use quarterly GDP from the IFS, line 99b. Data are available from 1979:3 to 1994:2. For Italy, the years 1991:4 to 1994:2 were missing and we used national debt figures from the Bolletino Mensile di Statistica, Institute Nazionale di Statistica, various issues, in order to construct surplus figures from the quarterly change in debt. A comparison for earlier years shows that surpluses constructed from these debt figures are very close to IFS figures.

Before estimating any VECM systems, we tested each time series for a unit root using augmented Dickey-Fuller tests at the 5% level. Results are available upon request. We employed Akaike's information criterion to select the appropriate lag lengths (see Ng and Perron, 1995) and used asymptotic critical values from the program of MacKinnon (1996). On the basis of our test results, we exclude Denmark from the system for nominal DM-based exchange rates because this variable seems to be $I(0)$. For deficit/GDP ratios, we exclude Spain for the same reasons and consider the model with and without Italy because the test statistic for Italy is very close to the 5% critical value. Except for inflation rates, which mostly seem to be $I(0)$, we are unable to reject the unit root hypothesis for any of the other series studied.

Table 7 reports results for the Trace and $\lambda_{\max}$ tests for ECU exchange rates, nominal DM exchange rates, real DM exchange rates, long-term interest rates, and deficit/GDP ratios. We set up a separate VECM for each one of these variables and used the Schwarz criterion to select the appropriate lag lengths. Each VECM involves the largest set of the 12 original EU countries for which data are available.

In order to decide which submodel to use, we tested the various submodels against each other using Likelihood Ratio tests, which are distributed as χ^2; see Johansen and Juselius (1990, Section 4.1). For all but one of the systems in Table 7, we found that Case 0 is appropriate. The exception is interest rates, for which Case 1* seems to be appropriate.

We discuss the results for exchange rates first. We included all 12 countries for ECU rates. For DM rates, Germany had to be deleted, and we also deleted Denmark because the series appears to be $I(0)$ and Luxembourg because its exchange rate was fixed to Belgium's. Thus 9 countries were included in the system for DM rates. Based on the P values in Table 7, we decisively reject the null hypothesis of no cointegration (which states that r, the number of cointegrating vectors, is equal to zero) for both nominal exchange rate systems. If we use a 5% level for all tests, the Trace test leads us to find five cointegrating vectors for the DM rates and eight vectors for the ECU rates. With the $\lambda_{\max}$ test, we find fewer cointegrating vectors, three for the DM rates and four for the ECU rates. We conclude that nominal exchange rates among the EU countries are cointegrated no matter which exchange rate definition is used.

For the real DM exchange rate, we had to exclude Ireland because the CPI was not available on a monthly basis, leaving us with real DM exchange rates for 10 EU countries. For this system, the Trace test detects two cointegrating vectors, but the $\lambda_{\max}$ test detects none. For long-term interest rates, we had to exclude Greece because data were not available, leaving us with 11 countries. For this system, neither test found cointegration at the 5% level, although the $\lambda_{\max}$ test came close. Finally, the last column of Table 7 reports results for the deficit/GDP ratio for France, Germany, the Netherlands, and the U.K. Spain and Italy were excluded because their series appeared to be $I(0)$. Adding Italy (for which the unit root test was inconclusive) did not affect the test results. We find that both cointegration tests detect one cointegrating vector for this system.

In Table 8, we examine subgroups of EU countries for the real DM exchange rates and the long-term interest rates that we found in Table 7 not to be cointegrated among the full set of EU members. The largest subgroup that leads to cointegration for real DM exchange rates consists of seven countries: Belgium, Denmark, France, Greece, Luxembourg, the Netherlands, and the U.K. Adding Portugal or Spain does not lead to cointegration any more. For long-term interest rates, we find cointegration among eight countries: Belgium, Denmark, France, Germany, Ireland, Luxembourg, the Netherlands, and the U.K.

Our empirical results support the view that the EMU may not be feasible for all 12 original EU countries if fiscal and monetary policies are not aligned further. However, the EMU may be feasible among the following countries: Belgium, Denmark, France, Germany, Luxembourg, the Netherlands, and the U.K. Due to lack of data, we cannot tell whether Greece and Ireland should be included in this group. On the other hand, Italy, Portugal, and Spain apparently have to align their economic policies further with the other EU countries in order to become viable EMU members.

7. 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. Although the paper contains tables of critical values, which are far more accurate than those previously available, the results consist chiefly of tables of estimated asymptotic quantiles, along with a computer program that uses these to calculate critical values and P values. Both of these are available via the Internet.

In the final section of the paper, we present, as an empirical example, an analysis of cointegration among the countries of the European Union. The ability to compute asymptotic P values makes it possible to present results for several VECM systems in only two, easily assimilated tables. These results support the view that the EMU may be feasible only for a subset of the EU countries.

Appendix

The tables of asymptotic quantiles from the response surface regressions and the associated computer programs may be obtained via ftp or via the World Wide Web. For the former, ftp to

qed.econ.queensu.ca

and then go to the directory `pub/uroot/johtest`. This directory contains the Fortran source files `johdist.f` and `johrouts.f`, a compiled version of `johdist` for DOS-based personal computers with at least 4 MB of memory and a numeric coprocessor, and zipped files that contain the estimated quantiles of the asymptotic distribution functions.

World Wide Web users may obtain the same files by pointing their web browsers to

http://qed.econ.queensu.ca/pub/faculty/mackinnon

and then following the appropriate link from the first author's personal home page.

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.

Bayoumi, T., and M. P. Taylor (1995), "Macroeconomic shocks, the ERM and tripolarity", *Review of Economics and Statistics*, **77**, 321–331.

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.

Delors Report (1989), *Report on Economic and Monetary Union in the European Community*, Committee for the Study of Economic and Monetary Union, Luxembourg: Office for Official Publications of the European Communities.

Eichengreen, B., and C. Wyplosz (1993), " The Unstable EMS", *Brookings Papers on Economic Activity*, **1**, 51–125.

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

Ericsson, N. R. (1991), "Monte Carlo methodology and the finite sample properties of instrumental variables statistics for testing nested and non-nested hypotheses", *Econometrica*, **59**, 1249–1277.

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

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

Hendry, D. F. (1984), "Monte Carlo experimentation in econometrics", Ch. 16 in *Handbook of Econometrics*, Vol. II, eds. Z. Griliches and M. D. Intriligator, Amsterdam, North-Holland.

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, Oxford University Press.

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.

Maastricht Treaty (1992), *Treaty on Economic Union*, Luxembourg: Office for Official Publications of the European Communities.

MacDonald, R., and M. P. Taylor (1991), "Exchange rates, policy convergence, and the European Monetary System", *Review of Economics and Statistics*, **73**, 553–558.

MacKinnon, J. G. (1991), "Critical values for cointegration tests", in *Long-run Economic Relationships: Readings in Cointegration*, eds. R. F. Engle and C. W. J. Granger, Oxford: Oxford University Press, 267–276.

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*, forthcoming.

Marsaglia, G., and T. A. Bray (1964), "A convenient method for generating normal variables", *SIAM Review*, **6**, 260–264.

Nielsen, B. (1996), "Bartlett correction of the unit root test in autoregressive models," *Biometrika*, forthcoming.

Ng, S., and P. Perron (1995), "Unit root tests in ARMA models with data dependent methods for the selection of the truncation lag", *Journal of the American Statistical Association*, **90**, 268–281.

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.

Phillips, P. C. B. (1991), "Optimal inference in cointegrated systems", *Econometrica*, **59**, 283–306.

von Hagen, J., and M. J. Neumann (1994), "Real exchange rates within and between currency areas: How far away is the EMU?", *Review of Economics and Statistics*, **76**, 236–244.

Table 1. Critical Values for Case 0

	λ_{max}					Trace				
$p-r$	1%	2%	5%	10%	20%	1%	2%	5%	10%	20%
1	6.94	5.71	4.13	2.98	1.88	6.94	5.71	4.13	2.98	1.89
2	15.09	13.45	11.22	9.47	7.62	16.36	14.65	12.32	10.47	8.51
3	22.25	20.38	17.80	15.72	13.45	29.51	27.33	24.28	21.78	19.01
4	29.06	27.02	24.16	21.84	19.28	46.57	43.92	40.17	37.03	33.49
5	35.72	33.52	30.44	27.92	25.10	67.64	64.53	60.06	56.28	51.96
6	42.23	39.90	36.63	33.93	30.91	92.71	89.13	83.94	79.53	74.44
7	48.66	46.22	42.77	39.91	36.70	121.74	117.70	111.78	106.74	100.88
8	55.04	52.48	48.88	45.89	42.52	154.80	150.27	143.67	138.00	131.38
9	61.35	58.71	54.97	51.85	48.32	191.83	186.81	179.52	173.23	165.84
10	67.65	64.92	61.03	57.80	54.14	232.84	227.41	219.41	212.47	204.30
11	73.89	71.07	67.08	63.73	59.94	278.00	272.01	263.26	255.68	246.78
12	80.12	77.22	73.09	69.65	65.73	326.96	320.57	311.13	302.90	293.21

Table 2. Critical Values for Case 1*

	λ_{max}					Trace				
$p-r$	1%	2%	5%	10%	20%	1%	2%	5%	10%	20%
1	12.76	11.23	9.16	7.56	5.88	12.76	11.23	9.16	7.56	5.88
2	20.16	18.37	15.89	13.91	11.77	25.08	23.07	20.26	17.98	15.48
3	27.07	25.07	22.30	20.05	17.59	41.20	38.72	35.19	32.27	28.99
4	33.73	31.59	28.59	26.12	23.40	61.27	58.31	54.08	50.53	46.46
5	40.29	38.01	34.81	32.17	29.21	85.34	81.90	76.97	72.77	67.94
6	46.75	44.35	40.96	38.16	35.02	113.42	109.51	103.84	99.02	93.42
7	53.12	50.61	47.07	44.13	40.81	145.40	141.03	134.68	129.23	122.85
8	59.51	56.89	53.19	50.11	46.64	181.51	176.70	169.61	163.50	156.36
9	65.79	63.07	59.24	56.05	52.44	221.45	216.17	208.45	201.69	193.80
10	72.10	69.26	65.30	61.99	58.25	265.53	259.75	251.27	243.96	235.28
11	78.29	75.42	71.33	67.93	64.05	313.75	307.37	298.17	290.17	280.73
12	84.51	81.56	77.38	73.85	69.86	365.64	358.96	348.99	340.38	330.19

Table 3. Critical Values for Case 1

	λ_{max}					Trace				
$p-r$	1%	2%	5%	10%	20%	1%	2%	5%	10%	20%
1	6.63	5.41	3.84	2.71	1.64	6.63	5.41	3.84	2.71	1.64
2	18.52	16.72	14.26	12.30	10.19	19.94	18.07	15.50	13.43	11.19
3	25.86	23.88	21.13	18.89	16.44	35.46	33.12	29.80	27.07	24.01
4	32.71	30.58	27.58	25.12	22.41	54.68	51.87	47.86	44.49	40.65
5	39.37	37.08	33.88	31.24	28.30	77.82	74.54	69.82	65.82	61.21
6	45.87	43.47	40.08	37.28	34.15	104.96	101.20	95.75	91.11	85.74
7	52.31	49.78	46.23	43.29	39.98	135.97	131.76	125.61	120.37	114.23
8	58.67	56.06	52.36	49.29	45.82	171.09	166.38	159.53	153.63	146.74
9	64.99	62.28	58.43	55.24	51.63	210.06	204.93	197.37	190.88	183.24
10	71.26	68.46	64.51	61.20	57.45	253.24	247.54	239.25	232.11	223.72
11	77.49	74.63	70.53	67.13	63.26	300.29	294.13	285.14	277.38	268.17
12	83.70	80.79	76.58	73.06	69.08	351.25	344.66	334.98	326.53	316.63

Table 4. Critical Values for Case 2*

	λ_{max}					Trace				
$p-r$	1%	2%	5%	10%	20%	1%	2%	5%	10%	20%
1	16.55	14.85	12.52	10.67	8.69	16.55	14.85	12.52	10.67	8.69
2	23.97	22.05	19.39	17.23	14.88	31.16	28.95	25.87	23.34	20.54
3	30.83	28.75	25.82	23.44	20.81	49.36	46.69	42.91	39.75	36.17
4	37.49	35.25	32.12	29.54	26.67	71.47	68.36	63.88	60.09	55.73
5	44.02	41.66	38.33	35.58	32.52	97.60	93.99	88.80	84.38	79.26
6	50.47	47.99	44.50	41.60	38.34	127.71	123.61	117.71	112.65	106.77
7	56.85	54.26	50.59	47.56	44.15	161.72	157.17	150.56	144.87	138.24
8	63.17	60.49	56.71	53.55	49.98	199.81	194.80	187.47	181.16	173.75
9	69.44	66.68	62.75	59.49	55.79	241.74	236.27	228.31	221.36	213.20
10	75.69	72.86	68.81	65.44	61.61	287.87	281.95	273.19	265.63	256.70
11	81.94	79.00	74.84	71.36	67.41	337.97	331.58	322.06	313.86	304.15
12	88.11	85.15	80.87	77.30	73.23	392.01	385.14	374.91	366.11	355.65

Table 5. Critical Values for Case 2

	λ_{max}					Trace				
$p - r$	1%	2%	5%	10%	20%	1%	2%	5%	10%	20%
1	6.63	5.41	3.84	2.71	1.64	6.63	5.41	3.84	2.71	1.64
2	21.74	19.82	17.15	15.00	12.66	23.15	21.17	18.40	16.16	13.71
3	29.26	27.16	24.25	21.87	19.25	41.08	38.56	35.01	32.06	28.74
4	36.19	33.95	30.82	28.24	25.38	62.52	59.54	55.24	51.65	47.53
5	42.86	40.49	37.16	34.42	31.36	87.78	84.32	79.34	75.10	70.20
6	49.41	46.93	43.42	40.53	37.28	116.99	113.05	107.34	102.47	96.81
7	55.81	53.25	49.58	46.56	43.15	150.08	145.69	139.28	133.79	127.37
8	62.17	59.51	55.73	52.58	49.02	187.20	182.31	175.16	169.07	161.90
9	68.50	65.73	61.81	58.53	54.85	228.23	222.91	215.12	208.36	200.43
10	74.74	71.91	67.90	64.53	60.70	273.37	267.54	259.02	251.63	242.94
11	81.07	78.14	73.94	70.46	66.51	322.41	316.14	306.90	298.89	289.45
12	87.23	84.24	79.97	76.41	72.35	375.30	368.52	358.72	350.12	339.88

Table 6. Comparisons of Critical Values for $p - r = 1$

	1%	5%	10%
Case 0:			
Benchmark	6.9383	4.1293	2.9776
Our Estimate	6.9414	4.1300	2.9760
O-L Estimate	6.51	3.84	2.86
Johansen Est.	7.02	4.14	2.98
Case 1:			
Benchmark	6.6349	3.8415	2.7055
Our Estimate	6.6296	3.8435	2.7074
O-L Estimate	6.65	3.76	2.69
Case 2:			
Benchmark	6.6349	3.8415	2.7055
Our Estimate	6.6391	3.8434	2.7058
O-L Estimate	6.40	3.74	2.57

Table 7. *P* Values for Tests Using Largest Set of Countries

$p-r$	ECU Ex. Rates Trace	ECU Ex. Rates $\lambda_{\max}$	Nominal DM Ex. Rates Trace	Nominal DM Ex. Rates $\lambda_{\max}$	Real DM Ex. Rates Trace	Real DM Ex. Rates $\lambda_{\max}$	Interest Rates Trace	Interest Rates $\lambda_{\max}$	Deficit/GDP Ratios Trace	Deficit/GDP Ratios $\lambda_{\max}$
12	.0001	.0001								
11	.0001	.0011					.0959	.0549		
10	.0001	.0158			.0036	.2858	.4916	.3185		
9	.0001	.0222	.0001	.0010	.0166	.1804	.8242	.9437		
8	.0002	.1675	.0001	.0263	.0856	.1822	.8679	.8833		
7	.0014	.1942	.0001	.0251	.3132	.5745	.9253	.8186		
6	.0060	.1609	.0030	.0948	.4725	.6309	.9738	.9909		
5	.0258	.0570	.0211	.1129	.6142	.9103	.9571	.9957		
4	.2049	.4134	.1041	.1783	.5056	.7870	.8818	.9850	.0001	.0001
3	.2954	.4203	.3038	.2843	.4233	.4949	.7159	.8404	.0990	.1363
2	.3651	.2850	.5784	.5435	.4989	.6230	.6163	.6131	.3171	.4275
1	.9804	.9804	.6027	.6027	.2640	.2640	.6228	.6228	.2237	.2237
Case:	0		0		0		1*		0	

Note: A reported value of .0001 indicates a P value of less than .0001. The value reported at the top of each column is for $r = 0$, so that $p - r = p$, where p is the number of countries included. The value reported in the second line of each column is for $p - r = p - 1$, the third for $p - r = p - 2$, and so on until $p - r = 1$ in the last line.

Table 8. *P* Values for Tests Using Subsets of Countries

$p-r$	Real DM Ex. Rates Trace	Real DM Ex. Rates $\lambda_{\max}$	Interest Rates Trace	Interest Rates $\lambda_{\max}$
8			.0136	.0097
7	.0085	.0356	.2759	.2473
6	.1151	.0298	.6301	.6856
5	.7050	.7205	.7611	.9488
4	.7959	.5522	.6424	.7287
3	.9570	.9641	.6660	.8240
2	.8474	.7886	.4847	.4327
1	.9429	.9429	.6527	.6527
Case:	0		0	

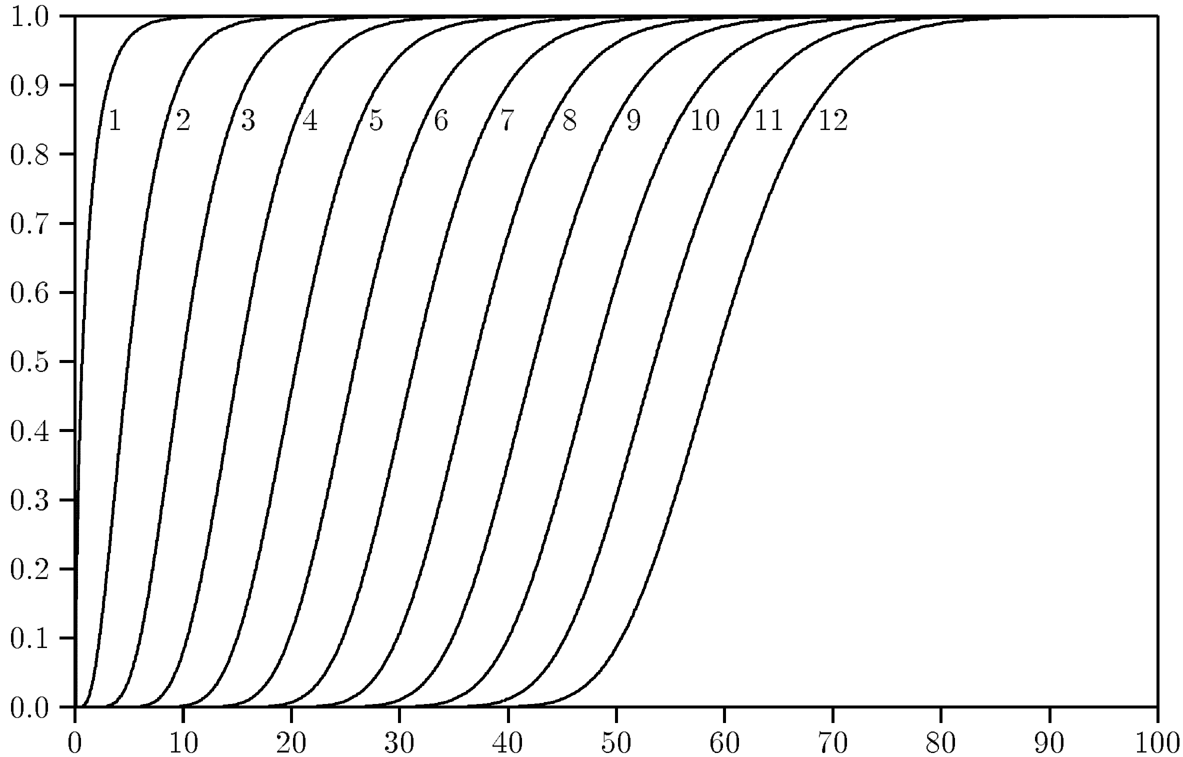

Figure 1. Asymptotic distributions of $\lambda_{\max}$ tests, Case 0

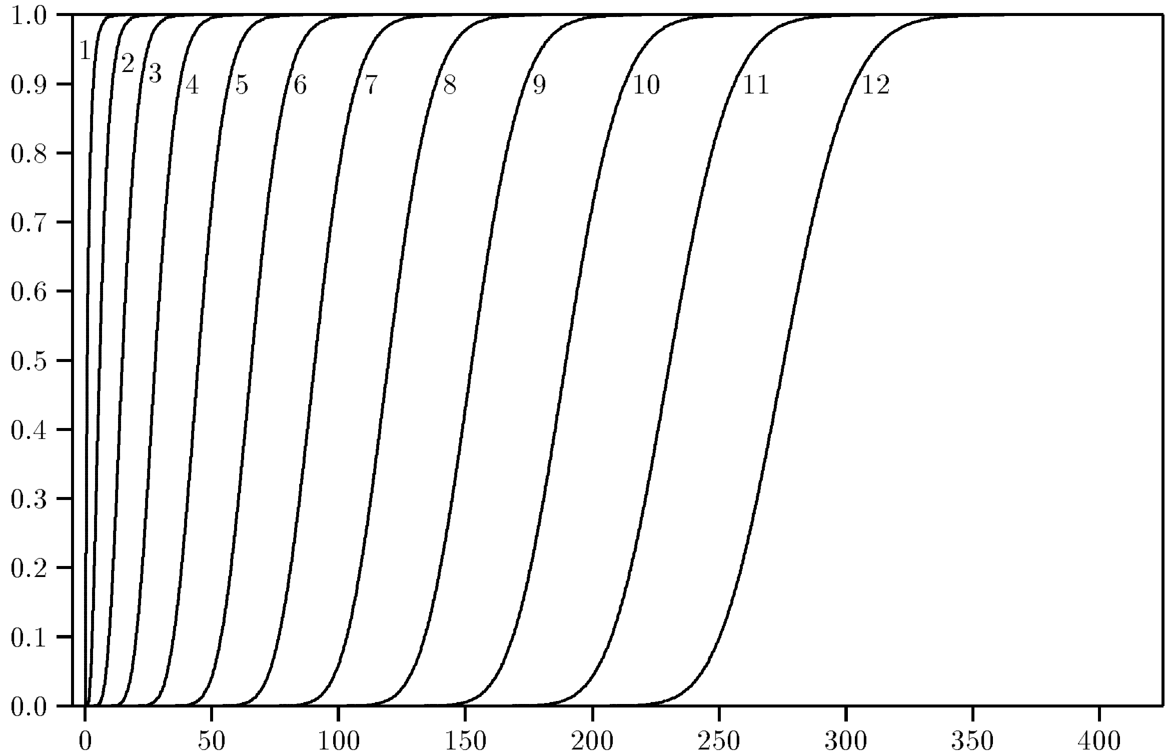

Figure 2. Asymptotic distributions of Trace tests, Case 0

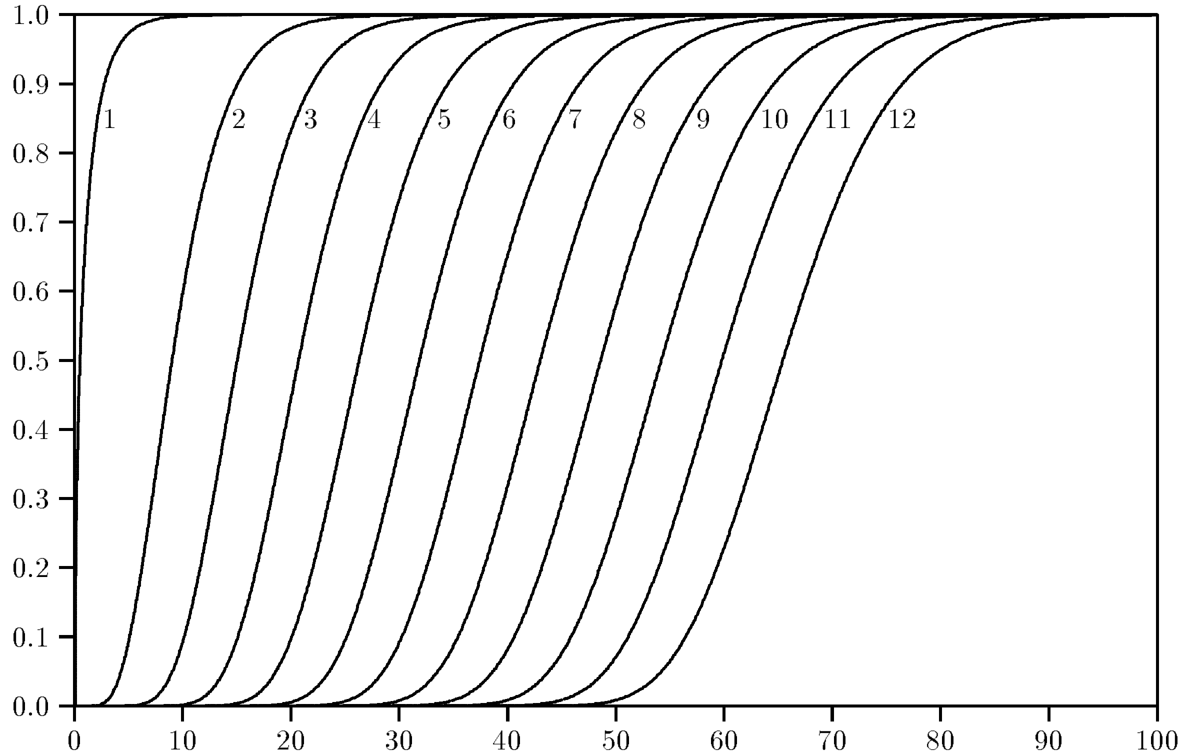

Figure 3. Asymptotic distributions of $\lambda_{\max}$ tests, Case 2

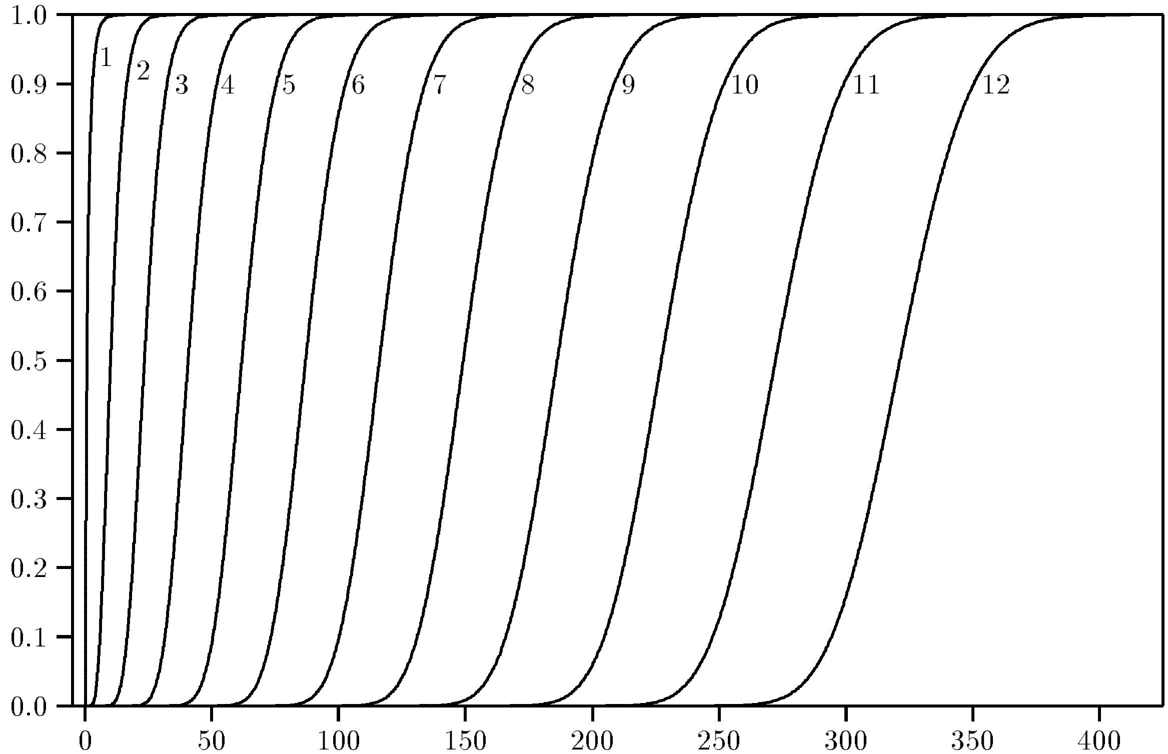

Figure 4. Asymptotic distributions of Trace tests, Case 2