Queen's Economics Department Working Paper No. 861

Approximate Asymptotic Distribution Functions for Unit Roots and Cointegration Tests

James G. MacKinnon

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

11-1992

Discussion Paper #861

Approximate Asymptotic Distribution Functions for Unit Root and Cointegration Tests

by

James G. MacKinnon
Queen's University
Kingston, Ontario
K7L 3N6

November 1992

This research was supported, in part, by the Social Sciences and Humanities Research Council of Canada. Earlier versions of this paper were presented in workshops at Queen's University and Wilfred Laurier University. I am grateful to participants in those workshops and to Russell Davidson, Allan Gregory, and Colin Telmer for comments.

Approximate Asymptotic Distribution Functions for Unit Root and Cointegration Tests

by

James G. MacKinnon

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

Abstract

This paper uses Monte Carlo experiments and regression methods to calculate approximate asymptotic distribution functions for a number of well-known unit root and cointegration test statistics. These allow empirical workers to calculate approximate P values for these tests. The results of the paper are based on a very extensive set of Monte Carlo experiments, which yield finite-sample critical values for a number of sample sizes. Response surface regressions are then used to obtain asymptotic critical values for a large number of different test sizes. Finally, regression methods are used to estimate approximate distribution functions with simple functional forms.

This research was supported, in part, by the Social Sciences and Humanities Research Council of Canada. Earlier versions of this paper were presented in workshops at Queen's University and Wilfred Laurier University. I am grateful to participants in those workshops and to Russell Davidson, Allan Gregory, and Colin Telmer for comments.

November, 1992

1. Introduction

Tests of the null hypothesis that a time-series process has a unit root have been widely used in recent years. The most common of these tests, based on the work of Dickey and Fuller (1979) and Said and Dickey (1984), are known as Dickey-Fuller (DF) tests and Augmented Dickey-Fuller (ADF) tests, respectively; see Dickey, Bell, and Miller (1986) for an illuminating exposition. One problem with these tests is that they have nonstandard asymptotic distributions, for which only a few critical values have been tabulated; Fuller (1976) is the classic reference. The aim of this paper is to calculate approximate asymptotic distribution functions for these and a number of related tests. Using these, empirical workers can easily calculate approximate P values for a large number of commonly used tests.

The simplest unit root tests are DF tests. Suppose we wish to test the null hypothesis that the series y_t has a unit root. Then DF tests may be based on OLS estimates of any of the following regressions:

$$\Delta y_t = (\alpha - 1)y_{t-1} + u_t \tag{1}$$

$$\Delta y_t = \beta_0 + (\alpha - 1)y_{t-1} + u_t \tag{2}$$

$$\Delta y_t = \beta_0 + \beta_1 t + (\alpha - 1)y_{t-1} + u_t \tag{3}$$

$$\Delta y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + (\alpha - 1)y_{t-1} + u_t, \tag{4}$$

where $\Delta y_t \equiv y_t - y_{t-1}$, t is a linear time trend, u_t is an error term, and α is a parameter that equals unity under the null hypothesis. There are two types of DF tests, τ tests and z tests. For the former, the test statistic is the ordinary t statistic for $\alpha - 1$ to equal zero, and for the latter it is T (the sample size) times $\hat{\alpha} - 1$. I shall refer to the τ statistics based on equations (1) through (4) as τ_{nc}, τ_c, τ_{ct}, and τ_{ctt}, respectively, and to the corresponding z statistics as z_{nc}, z_c, z_{ct}, and z_{ctt}. The subscripts stand for "no constant," "constant," "constant and trend," and "constant, trend, and trend squared," respectively. This notation is not entirely standard, but it is much easier to remember than other notations that are commonly used.

Each of equations (1) through (4) will be appropriate under different assumptions. Equation (1) makes sense for $\alpha < 1$ only if y_t has (population) mean zero. It is therefore hard to imagine ever using (1) with economic time series data. The other three equations are less restrictive. For each of them, the β_j coefficient with the largest value of j must be equal to zero under the null hypothesis, since otherwise the series y_t would be of a different order under the null and alternative hypotheses. Thus (2) allows y_t to have a nonzero mean, (3) allows it to have a trend, and (4) allows it to have a trend that changes over time, under both the null and alternative

hypotheses. Equation (4) is less commonly encountered than equations (2) and (3), but it may well be a plausible specification in some cases, as Ouliaris, Park, and Phillips (1989) have argued.

Each of the eight test statistics defined above has a different, nonstandard asymptotic distribution that does not depend on the u_t's being either normally distributed or homoskedastic. The assumption of serial independence is essential, however. The easiest way to relax this assumption is to use ADF τ tests, in which enough lags of Δy_t are added to equations (1) through (4) to whiten the residuals. The τ statistics, computed as ordinary t statistics, remain asymptotically valid in the presence of serial correlation when this is done, provided the number of lags of Δy_t is allowed to increase at an appropriate rate. The usual z statistics are not valid, however, and since it is relatively difficult to compute valid ones, ADF z statistics are rarely used. A second approach is to use "nonparametric" τ or z tests, as proposed by Phillips (1987) and Phillips and Perron (1988). These tests turn out to have exactly the same asymptotic distributions as the corresponding DF and ADF tests.

Engle and Granger (1987) showed that tests of the null hypothesis that two or more integrated time series are not cointegrated can be performed in much the same way as unit root tests. Suppose that $\boldsymbol{Y}$ denotes a $T \times k$ matrix of observations on k time series that are believed to be $I(1)$ and may be cointegrated. Then if $\boldsymbol{y}_1$ denotes one column of $\boldsymbol{Y}$, $\boldsymbol{Y}_1$ denotes the remaining $k-1$ columns, and $\boldsymbol{X}$ denotes a matrix of nonstochastic regressors such as a constant and possibly one or more trend terms, we can always use OLS to estimate the equation

$$\boldsymbol{y}_1 = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Y}_1\boldsymbol{\eta}_1 + \boldsymbol{\nu}. \tag{5}$$

If the variables in $\boldsymbol{Y}$ are cointegrated, equation (5) is a cointegrating equation and the error vector $\boldsymbol{\nu}$ should be stationary. If they are not cointegrated, however, $\boldsymbol{\nu}$ must have a unit root. Thus the null hypothesis of noncointegration may be tested by using a DF or ADF test on the residuals from OLS estimation of (5). For the former case, the test regression is simply

$$\Delta\hat{\nu}_t = (\alpha - 1)\hat{\nu}_{t-1} + \text{residual}, \tag{6}$$

where $\hat{\nu}_t$ denotes the t^{th} residual from OLS estimation of equation (5), and $\Delta\hat{\nu}_t = \hat{\nu}_t - \hat{\nu}_{t-1}$. Notice that $\boldsymbol{X}_t$ is not included in this regression. Including it would make no difference asymptotically and little difference in finite samples. Tests based on (6) are called Engle-Granger or EG tests. When there is serial correlation, lagged values of $\Delta\hat{\nu}_t$ must be added to the right-hand side of this equation to whiten the residuals, and the resulting tests are called Augmented Engle-Granger or AEG tests.

The asymptotic distributions of the EG and AEG τ tests computed in this way depend on k, which is defined as one more than the number of elements of $\boldsymbol{\eta}_1$ that have to be estimated. This may be less than the number of possibly cointegrated variables if some elements of $\boldsymbol{\eta}_1$ are known. They also depend on the form of the matrix $\boldsymbol{X}$, which will typically consist of a constant, a constant and a linear trend, or a constant, a linear trend, and a quadratic trend, by analogy with (2) through (4). These asymptotic distributions are not the same as those of the DF and ADF tests, unless no elements of $\boldsymbol{\eta}_1$ have to be estimated, so that $k = 1$. Phillips and Ouliaris (1990) have proposed nonparametric τ and z tests based on regression (6). They have the same asymptotic distributions as the corresponding EG and AEG tests.

It is clear from the above discussion that the asymptotic distributions of the statistics which I will call $\tau_c(k)$, $z_c(k)$, $\tau_{ct}(k)$, $z_{ct}(k)$, $\tau_{ctt}(k)$, and $z_{ctt}(k)$ for $k = 1, 2, \ldots$, are of considerable theoretical and practical interest. Those of the statistics $\tau_{nc}(k)$ and $z_{nc}(k)$ are of theoretical but not practical interest. Nevertheless, there exists at present no easy way for applied workers to obtain these distributions. A few tables of accurate critical values for the unit root case may be found in Fuller (1976), while Engle and Yoo (1987) and Phillips and Ouliaris (1990) provide some rather inaccurate critical values for the cointegration case. MacKinnon (1991) provides reasonably accurate critical values for the τ_c and τ_{ct} tests only. It is possible to compute the asymptotic distributions of z statistics analytically, as Nabeya and Tanaka (1990) do for the unit root case, but the methods employed are not at all easy.

In this paper, I obtain simple approximations to the asymptotic distributions of the statistics $\tau_c(k)$, $z_c(k)$, $\tau_{ct}(k)$, $z_{ct}(k)$, $\tau_{ctt}(k)$, and $z_{ctt}(k)$ for $k = 1, 2, \ldots, 6$. Results are also obtained for $\tau_{nc}(k)$ and $z_{nc}(k)$, but since those test statistics are of no practical interest, they are not reported to save space. The first step is to use a very extensive series of Monte Carlo experiments to estimate a large number of points on the finite-sample distributions of the test statistics for a number of finite sample sizes. Then, following MacKinnon (1991), response surface regressions are used to obtain the corresponding points on the asymptotic distributions of the test statistics. All of this will be discussed in Section 2. Asymptotic distribution functions for many of the statistics will be presented graphically in Section 3. Finally, simple approximations are fit to these points, in order to make it possible to calculate approximate P values. These approximations will be presented in Section 4.

The use of Monte Carlo experiments to obtain asymptotic distribution functions is somewhat unusual. However, this approach has several advantages over alternative approaches such as those used by Perron (1989) and

Nabeya and Tanaka (1990). The latter approaches both depend on analytical results that will be different for each of the 48 test statistics dealt with in this paper. The approach used by Nabeya and Tanaka works only for z statistics, and has been applied by them only to the $z_{nc}(1)$, $z_c(1)$, and $z_{ct}(1)$ statistics. The approach used by Perron requires rather heavy asymptotic theory, but still requires extensive Monte Carlo experiments to approximate the asymptotic distribution functions. In contrast, the Monte Carlo approach used in this paper works for any test statistic that converges to some asymptotic distribution at an appropriate rate. It is conceptually simple and reasonably easy to implement, provided enough computing resources are available.

It might well be argued that finite-sample distributions would be more useful than asymptotic ones. In fact, MacKinnon (1991) presented response surface estimates that can be used to obtain finite-sample critical values for any sample size, and the response surfaces estimated in this paper could be used in the same way. However, finite-sample critical values, like any finite-sample results for unit root and cointegration tests, depend critically on very strong assumptions about the error terms of the test regressions. As Gregory (1992) and others have shown, deviations from the assumptions of homoskedasticity and serial independence can affect the finite-sample distributions of unit root and cointegration tests quite severely, and the direction of the effects is not at all easy to predict. The only distributions that do not depend on nuisance parameters are the asymptotic ones that are calculated in this paper.

2. Monte Carlo Experiments

The results of this paper were obtained from a very extensive set of Monte Carlo experiments. Because the objective was to obtain accurate estimates of asymptotic distributions, the design of the experiments was somewhat unusual. In particular, the total number of replications was extremely large. In order to make it feasible to store and sort the experimental results for several different test statistics at once, however, the number of replications in a single experiment was limited to 50,000. This number is large enough that any bias in estimating quantiles should be negligible, but it is not nearly large enough for a single experiment to yield accurate results. For example, in 100 different experiments, each with 50,000 replications, estimates of the .01 critical value for the unit root τ_c test for $T = 100$ varied from -3.455 to -3.532, with a mean of -3.497 and a standard deviation of 0.017. Thus there is evidently quite a bit of experimental randomness in the results of a single such experiment.

Two different programs were used to run the experiments. All eight unit root test statistics (four τ tests and four z tests) were computed together using one program, and all forty cointegration test statistics (eight different tests for each of five values of k) were computed together using another. The data for both sets of experiments were generated by independent random walks with NID errors. For each sample size, 100 experiments were done for the unit root tests, and 50 experiments were done for the cointegration tests. The number of experiments was less in the latter case because computation costs were much greater, and experience with the unit root case had shown that accuracy would still be adequate.

The choice of how many sample sizes, and which ones, to use in the experiments was somewhat arbitrary. Small values of T involve lower computational costs than large ones but provide less information about the asymptotic distributions of interest. At least three different values of T must be used, because the response surfaces have three parameters. It is highly desirable to use more than three values of T, of course, since that makes it possible to perform specification tests. However, it would have been inconvenient to use a very large number. In the end, 14 different values of T were used: 50, 60, 75, 100, 125, 150, 200, 250, 300, 400, 500, 750, 1000, and 1250. Thus the total number of replications was 70 million for the unit root tests and 35 million for the cointegration tests.

The amount of computation required to perform all these experiments was substantial but by no means unreasonable by today's standards. Only about 190 hours of CPU time were required on an IBM RS/6000 Model 550 workstation. This is roughly equivalent to 2000 hours on a 486/33 personal computer or 6000 hours on a 386/33. Disk storage could have been a greater problem than processor time, because storing 8 times 70 million unit root test statistics plus 40 times 35 million cointegration test statistics would have required an enormous amount of disk space. Therefore, no attempt was made to store the actual test statistics. Instead, for each experiment of 50,000 replications, 199 estimated quantiles $(.005, .010, \ldots, .990, .995)$ were calculated and stored.[1] This strategy dramatically reduced the disk storage requirements to only about 70 megabytes.

The estimated finite-sample quantiles from the Monte Carlo experiments were used to estimate response surfaces in which the quantiles of the asymptotic distributions of the various test statistics appear as parameters. Consider the estimation of the p^{th} quantile for some test statistic. Let $q^p(T_i)$ denote the estimate of that quantile based on the i^{th} experiment,

[1] The $.001, .002, \ldots, .009$ quantiles were also calculated, because the left-hand tail is the most interesting part of the distribution, but these were not used in estimating the approximate distribution functions discussed in Section 4.

for which the sample size is T_i. Then the response surface regressions have the form

$$q^p(T_i) = \beta_\infty + \beta_1 T_i^{-1} + \beta_2 T_i^{-2} + \varepsilon_i. \tag{7}$$

The first parameter here, β_∞, is the p^{th} quantile of the asymptotic distribution. It is what we are trying to estimate. The other two parameters allow the finite-sample distributions to differ from the asymptotic ones and to approach them as the sample size increases. The functional form (7) was determined empirically in MacKinnon (1991). As will be explained below, there is strong evidence that it fits well. However, various other functional forms, involving terms such as $T_i^{-1/2}$ and $T_i^{-3/2}$, were also tried. These sometimes performed about as well as (7), but never performed substantially better. In many cases, the estimate of β_2 was small and insignificant, but it was retained for all the response surfaces because it was highly significant for some of them.

Equation (7) was estimated 199 times for each of 48 different test statistics. There were 1400 observations for each of the unit root tests and 700 for each of the cointegration tests. Although the ε_i's are certainly correlated for adjacent values of p, there is no reason not to use single-equation estimation methods, because the same regressors always appear on the right-hand side of (7). Because each experiment involved so many replications, the random error terms ε_i, which arise from experimental error, were quite small, and the response surfaces therefore fit extremely well. In this circumstance, almost any sensible estimation method would have yielded essentially the same results. Because there was some evidence of heteroskedasticity (quantiles tended to be estimated more precisely for larger values of T), I chose to use a form of GMM estimation (Hansen, 1982). The estimation technique that was used is a variant of the one proposed by Cragg (1983). Using matrix notation, we can rewrite (7) as

$$\boldsymbol{q} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \boldsymbol{\Omega},$$

where $\boldsymbol{q}$ is a vector of 700 or 1400 observations on $q^p(T_i)$, $\boldsymbol{X}$ is a matrix with three columns having typical elements 1, T_i^{-1}, and T_i^{-2}, and $\boldsymbol{\Omega}$ is a diagonal covariance matrix. Let $\boldsymbol{Z}$ denote a matrix of 14 zero-one dummy variables, corresponding to the 14 different sample sizes. These may be thought of as instruments, although in this case the regressor matrix $\boldsymbol{X}$ lies entirely in the subspace spanned by the columns of $\boldsymbol{Z}$. Then Cragg's GMM estimator can be written as

$$\tilde{\boldsymbol{\beta}} = \left(\boldsymbol{X}'\boldsymbol{Z}(\boldsymbol{Z}'\hat{\boldsymbol{\Omega}}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{Z}(\boldsymbol{Z}'\hat{\boldsymbol{\Omega}}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{q}, \tag{8}$$

where $\hat{\Omega}$ is a diagonal matrix, the principal diagonal of which consists of the squared residuals from an OLS regression of $\boldsymbol{q}$ on $\boldsymbol{Z}$.

Because the columns of the $\boldsymbol{Z}$ matrix are dummy variables, the estimator (8) is extremely easy to compute. There are two steps. The first step is to regress $\boldsymbol{q}$ on $\boldsymbol{Z}$. This simply involves computing the sample means $\hat{q}_T^p$ of the $q^p(T_i)$'s for each of the 14 different values of T. At the same time, one computes the estimated standard error $\hat{\sigma}_T$ of each $\hat{q}_T^p$. The second step is to run a weighted least squares regression with 14 observations, in which the $\hat{q}_T^p$'s are regressed on 1, T^{-1}, and T^{-2}, using the inverses of the $\hat{\sigma}_T$'s as weights. This regression may be written as

$$\hat{q}_T^p/\hat{\sigma}_T = \beta_\infty(1/\hat{\sigma}_T) + \beta_1(T^{-1}/\hat{\sigma}_T) + \beta_2(T^{-2}/\hat{\sigma}_T) + v_T. \tag{9}$$

It is easy to verify that this procedure yields $\tilde{\boldsymbol{\beta}}$ as defined in (8). The covariance matrix of $\tilde{\boldsymbol{\beta}}$ may then be estimated by $1/s^2$ times the usual covariance matrix from OLS estimation of (9), where s^2 is the OLS estimate of the variance of v_T. Since that variance is actually unity, it makes no sense to estimate it in this case.

In addition to accounting for heteroskedasticity, this GMM estimation procedure automatically generates a statistic for testing the specification of the response surface equation (7). The objective function that is minimized by (8) is the quadratic form

$$(\boldsymbol{q} - \boldsymbol{X\beta})'\boldsymbol{Z}(\boldsymbol{Z}'\hat{\Omega}\boldsymbol{Z})^{-1}\boldsymbol{Z}'(\boldsymbol{q} - \boldsymbol{X\beta}). \tag{10}$$

The value of this objective function is just the sum of squared residuals from regression (9). Standard results about GMM estimation imply that, under the null hypothesis that (7) is a correct specification, expression (10) is asymptotically distributed as $\chi^2(11)$. There are 11 degrees of freedom because there are 14 instruments and 3 parameters.

The GMM specification test statistic (10) provides a simple way to test the functional form (7) against the very general alternative that the conditional mean q_T^p varies in any imaginable way with T, while allowing the variance of $\hat{q}_T^p$ around that mean, σ_T^2, to differ for every T. Because the experiments involved so many replications, the σ_T^2's were very small, and these tests were consequently very powerful. To illustrate this point, I deliberately introduced an error into the data for $T = 500$ for $\tau_{ct}(1)$. When all the .05 estimated critical values were increased by 0.01 for this sample size only, the value of the GMM test statistic rose from 3.16 to 125.79. The former value is not significant at any level, since it is well below the mean of the $\chi^2(11)$ distribution, while the latter exceeds any imaginable critical value for that distribution. Thus the GMM specification test appears to be able to detect even quite small discrepancies between the data and the

response surface postulated in (7). The results of the GMM specification tests for all $48 \times 199 = 9552$ sets of estimates will be discussed below.

The Monte Carlo experiments used a great many pseudo-random numbers, and I have not yet said anything about the random number generator that was used. Initially, I simply used a good multiplicative congruential generator, one of the ones suggested by L'Ecuyer (1988). This generator worked satisfactorily when I used it in MacKinnon (1991). However, when (7) was estimated using a full set of Monte Carlo results, the GMM test statistics were disturbingly high. Most of them exceeded 24.73, the .01 critical value for the $\chi^2(11)$ distribution, and many of them exceeded 31.26, the .001 critical value. I initially thought that the response surface regression might not fit satisfactorily for small values of T. However, omitting smaller values of T from the regression did not improve the results. Instead, to my surprise, omitting larger values of T did improve them.

The problem, of course, was that the random number generator I was using was not capable of generating enough different pseudo-random numbers. Like all good multiplicative congruential generators that use 32-bit signed integers, it had a period of about 2.15×10^9. After that many numbers had been generated, the same sequence was reappearing. But for $T = 1250$, the largest sample size, the unit root experiments required 6.23×10^9 numbers, and the cointegration experiments required 18.75×10^9 numbers. Thus, for the larger sample sizes, the results were not independent across the 50 or 100 experiments, and the $\hat{q}_T^p$'s were not as accurate as their estimated standard errors, the $\hat{\sigma}_T$'s, indicated. As a consequence, the GMM test statistics were too large.

To test this explanation, I redid some of the unit root experiments using a much inferior random number generator, one of the portable ones with a period of less than 10^6 discussed in Press, Flannery, Teukolsky, and Vetterling (1986, Chapter 7). Even though the number of replications was relatively modest (50 experiments for each sample size, with only 10,000 replications each), I obtained GMM test statistics that were absolutely enormous, between 10^3 and 10^4. This result confirmed that the GMM test statistics have excellent power to detect a poor random number generator. It also showed that the generator I used originally, although inadequate, was not nearly as bad as it could have been.

The solution I adopted was to use two different pseudo-random number generators, combined in a somewhat unusual way. Both generators were taken from L'Ecuyer (1988). One had a multiplier of 40,692 and a modulus of 2,147,483,399, and the other had a multiplier of 40,014 and a modulus of 2,147,483,563. 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. A modified Box-Muller

procedure was then used to transform these two uniform variates into two $N(0,1)$ variates.[2] Because each generator has a different modulus, the fact that each sequence of uniform variates will recur after roughly 2.147×10^9 iterations does not imply that the same sequence of $N(0,1)$ variates will do so, since the uniform variates from the two generators will be paired up differently each time the same sequences of uniforms reappear. The final results of the GMM specification tests, which will be discussed shortly, confirm that the combined generator does indeed work well.

Because there were 48 different test statistics and 199 quantiles for each, the response surface regression (7) was estimated 9552 times. It would not be possible to present all these results here. However, to give the flavor of the results, here are estimates for two different test statistics at $p = .05$:

$$\begin{aligned} \text{For } \tau_{ct}(1): \quad q^{.05} &= -3.41064 - 4.4755\, T^{-1} - 4.9224\, T^{-2} \\ &\qquad (0.00045) \quad (0.1389) \qquad (7.1711) \\ \text{SSR} &= 3.155 \end{aligned}$$

$$\begin{aligned} \text{For } z_{ct}(1): \quad q^{.05} &= -21.7100 + 128.134\, T^{-1} - 483.787\, T^{-2} \\ &\qquad (0.0055) \quad (1.545) \qquad (76.785) \\ \text{SSR} &= 7.978 \end{aligned}$$

Estimated standard errors are in parentheses, and SSR is the sum of squared residuals from the weighted least squares regression (9), which is the GMM test statistic. Notice that T^{-2} is insignificant in the regression for $\tau_{ct}(1)$ but highly significant in the regression for $z_{ct}(1)$. This term was generally more significant for the z tests than for the τ tests, although it was often significant for the latter as well.

It is possible to calculate asymptotic critical values analytically for $z(1)$ tests, and a few such values have been published in Nabeya and Tanaka (1990). It is of interest to compare these true values with the estimated ones obtained from regression (7). As can be seen from Table 1, the estimated critical values are extremely close to the true ones, and the magnitudes of the discrepancies between them are entirely consistent with the estimated standard errors of the former. This appears to provide strong evidence that the technique used in this paper works well, at least for the $z(1)$ tests.

---

[2] This procedure is described in Press, Flannery, Teukolsky, and Vetterling (1986, Chapter 7). Note that not every pair of uniform variates yields a pair of normal variates. Approximately 22% of the time, the latter cannot be computed and the uniform variates have to be discarded. Thus the combined generator will produce about 3.35×10^9 $N(0,1)$ variates before the individual generators start to cycle.

Table 1. True and Estimated Critical Values for z Tests

	.01 True	.01 Est.	.05 True	.05 Est.	.10 True	.10 Est.
z_{nc}	−13.6954	−13.6827 (0.0096)	−8.0391	−8.0352 (0.0038)	−5.7137	−5.7112 (0.0025)
z_c	−20.6259	−20.6207 (0.0105)	−14.0936	−14.0927 (0.0047)	−11.2506	−11.2469 (0.0033)
z_{ct}	−29.3586	−29.3763 (0.0114)	−21.7112	−21.7100 (0.0055)	−18.2453	−18.2388 (0.0039)

The most convincing evidence that the response surface regression (7) fits acceptably well comes from the GMM specification test statistics. Since there are 9552 of these, it is not feasible to report them individually. Instead, they are summarized in Table 2. Asymptotically, these statistics should follow the $\chi^2(11)$ distribution. Under that distribution, they should have mean 11 and variance 22, and they should exceed their nominal .05 critical value (which is 19.675) 5% of the time.

Table 2. Descriptive Statistics for GMM Statistics

Cases	Number	Mean	Variance	Rejections (.05)
All test statistics	9552	12.04	26.60	.083
Unit root statistics	1592	10.98	22.38	.077
$\tau(1)$ statistics	796	10.88	19.69	.065
$z(1)$ statistics	796	11.08	25.05	.089
All cointegration statistics	7960	12.25	27.17	.085
$\tau(k)$ statistics, $k \geq 2$	3980	11.61	23.10	.061
$z(k)$ statistics, $k \geq 2$	3980	12.89	30.43	.108

It can be seen from the table that the GMM statistics conform remarkably well, if not quite perfectly, to the $\chi^2(11)$ distribution. If the statistics had been independent, it would have been easy to test whether the numbers reported in the table differ significantly from the corresponding theoretical values. However, they are emphatically not independent. First of all, for any given test statistic, for example, the $\tau_c(2)$ test, the GMM statistics tended to be very similar for neighboring values of p. Secondly, one set of random numbers was used to generate all 8 unit root tests, and another set was used to generate all 40 cointegration tests, so that there must also

be some dependence across test statistics. In these circumstances, conventional tests on the means, variances, or rejection frequencies of the GMM test statistics would be invalid, and would be likely to overreject the null hypothesis.

At least part of the (relatively modest) discrepancy between the empirical and theoretical distributions of the GMM statistics can be attributed to finite-sample effects. A small Monte Carlo experiment demonstrated that the GMM statistics for regression (7) could be expected to overreject the null hypothesis to some extent, even when it was true, especially when there were only 700 observations. This effect might well be great enough to account for the discrepancies observed for the unit root tests and for the $\tau(k)$ tests for $k = 2, \ldots, 6$. It was not large enough to account for the discrepancies observed for the $z(k)$ tests for $k = 2, \ldots, 6$, however. In the case of the latter, there does appear to be some evidence that the functional form of (7) may not be quite correct. However, it still provides an extremely good approximation. If it did not, the GMM test statistics would on average have been very much larger than they were.

3. Asymptotic Distributions

The primary purpose of this paper is to obtain simple formulas that may be used to compute approximate P values. That will be done in the next section. The results from the response surface regressions of Section 2 may be used to plot the asymptotic distributions of the various unit root and cointegration test statistics. The shapes of these distributions are of considerable interest in their own right, and knowing what they look like will be useful for obtaining approximations. Therefore, this section presents results from some of the response surface regressions in graphical form.

Each of Figures 1 through 4 plots the cumulative distribution functions of several unit root or cointegration test statistics. These are obtained by joining the estimates of β_∞ from the response surface regressions (7) for all available values of p.[3] No smoothing at all was done. Nevertheless, because the estimates of β_∞ are so accurate, all the curves appear perfectly smooth to the naked eye.

Figure 1 shows the asymptotic c.d.f.'s of the four unit root τ statistics. For comparison, the c.d.f. of the standard normal distribution is also shown. Two things are evident from the figure. First, as we know from the critical

[3] In addition to the 199 values of p from .005 through .995, the values $p = .001, .002, \ldots, .010$ were used in drawing all the figures. This means that the left-hand tail is slightly more accurate than the right-hand tail for all the distributions.

values that have been tabulated, adding more nonstochastic regressors to the DF test regression causes the asymptotic c.d.f. to move to the left. Second, the distribution of τ_{nc} is somewhat asymmetrical. As more regressors are added, the c.d.f.'s become progressively more symmetrical and steeper. Thus the distribution of τ_{ctt} looks as if it might be approximately normal, with mean roughly -2.63 and standard deviation roughly 0.72. In fact, as we shall see in the next section, this is a pretty good approximation, except in the tails.

Figure 2 shows the asymptotic c.d.f.'s of the four unit root z statistics. These look very different from their τ counterparts, since they are grossly asymmetrical and have very long left-hand tails. As more regressors are added to the test regression, all the c.d.f.'s move to the left, and they become somewhat less asymmetrical, but they also become flatter rather than steeper. Moreover, it is clear that a normal approximation will never be a good one.

Figure 3 shows the asymptotic c.d.f.'s of the six statistics $\tau_{ct}(1)$ through $\tau_{ct}(6)$. These look very similar to each other, except that they move steadily to the left as k increases. The effect of adding another stochastic regressor is qualitatively similar, but not identical, to the effect of adding another nonstochastic one. Much the same can be said about the asymptotic c.d.f.'s of the six statistics $z_{ct}(1)$ through $z_{ct}(6)$, which are shown in Figure 4. These are highly asymmetrical for all values of k and have very long left-hand tails.

4. Convenient Approximations

In this section, I obtain simple formulas that may be used to compute asymptotic P values for the c, ct, and ctt varieties of the τ and z tests for $k = 1, \ldots, 6$. To save space, I omit the nc varieties, which are of no practical interest. There are, in principle, many different ways to use estimated quantiles to obtain approximate c.d.f.'s. The one that I use seems to work reasonably well, but I do not claim that it is optimal in any sense.

Figures 1 and 3 suggest that, if one wishes to approximate the asymptotic distributions of τ statistics, the normal distribution is a good place to start. Suppose that X is a random variable distributed as $N(\mu, \sigma^2)$. Then, if $p \equiv \Pr(X < x)$, we know that

$$p = \Phi\big((x - \mu)/\sigma\big) = \Phi\big(\gamma_0 + \gamma_1 x\big), \tag{11}$$

where $\Phi(\cdot)$ denotes the cumulative standard normal distribution function, $\gamma_0 \equiv -\mu/\sigma$, and $\gamma_1 \equiv 1/\sigma$. Suppose we knew two or more points (x_j, p_j)

on the c.d.f. of X. Then we could use (11) to calculate the parameters of the distribution. The easiest way to do so would be to write

$$\Phi^{-1}(p_j) = \gamma_0 + \gamma_1 x_j. \tag{12}$$

The parameters γ_0 and γ_1 could then be found by solving two linear equations in two unknowns, the two equations being equation (12) for $j = 1, 2$.

The preceding discussion has made two assumptions that are untenable in the present case. The first of these is that the functional form of the c.d.f. to be estimated is precisely that of the normal distribution. One obvious way to relax this assumption is to add additional terms in powers of x_j to equation (12). The third term would be $\gamma_2 x_j^2$, the fourth term would be $\gamma_3 x_j^3$, and so on. Ideally, the number of such additional terms should be quite small.

Another possibility would be to use some sort of spline regression; see Poirier (1976) and Eubank (1984). The idea would be to divide the domain of X into a number of segments and estimate different approximations for each segment, constraining them to be equal at the "knots" where the segments meet. The problem with this spline approach is that there may well be quite a few parameters to estimate, and it would therefore be very difficult to write down the results for 36 different test statistics in a compact form.

The second untenable assumption is that points on the c.d.f. of interest are known exactly. Although the Monte Carlo experiments discussed in Section 2 yielded remarkably accurate estimates of 199 points on the c.d.f.'s of each of the τ statistics, these estimates do contain some experimental error. Instead of points (x_j, p_j), we actually have points (x_j^*, p_j), where $x_j^* = x_j + e_j^*$ and e_j^* is an error term. Thus, if equation (12) held exactly, we would have

$$\Phi^{-1}(p_j) = \gamma_0 + \gamma_1 x_j^* - \gamma_1 e_j^* = \gamma_0 + \gamma_1 x_j^* + e_j. \tag{13}$$

This has the form of a linear regression model. Unfortunately, it is a model that has several statistical problems. First, there will inevitably be some heteroskedasticity, because the e_j^*'s do not have constant variance. However, since x_j^* is just the estimate of β_∞ from equation (9) for the p_j^{th} quantile, consistent estimates of the variances of the e_j^*'s are available. Second, there will inevitably be serial correlation, because the estimates of β_∞ for nearby values of p will surely be correlated. This problem is harder to deal with, because it is not obvious that the serial correlation would be expected to follow any standard ARMA process.

A third problem is that the regressand of equation (13) is nonstochastic, while the regressor x_j^* is stochastic and correlated with the error

term e_j. At first glance, this seems to be an extreme form of the classical errors-in-variables problem. Nevertheless, because the standard errors of the e_j^*'s are extremely small, we can simply ignore this problem. The estimated standard errors vary a certain amount, tending to be larger for values of p close to 0 and 1, but the largest ones are less than 0.0022, and most of them are between 0.0002 and 0.0005. Let $\boldsymbol{y}$ denote the vector of observations on $\Phi^{-1}(p_j)$, $\boldsymbol{x}^*$ the vector of observations on x_j^*, $\boldsymbol{x}$ the corresponding vector of unobserved x_j's, $\boldsymbol{M}_\iota$ the projection matrix that takes deviations from the mean, and $\boldsymbol{e}$ the vector of error terms. Then the OLS estimate of γ_1 from regression (13) is

$$\begin{aligned}\hat{\gamma}_1 &= \bigl((\boldsymbol{x}^*)'\boldsymbol{M}_\iota\boldsymbol{x}^*\bigr)^{-1}(\boldsymbol{x}^*)'\boldsymbol{M}_\iota\boldsymbol{y} \\ &= \bigl(\boldsymbol{x}'\boldsymbol{M}_\iota\boldsymbol{x} + \boldsymbol{e}'\boldsymbol{M}_\iota\boldsymbol{e} + 2\boldsymbol{x}'\boldsymbol{M}_\iota\boldsymbol{e}\bigr)^{-1}\bigl(\boldsymbol{x}'\boldsymbol{M}_\iota\boldsymbol{y} + \boldsymbol{e}'\boldsymbol{M}_\iota\boldsymbol{y}\bigr).\end{aligned}$$

It is easy to evaluate the second line here, and its counterpart for γ_0, for sensible values of the parameters and the variance of the e_j's. When one does so, one finds that any bias in the estimates of γ_0 and γ_1 is tiny, generally less than 0.00001 in absolute value for both parameters. Essentially, this is because $\boldsymbol{e}'\boldsymbol{M}_\iota\boldsymbol{e}$ is very much smaller than $\boldsymbol{x}'\boldsymbol{M}_\iota\boldsymbol{x}$. Thus it appears to be quite safe to ignore the errors-in-variables problem and simply estimate some variant of equation (13).

The fact that the error terms are so small can be used to justify ignoring the other two problems as well. Taking heteroskedasticity and serial correlation into account would yield more efficient estimates, of course. But efficiency is not much of a concern when the error terms are tiny. Nevertheless, I did in fact take heteroskedasticity into account by using weighted least squares, with weights equal to the inverses of the estimated standard errors of the e_j^*'s. This is approximately correct even if higher-order terms are added to the right-hand side of regression (13), as long as the coefficients on those higher-order terms are small relative to γ_1.

Even though the asymptotic distributions of the τ statistics appear to be approximately normal (see Figures 1 and 3), estimation of (13) does not yield entirely satisfactory results. For example, letting τ denote the value of the test statistic and p the corresponding P value, here are some results for $\tau_c(1)$:

$$\begin{aligned}(14)\qquad \Phi^{-1}(p) &= \underset{(0.0074)}{-1.8951} + \underset{(0.0043)}{1.2236}\,\tau \\ R^2 &= 0.9977 \qquad \text{DW} = 0.002\end{aligned}$$

$$(15) \qquad \Phi^{-1}(p) = \underset{}{-1.7760} + \underset{(0.0066)}{1.0448}\,\tau - \underset{(0.0083)}{0.0578}\,\tau^2 \qquad \underset{(0.0026)}{}$$

$$R^2 = 0.9993 \qquad \text{DW} = 0.006$$

$$(16) \qquad \Phi^{-1}(p) = \underset{}{-1.7325} + \underset{(0.0039)}{0.8898}\,\tau - \underset{(0.0079)}{0.1836}\,\tau^2 - \underset{(0.0056)}{0.0282}\,\tau^3 \qquad \underset{(0.0012)}{}$$

$$R^2 = 0.9998 \qquad \text{DW} = 0.037$$

On the one hand, these equations all fit remarkably well, as evidenced by extremely high R^2's (in terms of the weighted data). On the other hand, they also have extremely low Durbin-Watson statistics, which suggests that they are severely misspecified. The first two equations are certainly misspecified, since τ^2 and τ^3 are both highly significant in the third. So is the third, because higher powers of τ are highly significant when added to it, and so, by a similar argument, are models that are much more complicated than the third.

Although equations (14) through (16) are not satisfactory from a statistical point of view, they might be perfectly satisfactory for some purposes. Recall that equations like these will ultimately be used in the form of equation (11), with p expressed as $\Phi(\gamma_0 + \gamma_1\tau + \cdots)$. The maximum absolute error in estimating p (over the 199 values used) is 0.0143 for equation (14), 0.0072 for equation (15), and 0.0046 for equation (16). Errors of this magnitude might be entirely acceptable for many purposes. Results for equations similar to (16) are therefore presented in Table 3. The column headed $|p - \hat{p}|$ shows the maximum absolute error in estimating p. The maximum error becomes smaller as k increases and as the number of nonstochastic regressors increases. This confirms the visual impression, from Figures 1 and 3, that the asymptotic distributions of τ statistics become closer to the normal distribution in both cases.

One problem with equations like (16) as approximations to c.d.f.'s is that, because they involve powers of τ, they inevitably break down for sufficiently extreme values of τ. As $\tau \to -\infty$, p should tend to 0, and as $\tau \to +\infty$, p should tend to 1. But that is not what happens with the approximations of which the coefficients appear in Table 3. Instead, there is a fairly large negative value, $\tau_{\min}$, at which p achieves a minimum, say $p_{\min}$, and a positive value, $\tau_{\max}$, at which p achieves a maximum. The values of $p_{\min}$, $\tau_{\min}$, and $\tau_{\max}$ are shown in the last three columns of Table 3. These approximations should obviously not be used when $\tau < \tau_{\min}$ or $\tau > \tau_{\max}$. In the former case, all we know about p is that it is close to 0, and in the latter case, all we know is that it is close to 1.

Table 3. Approximate CDF's for τ Tests

Statistic	γ_0	γ_1	γ_2	γ_3	$\lvert p - \hat{p} \rvert$	$p_{\min}$	$\tau_{\min}$	$\tau_{\max}$
$\tau_c(1)$	1.7325	0.8898	−0.1836	−0.02820	.0046	.00002	−6.07	1.73
$\tau_c(2)$	2.2092	0.6808	−0.2705	−0.03833	.0028	.00039	−5.74	1.03
$\tau_c(3)$	2.7246	0.6720	−0.2545	−0.03256	.0020	.00026	−6.30	1.09
$\tau_c(4)$	3.2776	0.7667	−0.2066	−0.02452	.0018	.00007	−7.09	1.47
$\tau_c(5)$	3.8227	0.8783	−0.1617	−0.01817	.0015	.00001	−7.96	2.02
$\tau_c(6)$	4.3062	0.9499	−0.1353	−0.01455	.0011	.00000	−8.70	2.50
$\tau_{ct}(1)$	2.6130	0.7831	−0.2828	−0.04285	.0022	.00091	−5.51	1.11
$\tau_{ct}(2)$	3.0348	0.8084	−0.2317	−0.03125	.0019	.00029	−6.31	1.37
$\tau_{ct}(3)$	3.4954	0.8754	−0.1840	−0.02271	.0016	.00005	−7.19	1.79
$\tau_{ct}(4)$	3.9904	0.9717	−0.1408	−0.01650	.0015	.00001	−8.11	2.42
$\tau_{ct}(5)$	4.4318	1.0233	−0.1183	−0.01317	.0011	.00000	−8.90	2.91
$\tau_{ct}(6)$	4.8639	1.0739	−0.1005	−0.01082	.0009	.00000	−9.63	3.44
$\tau_{ctt}(1)$	3.3784	0.9197	−0.2238	−0.03180	.0021	.00041	−6.24	1.55
$\tau_{ctt}(2)$	3.8109	1.0131	−0.1605	−0.02126	.0016	.00005	−7.23	2.20
$\tau_{ctt}(3)$	4.2292	1.0763	−0.1225	−0.01526	.0014	.00001	−8.21	2.86
$\tau_{ctt}(4)$	4.6461	1.1291	−0.0973	−0.01163	.0011	.00000	−9.12	3.55
$\tau_{ctt}(5)$	5.0308	1.1549	−0.0848	−0.00970	.0009	.00000	−9.86	4.03
$\tau_{ctt}(6)$	5.4153	1.1863	−0.0736	−0.00820	.0008	.00000	−10.55	4.57

Using the results in Table 3 to compute asymptotic P values is very easy. Given a value of one of the test statistics, say τ, one first verifies that it is less than $\tau_{\max}$ and greater than $\tau_{\min}$ and then calculates

$$h = \gamma_0 + \gamma_1\tau + \gamma_2\tau^2 + \gamma_3\tau^3$$

using the appropriate set of γ_i's from the table. One then calculates p as $\Phi(h)$, the standard normal c.d.f. evaluated at h. In effect, we are computing the P value for a standard one-tailed test of the null hypothesis that $h \geq 0$, where h follows the standard normal distribution.

It is possible to make the errors in estimating p for the left-hand tail of the distribution very much smaller at the cost of increasing them elsewhere. Since the left-hand tail is the part of the distribution that is primarily of interest, it may make sense to do so. Trial and error suggested that if only the first 30 observations (for $p = .005$ to $p = .150$) were used, very good results could be obtained from a model like (15) with only a constant, τ, and τ^2 as regressors. The results were good in the sense that R^2 was essentially unity, and the maximum absolute error in estimating p within

Table 4. Approximate CDF's for τ Tests when p is Small

Statistic	γ_0	γ_1	γ_2	p^*	τ^*	$\tau_{\min}$
$\tau_c(1)$	2.1659	1.4412	0.03827	.495	−1.586	−18.83
$\tau_c(2)$	2.9200	1.5012	0.03980	.380	−2.285	−18.86
$\tau_c(3)$	3.4699	1.4856	0.03164	.295	−2.877	−23.48
$\tau_c(4)$	3.9673	1.4777	0.02632	.255	−3.330	−28.07
$\tau_c(5)$	4.5509	1.5338	0.02954	.375	−3.399	−25.96
$\tau_c(6)$	5.1399	1.6036	0.03445	.480	−3.498	−23.27
$\tau_{ct}(1)$	3.2512	1.6047	0.04959	.255	−2.657	−16.18
$\tau_{ct}(2)$	3.6646	1.5419	0.03645	.265	−2.998	−21.15
$\tau_{ct}(3)$	4.0983	1.5173	0.02990	.250	−3.372	−25.37
$\tau_{ct}(4)$	4.5844	1.5338	0.02880	.360	−3.447	−26.63
$\tau_{ct}(5)$	5.0722	1.5634	0.02947	.480	−3.510	−26.53
$\tau_{ct}(6)$	5.5300	1.5914	0.03039	.490	−3.763	−26.18
$\tau_{ctt}(1)$	4.0002	1.6580	0.04829	.280	−3.034	−17.17
$\tau_{ctt}(2)$	4.3534	1.6016	0.03795	.315	−3.275	−21.10
$\tau_{ctt}(3)$	4.7343	1.5768	0.03240	.310	−3.582	−24.33
$\tau_{ctt}(4)$	5.2140	1.6077	0.03345	.535	−3.436	−24.03
$\tau_{ctt}(5)$	5.6481	1.6274	0.03345	.920	−2.760	−24.33
$\tau_{ctt}(6)$	5.9296	1.5929	0.02822	.325	−4.343	−28.22

the sample was extremely small (less than 0.0001 in all cases). The DW statistics were always less than two, but were consistent with the amount of serial correlation that could be expected to arise from the design of the Monte Carlo experiments; they were never anything like as tiny as those reported above. In many cases, τ^3 was still significant when added to these equations, but not massively significant as it is in equation (16) above. Moreover, adding it had little effect on the fitted values within the sample and often made the equation fit worse outside the sample.

The results obtained from this approach are presented in Table 4. These results are extremely accurate for small values of p, including values less than .005 and greater than .150. Because $\tau_{\min}$ is always very much smaller in Table 4 than it was in Table 3, it is clear that these approximations should be reasonably accurate even for very small values of p. On the other hand, this approach can yield quite inaccurate results for some large values of p. The column labelled p^* in Table 4 contains the largest value of p for which the absolute error is less than 0.001, and the corresponding value of τ is found in the column labelled τ^*. In most cases, practitioners should use the approximations in Table 4, because they are simpler and

Table 5. Approximate CDF's for z Tests

Statistic	γ_0	γ_1	$\gamma_2 \times 10^2$	$\gamma_3 \times 10^3$	$\gamma_4 \times 10^5$	$\lvert p - \hat{p} \rvert$	$p_{\min}$	$z_{\min}$
$z_c(1)$	1.7157	0.5536	4.5518	2.2466	4.2537	.0045	.0084	−22.03
$z_c(2)$	2.2315	0.4164	2.2550	0.7765	1.0572	.0039	.0088	−30.87
$z_c(3)$	2.7220	0.3520	1.4065	0.3653	0.3819	.0024	.0070	−40.19
$z_c(4)$	3.1785	0.3162	1.0217	0.2171	0.1876	.0020	.0060	−48.61
$z_c(5)$	3.5856	0.2893	0.7802	0.1393	0.1019	.0015	.0047	−57.36
$z_c(6)$	3.9485	0.2675	0.6167	0.0947	0.0600	.0011	.0036	−66.16
$z_{ct}(1)$	2.7119	0.4594	2.3747	0.7488	0.9333	.0026	.0076	−32.85
$z_{ct}(2)$	3.0557	0.3899	1.6247	0.4241	0.4406	.0031	.0080	−39.86
$z_{ct}(3)$	3.3848	0.3374	1.1303	0.2430	0.2108	.0026	.0071	−47.95
$z_{ct}(4)$	3.7066	0.2996	0.8238	0.1477	0.1080	.0018	.0053	−57.01
$z_{ct}(5)$	4.0233	0.2730	0.6369	0.0981	0.0622	.0014	.0040	−65.84
$z_{ct}(6)$	4.3238	0.2522	0.5087	0.0684	0.0381	.0011	.0027	−75.09
$z_{ctt}(1)$	3.4216	0.4170	1.6939	0.4203	0.4153	.0027	.0075	−41.18
$z_{ctt}(2)$	3.6844	0.3631	1.2347	0.2622	0.2231	.0026	.0071	−48.29
$z_{ctt}(3)$	3.9298	0.3179	0.8967	0.1615	0.1177	.0018	.0055	−56.75
$z_{ctt}(4)$	4.1861	0.2853	0.6822	0.1059	0.0672	.0015	.0044	−65.39
$z_{ctt}(5)$	4.4458	0.2607	0.5377	0.0730	0.0419	.0012	.0033	−74.23
$z_{ctt}(6)$	4.6932	0.2403	0.4306	0.0512	0.0253	.0012	.0021	−84.21

more accurate for small values of p. However, if $\tau > \tau^*$ and accurate results are desired, it will generally be better to use the results in Table 3.

It was considerably harder to model the asymptotic c.d.f.'s of the z statistics than those of the τ statistics. This might be expected in view of the rather strange shapes of the former that are evident in Figures 2 and 4. Results based on all 199 observations are presented in Table 5. These are similar to those in Table 3, except that there are five regressors (the constant, z, z^2, z^3, and z^4) rather than four. Note that the values of γ_2, γ_3, and γ_4 printed in the table have been multiplied by 10^2, 10^3, and 10^5, respectively, in order to preserve precision while conserving space. As before, the column headed $|\hat{p} - p|$ contains the maximum absolute error in estimating p within the range from .005 to .995.

The approximations in Table 5 perform very badly for small values of z. What happens is that $\hat{p}$ starts to increase rapidly as z declines below some value. As a result, the smallest value of $\hat{p}$ that can be obtained is significantly greater than zero. This value is shown in the column headed

Table 6. Approximate CDF's for z Tests when p is Small

Test Stat.	δ_0	δ_1	δ_2	δ_3	p^*	z^*
$z_c(1)$	2.2142	−1.7863	0.3283	−0.07727	.220	−7.96
$z_c(2)$	1.1662	0.1814	−0.3671		.215	−13.07
$z_c(3)$	6.6584	−4.3486	1.0471	−0.15011	.210	−18.14
$z_c(4)$	4.6795	−2.0163	0.3368	−0.08044	.250	−21.67
$z_c(5)$	1.7428	0.9638	−0.5212		.240	−26.47
$z_c(6)$	2.0856	1.0550	−0.5424		.275	−29.83
$z_{ct}(1)$	4.6476	−2.8932	0.5832	−0.09990	.245	−13.46
$z_{ct}(2)$	7.2453	−4.7021	1.1270	−0.15665	.250	−17.65
$z_{ct}(3)$	5.7487	−2.8370	0.5578	−0.10078	.335	−19.78
$z_{ct}(4)$	1.6604	1.0375	−0.5338		.245	−27.08
$z_{ct}(5)$	2.0060	1.1197	−0.5532		.310	−29.67
$z_{ct}(6)$	2.1161	1.3046	−0.5846		.275	−34.58
$z_{ctt}(1)$	4.4599	−1.8635	0.2126	−0.06070	.345	−16.27
$z_{ctt}(2)$	2.0864	0.5594	−0.4626		.230	−23.34
$z_{ctt}(3)$	2.0062	0.8907	−0.5171		.245	−27.19
$z_{ctt}(4)$	2.3870	0.9467	−0.5324		.355	−28.05
$z_{ctt}(5)$	2.1998	1.2828	−0.5834		.275	−35.02
$z_{ctt}(6)$	2.1803	1.5182	−0.6206		.275	−39.17

$p_{\min}$, and the column headed $z_{\min}$ shows the corresponding value of z. One should evidently never use these approximations if $z < z_{\min}$.

As with the τ tests, it is possible to model the left-hand tails of the asymptotic distributions of the z tests quite accurately by using only the observations for $p \leq .150$. Figures 2 and 4 suggest that the normal approximation is not a very good one. Therefore, I used approximations of the form

$$p = \Phi\big(\delta_0 + \delta_1 \log|z| + \delta_2 \log|z|^2 + \delta_3 \log|z|^3\big), \tag{17}$$

which allow the left-hand tail to be considerably longer than that of the normal distribution. This approximation does not work well for z close to zero, but it does seem to work well for the left-hand tails. Equation (17) was estimated by regressing $\Phi^{-1}(p)$ on a constant and powers of $\log|z|$. As before, this was a weighted least squares regression. If one assumes that z is measured with error and that δ_2 and δ_3 are small relative to δ_1, the error term has variance δ_1^2/z^2 times the variance of the measurement error. Therefore, the weight for observation j was equal to z_j divided by the standard error of e_j^*.

The results of estimating equation (17) for the z_c, z_{ct}, and z_{ctt} statistics are presented in Table 6. To make these results easier to use, the coefficient δ_3 was set to zero whenever doing so did not appreciably affect the quality of the approximation. In all cases, the maximum absolute error in estimating p is no greater than 0.0001 for $.005 \leq p \leq .150$. As before, the column labelled p^* contains the largest value of p for which the absolute error is less than 0.001, and the corresponding value of z is found in the column labelled z^*. It would be unwise to use these approximations if z were much greater than z^*. There is no column labelled $z_{\min}$ because these approximations behave as they should in the left-hand tail; that is, $p \to 0$ as $z \to -\infty$. When z is very large and negative, the approximations will say, correctly, that p is very close to zero.

5. Conclusion

The results of this paper make it very easy to compute asymptotic P values for a number of common unit root and cointegration tests. Although these results were obtained by means of Monte Carlo experiments, they appear to be more than accurate enough for most purposes. For each of 36 test statistics, two different approximations are given: one that is reasonably accurate for $.005 \leq p \leq .995$ and one that is very accurate for $p \leq .150$. For the former, any experimental error is swamped by the errors of approximation that are inevitable when one tries to use relatively simple functional forms. For both approximations, the discrepancies between the (unknown) finite-sample distributions of various test statistics and their asymptotic distributions will, in practice, generally be greater than any inaccuracies that may have arisen from either source of error.

References

Cragg, J. G. (1983). "More efficient estimation in the presence of heteroskedasticity of unknown form," *Econometrica*, 51, 751–63.

Dickey, D. A., W. R. Bell, and R. B. Miller (1986). "Unit roots in time series models: tests and implications," *The American Statistician*, 40, 12–26.

Dickey, D. A., and W. A. Fuller (1979). "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, 74, 427–31.

Engle, R. F., and C. W. J. Granger (1987). "Co-integration and error correction: representation, estimation and testing," *Econometrica*, 55, 251–76.

Engle, R. F., and B. S. Yoo (1987). "Forecasting and testing in co-integrated systems," *Journal of Econometrics*, 35, 143–59.

Eubank, R. L. (1984). "Approximate regression models and splines," *Communications in Statistics: Theory and Methods*, 13, 433–484.

Fuller, W. A. (1976). *Introduction to Statistical Time Series*, New York, John Wiley & Sons.

Gregory, A. W. (1992). "Testing for cointegration in linear quadratic models," Queen's University Institute for Economic Research Discussion Paper No. 811.

Hansen, L. P. (1982). "Large sample properties of generalized method of moments estimators," *Econometrica*, 50, 1029–54.

L'Ecuyer, P. (1988). "Efficient and portable random number generators," *Communications of the ACM*, 31, 742–51.

MacKinnon, J. G. (1991). "Critical values for cointegration tests," Ch. 13 in *Long-run Economic Relationships: Readings in Cointegration*, eds. R. F. Engle and C. W. J. Granger, Oxford, Oxford University Press.

Nabeya, S., and K. Tanaka (1990). "A general approach to the limiting distribution for estimators in time series regression with nonstable autoregressive errors," *Econometrica*, 58, 145–63.

Ouliaris, S., J. Y. Park, and P. C. B. Phillips (1989). "Testing for a unit root in the presence of a maintained trend," Ch. 1 in *Advances in Econometrics*, ed. B. Raj, Boston, Klumer Academic Publishers.

Perron, P. (1989). "The great crash, the oil price shock, and the unit root hypothesis," *Econometrica*, 57, 1361–1401.

Phillips, P. C. B. (1987). "Time series regression with a unit root," *Econometrica*, 55, 277–301.

Phillips, P. C. B., and S. Ouliaris (1990). "Asymptotic properties of residual based tests for cointegration," *Econometrica*, 58, 165–93.

Phillips, P. C. B., and P. Perron (1988). "Testing for a unit root in time series regression," *Biometrika*, 75, 335–46.

Poirier, D. J. (1976). *The Econometrics of Structural Change, with Special Emphasis on Spline Functions,* Amsterdam, North-Holland.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1986). *Numerical Recipes*, Cambridge, Cambridge University Press.

Said, E. S., and D. A. Dickey (1984). "Testing for unit roots in autoregressive-moving average models of unknown order," *Biometrika*, 71, 599–607.

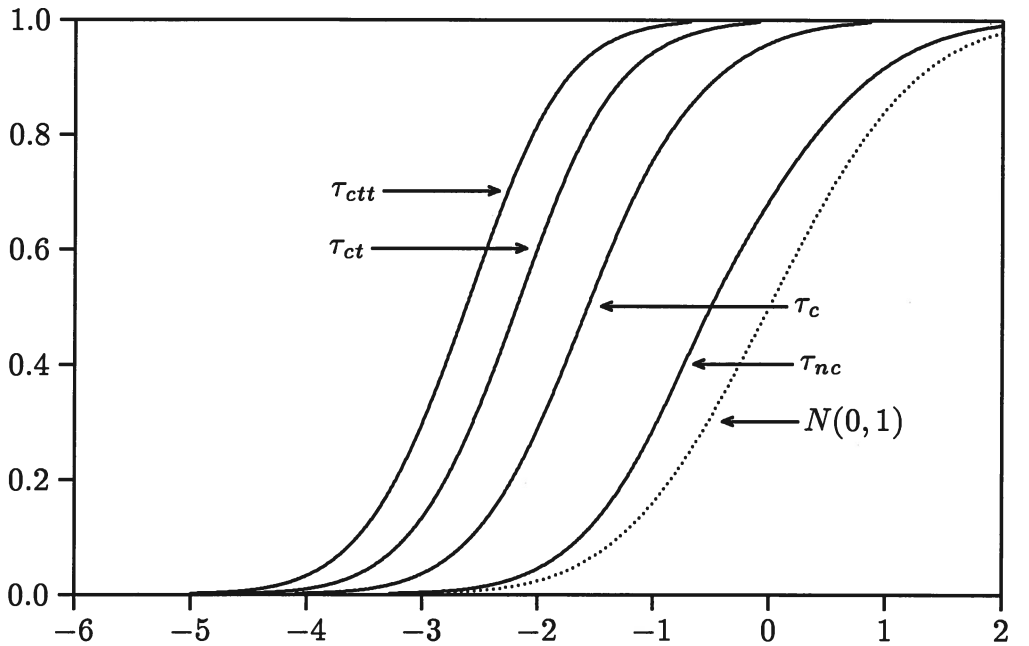

Figure 1. Asymptotic Distributions of $\tau(1)$ Statistics

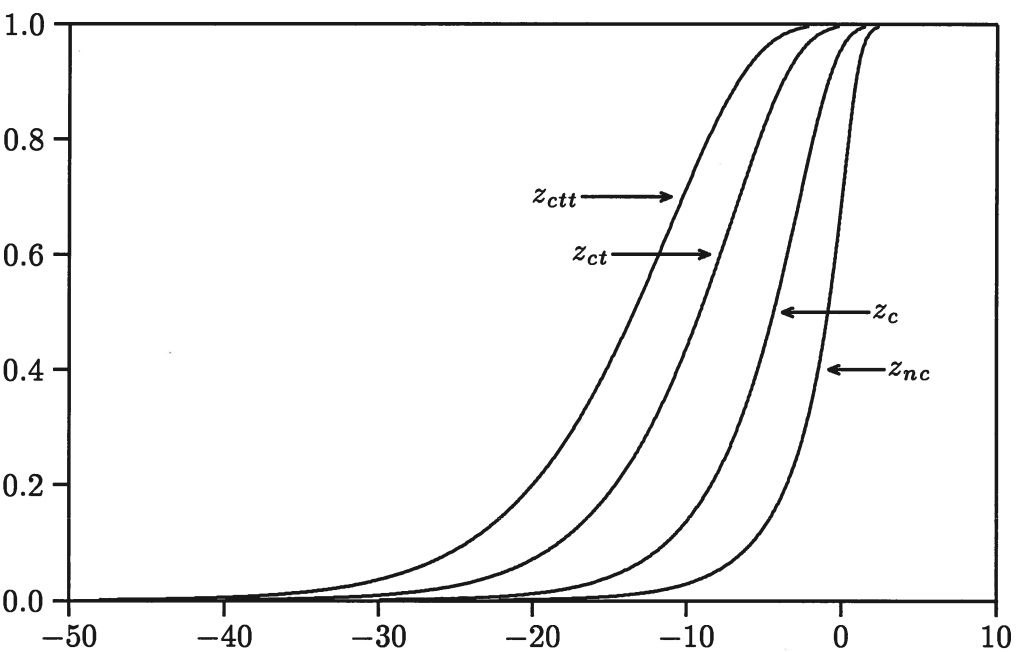

Figure 2. Asymptotic Distributions of $z(1)$ Statistics

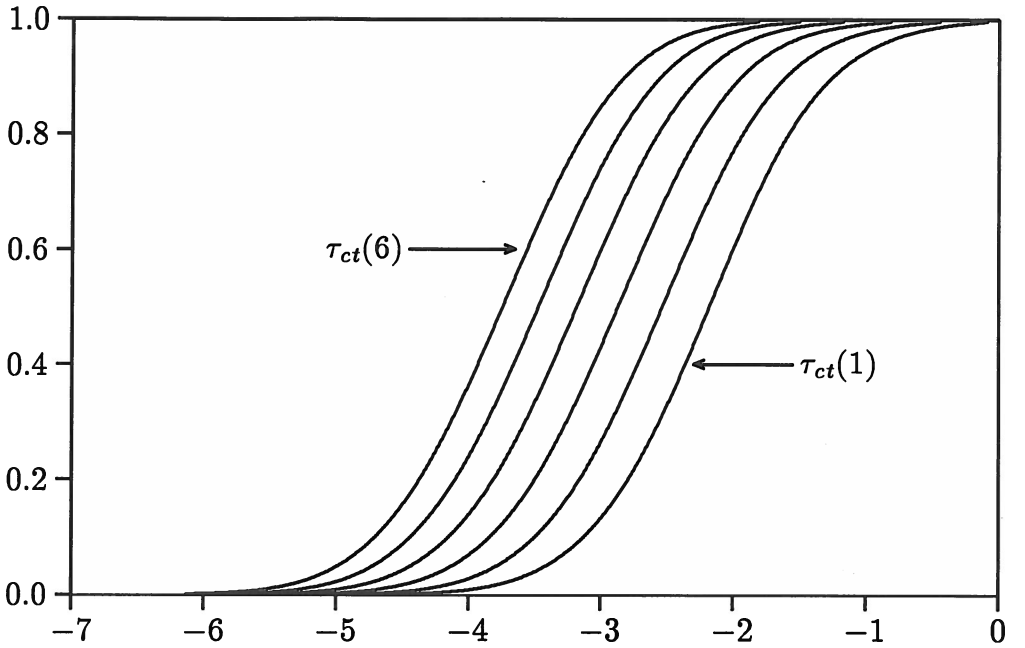

Figure 3. Asymptotic Distributions of $\tau_{ct}(1)$ through $\tau_{ct}(6)$

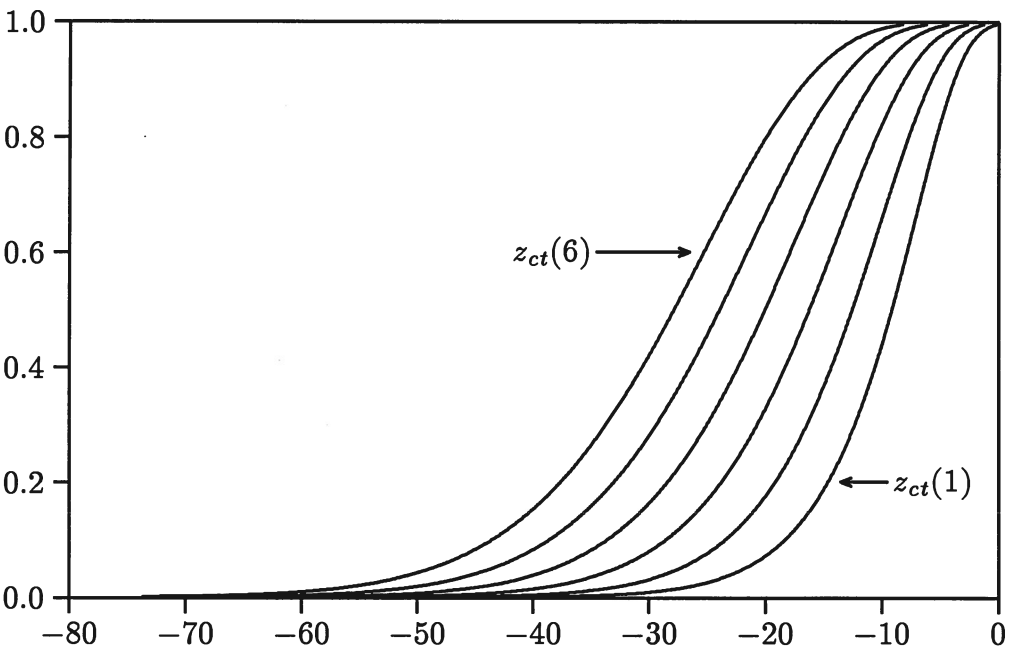

Figure 4. Asymptotic Distributions of $z_{ct}(1)$ through $z_{ct}(6)$