COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale University
New Haven, Connecticut 06520

Cowles Foundation Discussion Paper No. 933

TESTING FOR A UNIT ROOT IN THE PRESENCE OF DETERMINISTIC TRENDS

by

Peter Schmidt

and

Peter C.B. Phillips

October 1989

TESTING FOR A UNIT ROOT IN THE PRESENCE OF DETERMINISTIC TRENDS

Peter Schmidt
Michigan State University

and

Peter C. B. Phillips
Yale University

August, 1989

Revised: October, 1989

The financial support of the National Science Foundation is gratefully acknowledged.

1. INTRODUCTION

The most commonly used tests of the null hypothesis of a unit root in an observed time series are derivatives of the Dickey-Fuller tests (Dickey (1976), Fuller (1976), Dickey and Fuller (1979)). The Dickey-Fuller tests were developed for simple Gaussian random walks and the derivative procedures (notably Said and Dickey (1984), Phillips (1987) and Phillips and Perron (1988)) are intended to detect the presence of a unit root in a general integrated process of order one (I(1) process). The Dickey-Fuller tests are based on the regression of the observed variable (say, y) on its one-period lagged value, with the regression sometimes including an intercept and time trend; that is, they are based on regressions of the form:

(1) $$y_t = \beta y_{t-1} + \varepsilon_t$$

(2) $$y_t = \alpha + \beta y_{t-1} + \varepsilon_t$$

(3) $$y_t = \alpha + \beta y_{t-1} + \delta t + \varepsilon_t ,$$

for $t = 1, 2, \ldots, T$. The $\hat{\rho}$, $\hat{\rho}_\mu$, and $\hat{\rho}_\tau$ tests are based on the statistic $T(\hat{\beta} - 1)$, where $\hat{\beta}$ is the OLS estimator of β in (1), (2) and (3) respectively, while the $\hat{\tau}$, $\hat{\tau}_\mu$ and $\hat{\tau}_\tau$ tests are based on the t-statistics for the hypothesis $\beta = 1$ in the same three regressions. The former are coefficient tests, and the latter are t-ratio tests. Both types of test have time series extensions by the semiparametric correction method of Phillips (1987) and Phillips and Perron (1988). Only the t-ratio test is extended in the long autoregression method of Said and Dickey (1984).

Following the empirical work of Nelson and Plosser (1982), a common motivation for testing for a unit root is to test the hypothesis that a series is difference stationary against the alternative that it is trend stationary. That is, one wishes to test for a unit root in the presence of deterministic trend. Economists are especially interested in such tests because under the alternative hypothesis of stationarity time series exhibit trend reversion characteristics, whereas under the null they do not. Unfortunately, the Dickey-Fuller tests are not well designed for testing trend reversion. This is most obvious in the case of the tests based on regression (1), which does not allow for trend under either the null hypothesis or the alternative. Regression (2) allows for trend under the null, since when $\beta = 1$ the solution for y_t includes the deterministic trend term αt. However, the tests based on (2) are still not very suitable in the presence of trend, for two reasons. First, the distributions of the test statistics (even under the null hypothesis) depend on the nuisance parameter α (Evans and Savin (1984), Nankervis and

Savin (1985), Schmidt (1989), Guilkey and Schmidt (1989)). Second, regression (2) does not allow for trend under the alternative that $\beta < 1$, and therefore tests based on (2) are inconsistent against trend stationary alternatives (West (1987)). Finally, regression (3) allows for deterministic trend under both the null hypothesis and the alternative, but it does so in a clumsy way. Under the null hypothesis the model allows for linear trend even with $\delta = 0$ (and non-zero δ generates quadratic trend), whereas under the alternative linear trend requires non-zero δ. Thus the test allows for linear trend under the alternative by including in the regression a variable that is irrelevant under the null. This might be expected to result in some loss of power in finite samples, a conjecture supported by Monte Carlo evidence reported in this paper. Furthermore, the distributions of the test statistics based on (3) are now independent of α but they depend on δ under both the null and the alternative hypotheses.

This paper provides a new unit root test based on the alternative parameterization

$$(4) \qquad y_t = \psi + \xi t + X_t , \quad X_t = \beta X_{t-1} + \varepsilon_t ,$$

which has previously been considered by Bhargava (1986). Once again the unit root corresponds to $\beta = 1$. This parameterization allows for trend under both the null and the alternative, without introducing any parameters that are irrelevant under either. Indeed, one of the attractions of this parameterization is that the meaning of the nuisance parameters ψ and ξ does not depend on whether the unit root hypothesis is true: ψ represents level and ξ represents deterministic trend, whether $\beta = 1$ or not. As noted above, this is not so in the Dickey-Fuller parameterizations. For example, in (2) the parameter α represents trend when $\beta = 1$, but it determines level when $\beta < 1$ (since y is then stationary around the level $\alpha/(1-\beta)$). In (3), similarly, when $\beta = 1$, α reprsents linear trend and δ represents quadratic trend, whereas when $\beta < 1$, α determines level and δ represents linear trend.

The new test is extracted from the score or LM principle under the assumption that the ε_t are iid $N(0, \sigma_\varepsilon^2)$, but our asymptotics hold under more general assumptions about the errors. Two forms of the test (a coefficient test and a t-test) are derived. A valuable property of these tests is that their distributions under both the null and alternative hypotheses are independent of the nuisance parameters ψ, ξ and σ_ε. Thus the difficulties for the Dickey-Fuller tests caused by nuisance parameters representing deterministic trend do not arise here.

The plan of the paper is as follows. Section 2 derives the new test statistics and compares them to the Dickey-Fuller $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ statistics. Section 3 gives results on the finite sample distributions of the statistics under the assumption of iid errors. Section 4 provides the asymptotic distribution of the statistics under more general error assumptions, and gives extensions along the lines of Phillips (1987) and Phillips and Perron (1988) that are asymptotically robust to error autocorrelation and heteroskedasticity. Section 5 extends the tests to the case of deterministic trend that follows a higher order polynomial in time. Section 6 provides some Monte Carlo evidence on the power of the tests. Finally, Section 7 contains our conclusions.

2. NEW UNIT ROOT TESTS

We begin with the model as given in (4) above, where the errors ε_t are assumed to be iid $N(0, \sigma_\varepsilon^2)$ and where the initial condition X_0 is taken as fixed. We wish to derive the LM test of the hypothesis $\beta = 1$ in this model. The derivation is given in Appendix 1, and here we will give only a brief summary. The restricted MLE's (that is, the MLE's when we impose $\beta = 1$) of ξ and $\psi_X = \psi + X_0$ are as follows:

(5) $$\tilde{\xi} = \text{mean } \Delta y = (y_T - y_1)/(T-1)$$

(6) $$\tilde{\psi}_X = = y_1 - \tilde{\xi} .$$

(The parameters ψ and X_0 are identified separately under the alternative hypothesis, but not under the null hypothesis that β equals one.) Note that, as expected, the estimate of ξ comes from estimation of (4) in differences. Now define the "residuals"

(7) $$\tilde{S}_{t-1} = y_{t-1} - \tilde{\psi}_X - \tilde{\xi}(t-1) , \quad t = 2,...,T .$$

These are the (lagged) residuals from the model (4) in levels, but where the parameters have been estimated from the model in differences. The score vector evaluated at the restricted MLE's is proportional to

(8) $$\sum_{t=2}^{T} (\Delta y_t - \tilde{\xi})\tilde{S}_{t-1} .$$

This is the numerator of the estimated regression coefficient of $\tilde{S}_{t-1}$ in the regression

(9) $$\Delta y_t = \text{intercept} + \phi\tilde{S}_{t-1} + \text{error} \quad (t = 2, ..., T) .$$

Denote the least squares estimate of ϕ by $\tilde{\phi}$. We then define the test statistics

$$\tilde{\rho} = T\tilde{\phi} \tag{10}$$

$$\tilde{\tau} = \text{usual t-statistic for } \phi = 0 \text{ in (9).} \tag{11}$$

It is instructive to compare these statistics to the Dickey-Fuller statistics $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ based on regression (3). This is a regression of y on intercept, time trend and lagged y ; equivalently, it is a regression of Δy on the same variables. By the standard algebra of least-squares regression, it follows that this is in turn equivalent to the regression

$$\Delta y_t = \text{intercept} + \rho\hat{S}_{t-1} + \text{error} \quad (t = 2, ..., T), \tag{12}$$

where $\hat{S}_{t-1}$ is the residual from an *ordinary least squares* regression of y_{t-1} on an intercept and time trend. We then have $\hat{\rho}_\tau$ as the estimated coefficient of $\hat{S}_{t-1}$ in (12), and $\hat{\tau}_\tau$ as the t-statistic for the hypothesis $\rho = 0$. Comparing (12) to (9), the only difference between the new tests $\tilde{\rho}$ and $\tilde{\tau}$ and the Dickey-Fuller tests $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ is the nature of the residual upon which Δy_t is regressed. Both $\tilde{S}_{t-1}$ in (9) and $\hat{S}_{t-1}$ in (12) are residuals in the levels equation for y_{t-1}, but the parameters used to calculate the residuals are estimated differently: the parameters used to calculate $\tilde{S}_{t-1}$ are estimated from the model in differences, while the parameters used to calculate $\hat{S}_{t-1}$ are estimated from the model in levels. Given that y is I(1) under the null hypothesis, the regression of y_{t-1} on intercept and time *in levels* is spurious and thus the residual process $\hat{S}_{t-1}$ is I(1). $\tilde{S}_{t-1}$ is also a detrended I(1) process but the trend is estimated using the information that $\beta = 1$.

3. FINITE SAMPLE DISTRIBUTION THEORY

The finite sample distribution of the test statistics $\tilde{\rho}$ and $\tilde{\tau}$ are complicated and will be tabulated by simulation. However, we first demonstrate the simple but important fact that the distribution of the test statistics under the null hypothesis is independent of the nuisance parameters ψ, ξ, X_0 and σ_ε. Thus the distributions of the test statistics under the null hypothesis depend only on the sample size (T).

To prove this result, we define the partial sum

$$S_t = \sum_{j=1}^{t} \varepsilon_j \tag{13}$$

and note the solution for y_t under the null hypothesis:

$$y_t = \psi_X + \xi t + S_t. \tag{14}$$

The restricted MLE's given by (5) and (6) satisfy

(15) $$\hat{\xi} = \xi + \bar{\varepsilon}$$

(16) $$\tilde{\psi}_X = y_1 - \tilde{\xi} = \psi_X + (\varepsilon_1 - \bar{\varepsilon}) \; .$$

Inserting (13)–(16) into the definition (7) of $\tilde{S}_{t-1}$, we obtain

(17) $$\tilde{S}_{t-1} = S_{t-1} - (t-2)\bar{\varepsilon} - \varepsilon_1 \; .$$

This does not depend on ξ , ψ or X_0 . Now consider the regression of $\Delta y_t = \xi + \varepsilon_t$ on intercept and $\tilde{S}_{t-1}$. The intercept absorbs ξ and both the estimated coefficient of $\tilde{S}_{t-1}$ (i.e., $\tilde{\phi}$) and its estimated standard error are independent of ξ (as well as ψ and X_0). Finally, the scale factor σ_ε also cancels out of all expressions for $\tilde{\rho}$ and $\tilde{\tau}$, so that their distributions are independent of σ_ε as well as the other nuisance parameters.

Bhargava (1986) has previously derived a test of the unit root hypothesis based on our model (4) using the theory of invariance to ensure that the test statistic's distribution does not depend on nuisance parameters. King (1981) and Dufour and King (1989) have similarly used the theory of invariance to yield test statistics independent of nuisance parameters. Their tests are designed to be point optimal and hence of rather different form than this paper's tests or Bhargava's.

Critical values for the test statistics $\tilde{\rho}$ and $\tilde{\tau}$ are given in Tables 1A. These are calculated by a direct simulation using 50000 replications. Random deviates were generated by the routines GASDEV and RAN3 of Press, Flannery, Teukolsky and Vetterling (1986); more detail on this random number generation scheme can be found in Guilkey and Schmidt (1989). We note in passing that the lower tail critical values are smaller in absolute value than the corresponding lower tail critical values for the Dickey-Fuller $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ tests.

Under the alternative hypothesis that β is not equal to one, the distributions of $\tilde{\rho}$ and $\tilde{\tau}$ are independent of ψ , ξ and σ_ε , but they depend on $X_0^* = X_0/\sigma_\varepsilon$. (They also depend on β and T , of course.) The proof of this assertion follows the same lines as above, but is somewhat more involved, and is given in Appendix 2. Some simulation evidence on the powers of the tests will be given in Section 5.

4. ASYMPTOTICS

Following Phillips (1987) and Phillips and Perron (1988), we can relax the assumption that the ε_t are iid by considering the asymptotic distribution of the test statistics and correcting for serial dependence. We assume the same regularity conditions as Phillips and Perron (1988, p. 336); these put some limits on the degree of heterogeneity and autocorrelation allowed in the ε sequence but are otherwise fairly general. We define the two nuisance parameters

$$(18) \qquad \sigma_\varepsilon^2 = \lim_{T\to\infty} T^{-1}E\left(\sum_{t=1}^{T} \varepsilon_t^2\right)$$

$$(19) \qquad \sigma^2 = \lim_{T\to\infty} T^{-1}E(S_T^2)$$

(where S_T is the partial sum as defined in (13) above). We also define the ratio $\omega^2 = \sigma_\varepsilon^2/\sigma^2$. Then as shown in Appendix 3 the following limit theory holds for the test statistics $\tilde{\rho}$ and $\tilde{\tau}$:

$$(20) \qquad \tilde{\rho} \to -\left[2\int_0^1 \underline{V}^2\right]^{-1}\omega^2$$

$$(21) \qquad \tilde{\tau} \to -(1/2)\left[\int_0^1 \underline{V}^2\right]^{-1/2}\omega .$$

Here $V(r)$ is a standard Brownian bridge on the interval [0,1] and $\underline{V}(r)$ is the demeaned Brownian bridge

$$(22) \qquad \underline{V}(r) = V(r) - \int_0^1 V(r)dr .$$

The symbol " $\to$ " in (20) and (21) signifies weak convergence of the associated probability measures.

It is easy to relate the limit formulae to those of earlier work. Consider the integral

$$
\begin{aligned}
(23) \qquad \int_0^1 VdV &= \sigma^2\int_0^1 (W(r) - rW(1))(dW(r) - drW(1)) \\
&= \sigma^2\{\int_0^1 WdW - (\int_0^1 rdW)W(1) - (\int_0^1 Wdr)W(1) + (\int_0^1 rdr)W(1)^2\} \\
&= \sigma^2\{\int_0^1 WdW - W(1)^2 + (\int_0^1 Wdr)W(1) - (\int_0^1 Wdr)W(1) + (1/2)W(1)^2\} \\
&= \sigma^2\{\int_0^1 WdW - (1/2)W(1)^2\} \\
&= (1/2)\sigma^2\{(W(1)^2 - 1) - W(1)^2\} \\
&= -\sigma^2/2 .
\end{aligned}
$$

Also note that

$$(24) \qquad \int_0^1 \underline{V} dV = \int_0^1 (V(r) - \int_0^1 V) dV$$
$$= \int_0^1 V dV - (\int_0^1 V) V(1)$$
$$= \int_0^1 V dV$$

since $V(1) = 0$. Thus, in place of (20) we have

$$(25) \qquad \tilde{\rho} \to \{\sigma^2 \int_0^1 \underline{V} dV + (1/2)(\sigma^2 - \sigma_\varepsilon^2)\} / \{\sigma^2 \int_0^1 \underline{V}^2 dr\}$$

which is entirely analogous to the limit representations given in Phillips (1987) and Ouliaris, Park and Phillips (1988). The difference is that the standard Brownian motion $W(r)$ that appears in the formulae of the latter papers is replaced by the Brownian bridge $V(r)$. The reason for this difference is easy to explain. The new test statistics $\tilde{\rho}$ and $\tilde{\tau}$ are based on the least squares regression coefficient $\tilde{\varphi}$ in (9). This is the coefficient of the detrended process $\tilde{S}_{t-1}$ given in (17). As shown in Appendix 3, standardized by $T^{-1/2}$ this process converges weakly to a Brownian bridge. This is as we would expect because

$$(26) \qquad T^{-1/2} \tilde{S}_T = -T^{-1/2} (\varepsilon_1 - \bar{\varepsilon}) \to_p 0 ,$$

and as $T \to \infty$, $T^{-1/2} \tilde{S}_T$ is itself effectively tied down to the origin. Such behavior does not arise in the case of tests that are based on the coefficients of y_{t-1} in the Dickey-Fuller regressions (1)-(3).

The asymptotic formulae (20) and (21) require only simple corrections to remove the effects of dependent and heterogeneous errors. Multiplying $\tilde{\rho}$ by a consistent estimate of $1/\omega^2 = (\sigma^2/\sigma_\varepsilon^2)$ yields a corrected test statistic whose asymptotic distribution is identical to the asymptotic distribution that $\tilde{\rho}$ would have under iid errors, so that the critical values given in Table 1 are asymptotically correct. Similarly, multiplying $\tilde{\tau}$ by a consistent estimate of $1/\omega = (\sigma/\sigma_\varepsilon)$ yields a corrected test statistic for which the critical values in Table 1 are asymptotically correct. These corrections are very simple in comparison with the corrections given in Phillips (1987) and Phillips and Perron (1989) for the Dickey-Fuller tests.

Estimation of σ^2 and σ_ε^2 can be performed along the lines suggested in Phillips (1987), Phillips and Ouliaris (1987), and Phillips and Perron (1988). In particular, the arguments given in Phillips and Ouliaris (1987) apply and the consistency of both tests requires that σ^2 and σ_ε^2 be estimated from regression residuals rather than first differences. Thus, let $\hat{\varepsilon}_t$ be the residuals from a least squares regression on (3). Then by arguments analogous to those of Theorem 4.2 in Phillips (1987) we find that the following estimates are consistent for the variance parameters under the null:

(27) $$s^2 = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t^2 \to_p \sigma_\varepsilon^2$$

and

(28) $$s^2(\ell) = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t^2 + 2T^{-1}\sum_{s=1}^{\ell}\sum_{t=s+1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_{t-s} \to_p \sigma^2 .$$

In the case of $s^2(\ell)$ we require that the lag truncation parameter $\ell \to \infty$ as $T \to \infty$. The rate $\ell = o(T^{1/2})$ will usually be satisfactory, as for the case of stationary sequences ε_t (see Andrews (1989)).

With these estimates we construct $\hat{\omega}^2 = s^2/s^2(\ell)$ and the test statistics

(29) $$Z(\rho) = \tilde{\rho}/\hat{\omega}^2 , \; Z(\tau) = \tilde{\tau}/\hat{\omega} .$$

Under the null hypothesis we have

(30) $$Z(\rho) \to -\left(2\int_0^1 \underline{V}^2\right)^{-1} , \; Z(\tau) \to -(1/2)\left(\int_0^1 \underline{V}^2\right)^{-1/2} .$$

These limit distributions are free of nuisance parameters and they are negative almost surely. Under the alternative hypothesis that $|\beta| < 1$ we find that

$$Z(\rho) = O_p(T) , \; Z(\tau) = O_p(T^{1/2})$$

as in Theorem 5.1 of Phillips and Ouliaris (1987). Thus, the statistics diverge under the alternative and the two tests are consistent, but at different rates as $T \to \infty$.

REMARK. As observed above, the construction of a consistent test requires the use of regression residuals rather than first differences. This means that a regression such as (3) is needed, at least at this stage, to remove nuisance parameters. Although this does not cause any loss in asymptotic local power, it seems likely that it will have finite sample effects in terms of some size distortion and power loss.

5. EXTENSIONS TO HIGHER ORDER POLYNOMIAL TRENDS

We now wish to replace the linear deterministic trend in (4) with a higher order polynomial trend. To do so, we first consider the more general model

(31) $$y_t = \alpha + Z_t\delta + X_t \, , \; X_t = \beta X_{t-1} + \varepsilon_t \, ,$$

where Z_t is at this point a general row vector of explanatory variables. The null hypothesis is $\beta = 1$, as before, and to construct the LM statistic we need to consider the differenced version of (31), namely

(32) $$\Delta y_t = \Delta Z_t\delta + u_t$$

(where $u_t = \varepsilon_t$ under the null hypothesis). Define the restricted MLE's: $\tilde{\delta}$ = OLS estimate of δ from (32), $\tilde{\psi}_X = y_1 - Z_1\tilde{\delta}$; and define

(33) $$\tilde{S}_t = y_t - \tilde{\psi}_X - Z_t\tilde{\delta} \, .$$

Finally, run the regression

(34) $$\Delta y_t = \Delta Z_t\gamma + \phi\tilde{S}_{t-1} + \text{error} \, .$$

We again define $\tilde{\rho} = T\tilde{\phi}$, where $\tilde{\phi}$ is the least squares estimate of ϕ in (34), and $\tilde{\tau}$ = usual t statistic for the hypothesis $\phi = 0$. Essentially the same algebra as in Appendix 1 shows that $\tilde{\tau}$ is the LM statistic for the hypothesis $\beta = 1$ when the ε_t are iid normal.

We are specifically interested in the case $Z_t = [t, t^2, ..., t^p]$, so that the model allows a p^{th} order polynomial time trend. In this case the differenced model is equivalent to a $(p-1)^{th}$ order time trend, and it is convenient to rewrite the above general expressions as follows. The model is

(35) $$y_t = \sum_{j=0}^{p} a_j t^j + X_t \, , \; X_t = \beta X_{t-1} + \varepsilon_t \, .$$

The differenced model can be written as

(36) $$\Delta y_t = \sum_{j=0}^{p-1} b_j t^j + u_t \, .$$

Define $\tilde{u}_1 = 0$ and $\tilde{u}_t$ = OLS residual from (36), $t = 2, ..., T$. Then it is easy to show that $\tilde{S}$ as defined in (33) can be calculated as the partial sum of the $\tilde{u}$:

(37) $$\tilde{S}_t = \sum_{k=1}^{t} \tilde{u}_k ,$$

and the regression (34) that defines the test statistics is simply

(38) $$\Delta y_t = \sum_{j=0}^{p-1} c_j t^j + \phi \tilde{S}_{t-1} + \text{error} .$$

From the limit theory given in Appendix 3(ii) we have

(39) $$\tilde{\rho} \to -\left(2\int_0^1 \underline{V}_p^2\right)^{-1}\omega^2 , \quad \tilde{\tau} \to -(1/2)\left(\int_0^1 \underline{V}_p^2\right)^{-1/2}\omega .$$

In the above formulae $\underline{V}_p(r)$ is a detrended p-level Brownian bridge; i.e.

(40) $$\underline{V}_p(r) = V_p(r) - \sum_{j=0}^{p-1} \hat{\alpha}_j r^j$$

and

(41) $$\hat{\alpha} = \operatorname{argmin}_\alpha \int_0^1 \left(V_p(r) - \sum_{j=0}^{p-1} \alpha_j r^j\right)^2 dr .$$

Here $V_p(r)$ is a Gaussian process which can be defined in terms of standard Brownian motion $W(r)$ as follows:

$$V_p(r) = W(r) - \left(\int_0^1 dW(s) g(s)'\right) Q^{-1} q(r) ,$$

where $g(s)' = (1, s, ..., s^{p-1})$, Q is $p \times p$ with $(i,j)^{th}$ element $q_{ij} = 1/(i+j-1)$ and $q(r)$ is $p \times 1$ with i^{th} element r^i/i. As noted in Appendix 3(ii), $V_p(r)$ is tied down in the [0,1] interval with

$$V_p(0) = V_p(1) = 0$$

just like a Brownian bridge. In fact when $p = 1$ we have

$$V_1(r) = V(r) ,$$

a simple Brownian bridge; and when $p = 0$ we have $V_0(r) = W(r)$, a standard Brownian motion. Also as shown in (A3.8) of Appendix 3(ii), $V_p(r)$ is the weak limit of a standardized partial sum of detrended innovations. Thus, writing $\underline{\varepsilon}_t$ as the residual in the regression of ε_t on a time trend of order $p-1$, viz.

$$\varepsilon_t = \sum_{j=0}^{p-1} \hat{\delta}_j t^j + \underline{\varepsilon}_t ,$$

we have

$$T^{-1/2} \sum_{t=1}^{[Tr]} \underline{\varepsilon}_t \rightarrow V_p(r) .$$

The detrended process $\underline{V}_p(r)$ is most easily interpreted as a Hilbert projection in $L_2[0,1]$ of the process $V_p(r)$ on the orthogonal complement of the space spanned by the trend functions $\{r^j; j = 0, 1, ..., p-1\}$.

The nuisance parameter ω , or variance ratio ω^2 , that appears in the limit formulae (39) may be eliminated by transformation as discussed in the preceding section, leading to test statistics $Z(\rho)$ and $Z(\tau)$.

Tables 1B-1D give the critical values for the test statistics $\tilde{\rho}$ and $\tilde{\tau}$, for $p = 2, 3$ and 4 (where p is the order of the deterministic polynomial trend in the model (35)). These are calculated as described in Section 3.

6. POWER OF THE TESTS

In this section we perform some Monte Carlo experiments to compare the power of the tests proposed in this paper with the power of the Dickey-Fuller tests. These tests rely on different parameterizations, and so we begin by considering how the parameters in one parameterization relate to those in another.

We consider the three parameterizations:

(4) $$y_t = \psi + \xi t + X_t , \; X_t = \beta X_{t-1} + \varepsilon_t$$

(3) $$y_t = \alpha + \beta y_{t-1} + \delta t + \varepsilon_t$$

(42) $$y_t^* = \gamma + \beta y_{t-1}^* + \Phi t + \varepsilon_t^* .$$

Here (3) and (4) are as previously considered. Equation (42) gives a standardization of (3) used previously by Nankervis and Savin (1985), DeJong *et al.* (1988) and others. In (42) we have

(43) $$y_t^* = (y_t - y_0)/\sigma_\varepsilon , \; \varepsilon_t^* = \varepsilon_t/\sigma_\varepsilon ,$$
$$\gamma = [\alpha + y_0(\beta-1)]/\sigma_\varepsilon , \; \Phi = \delta/\sigma_\varepsilon .$$

It is important to note that the parameterizations (3), (4) and (42) are *completely equivalent* when β is not equal to one. From any one of these equations we can derive the other two, and there is a clear mapping from any set of parameters to any other set. For example, if we start with equation (4), subtract βy_{t-1} from both sides, and do some algebra, we obtain

(44) $$y_t = [\psi(1-\beta) + \xi\beta] + \xi(1-\beta)t + \beta y_{t-1} + \varepsilon_t ,$$

so that the parameters of equation (3) in terms of the parameters of equation (4) are given by

(45) $$\alpha = [\psi(1-\beta) + \xi\beta] , \quad \delta = \xi(1-\beta) .$$

Similarly, substituting these expressions into equation (43) we obtain

(46) $$\gamma = \xi^*\beta + X_0^*(\beta-1) , \quad \Phi = \xi^*(1-\beta) ,$$

where as before $X_0^* = X_0/\sigma_\varepsilon$ and $\xi^* = \xi/\sigma_\varepsilon$. The complete set of relationships between the parameters in (3), (4) and (42) is given in Appendix 4.

It is interesting to note in passing the only substantive difference between our parameterization (4) and the Dickey-Fuller parameterization (3): equation (4) implies equation (3) but it also implies that the parameter δ in (3) should equal zero when β equals one. (This is so because $\delta = \xi(1-\beta)$, as given in (45).) We earlier argued that (3) is an awkward parameterization precisely because it allows for trend under the alternative hypothesis by including a variable that is irrelevant under the null. On the other hand, starting with the parameterization (4) forces the coefficient of this variable to equal zero under the null, as it should if trend is linear.

In the Monte Carlo experiment we consider the performance of six tests: the new tests $\tilde{\rho}$ and $\tilde{\tau}$, and the Dickey-Fuller tests $\hat{\rho}_\mu$, $\hat{\tau}_\mu$, $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$. The focus of the experiments will be on power against trend stationary alternatives, and we do not expect the $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests to have much power against such alternatives. These tests would be expected to be more powerful than the other four tests against level stationary alternatives, and it is interesting to see how much trend or how large a sample size it takes before the $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests are dominated by the other tests. However, the main focus of the experiments is to compare the power of the new tests $\tilde{\rho}$ and $\tilde{\tau}$ with the power of the Dickey-Fuller $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ tests.

The results of our experiments are given in Tables 2-4. The results were generated by a simulation using 20000 replications; the random number generator was described in Section 3. The tables give percentages of rejections for 5% lower tail tests; other significance levels would tell essentially the same story.

The parameters that are relevant are sample size (T) , β , standardized trend (ξ^*) and standardized intitial condition (X_0^*) , though we also present the corresponding values of the Nankervis-Savin parameters (γ

and Φ) for ease of comparison to other studies. An important advantage of the parameterization of this paper is that it simplifies the comparisons of the powers of the tests of this paper and the Dickey-Fuller tests. Under the null hypothesis that $\beta = 1$, the tests $\tilde{\rho}$, $\tilde{\tau}$, $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ have distributions that are independent of the nuisance parameters ψ, ξ^*, X_0^* and σ_ε, while the distributions of $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ depend on ξ^* but are independent of ψ, X_0^* and σ_ε^2. When β is not equal to one, we proved (Appendix 2) that the distributions of the tests $\tilde{\rho}$ and $\tilde{\tau}$ depend only on T, β and X_0^*; for given T, β and X_0^*, they are independent of ψ, ξ^* and σ_ε. Interestingly, though we have not proved that it is so, it became clear in our simulations that the same is true of the Dickey-Fuller tests $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$. This is somewhat surprising because, for a given T and β, the power of the Dickey-Fuller tests is known to depend on both γ and Φ in parameterization (42); for example, see DeJong *et al.* (1988). However, the apparent dependence of the power of these tests on two parameters is just a symptom of using an inconvenient parameterization. From Appendix 4 we have

(47) $$X_0^* = [\beta\Phi - \gamma(1-\beta)]/(1-\beta)^2 ,$$

and values of γ and Φ that imply the same value of X_0^* also imply the same power for the $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ tests. (DeJong *et al.* (1988) have proved a special case of this result. They show that power is constant for values of γ and Φ that satisfy $\gamma = \Phi\beta/(1-\beta)$. From (47), these values of γ and Φ correspond to $X_0^* = 0$.) Thus the parameterization of this paper is convenient in terms of analyzing the properties of existing tests, as well as in terms of yielding new tests.

Our first experiment, called Experiment 1 in Table 2, studies the size of the various tests under the null hypothesis. We set $T = 100$, $\beta = 1$, $X_0^* = 0$ (its value is irrelevant anyway), and varied ξ^*. Specifically, we considered values 0, .02, .05, .10, .20 and .50 for ξ^*. All tests except $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ should have size equal to the nominal critical value (.05), and this is so apart from randomness. (With 20000 replications, a 95% confidence interval around .05 is approximately [.047, .053].) When $\xi^* = 0$ the $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests should also have size equal to the nominal critical value, and they do apart from randomness. However, the size of the $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests should decrease to zero as ξ^* increases, for fixed T, or as T increases for fixed ξ^*. For $T = 100$, we can see in Table 2 that the size of these tests does go to zero as ξ^* increases; it is nearly zero for ξ^* as large as .50. Results presented in Table 4 for $T = 200$ and $T = 500$ (Experiments 5A and 5B) confirm that the size distortion of the $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests is larger for larger sample sizes; the larger T, the smaller the value of ξ^* required to produce substantial size distortions, and conversely.

The values of ξ^* considered here are empirically relevant. Recall that ξ^* is standardized trend, equal to ξ/σ_ε. We can estimate ξ by $\tilde{\xi}$ = mean $\Delta y = (y_T - y_1)/(T-1)$; this is the MLE subject to $\beta = 1$, but it is a consistent estimate of ξ even if β is not equal to one. Similarly, imposing the unit root, the MLE of σ_ε^2 is the empirical variance of the Δy's, and this is a consistent estimate even if y is stationary. Thus a consistent estimate of ξ^* is just the mean of the Δy's divided by the standard deviation of the Δy's . Schmidt (1989) provides values of this measure (which he calls "standardized drift") for the Nelson-Plosser data, and values in the range [.2, .5] are the norm. Smaller values of ξ^* might be expected in higher-frequency data, or in financial data, but it is nevertheless clear that the size distortions for the $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests are potentially very serious for values of sample size and trend encountered in economic data.

Experiment 2, reported in Table 2, explores the effect of the intial condition X_0^* on the power of the various tests. Consider first experiment 2A, which sets $T = 100$, $\beta = .90$, and $\xi^* = 0$. (The value of ξ^* matters only for $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$.) We consider values of X_0^* ranging from -10 to 10. Some interesting patterns emerge. Apart from randomness, the effects of X_0^* on power are symmetric about $X_0^* = 0$. As X_0^* increases in absolute value, the power of the $\hat{\tau}_\mu$ and $\hat{\tau}_\tau$ tests increases monotonically while the power of the other four tests decreases. The $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests are more powerful than the other four tests, because experiment 2A has $\xi^* = 0$ (no trend). A more interesting comparison is between the powers of the Dickey-Fuller $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ tests and the $\tilde{\rho}$ and $\tilde{\tau}$ tests. This comparison depends only on the value of X_0^*. For $|X_0^*| \leq 2$, the new tests are more powerful than the Dickey-Fuller $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ tests, while the opposite is true for $|X_0^*| \geq 5$. The standardized initial condition is a hard parameter to get a feel for; it is not even identified under the null hypothesis. However, a value of X_0 that is five standard deviations (of ε) away from zero is rather unlikely. Our summary of these results is that the new $\tilde{\rho}$ and $\tilde{\tau}$ tests dominate the Dickey-Fuller $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ tests except for values of X_0^* that are unreasonably large, but the reader is of course free to draw his own conclusions.

We note also that the $\tilde{\rho}$ and $\tilde{\tau}$ tests have very similar power. The power of the $\hat{\rho}_\tau$ test is higher than the power of the $\hat{\tau}_\tau$ test (the conventional wisdom), except when X_0^* is unreasonably large, in which case this ranking reverses.

Experiments 2B and 2C also vary X_0^*, holding constant $T = 100$ and $\beta = .90$, but they use nonzero values of ξ^*, namely $\xi^* = .02$ and .05. As noted above, the value of ξ^* affects only the powers of the

$\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests. A value of about $\xi^* = .05$ is sufficient to make the $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests less powerful than the other four tests.

Experiment 3, made up of parts 3A-3D, is reported in Table 3. This experiment systematically varies the trend parameter ξ^*, for various values of X_0^* (0, -2 and -5) and β (.90 and .95). Once again the results are consistent with the conclusion that $\xi^* > .05$ is sufficient to make the power of the $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests less than the power of the tests that explicitly allow for trend. It should be stressed that, empirically, this is a very small value of ξ^*, and these results argue strongly that it is a mistake to apply the $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests to data with noticable trend.

Experiment 4, made up of parts 4A - 4C, is also reported in Table 3. This experiment varies β over the range from one to .80, for $X_0^* = 0$ and -2, and for $\xi^* = 0$ and .05. There are no surprises in the results. Over this range of X_0^*, we expect the $\tilde{\rho}$ and $\tilde{\tau}$ tests to be more powerful than the $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ tests, and they are. We expect the $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests to be more powerful than the $\tilde{\rho}$ and $\tilde{\tau}$ tests for $\xi^* = 0$, and less powerful for $\xi^* = .05$, and they are. A result worth noting, however, is that the loss in power from using the $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ test in the presence of non-zero trend is larger when β is farther from unity.

Table 4 gives the results of some experiments done for $T = 200$ and 500. We have already commented on the results of Experiment 5, which considers the size of the tests under the null hypothesis. The other experiments consider power under the alternative: Experiment 6 varies X_0^* for $T = 200$, $\xi^* = 0$ and $\beta = .90$; Experiment 7 varies ξ^* for $T = 200$ and 500, $X_0^* = 0$ and $\beta = .90$ and .95; and Experiment 8 varies β for $T = 200$ and 500 and $X_0^* = \xi^* = 0$. The results are in line with the previous discussion and so their detailed analysis is left to the reader.

Two clear conclusions emerge from our experiments, and bear repeating. First, the Dickey-Fuller $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ tests are more powerful than the other tests we considered against level-stationary alternatives, but they have little power against trend-stationary alternatives and therefore should not be used in the presence of discernible trend. Second, the $\tilde{\rho}$ and $\tilde{\tau}$ tests proposed in this paper are about equally powerful, and they are more powerful than the Dickey-Fuller $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ tests except when the initial condition term (X_0^*) is unreasonably large.

7. CONCLUDING REMARKS

This paper has proposed two new tests of the unit root hypothesis. They are based on a different parameterization than the Dickey-Fuller tests. The choice of a parameterization is to some extent a matter of taste. However, the parameterization we use has two important advantages. First, the meaning of the parameters governing level and trend is independent of whether or not the unit root hypothesis is true. Second, the analysis of the distributional properties of both the new tests and the Dickey-Fuller $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ tests is simplified.

Although the new tests were not derived on the basis of considerations of invariance, they do have the property that their distributions under the null hypothesis are independent of the nuisance parameters reflecting level, trend and variance. Because they were derived as LM tests, they should be expected to have good local power properties. Our simulation results indicate that the comparison of their power to the power of the Dickey-Fuller $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ tests hinges on an initial conditions parameter (X_0^*), and that they should be more powerful than the Dickey-Fuller tests except for unreasonably large values of this parameter.

The LM test procedure used in this paper can be used in other settings, with appropriate modifications to the distributional theory. Essentially we have tested the hypothesis that the error in a regression has a unit root, where the regressors form a deterministic trend. This could be extended to accommodate stochastic trends. Thus, if the regressor is an I(1) variable rather than a deterministic trend, the unit root test becomes a cointegration test. In both cases the change in the nature of the regressors will change the asymptotic distribution of the test statistics, but in ways that are presumably straightforward.

Another reasonable extension of this paper would be to follow Said and Dickey (1984) in developing the asymptotic theory for "augmented" versions of the tests proposed here. It is intuitively reasonable that, as in the case of the Dickey-Fuller tests, we can correct (asymptotically) for the effects of autocorrelated errors by the inclusion of lagged values of Δy in the regression (9) that generates the test statistics. The necessary asymptotic theory is not worked out here but the limit theory will correspond to that given in Sections 4 and 5 for the $Z(\tau)$ statistic.

APPENDIX 1

DERIVATION OF THE LM TEST

We begin with the model as given in equation (4) of the main text. It implies

$$\text{(A1.1)} \qquad \begin{aligned} y_1 &= \psi + \beta X_0 + \xi + \varepsilon_1 \\ y_t &= \beta y_{t-1} + \psi(1-\beta) + \xi(t + \beta - t\beta) + \varepsilon_t \quad t = 2, \ldots, T \,. \end{aligned}$$

We assume that the ε_t $(t = 1, \ldots, T)$ are iid $N(0, \sigma^2)$ and we treat the initial condition X_0 as fixed. Since the Jacobian from $(\varepsilon_1, \ldots, \varepsilon_T)$ to $(y_1, \ldots, y_T)$ is unity, we obtain the log likelihood

$$\text{(A1.2)} \qquad \ln L = \text{constant} - \frac{T}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \text{SSE} \,,$$

where

$$\text{(A1.3)} \qquad \text{SSE} = (y_1 - \psi - \beta X_0 - \xi)^2 + \sum_{t=2}^{T} \left[(y_t - \beta y_{t-1}) - \psi(1-\beta) - \xi(t+\beta-t\beta)\right]^2 .$$

At the maximum $\tilde{\sigma}^2 = \text{SSE}/T$ and so the concentrated log likelihood is

$$\text{(A1.4)} \qquad \ln L^* = \text{const.} - \frac{T}{2} \ln(\text{SSE}/T) \,.$$

To derive the MLE's subject to the restriction $\beta = 1$, we note that, when $\beta = 1$, SSE simplifies to

$$\text{(A1.5)} \qquad \text{SSE}_R = [y_1 - \psi_X - \xi]^2 + \sum_{t=2}^{T} (\Delta y - \xi)^2 \,, \quad \psi_X = \psi + X_0$$

and this is minimized by the restricted MLE's

$$\text{(A1.6)} \qquad \begin{aligned} \tilde{\xi} &= \overline{\Delta y} = (y_T - y_1)/(T-1) \\ \tilde{\psi}_X &= y_1 - \tilde{\xi} = (Ty_1 - y_T)/(T-1) \,, \end{aligned}$$

as given in equations (5) and (6) of the main text.

To derive the LM test we need to calculate the efficient score, evaluated at $\tilde{\beta} = 1$:

$$\text{(A1.7)} \qquad \frac{\partial \ln L^*(\sim)}{\beta} = \frac{-1}{2\sigma^2} \frac{\partial \text{SSE}}{\partial \beta} .$$

It is easy to calculate that

(A1.8) $$\frac{\partial SSE}{\partial \beta} = -2X_0(y_1 - \psi - \beta X_0 - \xi) - 2\sum_{t=2}^{T}[y_{t-1} - \psi - \xi(t-1)]$$

$$\cdot[(y_t - \beta y_{t-1}) - \psi(1-\beta) - \xi(t+\beta-t\beta)] .$$

Define

(A1.9) $$\tilde{S}_{t-1} = y_{t-1} - \tilde{\psi}_X - \tilde{\xi}(t-1) ,$$

as in equation (7) of the main text. Then (A1.8), evaluated at the restricted MLE's, becomes

(A1.10) $$\frac{\partial SSE}{\partial \beta} = -2\sum_{t=2}^{T}(\Delta y_t - \tilde{\xi})\tilde{S}_{t-1} ,$$

and the score becomes

(A1.11) $$\frac{\partial \ln L^*(\sim)}{\partial \beta} = \frac{1}{\tilde{\sigma}^2}\sum_{t=2}^{T}(\Delta y_t - \tilde{\xi})\tilde{S}_{t-1} = \frac{1}{\tilde{\sigma}^2}\sum_{t=2}^{T}(\Delta y_t - \tilde{\xi})(\tilde{S}_{t-1} - \bar{\tilde{S}}) .$$

(The last equality holds since $(\Delta y_t - \tilde{\xi})$ sums to zero. Therefore, without loss of generality, we can treat the $\tilde{S}_{t-1}$ as centered at zero.)

The term $\Sigma(\Delta y_t - \tilde{\xi})\tilde{S}_{t-1}$ in (A1.11) is the numerator of the estimated coefficient (say $\tilde{\phi}$) in the regression

(A1.12) $$\Delta y_t = \text{intercept} + \phi\tilde{S}_{t-1} + \text{error} ,$$

as given also by equation (9) of the main text.

To construct the LM test, we also need the information matrix. We calculate

(A1.13) $$\frac{\partial^2 \ln L^*}{\partial \beta^2} = \frac{-1}{\sigma^2}\{X_0^2 + \sum_{t=2}^{T}[y_{t-1} - \psi - \xi(t-1)]^2\} .$$

Evaluating (A1.13) at the restricted MLE's, and ignoring X_0 (which will be negligible asymptotically), we have

(A1.14) $$\frac{\partial^2 \ln L^*(\sim)}{\partial \beta^2} = \frac{1}{\tilde{\sigma}^2}\sum_{t=2}^{T}\tilde{S}_{t-1}^2 .$$

We show below that the information matrix is asymptotically block diagonal between β and $[\psi_X, \xi]$. Therefore the LM statistic becomes

$$\text{(A1.15)} \qquad LM = \left(\frac{\partial \ln L^*(\sim)}{\partial \beta}\right)^2 \Big/ \frac{\partial^2 \ln L^*(\sim)}{\partial \beta^2}$$

and using (A1.11) and (A1.14) we have

$$\text{(A1.16)} \qquad LM = \frac{\left[\sum_{t=2}^{T} (\Delta y_t - \tilde{\xi})\tilde{S}_{t-1}\right]^2}{\tilde{\sigma}^2 \sum_{t=2}^{T} \tilde{S}_{t-1}^2}.$$

This is the t-statistic for the hypothesis $\phi = 0$ in the regression (A1.12) above.

Finally, it remains to show that the information matrix is block diagonal. A straightforward calculation yields

$$\text{(A1.17)} \qquad \frac{\partial^2 \ln L^*(\sim)}{\partial \beta \partial \xi} = \frac{-1}{\tilde{\sigma}^2}\{X_0^2 + \sum_{t=2}^{T} [(y_t - \tilde{\xi})(t-1) + \tilde{S}_{t-1}]\}$$

$$\text{(A1.18)} \qquad \frac{\partial^2 \ln L^*(\sim)}{\partial \beta \partial \psi_X} = -X_0^2/\tilde{\sigma}^2 .$$

The appropriate normalization for the information matrix is T^{-2}, since T^{-2} times (A1.14) approaches a limiting distribution, in light of the convergence

$$\text{(A1.19)} \qquad T^{-2} \sum_t S_{t-1}^2 \rightarrow \sigma^2 \int_0^1 W(r)^2 dr , \quad S_t = \sum_1^t \epsilon_t .$$

Here $W(r)$ is the standard Wiener process (Brownian motion) on [0,1]. It is obvious that T^{-2} times (A1.18) approaches zero. The same is true of (A1.17). We have

$$\text{(A1.20)} \qquad T^{-3/2} \sum_t t\epsilon_t \rightarrow \sigma\left[W(1) - \int_0^1 W(r)dr \right]$$

$$\text{(A1.21)} \qquad T^{-3/2} \sum_t S_{t-1} \rightarrow \sigma \int_0^1 W(r)dr .$$

Therefore $T^{-3/2}$ times (A1.17) has a limiting distribution, and T^{-2} times (A1.17) approaches zero.

APPENDIX 2

INVARIANCE RESULTS

We consider the regression coefficient

(A2.1) $$\tilde{\phi} = \sum_{t=2}^{T} (\tilde{S}_{t-1} - \bar{\tilde{S}})(\Delta y_t - \overline{\Delta y}) / \sum_{t=2}^{T} (\tilde{S}_{t-1} - \bar{\tilde{S}})^2$$

and its associated t-statistic. We wish to show that their finite sample distributions are independent of ψ , ξ and σ_ε , though they depend on $X_0^* = X_0/\sigma_\varepsilon$ when $\beta \neq 1$.

The model is as given by (4) of the main text. The solution for y_t is

(A2.2) $$y_t = \psi + \xi t + \beta^t X_0 + S_t(\beta) ,$$

where

(A2.3) $$S_t(\beta) = \sum_{j=0}^{t-1} \beta^j \varepsilon_{t-j} .$$

This implies

(A2.4) $$\Delta y_t = \xi + (\beta-1)\beta^{t-1}X_0 + \Delta S_t(\beta)$$
$$= \xi + (\beta-1)\beta^{t-1}X_0 + \varepsilon_t + (\beta-1)S_{t-1}(\beta) .$$

Furthermore,

(A2.5) $$\tilde{\xi} = \overline{\Delta y} = \xi - \beta X_0(1 - \beta^{T-1})/(T-1) + \bar{\varepsilon} + (\beta-1)\overline{S(\beta)} .$$

Therefore,

(A2.6) $$\Delta y_t - \overline{\Delta y} = [(\beta-1)\beta^t - \beta(1 - \beta^{T-1})/(T-1)]X_0 + (\varepsilon_t - \bar{\varepsilon}) + (\beta-1)[S_{t-1}(\beta) - \overline{S(\beta)}]$$

Thus $\Delta y_t - \overline{\Delta y}$ is independent of ψ and ξ . It depends on X_0 except when $\beta = 1$. Given T , β and X_0^* , it has the same scale as σ_ε .

We proceed similarly with the terms involving $\tilde{S}$. We have

(A2.7) $\tilde{\psi}_X = y_1 - \tilde{\xi}$

$$= \psi + \beta X_0[1 + (1-\beta^{T-1})/(T-1)] + (\varepsilon_1 - \bar{\varepsilon}) + (\beta-1)\overline{S(\beta)}$$

(A2.8) $\tilde{S}_{t-1} = y_{t-1} - \tilde{\psi}_X - \tilde{\xi}(t-1)$

and, using (A2.2), (A2.5) and (A2.7),

(A2.9) $\tilde{S}_{t-1} = S_{t-1}(\beta) + \beta X_0[\beta^{t-2} - 1 + (t-2)(1-\beta^{T-1})/(T-1)] - \varepsilon_1 - (t-2)\bar{\varepsilon} - (t-2)(\beta-1)\overline{S(\beta)}$.

From (A2.9), it is clear that $\tilde{S}_{t-1}$ is independent of ψ and ξ . It depends on X_0 unless $\beta = 1$; and, given T , β and X_0^* , it has the same scale as σ_ε .

Since both $\tilde{\varphi}$ and its associated t-statistic depend only on $\Delta y_t - \overline{\Delta y}$ and $\tilde{S}_{t-1}$ $(t = 2, ..., T)$, they are independent of ψ and ξ , but their distribution depends on X_0 when $\beta \neq 1$. Also, since the scale factor σ_ε enters the numerator and denominator of $\tilde{\varphi}$ and the t-statistic in exactly the same way, the distributions of $\tilde{\varphi}$ and the t-statistic (for given T , β and X_0^*) are also independent of σ_ε .

APPENDIX 3

ASYMPTOTIC THEORY

(i) Linear Trends

We employ the functional limit theory used in Phillips and Perron (1988) and some of the subsidiary limit results on partial sums given there. Let $W(r)$ be a standard Brownian motion on the [0,1] interval, $V(r) = W(r) - rW(1)$ be a standard Brownian bridge and define $\underline{V}(r) = V(r) - \int_0^1 V(r)dr = W(r) + \left(\frac{1}{2} - r\right)W(1) - \int_0^1 W(r)dr$. Observe that in $L_2[0,1]$, $\underline{V}(r)$ is the projection of $V(r)$ on the orthogonal complement of the constant function. Thus $\underline{V}(r)$ is simply a demeaned Brownian bridge.

From (17) we find that $\tilde{S}_{t-1} = \Sigma_2^t(\varepsilon_s - \bar{\varepsilon})$ so that

$$T^{-1/2}\tilde{S}_{[Tr]} = T^{-1/2}S_{[Tr]} - ([Tr]/T)T^{-1/2}S_T$$
$$\rightarrow \sigma(W(r) - rW(1)) = \sigma V(r)$$

With this in hand it is easy to see that

$$\text{(A3.1)} \qquad T^{-2}\textstyle\sum_1^T(\tilde{S}_{t-1} - \bar{\tilde{S}})^2 \rightarrow \sigma^2\int_0^1 \underline{V}(r)^2 dr .$$

Further

$$T^{-1}\textstyle\sum_1^T(\tilde{S}_{t-1} - \bar{\tilde{S}})\varepsilon_t = T^{-1}\sum_1^T(\tilde{S}_{t-1} - \bar{\tilde{S}})(\varepsilon_t - \bar{\varepsilon})$$
$$= (1/2)\{T^{-1}(\textstyle\sum_1^T(\varepsilon_t - \bar{\varepsilon}))^2 - T^{-1}\sum_1^T(\varepsilon_t - \bar{\varepsilon})^2\}$$
$$\text{(A3.2)} \qquad \rightarrow_p (-1/2)\sigma_\varepsilon^2 .$$

Now

$$\tilde{\rho} = \{T^{-2}\textstyle\sum_1^T(\tilde{S}_{t-1} - \bar{\tilde{S}})^2\}^{-1}\{T^{-1}\sum_1^T(\tilde{S}_{t-1} - \bar{\tilde{S}})\varepsilon_t\}$$
$$\text{(A3.3)} \qquad \rightarrow_p (-1/2)\sigma_\varepsilon^2/\sigma^2\int_0^1 \underline{V}(r)^2 dr$$

by joint convergence of the numerator and denominator.

Next observe that

$$\text{(A3.4)} \qquad \tilde{\tau} = \{T^{-2}\textstyle\sum_1^T(\tilde{S}_{t-1} - \bar{\tilde{S}})^2\}^{1/2}\tilde{\rho}/s$$

where s is the estimated standard error of the regression (9). Since $s \to_p \sigma_\varepsilon$ we obtain from (A3.1), (A3.3) and (A3.4) the following limit for $\tilde{\tau}$:

$$\tilde{\tau} \to (-1/2)(\sigma_\varepsilon/\sigma)\left[\int_0^1 \underline{V}(r)^2 dr\right]^{-1/2} .$$

(ii) Higher Order Trends

As in the linear trend case, the asymptotics are determined by the behavior of the partial sum process $\tilde{S}_t$ in (37). Note that from the regression

$$\Delta y_t = \sum_{j=0}^{p-1} \tilde{b}_j t^j + \tilde{u}_t \tag{A3.5}$$

we have, under the null hypothesis,

$$\tilde{u}_t = \varepsilon_t - \sum_{j=0}^{p-1} (\tilde{b}_j - b_j) t^j . \tag{A3.6}$$

Let X be the trend regressor matrix in (A3.5) and define

$$D_T = \text{diag}(T, T^3, \ldots, T^{2p-3}) .$$

Then, in conventional regression notation, we have

$$\begin{aligned} D_T^{1/2}(\tilde{b} - b) &= \left[D_T^{-1/2} X' X D_T^{-1/2}\right]^{-1}(D_T^{-1/2} X' \varepsilon) \\ &\to \sigma Q^{-1} \int_0^1 g(r) dW(r) \end{aligned} \tag{A3.7}$$

where $Q = (q_{ij})$ with $q_{ij} = 1/(i+j-1)$ and $g(r)' = (1, r, \ldots, r^{p-1})$. The weak convergence to (A3.7) follows in a straightforward way using the methods in Phillips (1987). Next, we observe that

$$T^{-j-1} \sum_{t=1}^{[Tr]} t^j \sim [Tr]^{j+1}/T^{j+1}(j+1) \to r^{j+1}/(j+1) .$$

Hence

$$T^{-1/2}\sum_{t=1}^{[Tr]}\tilde{u}_t = T^{-1/2}S_{[Tr]} - \sum_{j=0}^{p-1}(\tilde{b}_j - b_j)(T^{-1/2}\sum_{t=1}^{[Tr]} t^j)$$

$$= T^{-1/2}S_{[Tr]} - (D_T^{1/2}(\tilde{b} - b))' D_T^{-1/2}(T^{-1/2}\sum_1^{[Tr]} t^j)$$

$$\to \sigma W(r) - \sigma(\int_0^1 dW(s)g(s)')Q^{-1}q(r)$$

$$= \sigma\{W(r) - (\int_0^1 dW(s)g(s)')Q^{-1}q(r)\}$$

(A3.8) $$= \sigma V_p(r)\ ,$$

where $q(r)$ is $p \times 1$ with j'th element r^j/j. We shall call the Gaussian process $V_p(r)$ a p-level Brownian bridge. Like a conventional Brownian bridge, this process is tied down on the [0,1] interval since $V_p(0) = 0$ and

$$V_p(1) = W(1) - \int_0^1 dW(s)g(s)'Q^{-1}q(1)$$

$$= W(1) - \int_0^1 dW(s)g(s)'e_1$$

$$= W(1) - \int_0^1 dW(s)$$

$$= 0$$

where we use the fact that $Q^{-1}q(1) = e_1$, the first unit vector.

From (38) in conventional regression notation we have

(A3.9) $$\hat{\phi} = (\tilde{S}'Q_X\tilde{S})^{-1}(\tilde{S}'Q_X\Delta y)\ ,$$

where Q_X denotes the usual projection matrix on the orthogonal complement of the range of X. Now

(A3.10) $$T^{-2}\tilde{S}'Q_X\tilde{S} \to \sigma^2\int_0^1 \underline{V}_p(r)^2 dr$$

where

$$\underline{V}_p(r) = V_p(r) - \sum_{j=0}^{p-1}\hat{\alpha}_j r^j$$

and

$$\hat{\alpha} = \mathrm{argmin}_\alpha \int_0^1 \Big(V_p(r) - \sum_{j=0}^{p-1}\alpha_j r^j\Big)^2 dr\ .$$

The process $\underline{V}_p(r)$ is a detrended p-level Brownian bridge and is the asymptotic equivalent of the regression projection $Q_X\tilde{S}$.

We also find that

$$T^{-1}\tilde{S}'Q_X\Delta y = T^{-1}\tilde{S}'Q_X\varepsilon = T^{-1}\tilde{S}'\underline{\varepsilon}$$

where $\underline{\varepsilon} = Q_X\varepsilon$. But from (A3.6) we have

$$\tilde{S}_t = \sum_1^t\{\varepsilon_s - \sum_0^{p-1}(\tilde{b}_j - b_j)s^j\} = \sum_1^t \underline{\varepsilon}_s .$$

Thus

$$T^{-1}\tilde{S}'\underline{\varepsilon} = T^{-1}\sum_1^t\tilde{S}_{t-1}\,\underline{\varepsilon}_t$$

$$= (2T)^{-1}\{(\sum_1^t \underline{\varepsilon}_t)^2 - \sum_1^t \underline{\varepsilon}_t^2\}$$

(A3.11) $$\to -(1/2)\sigma_\varepsilon^2 .$$

From (A3.9)-(A3.11) we deduce that

$$\tilde{\rho} = T\hat{\phi} \to -(1/2)\sigma_\varepsilon^2/(\sigma^2\int_0^1\underline{V}_p^2) = -\left[2\int_0^1\underline{V}_p^2\right]^{-1}\omega^2 ,$$

and in a similar fashion the limit of the t ratio $\tilde{\tau}$ in (38) is found to be

$$\tilde{\tau} \to -(1/2)\sigma_\varepsilon^2/\sigma_\varepsilon\left[\sigma^2\int_0^1\underline{V}_p^2\right]^{1/2} = -(1/2)\left[\int_0^1\underline{V}_p^2\right]^{-1/2}\omega .$$

APPENDIX 4

ALTERNATIVE PARAMETERIZATIONS

(SP) $y_t = \psi + \xi t + X_t,\ X_t = \beta X_{t-1} + \varepsilon_t$

(DF) $y_t = \alpha + \delta t + \beta y_{t-1} + \varepsilon_t$

(NS) $y_t^* = \gamma + \Phi t + \beta y_{t-1}^* + \varepsilon_t^*,\ y_t^* = (y_t - y_0)/\sigma_\varepsilon,\ \varepsilon_t^* = \varepsilon_t/\sigma_\varepsilon$.

Parameters of (DF) in terms of (SP)

(A4.1)
$$\alpha = \psi - \psi\beta + \xi\beta$$
$$\delta = \xi(1-\beta)$$
$$y_0 = \psi + X_0$$

Parameters of (NS) in terms of (SP)

(A4.2)
$$\gamma = [\xi\beta + X_0(\beta-1)]/\sigma_\varepsilon$$
$$\Phi = \xi(1-\beta)/\sigma_\varepsilon$$

Parameters of (SP) in terms of (DF)

(A4.3)
$$\psi = [\alpha(1-\beta) - \beta\delta]/(1-\beta)^2$$
$$\xi = \delta/(1-\beta)$$
$$X_0 = y_0 - [\alpha(1-\beta) - \beta\delta]/(1-\beta)^2$$

Parameters of (NS) in terms of (DF)

(A4.4)
$$\gamma = [\alpha + y_0(\beta-1)]/\sigma_\varepsilon$$
$$\Phi = \delta/\sigma_\varepsilon$$

Parameters of (SP) in terms of (NS)

(A4.5)
$$\psi = y_0 + \sigma_\varepsilon^2[\gamma(1-\beta) - \beta\Phi]/(1-\beta)^2$$
$$\xi = \sigma_\varepsilon\Phi/(1-\beta)$$
$$X_0 = \sigma_\varepsilon[\beta\Phi - \gamma(1-\beta)]/(1-\beta)^2$$

Parameters of (DF) in terms of (NS)

(A4.6) $$\alpha = \sigma_\varepsilon \gamma + y_0(1-\beta)$$

$$\delta = \sigma_\varepsilon \Phi$$

REFERENCES

Andrews, D. W. K. (1989), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," Cowles Foundation Discussion Paper No. 877R.

Bhargava, A. (1986), "On the Theory of Testing for Unit Roots in Observed Time Series," *Review of Economic Studies*, 52, 369-384.

DeJong, D.N., J.C. Nankervis, N.E. Savin and C.H. Whiteman (1988), "Integration Versus Trend Stationarity in Macroeconomic Time Series," Working Paper 88-27, Department of Economics, University of Iowa.

Dickey, D.A. (1976), "Estimation and Hypothesis Testing in Nonstationary Time Series," unpublished Ph.D. dissertation, Iowa State University.

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, 427-431.

Dufour, J-M and M.L. King (1989), "Optimal Invariant Tests for the Autocorrelation Coefficient in Linear Regressions with Stationary and Nonstationary AR(1) Errors," unpublished manuscript.

Evans, G.B.A. and N.E. Savin (1984), "Testing for Unit Roots: 2," *Econometrica*, 52, 1241-1269.

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

Guilkey, D.K. and P. Schmidt (1989), "An Interpolation Test for a Unit Root in the Presence of Drift," unpublished manuscript.

King, M.L. (1981), "Testing for Random Walk Disturbances in the Linear Regression Model," unpublished manuscript.

Nankervis, J.C. and N.E. Savin (1985), "Testing the Autoregressive Parameter with the **t** Statistic," *Journal of Econometrics*, 27, 143-161.

Nankervis, J.C. and N.E. Savin (1987), "Finite Sample Distributions of **t** and **F** Statistics in an AR(1) Model with an Exogenous Variable," *Econometric Theory*, 3, 387-408.

Nelson, C.R. and C.I. Plosser (1982), "Trends Versus Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, 10, 139-162.

Ouliaris, S., J.Y. Park and P.C.B. Phillips (1988), "Testing for a Unit Root in the Presence of a Maintained Trend," working paper, Department of Economics, University of Maryland.

Phillips, P.C.B. (1987), "Time Series Regression with Unit Roots," *Econometrica*, 55, 277-301.

Phillips, P.C.B. and S. Ouliaris (1987), "Asymptotic Properties of Residual Based Tests for Cointegration," working paper 88-5, University of Maryland.

Phillips, P.C.B. and P. Perron (1988), "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75, 335-346.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling (1986), *Numerical Recipes: The Art of Statistical Computing*, Cambridge: Cambridge University Press.

Said, S.E. and D.A. Dickey, "Testing for Unit Roots in Autoregessive - Moving Average Models of Unknown Order," *Biometrika*, 1984, 599-608.

Schmidt, P. (1989), "Dickey-Fuller Tests with Drift," *Advances in Econometrics: Cointegration, Spurious Regressions, and Unit Roots*, forthcoming.

West, K.D. (1987), "A Note on the Power of Least Squares Tests for a Unit Root," *Economics Letters*, 24, 249-252.

TABLE 1A

CRITICAL VALUES FOR $\tilde{\tau}$

T	.01	.025	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.975	.99
25	-3.90	-3.50	-3.18	-2.85	-2.48	-2.24	-2.05	-1.89	-1.75	-1.60	-1.45	-1.28	-1.17	-1.08	-1.00
50	-3.73	-3.39	-3.11	-2.80	-2.46	-2.24	-2.06	-1.90	-1.76	-1.62	-1.47	-1.29	-1.16	-1.08	-0.99
100	-3.63	-3.32	-3.06	-2.77	-2.45	-2.24	-2.06	-1.90	-1.76	-1.62	-1.47	-1.29	-1.17	-1.07	-0.97
200	-3.61	-3.30	-3.04	-2.76	-2.45	-2.24	-2.06	-1.91	-1.77	-1.62	-1.47	-1.29	-1.16	-1.07	-0.97
500	-3.59	-3.29	-3.04	-2.76	-2.44	-2.22	-2.05	-1.90	-1.75	-1.61	-1.47	-1.29	-1.16	-1.07	-0.98
1000	-3.58	-3.28	-3.02	-2.75	-2.43	-2.22	-2.05	-1.90	-1.76	-1.62	-1.47	-1.29	-1.16	-1.07	-0.98
2000	-3.56	-3.27	-3.02	-2.75	-2.44	-2.22	-2.05	-1.90	-1.76	-1.62	-1.47	-1.29	-1.16	-1.07	-0.97

CRITICAL VALUES FOR $\tilde{\rho}$

T	.01	.025	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.975	.99
25	-20.4	-17.9	-15.7	-13.4	-10.9	-9.27	-8.02	-6.98	-6.05	-5.21	-4.35	-3.43	-2.89	-2.51	-2.16
50	-22.8	-19.6	-17.0	-14.3	-11.4	-9.62	-8.26	-7.12	-6.14	-5.24	-4.35	-3.37	-2.78	-2.39	-2.03
100	-23.8	-20.4	-17.5	-14.6	-11.6	-9.76	-8.34	-7.17	-6.17	-5.23	-4.32	-3.35	-2.75	-2.32	-1.92
200	-24.8	-20.9	-17.9	-14.9	-11.8	-9.86	-8.42	-7.23	-6.20	-5.22	-4.30	-3.31	-2.72	-2.30	-1.89
500	-25.3	-21.3	-18.1	-15.0	-11.8	-9.78	-8.32	-7.15	-6.14	-5.19	-4.29	-3.31	-2.70	-2.28	-1.90
1000	-25.3	-21.3	-18.1	-15.0	-11.8	-9.78	-8.32	-7.15	-6.14	-5.19	-4.29	-3.31	-2.70	-2.28	-1.90
2000	-25.2	-21.2	-18.1	-15.0	-11.8	-9.85	-8.39	-7.19	-6.15	-5.21	-4.30	-3.32	-2.68	-2.27	-1.87

TABLE 1B

CRITICAL VALUES FOR $\tilde{\tau}$, p = 2

T	.01	.025	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.975	.99
25	-4.52	-4.09	-3.77	-3.41	-3.03	-2.77	-2.57	-2.40	-2.23	-2.08	-1.90	-1.70	-1.56	-1.46	-1.35
50	-4.28	-3.93	-3.65	-3.34	-2.99	-2.75	-2.57	-2.40	-2.24	-2.09	-1.92	-1.71	-1.57	-1.46	-1.35
100	-4.16	-3.84	-3.60	-3.31	-2.97	-2.75	-2.56	-2.40	-2.25	-2.10	-1.93	-1.73	-1.58	-1.46	-1.35
200	-4.12	-3.81	-3.55	-3.28	-2.96	-2.74	-2.56	-2.41	-2.25	-2.10	-1.93	-1.72	-1.57	-1.46	-1.34
500	-4.08	-3.80	-3.55	-3.26	-2.94	-2.73	-2.55	-2.40	-2.24	-2.10	-1.93	-1.73	-1.58	-1.46	-1.34
1000	-4.06	-3.78	-3.53	-3.26	-2.94	-2.73	-2.55	-2.40	-2.25	-2.10	-1.93	-1.72	-1.57	-1.46	-1.34
2000	-4.06	-3.78	-3.52	-3.26	-2.94	-2.73	-2.55	-2.40	-2.25	-2.10	-1.93	-1.72	-1.58	-1.46	-1.34

CRITICAL VALUES FOR $\tilde{\rho}$, p = 2

T	.01	.025	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.975	.99
25	-24.6	-22.2	-20.1	-17.8	-15.2	-13.4	-12.0	-10.7	-9.57	-8.48	-7.31	-6.04	-5.15	-4.57	-3.99
50	-28.4	-25.1	-22.4	-19.5	-16.2	-14.1	-12.5	-11.1	-9.81	-8.62	-7.36	-5.98	-5.05	-4.38	-3.77
100	-30.4	-26.6	-23.7	-20.4	-16.8	-14.5	-12.8	-11.3	-9.98	-8.72	-7.46	-6.00	-5.03	-4.35	-3.70
200	-31.8	-27.5	-24.2	-20.7	-17.1	-14.7	-12.9	-11.4	-10.1	-8.76	-7.42	-5.93	-4.95	-4.25	-3.60
500	-32.4	-28.2	-24.8	-21.0	-17.1	-14.8	-12.9	-11.4	-9.98	-8.71	-7.40	-5.92	-4.96	-4.23	-3.58
1000	-32.5	-28.2	-24.6	-21.1	-17.2	-14.8	-13.0	-11.4	-10.1	-8.74	-7.39	-5.90	-4.89	-4.20	-3.53
2000	-32.6	-28.3	-24.7	-21.1	-17.2	-14.8	-13.0	-11.4	-10.1	-8.76	-7.44	-5.91	-4.95	-4.22	-3.58

TABLE 1C

CRITICAL VALUES FOR $\tilde{\tau}$, p = 3

T	.01	.025	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.975	.99
25	-5.07	-4.62	-4.26	-3.89	-3.48	-3.21	-2.99	-2.80	-2.63	-2.46	-2.28	-2.05	-1.90	-1.78	-1.65
50	-4.73	-4.37	-4.08	-3.77	-3.42	-3.17	-2.98	-2.81	-2.64	-2.48	-2.30	-2.07	-1.91	-1.78	-1.65
100	-4.59	-4.29	-4.03	-3.72	-3.39	-3.16	-2.97	-2.80	-2.65	-2.48	-2.31	-2.09	-1.93	-1.80	-1.67
200	-4.53	-4.24	-3.99	-3.69	-3.38	-3.15	-2.97	-2.80	-2.65	-2.49	-2.31	-2.08	-1.92	-1.79	-1.65
500	-4.50	-4.20	-3.96	-3.68	-3.36	-3.14	-2.96	-2.79	-2.64	-2.48	-2.30	-2.09	-1.93	-1.79	-1.65
1000	-4.49	-4.19	-3.95	-3.68	-3.36	-3.14	-2.96	-2.79	-2.64	-2.49	-2.31	-2.09	-1.92	-1.79	-1.65
2000	-4.44	-4.16	-3.93	-3.67	-3.35	-3.13	-2.95	-2.79	-2.64	-2.49	-2.31	-2.09	-1.92	-1.78	-1.65

CRITICAL VALUES FOR $\tilde{\rho}$, p = 3

T	.01	.025	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.975	.99
25	-28.1	-25.8	-23.8	-21.5	-18.8	-17.0	-15.4	-14.1	-12.8	-11.6	-10.3	-8.67	-7.58	-6.79	-5.99
50	-33.1	-29.8	-27.0	-24.0	-20.6	-18.2	-16.4	-14.9	-13.4	-12.0	-10.4	-8.68	-7.48	-6.57	-5.66
100	-36.3	-32.3	-29.1	-25.4	-21.5	-19.0	-16.9	-15.2	-13.7	-12.2	-10.6	-8.74	-7.48	-6.55	-5.68
200	-38.0	-33.6	-30.1	-26.1	-22.0	-19.3	-17.3	-15.4	-13.8	-12.3	-10.6	-8.66	-7.39	-6.45	-5.49
500	-39.1	-34.4	-30.6	-26.6	-22.3	-19.5	-17.3	-15.5	-13.8	-12.2	-10.6	-8.70	-7.42	-6.41	-5.42
1000	-39.5	-34.6	-30.8	-26.7	-22.3	-19.6	-17.4	-15.5	-13.9	-12.3	-10.6	-8.66	-7.33	-6.35	-5.42
2000	-39.7	-34.4	-30.6	-26.7	-22.4	-19.5	-17.4	-15.5	-13.9	-12.3	-10.6	-8.67	-7.32	-6.34	-5.42

TABLE 1D

CRITICAL VALUES FOR $\tilde{\tau}$, p = 4

T	.01	.025	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.975	.99
25	-5.57	-5.09	-4.70	-4.30	-3.86	-3.58	-3.35	-3.15	-2.97	-2.79	-2.59	-2.36	-2.19	-2.05	-1.92
50	-5.13	-4.78	-4.47	-4.15	-3.78	-3.54	-3.34	-3.16	-2.99	-2.81	-2.63	-2.39	-2.21	-2.07	-1.92
100	-4.99	-4.66	-4.39	-4.10	-3.75	-3.52	-3.33	-3.15	-2.98	-2.82	-2.63	-2.39	-2.22	-2.09	-1.93
200	-4.90	-4.60	-4.33	-4.06	-3.73	-3.51	-3.32	-3.16	-2.99	-2.83	-2.64	-2.41	-2.24	-2.09	-1.94
500	-4.85	-4.56	-4.31	-4.03	-3.71	-3.49	-3.31	-3.15	-2.99	-2.82	-2.64	-2.40	-2.23	-2.09	-1.95
1000	-4.83	-4.54	-4.31	-4.03	-3.71	-3.48	-3.30	-3.14	-2.98	-2.82	-2.64	-2.40	-2.22	-2.08	-1.93
2000	-4.81	-4.52	-4.29	-4.01	-3.71	-3.49	-3.30	-3.14	-2.98	-2.82	-2.64	-2.41	-2.23	-2.09	-1.94

CRITICAL VALUES FOR $\tilde{\rho}$, p = 4

T	.01	.025	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.975	.99
25	-31.0	-28.8	-26.9	-24.7	-22.0	-20.1	-18.5	-17.1	-15.8	-14.5	-13.0	-11.3	-10.0	-9.02	-8.09
50	-37.4	-34.1	-31.2	-28.1	-24.5	-22.1	-20.2	-18.4	-16.8	-15.2	-13.5	-11.4	-9.96	-8.87	-7.69
100	-41.8	-37.4	-34.0	-30.2	-26.0	-23.2	-21.0	-19.1	-17.2	-15.5	-13.7	-11.5	-9.94	-8.80	-7.60
200	-44.0	-39.2	-35.2	-31.2	-26.7	-23.8	-21.4	-19.5	-17.6	-15.8	-13.9	-11.6	-10.0	-8.77	-7.58
500	-45.3	-40.3	-36.2	-31.8	-27.1	-24.1	-21.7	-19.6	-17.7	-15.8	-13.9	-11.5	-9.92	-8.73	-7.57
1000	-45.7	-40.6	-36.6	-32.0	-27.2	-24.1	-21.7	-19.6	-17.7	-15.8	-13.9	-11.5	-9.84	-8.66	-7.44
2000	-45.8	-40.5	-36.4	-31.9	-27.3	-24.2	-21.7	-19.6	-17.7	-15.9	-13.9	-11.6	-9.92	-8.69	-7.46

TABLE 2

SIZE AND POWER, 5% LOWER TAIL TESTS, T = 100

Exp. No.	T	β	ξ^*	X_0^*	γ	Φ	$\hat{\tau}_\mu$	$\hat{\rho}_\mu$	$\hat{\tau}_\tau$	$\hat{\rho}_\tau$	$\tilde{\tau}$	$\tilde{\rho}$
1	100	1	0	0	0	0	.049	.048	.048	.050	.051	.052
1	100	1	.02	0	.02	0	.050	.046	.048	.050	.051	.052
1	100	1	.05	0	.05	0	.046	.045	.048	.050	.051	.052
1	100	1	.10	0	.10	0	.041	.034	.048	.050	.051	.052
1	100	1	.20	0	.20	0	.022	.008	.048	.050	.051	.052
1	100	1	.50	0	.50	0	.006	.000	.048	.050	.051	.052
2A	100	.90	0	-10	1	0	.686	.317	.304	.120	.032	.033
2A	100	.90	0	-5	.5	0	.413	.421	.211	.198	.161	.165
2A	100	.90	0	-2	.2	0	.334	.459	.191	.234	.248	.252
2A	100	.90	0	-1	.1	0	.322	.465	.188	.239	.260	.267
2A	100	.90	0	0	0	0	.321	.467	.186	.239	.264	.270
2A	100	.90	0	1	-.1	0	.321	.464	.187	.237	.259	.265
2A	100	.90	0	2	-.2	0	.333	.457	.191	.234	.243	.249
2A	100	.90	0	5	-.5	0	.410	.419	.215	.202	.164	.167
2A	100	.90	0	10	-1	0	.681	.319	.306	.124	.032	.033
2B	100	.90	.02	-5	.518	.002	.363	.302	.211	.198	.161	.165
2B	100	.90	.02	-2	.218	.002	.291	.371	.191	.234	.248	.252
2B	100	.90	.02	-1	.118	.002	.281	.393	.188	.239	.260	.267
2B	100	.90	.02	0	.018	.002	.281	.409	.186	.239	.264	.270
2C	100	.90	.05	-5	.545	.005	.219	.112	.211	.198	.161	.165
2C	100	.90	.05	-2	.245	.005	.154	.160	.191	.234	.248	.252
2C	100	.90	.05	-1	.145	.005	.144	.178	.188	.239	.260	.267
2C	100	.90	.05	0	.045	.005	.140	.198	.186	.239	.264	.270
2D	100	.95	.02	-5	.269	.001	.127	.102	.083	.085	.086	.088
2D	100	.95	.02	-2	.119	.001	.113	.149	.082	.095	.101	.104
2D	100	.95	.02	-1	.069	.001	.110	.162	.083	.097	.104	.106
2D	100	.95	.02	0	.019	.001	.113	.172	.082	.098	.105	.108

TABLE 3

SIZE AND POWER, 5% LOWER TAIL TESTS, T = 100

Exp. No.	T	β	ξ^*	X_0^*	γ	Φ	$\hat{\tau}_\mu$	$\hat{\rho}_\mu$	$\hat{\tau}_\tau$	$\hat{\rho}_\tau$	$\tilde{\tau}$	$\tilde{\rho}$
3A	100	.90	0	0	0	0	.321	.467	.186	.239	.264	.270
3A	100	.90	.02	0	.018	.002	.281	.409	.186	.239	.264	.270
3A	100	.90	.05	0	.045	.005	.140	.198	.186	.239	.264	.270
3A	100	.90	.10	0	.09	.01	.016	.014	.186	.239	.264	.270
3A	100	.90	.20	0	.18	.02	.000	.000	.186	.239	.264	.270
3B	100	.90	0	-2	.2	0	.334	.459	.191	.234	.248	.252
3B	100	.90	.02	-2	.218	.002	.291	.371	.191	.234	.248	.252
3B	100	.90	.05	-2	.245	.005	.154	.160	.191	.234	.248	.252
3B	100	.90	.10	-2	.29	.01	.025	.009	.191	.234	.248	.252
3B	100	.90	.20	-2	.38	.02	.000	.000	.191	.234	.248	.252
3C	100	.90	0	-5	.5	0	.413	.421	.211	.198	.161	.165
3C	100	.90	.02	-5	.518	.002	.363	.302	.211	.198	.161	.165
3C	100	.90	.05	-5	.545	.005	.219	.112	.211	.198	.161	.165
3C	100	.90	.10	-5	.059	.01	.059	.005	.211	.198	.161	.165
3C	100	.90	.20	-5	.68	.02	.003	.000	.211	.198	.161	.165
3D	100	.95	0	0	0	0	.120	.190	.082	.098	.105	.108
3D	100	.95	.02	0	.019	.001	.113	.172	.082	.098	.105	.108
3D	100	.95	.05	0	.0475	.0025	.084	.112	.082	.098	.105	.108
3D	100	.95	.10	0	.095	.005	.082	.098	.082	.098	.105	.108
3D	100	.95	.20	0	.19	.01	.003	.000	.082	.098	.105	.108
4A	100	1	0	0	0	0	.049	.048	.048	.050	.051	.052
4A	100	.95	0	0	0	0	.120	.190	.082	.098	.105	.108
4A	100	.90	0	0	0	0	.321	.467	.186	.239	.264	.270
4A	100	.80	0	0	0	0	.871	.950	.644	.734	.759	.765
4B	100	1	0	-2	0	0	.049	.046	.048	.050	.051	.052
4B	100	.95	0	-2	.1	0	.121	.181	.082	.095	.101	.104
4B	100	.90	0	-2	.2	0	.334	.459	.191	.234	.248	.252
4B	100	.80	0	-2	.4	0	.886	.952	.658	.732	.701	.707
4C	100	1	.05	0	.05	0	.046	.045	.048	.052	.051	.052
4C	100	.95	.05	0	.0475	.0025	.084	.112	.082	.098	.105	.108
4C	100	.90	.05	0	.045	.005	.140	.198	.186	.239	.264	.270
4C	100	.80	.05	0	.04	.01	.279	.400	.644	.734	.759	.765

TABLE 4

SIZE AND POWER, 5% LOWER TAIL TESTS, T = 200, 500

Exp. No.	T	β	ξ^*	X_0^*	γ	Φ	$\hat{\tau}_\mu$	$\hat{\rho}_\mu$	$\hat{\tau}_\tau$	$\hat{\rho}_\tau$	$\tilde{\tau}$	$\tilde{\rho}$
5A	200	1	0	0	0	0	.048	.048	.048	.048	.048	.050
5A	100	1	.02	0	.02	0	.048	.046	.048	.048	.048	.050
5A	200	1	.05	0	.05	0	.044	.038	.048	.048	.048	.050
5A	200	1	.10	0	.10	0	.034	.021	.048	.048	.048	.050
5A	200	1	.20	0	.20	0	.013	.002	.048	.048	.048	.050
5A	200	1	.50	0	.50	0	.004	.000	.048	.048	.048	.050
5B	500	1	0	0	0	0	.053	.051	.051	.051	.049	.051
5B	500	1	.02	0	.02	0	.046	.047	.051	.051	.049	.051
5B	500	1	.05	0	.05	0	.039	.031	.051	.051	.049	.051
5B	500	1	.10	0	.10	0	.020	.006	.051	.051	.049	.051
5B	500	1	.20	0	.20	0	.008	.000	.051	.051	.049	.051
6	200	.90	0	0	0	0	.858	.946	.617	.724	.751	.763
6	200	.90	0	-1	.1	0	.860	.947	.620	.720	.738	.752
6	200	.90	0	-2	.2	0	.868	.947	.627	.719	.707	.720
6	200	.90	0	-5	.5	0	.908	.949	.677	.710	.511	.526
7A	200	.90	0	0	0	0	.858	.946	.617	.724	.751	.763
7A	200	.90	.02	0	.018	.002	.615	.772	.617	.724	.751	.763
7A	200	.90	.05	0	.045	.005	.050	.076	.617	.724	.751	.763
7A	200	.90	.10	0	.09	.01	.000	.000	.617	.724	.751	.763
7B	500	.95	0	0	0	0	.967	.994	.819	.897	.910	.914
7B	500	.95	.02	0	.019	.001	.395	.577	.819	.897	.910	.914
7B	500	.95	.05	0	.0475	.0025	.000	.000	.819	.897	.910	.914
8A	200	1	0	0	0	0	.048	.048	.048	.048	.048	.050
8A	200	.95	0	0	0	0	.310	.460	.178	.233	.254	.266
8A	200	.90	0	0	0	0	.858	.946	.617	.724	.751	.763
8A	200	.80	0	0	0	0	1.00	1.00	.999	1.00	.997	.997
8B	500	1	0	0	0	0	.053	.051	.051	.051	.049	.051
8B	500	.95	0	0	0	0	.967	.994	.819	.897	.910	.914
8B	500	.90	0	0	0	0	1.00	1.00	1.00	1.00	1.00	1.00