Econometrics Journal (1998), volume **1**, pp. 27–43.

An ADF coefficient test for a unit root in ARMA models of unknown order with empirical applications to the US economy

ZHIJIE XIAO[1], PETER C.B. PHILLIPS[2]

[1]*Department of Economics, University of Illinois at Urbana–Champaign, USA*
[2]*Cowles Foundation for Research in Economics, Yale University, USA*
E-mail: `phillips@tempest.econ.yale.edu`

Received: June 1998

Summary This paper proposes an Augmented Dickey–Fuller (ADF) coefficient test for detecting the presence of a unit root in autoregressive moving average (ARMA) models of unknown order. Although the limit distribution of the coefficient estimate depends on nuisance parameters, a simple transformation can be applied to eliminate the nuisance parameter asymptotically, providing an ADF coefficient test for this case. When the time series has an unknown deterministic trend, we propose a modified version of the ADF coefficient test based on quasi-differencing in the construction of the detrending regression as in Elliott *et al*. (1996). The limit distributions of these test statistics are derived. Empirical applications of these tests for common macroeconomic time series in the US economy are reported and compared with the usual ADF t-test.

Keywords: *ADF test, ARMA process, Unit root test.*

1. INTRODUCTION

Tests for a unit root have attracted a considerable amount of work in the last ten years. One important reason is that these tests can help to evaluate the nature of the nonstationarity that many macroeconomic data exhibit. In particular, they help in determining whether the trend is stochastic, deterministic or a combination of both. Following Nelson and Plosser (1982), much empirical research has been done and evidence has accumulated that many macroeconomic variables have structures with a unit root. The literature on testing for a unit root is immense. The most commonly used tests for a unit root are the Dickey–Fuller test and the Phillips Z-tests. The Dickey–Fuller test (1979) is based on the regression of the observed variable on its one-period lagged value, sometimes including an intercept and time trend. In an important extension of Dickey and Fuller (1979), Said and Dickey (1984) show that the Dickey–Fuller t-test for a unit root, which was originally developed for AR representations of known order, remains asymptotically valid for a general ARMA process of unknown order. This t-test is usually called the Augmented Dickey–Fuller (ADF) test. An alternative semiparametric approach to detecting the presence of a unit root in general time series setting was proposed by Phillips (1987a) and extended in Phillips and Perron (1988). These tests are known as Phillips Z_α and Z_t tests. The Z -tests allow for a wide class of time series with heterogeneously and serially correlated errors.

The ADF test is a t-test in a long autoregression. Said and Dickey (1984) prove the validity of this test in general time series models provided the lag length in the autoregression increases

© Royal Economic Society 1998. Published by Blackwell Publishers Ltd, 108 Cowley Road, Oxford OX4 1JF, UK and 350 Main Street, Malden, MA, 02148, USA.

with the sample size at a rate less than $n^{1/3}$, where n = sample size. No such extension of the Dickey–Fuller coefficient test is recommended in their work, since even as the lag length goes to infinity, the coefficient estimate has a limit distribution that is dependent on nuisance parameters. However, the Z_α test is a coefficient based test with a nonparametric correction which successfully eliminates nuisance parameters. A similar idea can be applied to construct an ADF coefficient based test. In particular, the nuisance parameters can be consistently estimated and the coefficient estimate transformed to eliminate the nuisance parameters asymptotically, providing an ADF coefficient test with the same limit distribution as the original Dickey–Fuller coefficient test and the Z_α test.

The ADF coefficient test can also be extended to models with a deterministic trend. The traditional way in dealing with trending data is to detrend the data by ordinary least squares (OLS), and then apply a unit root test to the detrended time series. This is usually done by including a deterministic trend in the ADF regression model. Like the case with no deterministic trend, the limit distribution of the estimated autoregressive coefficient depends on a nuisance parameter and transformation is needed to eliminate this dependence. A natural way to compare tests is to examine their power in large samples. Recent study shows that power gains can be achieved when there is a deterministic trend and detrending is performed in a way that is efficient under the alternative hypothesis in constructing the test. Such detrending procedures use quasi-differenced (QD) data and were suggested in Elliott *et al*. (1996) to increase the power of unit root tests for models with deterministic trends. An analysis of the efficiency gain from this detrending procedure (which we call QD detrending) and its effects on test efficiency is given in Phillips and Lee (1996). As yet, few empirical applications of QD detrended unit root tests have appeared in the literature.

This paper develops an ADF-type coefficient based unit root test (called ADF_α) for ARMA models of unknown order, with a parametric correction that frees the limit distribution of the test statistic of nuisance parameters. We also extend this test to models with a deterministic trend and a modified ADF coefficient test based on QD detrending is developed. The limit distributions of the ADF coefficient test and its QD detrended version are the same as those of the Z_α test and QD detrended Z_α test. Empirical applications of these tests to the post-war quarterly US data and the extended Nelson–Plosser data are also reported. We compare the OLS detrended ADF_α test with the QD detrended ADF_α test, and examine the QD detrended ADF_α tests for different choices of c (the quasi-differencing parameter).

The paper is organized as follows. Section 2 develops the theory for the ADF coefficient test. The QD detrended ADF coefficient test and its limit theory are given in Section 3. Section 4 reports some empirical applications to a variety of macroeconomic time series data. Proofs of theorems are given in the Appendix. Our notation is standard. We use the symbol '$\Rightarrow$' to signify weak convergence of the associated probability measures. Continuous stochastic process such as the Brownian motion $B(r)$ on $(0, 1)$ are usually written simply as B and integrals like $\int$ are understood to be Lebesgue integrals over the interval $(0, 1)$, the measure of integration 'dr' being omitted for simplicity.

2. AN ADF COEFFICIENT TEST FOR A UNIT ROOT

Consider a time series

$$y_t = \alpha y_{t-1} + u_t, \quad t = 1, 2, \dots, n, \tag{1}$$

satisfying the following conditions:

© Royal Economic Society 1998

Assumption A1. *y_t is initialized at $t = 0$ by y_0, a random variable with finite variance.*

Assumption A2. *u_t is a stationary and invertible ARMA(p, q) process satisfying $a(L)u_t = b(L)\varepsilon_t$, where $\varepsilon_t = \text{i.i.d.}(0, \sigma^2)$, $a(L) = \sum_{j=0}^{p} a_j L^j$, $b(L) = \sum_{j=0}^{q} b_j L^j$, $a_0 = b_0 = 1$, and L is the lag operator.*

Assumption A3. *$n^{-1/2} \sum_{t=1}^{(nr)} u_t \Rightarrow B(r) = \omega W(r)$, $n^{-1/2} \sum_{t=1}^{(nr)} \varepsilon_t \Rightarrow B_\varepsilon(r) = \sigma W(r)$, where $\omega^2 = E(u_1^2) + 2\sum_{k=2}^{\infty} E(u_1 u_k) = \sigma^2\{b(1)/a(1)\}^2 =$ long run variance of u_t, and $W(r)$ is a standard Brownian motion.*

Notice that the limits of partial sums of u_t and ε_t depend on the same standard Brownian motion $W(r)$. From A2, we obtain the AR representation of u_t (e.g. Fuller (1976), Theorem 2.7.2)

$$\varepsilon_t = d(L)\, u_t = \sum_{j=0}^{\infty} d_j u_{t-j}, \qquad d_0 = 1.$$

Rewrite the above representation as

$$\varepsilon_t = d(L)\{\Delta y_t - (\alpha - 1)y_{t-1}\} = d(L)\Delta y_t - (\alpha - 1)d(L)y_{t-1},$$

and use a Beveridge–Nelson (BN) decomposition for $d(L)$ as $d(L) = d(1) + d^*(L)(1 - L)$, we obtain

$$\begin{aligned}\varepsilon_t &= d(L)\Delta y_t - (\alpha - 1)d^*(L)\Delta y_{t-1} - d(1)(\alpha - 1)y_{t-1} \\ &= \{d(L) - (\alpha - 1)Ld^*(L)\}\Delta y_t - (\alpha - 1)d(1)y_{t-1}.\end{aligned}$$

Define $\beta(L) = 1 - d(L) + (\alpha - 1)Ld^*(L) = \beta_1 L + \beta_2 L^2 + \cdots$, and $a = d(1)(\alpha - 1)$, we then have the following regression

$$\Delta y_t = a y_{t-1} + \beta(L)\Delta y_t + \varepsilon_t,$$

or

$$\Delta y_t = a y_{t-1} + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + \cdots + \varepsilon_t. \tag{2}$$

The null hypothesis of interest, which was $H_0 : \alpha = 1$ in (1), is equivalent to $H_0 : a = 0$. In place of the infinite AR regression (2), we consider the ADF regression model

$$\Delta y_t = a y_{t-1} + \beta_1 \Delta y_{t-1} + \cdots + \beta_k \Delta y_{t-k} + e_{tk}, \tag{3}$$

where e_{tk} is defined as $\Delta y_t - a y_{t-1} - \beta_1 \Delta y_{t-1} - \cdots - \beta_k \Delta y_{t-k}$. We use Z to denote the $n \times (k+1)$ matrix of explanatory variables and partition it in the following way: $Z = (y_{-1}, Z_k)$, where y_{-1} is the $n \times 1$ vector of lagged variables, and Z_k is the $n \times k$ matrix of observations of the k lagged difference variables $(\Delta y_{t-1}, \ldots, \Delta y_{t-k})$. Thus we have the following matrix representation

$$\Delta y = Z\beta + e_k,$$

where $\beta = (a, \beta_1, \ldots, \beta_k)'$, $e_k = (e_{k1}, \ldots, e_{kn})'$.

We shall be concerned with the limit behavior of the conventional least squares regression coefficient $\hat{a}$ for a in (3) given by

$$\hat{a} = (y_{-1}' P_z y_{-1})^{-1} y_{-1}' P_z \Delta y,$$

where $P_z = I - Z_k(Z_k' Z_k)^{-1} Z_k'$. We make the following condition on the expansion rate of the lag length k.

© Royal Economic Society 1998

Assumption A4. *$n^{-1/3}k \to 0$ as $n \to \infty$, and there exist $c > 0, r > 0$, such that $ck > n^{1/r}$.*

The limit distribution of $\hat{a}$ is given by the following theorem.

Theorem 1. *Under assumptions A1 to A4, when $\alpha = 1$,*

$$n\hat{a} \Rightarrow \frac{\sigma \int W dW}{\omega \int W^2}.$$

Remark 1. The limit distribution of the regression coefficient $\hat{a}$ depends on unknown scale parameters ω and σ, and thus the statistic $n\hat{a}$ can not be used directly for unit root testing. However, ω and σ can be consistently estimated, and there exists a simple transformation of the statistic $n\hat{a}$ which eliminates the nuisance parameters asymptotically. In particular, $\hat{\sigma}^2 = \sum \hat{e}_{tk}^2/n$ is a consistent estimator of σ^2, and ω^2 can be consistently estimated by the AR estimator $\hat{\omega}^2 = \hat{\sigma}^2/(1 - \sum \hat{\beta}_j)^2$. Thus, we define

$$ADF_\alpha = (\hat{\omega}/\hat{\sigma})n\hat{a}.$$

Under the null hypothesis that $\alpha = 1$, it is apparent that the modified coefficient test statistic

$$ADF_\alpha \Longrightarrow \frac{\int W\, dW}{\int W^2},$$

has the same limit distribution as that of the Phillips Z_α test and the original Dickey–Fuller coefficient test.

A minimal condition for a satisfactory statistical test is that it should be able to discriminate between the null hypothesis and a fixed alternative with probability one in large samples. The next theorem guarantees this property (for a definition of the spectral density, see the proof of Theorem 1 in the Appendix).

Theorem 2. *If y_t is generated by (1) with $\alpha < 1$, then $ADF_\alpha = O_p(n)$. The divergence rate is sharp in the sense that $n^{-1} ADF_\alpha \to_p c$ for some $c \neq 0$.*

Remark 2. As the sample size n increases, the test statistic ADF_α diverges faster under H_1 than the ADF t-ratio statistic, which is of order $O_p(n^{1/2})$. This suggests that the coefficient based statistic is likely to have higher power than the t-ratio statistic in large samples. Another way of comparing power of statistical tests is to look at their behavior under the local alternative. It can be verified that under $H_c : \alpha = 1 + c/n$, the limit theory for the two tests is as follows

$$ADF_\alpha \Rightarrow c + \left\{ \int J_c(r)^2\, dr \right\}^{-1} \int J_c(r)\, dW(r),$$

and

$$ADF_t \Rightarrow c\left\{ \int J_c(r)\, dr \right\}^{-1/2} + \left\{ \int J_c(r)^2\, dr \right\}^{-1/2} \int J_c(r)\, dW(r),$$

where $J_c(r) = \int_0^r e^{(r-s)c}\, dW(s)$, which are the same as those of the semiparametric Z_α and Z_t tests derived in Phillips (1987b). Although both ADF_α and ADF_t are $O_p(1)$ under the local

© Royal Economic Society 1998

Table 1. Size corrected power of ADF tests, 5% level.

		AR(1)		MA(1)	
	α	$\rho = 0.5$	$\rho = -0.5$	$\theta = 0.5$	$\theta = -0.8$
ADF_α	0.95	0.216	0.234	0.181	0.251
	0.90	0.458	0.524	0.478	0.566
	0.85	0.622	0.717	0.650	0.758
ADF_t	0.95	0.211	0.233	0.171	0.248
	0.90	0.448	0.522	0.472	0.565
	0.85	0.618	0.716	0.649	0.761

alternative for finite c, when $c \to -\infty$, it is easy to show using the results in Lemma 2 of Phillips (1987b) that $ADF_\alpha = O_p(|c|)$, whereas $ADF_t = O_p(|c|^{1/2})$. In fact, as $c \to -\infty$, it can be shown that $ADF_\alpha \sim (-2c)^{1/2} ADF_t \sim c + (-2c)^{1/2}\xi$, where ξ is $N(0, 1)$. These results, like those of the simple divergence rates, indicate that coefficient tests may be expected to have higher power than t-ratio based statistics, at least for large deviations from the null hypothesis.[1]

Table 1 reports some results from a Monte Carlo experiment on the finite sample power of an ADF_α test. In order to provide a comparison between the ADF_α and ADF_t test, we give here the size corrected power. The data generating process is (1) and four designs for the error structures are considered: AR(1) processes $u_t = \rho u_{t-1} + \varepsilon_t$ with $\rho = 0.5, -0.5$; and MA(1) processes $u_t = \varepsilon_t - \theta\varepsilon_{t-1}$ with $\theta = 0.5, -0.8$, with ε_t being i.i.d. $N(0, 1)$, and $n = 100$ in both cases. The number of repetition in the experiment is 15 000. These results corroborate the findings of other simulation experiments that coefficient based tests have relatively higher power.

3. AN EFFICIENTLY DETRENDED ADF COEFFICIENT TEST

Many macroeconomics variables, such as income, or consumption, are often characterized as 'integrated processes with drift', and can be expressed as the sum of a unit root process with zero-mean increments and a linear trend. Generally, if we allow for a deterministic trend in time series y_t, we have the following representation

$$y_t = \gamma' x_t + y_t^s \tag{4}$$
$$y_t^s = \alpha y_{t-1}^s + u_t \tag{5}$$

where u_t is defined as in the Appendix, Section A2 and x_t is the $p \times 1$ deterministic trend which satisfies the property that there exists a $p \times p$ scaling matrix D_n and a piecewise continuous limit

[1]These divergence rate arguments for local power, like those as $n \to \infty$, are only suggestive. As pointed out by the referees, ADF_t^3 is a directional t-test with greater divergence rate than the ADF_α test. This alternative test also has faster divergence for local power as $c \to -\infty$. However, on the basis of the analysis in Phillips and Park (1988), there are good reasons for wanting to avoid power statistics like ADF_t^3, as tests based on such functions can be expected to have greater size distortion than those based on ADF_t.

One can similarly analyse local power as $c \to 0$, giving the form of the power function in the immediate neighbourhood of the null hypothesis. In this case, both ADF_t and ADF_α differ from their respective null variates by a term of $O_p(c)$.

© Royal Economic Society 1998

trend function $X(r)$ such that $D_n x_{(nr)} \to X(r)$, as $n \to \infty$, uniformly in $r \in (0, 1)$. Without loss of generality and to simplify the proof, we consider the leading case of a linear trend where $x_t = (1, t)'$, $D_n = \text{diag}(1, n^{-1})$, and $X(r) = (1, r)'$.

The traditional way of removing this deterministic component in unit root tests is to run an OLS regression on an augmented equation. In the present case, this is simply

$$\Delta y_t = \mu' x_t + a y_{t-1} + \beta_1 \Delta y_{t-1} + \cdots + \beta_k \Delta y_{t-k} + e_{tk}, \tag{6}$$

and we can construct the test statistics from the above regression. Including the deterministic trend in the ADF regression equation is equivalent to detrending the time series y_t by OLS regression, and then applying the ADF regression to the detrended data $\hat{y}_t^s = y_t - \hat{\gamma}' x_t$. The limit distribution of the estimate $\widehat{a}$ of a in regression (6) is given as follows.

Theorem 3. *Under assumptions A1 to A4, when $\alpha = 1$,*

$$n\,\widehat{a} \Rightarrow (\sigma/\omega) \int W_X \, dW \Big/ \int W_X^2,$$

where W_X is the detrended Brownian motion

$$W_X(r) = W(r) - \left(\int W X'\right)\left(\int X X'\right)^{-1} X(r),$$

which depends on the limit trend function $X(r)$.

The corresponding detrended ADF coefficient test can then be constructed as

$$ADF_\alpha^\tau = \left(\frac{\hat{\omega}}{\hat{\sigma}}\right) n\,\widehat{a} \Rightarrow \int W_X \, dW \Big/ \int W_X^2.$$

Monte Carlo evidence indicates that unit root tests usually have low power against plausible trend stationary alternatives (see *inter alia*, Schwert (1989), Diebold and Rudebusch (1991), DeJong *et al.* (1992), Phillips and Perron (1988), Ng and Perron (1995), Stock (1995)). In recent years, much research effort has been devoted to the development of unit root tests with improved asymptotic properties. One of the mechanisms for increasing the efficiency of the unit root tests is related to point optimal test procedures. Under Gaussian or other distributional assumptions, the Neyman–Pearson lemma can be used to construct the most powerful test of a unit root against a simple alternative (see King (1988), Dufour and King (1991) and Elliott *et al.* (1996) among others). In the case with no deterministic trend, recent work on the topic shows that there are virtually no power gains from using this approach over the coefficient test. However, when the time series contains a deterministic trend, the power of unit root tests can be improved if we perform the detrending regression in a manner that is efficient under the alternative (Elliott *et al.* (1996)).

For alternatives that are distant from unit root, the Grenander–Rosenblatt (1957) theorem implies that OLS detrending like that in (4) will be asymptotically efficient. But a more relevant theory in the unit root case needs to give attention to alternatives that are closer to unity. Such alternative hypotheses can often be well modelled using local alternatives of the form

$$H_1' : \alpha = 1 + c/n,$$

© Royal Economic Society 1998

where c is a fixed constant. We can estimate the trend coefficient by taking quasi-differences on (4), and running a least squares regression of $\Delta_c y_t$ on $\Delta_c x_t$, where Δ_c is the quasi-difference operator defined as $\Delta_c = 1 - L - (c/n)L$, and L is the lag operator. If the fitted trend vector in this regression is $\tilde{\gamma}$, we compute the QD detrended series

$$y_t^* = y_t - \tilde{\gamma}' x_t,$$

which can now be used in the construction of unit root tests, just as in the case where there are no deterministic trends to be removed.

This detrending procedure is sometimes called GLS detrending in the literature because the regression based on quasi-differenced data has a stationary residual process under H_1' and thus is asymptotically equivalent to GLS estimation (e.g. Elliott *et al*. (1996)). It is more accurate to describe the procedure as detrending after quasi-differencing (see Phillips and Lee (1996) and Canjels and Watson (1997), for recent implementations) since full GLS is not used in the detrending regression, but only quasi-differencing. We therefore refer to the procedure as QD detrending.

To derive the asymptotics for the efficiently detrended ADF_α test, it is convenient to employ the following matrix notation,

$$\begin{aligned} X' &= (x_1, \ldots, x_t, \ldots, x_n), \\ y' &= (y_1, \ldots, y_t, \ldots, y_n), \\ \Delta_c X' &= (\Delta_c x_1, \ldots, \Delta_c x_t, \ldots, \Delta_c x_n), \\ \Delta_c y' &= (\Delta_c y_1, \ldots, \Delta_c y_t, \ldots, \Delta_c y_n). \end{aligned}$$

Then $\tilde{\gamma} = (\Delta_c X' \Delta_c X)^{-1} \Delta_c X' \Delta_c y$, and we have the following asymptotic result for the QD detrended series y_t^*.

Lemma 1. *Under assumptions A1 to A3 and when $\alpha = 1$,*

$$n^{-1/2} y^*_{(nr)} \Rightarrow \underline{B}_c(r) = \omega \underline{W}_c(r),$$

where

$$\begin{aligned} \underline{B}_c(r) &= B(r) - X(r)' \left\{ \int X_c(r) X_c(r)' \, dr \right\}^{-1} \left\{ \int X_c(r) \, dB(r) - c \int X_c(r) B(r) \, dr \right\} \\ &= \omega \left[W(r) - X(r)' \left\{ \int X_c(r) X_c(r)' \, dr \right\}^{-1} \left\{ \int X_c(r) \, dW(r) - c \int X_c(r) W(r) \, dr \right\} \right] \\ &= \omega \underline{W}_c(r) \end{aligned}$$

and

$$X_c(r) = (-c, 1 - cr)'.$$

Remark 3. When the deterministic trend includes a constant term, the invertibility of $\int X_c(r) X_c(r)' \, dr$ depends on c not equaling zero, since otherwise the constant term and the coefficient of the linear trend t are not identified.

The detrended data y_t^* can be used to construct an ADF_α test for a unit root by running the following regression

$$\Delta y_t^* = a y_{t-1}^* + \beta_1 \Delta y_{t-1}^* + \cdots + \beta_k \Delta y_{t-k}^* + e_{tk}^*. \tag{7}$$

© Royal Economic Society 1998

Let $\tilde{a}$ be the estimated coefficient of a in this regression. Then the QD detrended ADF_α statistic is

$$ADF_\alpha^* = (\hat{\omega}/\hat{\sigma})n\tilde{a}.$$

Theorem 4. *Under assumptions A1 to A4 and when* $\alpha = 1$*,*

$$ADF_\alpha^* \Rightarrow \frac{\int \underline{W}_c(r)\, dW(r)}{\int \underline{W}_c(r)^2}.$$

Remark 4. The limit distribution of the modified ADF_α test depends on both the trend function and the value of c that is used in the quasi-differencing filter. This limit distribution has the same form as that of a modified semiparametric Z_α test where we use efficiently detrended y in the construction of Z_α (for an analysis of the modified Z tests using quasi-differencing detrended data, see Ng and Perron (1995), among others).

For the choice of local parameter c, Elliott *et al*. (1996) suggested $c = -13.5$ for the linear trend case. This is the approximate value where the power functions are tangent to the power envelope at a power of 50%. We conducted a simulation experiment to examine the effect of the choice of c on the finite sample performance of the tests. Table 2 reports the size corrected power of the quasi-differecing detrended ADF_α tests for different choices of c, and compares them with the OLS detrended ADF_α test. The data generating process in the experiment is (1) and $n = 100$, but the disturbance term u_t is now i.i.d. $N(0, 1)$. Again, the number of repetition is 15 000. The Monte Carlo results in Table 2 indicate that for quite a wide range of choices of c, the QD detrended tests have reasonably good finite sample performance. The Monte Carlo results show that $c = -13.5$ is generally a good default choice, and we use it in the empirical analysis below.

4. EMPIRICAL APPLICATIONS

In this section, we apply the ADF_α test and its QD detrended version to the US macroeconomic time series to demonstrate the use of these tests, reassess previous empirical findings, compare different detrending procedures and examine the QD detrended tests for different choices of c. In particular, we examined the extended Nelson–Plosser data, the stock price data from Standard and Poor's series and the post-war quarterly US macroeconomics time series data. The general conclusion that many macroeconomic time series can be modelled by unit root processes is supported using these statistics. We report our empirical result on the extended Nelson–Plosser data and the post-war quarterly US macroeconomics data here. For an analysis of stock price data with these methods, readers are referred to an early version of the paper (Xiao and Phillips 1997).

4.1. The extended Nelson–Plosser data

The ADF_α test and efficient detrending QD prefilter were applied to the 14 time series of the US economy studied in Nelson and Plosser (1982), and extended by Schotman and Van Dijk

© Royal Economic Society 1998

Table 2. Effect of parameter c on the QD detrended ADF_α tests (size corrected power).

Part I

α	OLS detrended	QD detrended $c = -2.5$	QD detrended $c = -5$	QD detrended $c = -7.5$
0.975	0.0637	0.0679	0.0657	0.0653
0.95	0.0957	0.1079	0.1052	0.1065
0.925	0.1436	0.1694	0.1679	0.1665
0.9	0.2068	0.2503	0.2503	0.2637
0.875	0.2902	0.3465	0.3497	0.3517
0.865	0.3282	0.3858	0.3917	0.3946
0.85	0.3900	0.4477	0.4548	0.4585
0.825	0.4966	0.5487	0.5598	0.5680
0.8	0.6063	0.6330	0.6520	0.6635

Part II

α	QD detrended $c = -10$	QD detrended $c = -12.5$	QD detrended $c = -13.5$	QD detrended $c = -15$
0.975	0.0667	0.0658	0.0660	0.0660
0.95	0.1093	0.1091	0.1079	0.1072
0.925	0.1682	0.1681	0.1677	0.1674
0.9	0.2539	0.2520	0.2530	0.2520
0.875	0.3552	0.3553	0.3547	0.3538
0.865	0.3994	0.3990	0.3985	0.4000
0.85	0.4658	0.4678	0.4666	0.4665
0.825	0.5789	0.5828	0.5834	0.5843
0.8	0.6762	0.6815	0.6819	0.6836

Table 3. 5% Level critical values ($n = 100$).

	ADF_α test	ADF_t test
$c = -2.5$	−15.79	−2.81
$c = -5$	−17.15	−2.91
$c = -7.5$	−18.05	−2.97
$c = -10$	−18.71	−3.02
$c = -12.5$	−19.25	−3.07
$c = -13.5$	−19.47	−3.09
$c = -15$	−19.91	−3.11
OLS detrending	−20.7	−3.45

(1991). The starting dates for the series vary from 1860 for industrial production and consumer prices through to 1909 for GNP. All series terminate in 1970 in the original Nelson–Plosser data.

© Royal Economic Society 1998

Table 4. OLS detrended ADF_α and ADF_t tests.

Series	$\hat{\alpha}$	ADF_α	ADF_t	Series	$\hat{\alpha}$	ADF_α	ADF_t
CPI	0.998	−5.23	−1.4	Employment	0.854	−19.38	−3.28
GNP Def.	0.967	−6.44	−1.63	GNP/Cap.	0.81	−24.12[a]	−3.59[a]
Ind. Prod.	0.818	−25.8[a]	−3.68[a]	Interest rate	0.94	−6.01	−1.69
Money	0.936	−18.5	−2.89	Real GNP	0.812	−19.68	−3.05
Nom. GNP	0.938	−8.87	−2.03	Real wage	0.927	−8.49	−1.73
Stock price	0.916	−12.4	−2.42	Unemployment	0.772	−43.55[a]	−3.94[a]
Velocity	0.964	−4.62	−1.44	Nominal wage	0.933	−11.56	−2.43

[a] Values are smaller than the 5% level critical values.

Table 5. QD detrended ADF_α and ADF_t tests with a linear trend ($c = -10$).

Series	$\hat{\alpha}$	ADF_α	ADF_t	Series	$\hat{\alpha}$	ADF_α	ADF_t
CPI	0.99	–3.21	–1.04	Employment	0.88	–15.5	–2.76
GNP Def.	0.98	–3.62	–1.13	GNP/Cap.	0.86	–16.74	–2.88
Ind. Prod.	0.87	–17.4	–2.92	Interest rate	0.95	–5.58	–1.61
Money	0.94	–17.5	–2.87	Real GNP	0.87	–16.85	–2.9
Nom. GNP	0.95	–7.56	–1.85	Real wage	0.94	–6.96	–1.73
Stock price	0.95	–6.89	–1.71	Unemployment	0.77	–43.6[a]	–3.96[a]
Velocity	0.98	–2.43	–0.93	Nominal wage	0.95	–8.89	–2.04

[a] Values are smaller than the 5% level critical values.

Table 6. QD detrended ADF_α and ADF_t tests with a linear trend ($c = -13.5$).

Series	$\hat{\alpha}$	ADF_α	ADF_t	Series	$\hat{\alpha}$	ADF_α	ADF_t
CPI	0.99	–3.52	–1.07	Employment	0.88	–16.5	–2.86
GNP Def.	0.978	–4.11	–1.19	GNP/Cap.	0.85	–18.6	–3.05
Ind. Prod.	0.863	–19.1	–3.05	Interest rate	0.94	–5.7	–1.63
Money	0.937	–17.9	–2.89	Real GNP	0.85	–18.5	–3.04
Nominal GNP	0.944	–7.99	–1.89	Real wage	0.94	–7.56	–1.74
Stock price	0.946	–7.86	–1.81	Unemployment	0.77	–43.7[a]	–3.95[a]
Velocity	0.98	–2.67	–0.97	Nominal Wage	0.94	–9.6	–2.13

[a] Values are smaller than the 5% level critical values.

Schotman and Van Dijk extended all these 14 series to 1988. In their original study, Nelson and Plosser conducted the ADF_t test on these series and could not reject the unit root hypothesis at the 5% level of significance for all of the series except the unemployment rates. Perron (1988) arrived at similar conclusions using Z-tests.

We consider the null hypothesis that the variables are difference stationary ARMA processes versus the trend stationary alternatives. We use three detrending procedures for the ADF_α test:

© Royal Economic Society 1998

Table 7. 5% Level finite sample critical values ($n = 200$).

	ADF_α test	ADF_t test
$c = -10$	-17.00	-2.88
$c = -13.5$	-17.60	-2.92
$c = -20$	-18.43	-2.99
$c = -25$	-19.03	-3.05
OLS detrending	-21.20	-3.44

Table 8. OLS detrended tests on post-war quarterly US data.

Series	Estimated AR coefficient	ADF_α	ADF_t
Real GDP	0.97	-8.5	-1.94
Real investment	0.928	-37.28[a]	-3.84[a]
Real consumption	0.938	-14.77	-3.07
Employment	0.95	-18.58	-3.114

[a] Values are smaller than the 5% level critical values.

Table 9. QD detrended tests on post-war quarterly US data, $c = -10$.

Series	Estimated AR coefficient	ADF_α	ADF_t
Real GDP	0.98	-3.88	-1.17
Real investment	0.969	-14.67	-2.339
Real consumption	0.98	-4.336	-1.4
Employment	0.979	-10.11	-2.199

(T1): OLS detrending
(T2): QD detrending with the choice $c = -10$
(T3): QD detrending with the choice $c = -13.5$.
Thus, in the first test, we estimate the following ADF regression

$$\Delta y_t = a y_{t-1} + \beta_1 \Delta y_{t-1} + \cdots + \beta_k \Delta y_{t-k} + \gamma_0 + \gamma_1 t + e_t.$$

In the second and third tests, we run ADF regression (6) for the QD detrended data y_t^*. The value $c = -10$ was chosen because the sample sizes of the Nelson–Plosser series are around 100 (80–129) and estimates of autoregressive coefficients in economic time series are often around 0.9, corresponding to $1 + c/n$ for $n = 100$, $c = -10$. Also the c value for which local asymptotic power is 50% is approximately -13.5 for the case of a linear trend (Elliott *et al.*1996), so this value of c is another natural choice. To provide a basis for comparison, we also calculated the ADF_t statistics based on these three detrending procedures. Although theoretically the lag length of the ADF regression should grow at a rate $o(n^{1/3})$, this rate does not provide much information about lag length selection for specific sample sizes. Monte Carlo evidence shows that use of model selection methods are useful in this respect and provide some improvement in the finite sample

© Royal Economic Society 1998

Table 10. QD detrended tests on post-war quarterly US data, $c = -13.5$.

Series	Estimated AR coefficient	ADF_α	ADF_t
Real GDP	0.98	−4.43	−1.23
Real investment	0.96	−17.93	−2.55
Real consumption	0.977	−5.277	−1.55
Employment	0.976	−11.6	−2.36

Table 11. QD detrended tests on post-war quarterly US data, $c = -20$.

Series	Estimated AR coefficient	ADF_α	ADF_t
Real GDP	0.98	−5.1	−1.33
Real investment	0.956	−21.58[a]	−2.79
Real consumption	0.972	−6.528	−1.758
Employment	0.973	−13.26	−2.53

[a]Values are smaller than the 5% level critical values.

Table 12. QD detrended tests on post-war quarterly US data, $c = -25$.

Series	Estimated AR coefficient	ADF_α	ADF_t
Real GDP	0.98	−5.79	−1.44
Real investment	0.969	−14.67	−2.34
Real consumption	0.966	−7.926	−1.98
Employment	0.969	−14.8	−2.7

performance of the ADF_t test. We use the BIC criterion of Schwarz (1978) and Rissanen (1978) in selecting the appropriate lag length of the autoregression for all three data sets considered in this paper. Since critical values of these QD detrended tests are not available in the existing literature, we calculate the critical values of the QD detrended ADF_α and ADF_t tests corresponding to different choices of c values by simulation experiments with data generated by (1) with $\alpha = 1$ and Gaussian white noise u_t, based on 15 000 replications. Table 3 and Table 7 provides the finite sample critical values for the cases of $N = 100$ and $N = 200$ respectively.

Table 4 reports the values of the ADF tests based on OLS detrending. Table 5 and Table 6 give their values under QD detrending for $c = -10$ and $c = -13.5$. The estimated autoregressive coefficients are reported in the columns labelled '$\hat{\alpha}$'. The two tests based on $c = -10$ and -13.5 give qualitatively the same results. We are interested in testing whether or not the AR coefficient differs from unity. For most of the time series, we can not reject the null of unit root at the 5% level of significance. A few series exhibit values of ADF_α below the 5% level critical values. In particular, the unit root hypothesis is rejected for the unemployment series by all these tests (i.e. all three detrending procedures). For two series, per capita GNP and industrial production, unit

© Royal Economic Society 1998

roots were rejected in the OLS detrended test, but not rejected in the QD detrending procedure. However, the calculated test statistics are very close to the corresponding critical values. Thus the evidence is marginal for these two series. The ADF_t test gives qualitatively the same results. In conclusion, our results in Tables 4–6 are generally in accord with the findings in Nelson and Plosser (1982).

4.2. Post-war quarterly US data

In this section, we analysed some post-war quarterly US macroeconomic time series data. The data set consists of Real GDP, Real Investment, Real Consumption, and Employment. All these variables are from Citibase, over the period 1947:1–1993:4. The number of observations for these time series is 188. Table 7 gives the finite sample critical values for the case of $n = 200$. These critical values are calculated from simulation based on 15 000 replications. We tried the following detrending procedures for both ADF_α and ADF_t tests:

(T1): OLS detrending
(T2): QD detrending with the choice $c = -10$
(T3): QD detrending with the choice $c = -13.5$
(T4): QD detrending with the choice $c = -20$
(T5): QD detrending with the choice $c = -25$.

Tables 8–12 give the estimated test statistics and coefficients for these five detrending procedures. We can not reject the null hypothesis of a unit root in all these tests at the 5% level of significance for the consumption series, which, as argued in Hall (1978), should behave as a martingale. Thus, there is no evidence to reject the hypothesis that consumption behaves as a unit root process. We also find support for the hypothesis of a unit root in the series of real GDP, and employment in all these tests. For the series of real investment, the unit root hypothesis is rejected in the OLS detrended ADF_α and ADF_t tests. In QD detrending cases, when we choose $c = -13.5, -20$, the unit root hypothesis is rejected in the real investment series by the ADF_α test, but not by ADF_t test. For the values $c = -10, -25$, we can not reject a unit root hypothesis in any series. These results are generally in agreement with the conclusion of the extended Nelson–Plosser data that many macroeconomic time series are characterized by the presence of a unit root.

ACKNOWLEDGEMENTS

We would like to thank the coordinating editor, Neil Shephard, and the two referees for helpful comments on earlier versions of this paper. The paper was typed by the authors in Scientific Word 2.5 and computations were performed in Gauss. Xiao thanks the Cowles Foundation for support under a C.A. Anderson Fellowship and Phillips thanks NSF for research support under Grant No. SBR 94-22922.

© Royal Economic Society 1998

A. APPENDIX

A.1. Proof of Theorem 1

The limit distribution of $\hat{a}$ can be established in the following steps by using the BN (Beveridge and Nelson 1981) decomposition for the operators $a(L)$ and $b(L)$. Following the lines of Berk (1974), we use the standard Euclidean norm, $\|x\| = (x'x)^{1/2}$, of a column vector and use the matrix norm $\|B\| = \sup\{\|Bx\| : \|x\| < 1\}$.

Let $G_n = \operatorname{diag}(n^{-1}, n^{-1/2}, \ldots, n^{-1/2})$, then

$$G_n^{-1}(\hat{\beta} - \beta) = (G_n Z' Z G_n)^{-1} G_n Z' e_k.$$

If $k = o(n^{1/3})$ and k goes to ∞ with n, then, under the null hypothesis, we have:

(a) $k^{1/2}\|G_n Z' Z G_n - R_n\| \xrightarrow{p} 0$ and $k^{1/2}\|(G_n Z' Z G_n)^{-1} - R_n^{-1}\| \xrightarrow{p} 0$, as $n, k \to \infty$ (Said and Dickey1984), where

$$R_n = \operatorname{diag}\left[n^{-2}\{b(1)/a(1)\}^2 \sum S_{t-1}^2, \Gamma\right],$$

$$S_{t-1} = \sum_{j=1}^{t-1} \varepsilon_j,$$

$$\Gamma = (\gamma_{ij})_{n\times n}, \ \gamma_{ij} = E(u_i u_j);$$

(b) $\|G_n Z' e_k - G_n Z' \varepsilon\| = O_p(1/n)$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)'$ (Said and Dickey1984);

(c) $\|\hat{\beta} - \beta\|$ converges in probability to 0 (Said and Dickey1984).

Under H_0, we have

$$a(L) y_t = a(L) \sum_{j=1}^{t} u_j + O_p(1) = b(L) \sum_{j=1}^{t} \varepsilon_j + O_p(1).$$

Use the BN decomposition, giving

$$a(1) y_t = b(1) \sum_{j=1}^{t} \varepsilon_j + O_p(1) \tag{8}$$

the term $O_p(1)$ including finite linear combinations of the variates u_t and ε_t. Since $\sum_1^n \varepsilon_j$ is the $I(1)$ component in (8), we obtain

$$y_t = \frac{b(1)}{a(1)} S_t + O_p(1).$$

From (a), (b) and (c), the limit distribution of $G_n^{-1}(\hat{\beta} - \beta)$ is the same as that of $R_n^{-1} G_n Z' \varepsilon$. Thus the limit of $n\hat{a}$ is the same as that of the first element in $R_n^{-1} G_n Z' \varepsilon$, which is

$$\left(n^{-2} \sum y_{t-1}^2\right)^{-1} \left(n^{-1} \sum y_{t-1} \varepsilon_t\right) = \frac{a(1)}{b(1)} \left(n^{-2} \sum S_{t-1}^2\right)^{-1} \left(n^{-1} \sum S_{t-1} \varepsilon_t\right).$$

Notice that $n^{-2} \sum S_{t-1}^2 \Rightarrow \int B_\varepsilon^2 = \sigma^2 \int W^2$, $n^{-1} \sum S_{t-1} \varepsilon_t \Rightarrow \int B_\varepsilon dB_\varepsilon = \sigma^2 \int W dW$, and $\omega^2 = 2\pi f_{uu}(0) = \sigma^2 \{b(1)/a(1)\}^2$, where f_{uu} is the spectral density of u_t defined as $f_{uu}(\lambda) = (2\pi)^{-1} \sum_{n=-\infty}^{\infty} \gamma(h) e^{-ih\lambda}$, and thus

$$n\hat{a} \Rightarrow \frac{\sigma \int W dW}{\omega \int W^2}.$$

□

© Royal Economic Society 1998

A.2. Proof of Theorem 2

Under the alternative $H_1 : \alpha < 1$, notice that u_t is a stationary and invertible ARMA process, $f_{yy}(0) = \omega^2/\{2\pi(1-\alpha)\}^2 > 0$, y_t has a representation

$$\sum_{j=0}^{\infty} c_j y_{t-j} = e_t, c_0 = 1 \tag{9}$$

where $(e_t) = WN(0, \sigma_e^2)$. Following Fuller (1976), we can write (9) as

$$\Delta y_t = (\theta_1 - 1)y_{t-1} + \theta_2 \Delta y_{t-1} + \theta_3 \Delta y_{t-2} + \cdots + e_t$$

where $\theta_i = \sum_{j=i}^{\infty} c_j$ $(i = 2, 3, \ldots)$ and $\theta_1 = -\sum_{j=1}^{\infty} c_j$. Since $f_{yy}(0) > 0$ and y_t is stationary, $\theta_1 - 1 \neq 0$. In the ADF regression, as $k \to \infty$, we find that

$$\hat{a} \xrightarrow{p} \theta_1 - 1 \neq 0$$
$$\hat{\sigma}^2 \xrightarrow{p} \sigma_e^2 > 0$$
$$\hat{\omega}^2 \xrightarrow{p} 2\pi f_{uu}(0) > 0.$$

Thus, $ADF_\alpha = n(\hat{\omega}/\hat{\sigma})\hat{a} \sim n\{2\pi f_{uu}(0)/\sigma_e^2\}^{1/2}(\theta_1 - 1) = O_p(n)$. The divergence rate is sharp because $2\pi f_{uu}(0)/\sigma_e^2 > 0$ and $\theta_1 - 1 \neq 0$. □

A.3. Proof of Lemma 1

$$\begin{aligned} n^{-1/2} y_{(nr)}^{s*} &= n^{-1/2} y_{(nr)}^{s} - n^{-1/2} x_{(nr)}' (\Delta_c X' \Delta_c X)^{-1} \Delta_c X' \Delta y^s \\ &= n^{-1/2} y_{(nr)}^{s} - \{x_{(nr)} D_n\}(n^{-1} F_n \Delta_c X' \Delta_c X F_n)^{-1} F_n \Delta_c X' (n^{-1/2} \Delta y^s). \end{aligned}$$

Notice that $n^{-1/2} y_{(nr)}^s \Rightarrow B(r)$ and let $F_n = nD_n$, then

$$\begin{aligned} n^{-1} F_n \Delta_c X' \Delta_c X F_n &= n^{-1} F_n (\Delta X' \Delta X - n^{-1} c \Delta X' X_{-1} - n^{-1} c X_{-1}' \Delta X \\ &\quad + n^{-2} c^2 X_{-1}' X_{-1}) F_n \\ &\Rightarrow \int \{g(r)g(r)' - cX(r)g(r)' - cg(r)X(r)' + c^2 X(r)X(r)'\}\, dr \\ &= \int X_c(r) X_c(r)'\, dr \end{aligned}$$

and

$$\begin{aligned} F_n \Delta_c X' n^{-1/2} \Delta y^s &= n^{-1/2} F_n (\Delta X' \Delta y^s - n^{-1} c \Delta X' y_{-1}^s \\ &\quad - n^{-1} c X_{-1}' \Delta y^s + n^{-2} c^2 X_{-1}' y_{-1}^s) \\ &\Rightarrow \int \{g(r)\, dB(r) - cX(r)\, dB(r) - cg(r)B(r) + c^2 X(r)B(r)\} \\ &= \int X_c(r)\, dB(r) - c \int X_c(r) B(r)\, dr. \end{aligned}$$

Thus,

$$\begin{aligned} n^{-1/2} y_{(nr)}^{*} &\Rightarrow B(r) - X(r)' \left\{ \int X_c(r) X_c(r)'\, dr \right\}^{-1} \left\{ \int X_c(r)\, dB(r) - c \int X_c(r) B(r)\, dr \right\} \\ &= \underline{B}_c(r). \end{aligned}$$

□

© Royal Economic Society 1998

A.4. Proof of Theorem 3 and Theorem 4

The proofs of Theorems 3 and 4 are similar, so we show the argument for Theorem 4. We know that $\tilde{a} = (y^{*\prime}_{-1} P_M y^*_{-1})^{-1} y^{*\prime}_{-1} P_M \Delta y^*$, where $P_M = I - M(M'M)^{-1}M'$, M is the matrix of the k lagged difference variables $(\Delta y^*_{t-1}, \ldots, \Delta y^*_{t-k})$. We have

$$\begin{aligned} n^{-2} y^{*\prime}_{-1} P_M y^*_{-1} &= n^{-2} y^{*\prime}_{-1} \{I - M(M'M)^{-1}M'\} y^*_{-1} \\ &= n^{-2} y^{*\prime}_{-1} y^*_{-1} - n^{-1}(n^{-1} y^{*\prime}_{-1} M)(n^{-1}M'M)^{-1}(n^{-1}M' y^*_{-1}) \\ &= n^{-2} y^{*\prime}_{-1} y^*_{-1} + o_p(1) \\ &\Rightarrow \int \underline{B}_c(r)^2 \\ &= \omega^2 \int \underline{W}_c(r)^2. \end{aligned}$$

Notice that $P_M M = 0$,

$$n^{-1} y^{*\prime}_{-1} P_M \Delta y^* = n^{-1} y^{*\prime}_{-1} P_M (y^*_{-1} a + e^*_k)$$

where $e^*_k = (\ldots, e^*_{tk}, \ldots)$. Under the null that $a = 0$, $n^{-1} y^{*\prime}_{-1} P_M \Delta y^* = n^{-1} y^{*\prime}_{-1} P_M e^*_k$. Since u_t is stationary and invertible, there exists a sequence of real numbers d_j and numbers M and $0 < \lambda < 1$ such that $|d_j| < M\lambda^j$ (Fuller1976) and

$$\Delta y^s_t = a y^s_{t-1} - d_1 u_{t-1} - d_2 u_{t-2} - \cdots + \varepsilon_t.$$

We denote

$$e^*_t = \Delta y^*_t - a y^*_{t-1} + d_1 u^*_{t-1} + d_2 u^*_{t-2} + \cdots$$

and $e^* = (\ldots, e^*_t, \ldots)$, and y^*_t is the detrended time series and $u^*_t = \Delta y^*_t$. Notice that k is bounded below by a positive multiple of $n^{1/r}$ for some $r > 0$ and there are exponentially decreasing bounds on the d_i, by a similar argument to that of Said and Dickey (1984). Hence, we can show that

$$n^{-1} y^{*\prime}_{-1} P_M e^*_k - n^{-1} y^{*\prime}_{-1} P_M e^* = o_p(1)$$

and

$$n^{-1} y^{*\prime}_{-1} P_M \Delta y^* = n^{-1} y^{*\prime}_{-1} P_M \varepsilon + o_p(1) \Rightarrow \int \underline{B}_c(r)\, dB_\varepsilon(r) = \omega\sigma \int \underline{W}_c(r)\, dW(r).$$

It follows that

$$ADF^*_\alpha \Rightarrow \int \underline{W}_c(r)\, dW(r) \Big/ \int \underline{W}_c(r)^2. \qquad \square$$

REFERENCES

Berk, K. N. (1974). Consistent autoregressive spectral estimates. *Ann. Stat. 2*, 489–502.

Beveridge, S. and C. R. Nelson (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. *J. Monetary Economics 7*, 151–174.

Canjels, E. and M. W. Watson (1997). Estimating deterministic trends in the presence of serially correlated error. *Rev. Econ. Stat. 79*, 184–200.

DeJong, D. N., J. C. Nankervis, N. E. Savin and C. H. Whiteman (1988). Integration versus trend-stationarity in macroeconomic time series. Department of Economics, University of Iowa, Iowa City, IA.

DeJong, D. N. and C. H. Whiteman (1991). The temporal stability of dividends and stock prices, evidence from the likelihood function. *Am. Econ. Review 81*, 600–617.

© Royal Economic Society 1998

DeJong, D. N., J. C. Nankervis, N. E. Savin and C. H. Whiteman (1992). The power problems of unit root tests in time series with autoregressive errors. *J. Econometrics 53*, 323–43.

Dickey, D. A., and W. A. Fuller (1979). Distribution of estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc. 74*, 427–431.

Diebold, F.X. and G.D. Rudebusch (1991), On the power of Dickey–Fuller tests against fractional alternatives. *Econ. Lett. 35*, 155–160.

Dufour, J.M., and M.L. King (1991). Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary or nonstationary AR(1) errors. *J. Econometrics 47*, 115–143

Elliott, G., Rothenberg, T. J., and J. H. Stock (1996). Efficient tests for an autoregressive unit root. *Econometrica 64*, 813–836.

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

Grenander, U. and M. Rosenblatt (1957). *Statistical Analysis of Stationary Time Series*. New York: Wiley.

Hall, R. E. (1978). Stochastic implications of the life cycle permanent income hypothesis, theory and evidence. *J. Political Economy 86*, 971–988.

King, M.L. (1988). Towards a theory of point optimal testing. *Econometric Reviews 6*, 169–218

Lee, C. C., and P. C. B. Phillips (1994). Efficiency gains using QD over OLS under nonstationarity, mimeo, Yale University.

Nelson, C. R. and C. I. Plosser (1982). Trends and random walks in macroeconomic time series. *J. Monetary Economics 10*, 139–162.

Ng, S. and P. Perron (1995). Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *J. Am. Stat. Assoc. 90*, 268–281.

Perron, P. (1988). Trends and random walks in macroeconomic time series: further evidence from a new approach. *J. Economic Dynamics and Control 12*, 297–332.

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

Phillips, P. C. B. (1987b). Towards a unified asymptotic theory for autoregression. *Biometrika 74*, 535–547.

Phillips, P. C. B., and P. Perron (1988). Testing for a unit root in time series regression. *Biometrika 75*, 335–346.

Phillips, P. C. B. (1995). Nonstationary time series and cointegration. *J. Applied Econometrics 10*, 87–94.

Phillips, P. C. B. (1995). *A Primer on Unit Root Tests*, mimeo , Yale University.

Phillips, P. C. B. and C. C. Lee (1996). Efficiency gains from quasi-differencing under nonstationarity. In P. M. Robinson and M. Rosenblatt (Eds). *Essays in Memory of E. J. Hannan*. New York: Springer-Verlag.

Phillips, P. C. B. and J. Y. Park (1988). On the formulation of Wald tests of nonlinear restrictions. *Econometrica 56*, 1065–1085.

Rissanen, J. (1978). Modeling by shortest data description. *Automatica 14*, 465–471.

Said, S. E., and D. A. Dickey (1984). Testing for unit roots in autoregressive-moving average model of unknown order. *Biometrika 71*, 599–608.

Schotman, P. and H. K. Van Dijk (1991). On the Bayesian routes to unit root. *J. Applied Econometrics 6*, 387–402.

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat. 6*, 461–464.

Schwert, G. W. (1989). Tests for unit roots: A Monte Carlo investigation. *J. Business and Economic Statistics 7*, 147–160.

Stock, J. (1995). Unit roots, structural breaks and trends. *Handbook of Econometrics 4*, 2739–2841

Xiao, Z. and P.C.B. Phillips (1997). An ADF coefficient test for a unit root in ARMA models of unknown order with empirical applications to the US economy, Cowles Foundation Discussion Paper No. 1161.

© Royal Economic Society 1998