ISSN 0709-9231

**CAHIER 8632**

# Tests of Joint Hypotheses for Time Series Regression with a Unit Root

**by**

Pierre Perron[1]

[1] Department of Economics and C.R.D.E., Université de Montréal.

June 1986

This is a revised and extended version of Chapter 3 of the author's Ph.D. dissertation (Yale University). The author wishes to thank Peter C.B. Phillips for most helpful supervision and comments.

Cette étude a été publiée grâce à une subvention du fonds F.C.A.R. pour l'aide et le soutien à la recherche; ce cahier a également été publié par le Centre de recherche et développement en économique (Publication #2086).

## RÉSUMÉ - ABSTRACT

Cette recherche étudie des tests d'hypothèses multiples pour les régressions de séries chronologiques avec racine sur le cercle unitaire pour lesquelles les erreurs peuvent être dépendantes et hétérogènes. Nous considérons deux types de régression: une avec une constante et une variables dépendante retardée et l'autre avec un trend ajouté. Les statistiques étudiées sont les "tests-F" analysés originellement par Dickey et Fuller (1981) dans un contexte moins général. La distribution asymptotique est dérivée utilisant la théorie de convergence fonctionnelle. De nouveaux tests sont proposés qui ne requièrent que des valeurs critiques déjà compilées mais qui sont valides dans un contexte très général (incluant tous les processus ARMA d'ordre fini générés par des erreurs Gausiennes. Cette étude s'ajoute aux résultats d'hypothèse simple sur les coefficients de ces régressions dérivés par Phillips (1986) et Phillips et Perron (1986).



Mots-clés: racine unitaire; séries chronologiques; convergence fonctionnelle; processus de Weiner; tests d'hypothèses multiples


* * * * * *


This paper studies tests of joint hypotheses in time series regression with a unit root in which weakly dependent and heterogeneously distributed innovations are allowed. We consider two types of regression: one with a constant and lagged dependent variable, and the other with a trend added. The statistics studied are the regression "F-test" originally analysed by Dickey and Fuller (1981) in a less general framework. The limiting distributions are found using functional central limit theory. New test statistics are proposed which require only already tabulated critical values but which are valid in a quite general framework (including finite order ARMA models generated by Gaussian errors). This study extends the results on single coefficients derived in Phillips (1986a) and Phillips and Perron (1986).



Key words : non-stationarity; functional limit theory; Weiner process; asymptotic local power.




## 1. INTRODUCTION

Consider the following two autoregressive models

$$y_t = \mu + \alpha y_{t-1} + u_t \qquad (1)$$

$$t = 1, 2, \ldots$$

$$y_t = \mu + \beta(t-T/2) + \alpha y_{t-1} + u_t \qquad (2)$$

where $y_0$ is either a real constant (with probability one) or has a certain prespecified distribution. The null hypothesis to be considered is that $\alpha = 1$, i.e. the series $\{y_t\}_0^\infty$ has a unit root.

Let $\hat{\alpha}$ and $\hat{\mu}$ be the least squares estimators of the coefficients in regression model (1) and $t\hat{\alpha}$, $t\hat{\mu}$ their associated t-statistics. Similarly let $\tilde{\alpha}$, $\tilde{\mu}$, $\tilde{\beta}$, $t\tilde{\alpha}$, $t\tilde{\mu}$ and $t\tilde{\beta}$ be the corresponding statistics for regression model (2). Dickey and Fuller (1979, 1981) derived the limiting distribution of $T(\hat{\alpha} - 1)$, $T(\tilde{\alpha} - 1)$ and the various t-statistics under the maintained hypothesis that $\mu = 0$ in (1) and $\beta = 0$ (with $\mu$ a fixed constant) in (2) restricting the errors in both models to be i.i.d. $(0, \sigma^2)$. They also showed that the limiting distribution of the t-statistics is the same when the process generating $\{y_t\}_0^\infty$ is that of an AR(p) if (p - 1) lags of $\Delta y_t = y_t - y_{t-1}$ are added as regressors. Said and Dickey (1984) showed that the limiting distribution of $t\hat{\alpha}$ is also the same for a sequence $\{y_t\}$ generated by an ARMA (p,q) if the number of lags of first-differences added as regressors is an increasing function of the sample size.

Recently, Phillips (1986a) introduced a new approach valid under quite general conditions on the innovation sequence $\{u_t\}$ and which does not require the estimation of additional nuisance parameters. Phillips and Perron (1986) extended his results to the case of regression models (1) and (2) by considering the limiting distribution of the least squares estimators and their t-statistics, and providing a new set of test statistics.

As is well known, the limiting distribution of $\hat{\alpha}$ and $t\hat{\alpha}$ is not invariant with respect to $\mu$ (the drift) under the null hypothesis of a unit root (if $\mu \neq 0$, the limiting distribution is normal). Similarly the limiting distribution of $\tilde{\alpha}$ and $t\tilde{\alpha}$ is not invariant with respect to $\beta$ (the trend) though it is invariant with respect to $\mu$ if one is ready to assumne it fixed. Accordingly it is advisable to have statistics which permit tests of joint hypotheses, in particular for testing $H_0^1$: $(\mu, \alpha) = (0, 1)$ in model (1) and $H_0^3$: $(\mu, \beta, \alpha) = (\mu, 0, 1)$ in model (2). We shall also consider the null hypothesis $H_0^2$: $(\mu, \beta, \alpha) = (0, 0, 1)$ in model (2). The regression "F-test" for these hypotheses are

$$H_0^1:\ \Phi_1 = (2\hat{S}^2)^{-1}\left[T\,S_0^2 - (T-2)\,\hat{S}^2\right]$$
$$H_0^2:\ \Phi_2 = (3\tilde{S}^2)^{-1}\left[T\,S_0^2 - (T-3)\,\tilde{S}^2\right]$$
$$H_0^3:\ \Phi_3 = (2\tilde{S}^2)^{-1}\left[T\{S_0^2 - (\bar{Y}_{(0)} - \bar{Y}_{(-1)})^2\} - (T-3)\,\tilde{S}^2\right]$$

where $S_0^2 = T^{-1}\sum_{t=1}^{T}(y_t - y_{t-1})^2$, $\hat{S}^2 = (T-2)^{-1}\sum_{t=1}^{T}(y_t - \hat{\mu} - \hat{\alpha}y_{t-1})^2$

$\tilde{S}^2 = (T-3)^{-1}\sum_{t=1}^{T}(y_t - \tilde{\mu} - \tilde{\beta}(t - T/2) - \tilde{\alpha}\, y_{t-1})^2$ , $\bar{Y}_{(-i)} = T^{-1}\sum_{t=1}^{T} y_{t-i}$

These statistics were studied by Dickey and Fuller (1981) who derived their asymptotic distribution under the hypothesis of i.i.d. errors $\{u_t\}$. They also showed that the same asymptotic distribution holds for an AR(p) process with a unit root if (p - 1) extra lags of first differences are added as regressors. We shall be concerned here about the limiting distribution of $\Phi_1$, $\Phi_2$ and $\Phi_3$ under more general specifications of the innovation sequence $\{u_t\}$ (allowing, in particular, all finite ARMA processes with basic errors Gaussian).

The plan of the paper is as follows. Section 2 discusses the conditions imposed upon the innovation sequence $\{u_t\}$ and the functional central limit theorem which underlies the analysis. Section 3 derive the limiting distribution of $\Phi_1$, $\Phi_2$ and $\Phi_3$ under



the respective null hypotheses and section 4 considers transformations of these statistics which permit hypothesis testing in the more general structure but which requires only critical values already tabulated by Dickey and Fuller (1981). Section 5 contains a discussion of the local asymptotic power of the tests. Section 6 provides some applied examples. The proofs of the theorems are collected in a mathematical appendix.

## 2. ASSUMPTIONS AND THE INVARIANCE PRINCIPLE

This section describes the conditions to be imposed on the innovation sequence $\{u_t\}$ and presents the results concerning weak convergence to a Weiner process needed in the following section. The discussion is brief; a more thorough treatment can be found in Phillips (1986a) and Phillips and Perron (1986).

Let $\{u_t\}_1^\infty$ be a sequence of random variable in a probability space $(\Omega, B, P)$. We introduce the partial sums $S_t = \sum_{j=1}^{t} u_j$ and set $S_0 = 0$. Consider the following measure of temporal dependence

$$\alpha_m = \sup_n \sup_{j \geq n+m} \alpha(F_1^n, R_{n+m}^j)$$

where $$\alpha(F, G) = \sup_{f \in F,\, g \in G} |P(fg) - P(f)P(g)|$$

and $F_a^b$ denotes the $\sigma$-field generated by $\{u_a, u_{a+1}, \cdots, u_b\}$ and $R_a^b$ the $\sigma$-field generated by $\{u_a + u_{a+1} + \ldots + u_b = S_b - S_{a-1}\}$ for all $a \leq b$. A series is $\alpha$-mixing (or strong mixing) if $\alpha_m \downarrow 0$ as $m \uparrow \infty$. We now impose the following conditions on $\{u_t\}_1^\infty$ which characterize the extent of heterogeneity and dependence permitted in this study.

Assumption 2.1

(a) $E(u_t) = 0$ all t;

(b) $\sup_t E|u_t|^{\beta+\epsilon} < \infty$ for some $\beta > 2$ and $\epsilon > 0$;

(c) $\sigma^2 = \lim_{T\to\infty} T^{-1} E(S_T^2)$ exists and $\sigma^2 > 0$;

(d) $\{u_t\}_1^\infty$ is strong mixing with mixing numbers $\alpha_m$ that satisfy:

$$\Sigma_1^\infty \alpha_m^{1-2/\beta} < \infty$$

These conditions permit a wide variety of possible data generating mechanisms such as finite order ARMA models under very general conditions on the underlying errors. Note also that we do not insist that the series $\{u_t\}$ is stationary. However, if it is stationary, then condition (c) follows from (a), (b) and (d).

From the sequence of partial sums $\{S_t\}_1^T$ we construct the random elements

$$X_T(r) = \sigma^{-1}T^{-\frac{1}{2}}S_{[Tr]} = \sigma^{-1}T^{-\frac{1}{2}}S_{j-1}; \quad ((j-1)/T \leq r < j/T, \; j=1,\dots, T)$$

$$X_T(1) = \sigma^{-1}T^{-\frac{1}{2}}S_T$$

where [a] denotes the integer part of a. $X_T(r)$ lies in D[0,1], the space of all real-valued functions on the interval [0,1] that are right continuous and have finite left limits. D is endowed with the Skorokhod topology with Borel-$\sigma$-algebra under which D is separable and complete (see Billingsley, 1968, Section 14). The following lemma, due to Herrndorf (1984, Corrolary 1), will be used extensively.

<u>Lemma 2.2</u>: If $\{u_t\}_1^\infty$ is a sequence of random variables satisfying Assumption 2.1, then $X_T(r) \to w(r)$, a Weiner process, as $T \uparrow \infty$.

The notation "$\to$" is used to signify the weak convergence of the probability measure of $X_T(r)$ to the probability measure (here,



univariate Weiner measure) of the random function w(r). The result is a functional central limit theorem (CLT), i.e. a CLT on the function space D. Note that the Weiner process w(r) (also known as Brownian notion) has sample paths which lie almost surely (Weiner measure) in C = C[0,1], the space of all real-valued continuous functions on [0,1]. Moreover, w(r) is Gaussian with independent increments.

We shall also use the following result discussed in Billingsley (1968), Theorem 5.1:

<u>Lemma 2.3</u>: If h is any continuous functional on C[0,1] (continuous that is, except for at most a set of points $D_h$ for which $P(w \varepsilon D_h) = 0$), then $X_T(r) \to w(r)$ implies $h(X_T(r)) \to h(w(r))$.

Now, consider $\{y_t\}$ a stochastic process of length T + 1, generated in discrete time according to the following model:

a) $y_t = \alpha y_{t-1} + u_t$ $\quad (t = 1, \dots, T)$

b) $\alpha = 1$

(3) c) either i) $y_0 = k$, a constant with probability one

or ii) $y_0$ has a certain prespecified distribution (independent of T)

and d) $\{u_t\}$ satisfies assumption (2.1)

Note that the statistic $\Phi_3$ is invariant with respect to the introduction of a non-zero fixed drift $\mu$ if the stochastic process is generated by

(3.a') $y_t = \mu + \alpha y_{t-1} + u_t$

instead of (3.a). There is, however, no loss of generality in postulating $\mu = 0$ and we consider the limiting distribution of all test statistics under this latter hypothesis.

The statistics $\Phi_1$, $\Phi_2$ and $\Phi_3$ are functions of five basic elements which are themselves summations of functions of the variables in the sequence $\{y_t\}$. The following lemma gives the limiting distribution of these basic elements in terms of functionals of the Weiner process (see Phillips and Perron (1986) for a proof).

<u>Lemma 2.4</u>: If $\{y_t\}_0^T$ is a sequence of random variables generated according to model (3), then as $T \uparrow \infty$:

a) $T^{-3/2} \sum_{t=1}^{T} y_t \rightarrow \sigma\int_0^1 w(r)dr$

b) $T^{-2} \sum_{t=1}^{T} y_t^2 \rightarrow \sigma^2\int_0^1 w(r)^2 dr$

c) $T^{-5/2} \sum_{t=1}^{T} t y_{t-1} \rightarrow \sigma\int_0^1 r w(r)dr$

d) $T^{-1} \sum_{t=1}^{T} u_t^2 \underset{a \cdot s}{\longrightarrow} \sigma_u^2$

e) $T^{-1} \sum_{t=1}^{T} u_t\, y_{t-1} \rightarrow (\sigma^2/2)(w(1)^2 - \sigma_u^2/\sigma^2)$

where $\sigma_u^2 \equiv \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} E(u_t^2)$

## 3. THE LIMITING DISTRIBUTION OF THE STATISTICS

The following theorem (proved in the appendix) characterizes the limiting distribution of the statistics $\Phi_1$, $\Phi_2$ and $\Phi_3$.

<u>Theorem 3.1</u>: If $\{y_t\}_0^T$ is a sequence of random variables generated according to model (3), then as $T \uparrow \infty$:

a) $\Phi_1 \rightarrow (\tfrac{1}{2})(\sigma^2/\sigma_u^2)[w(1)^2 + \{\int_0^1 w(r)^2 dr - (\int_0^1 w(r)dr)^2\}^{-1}\{\tfrac{1}{2}(w(1)^2 - \sigma^2/\sigma_u^2) - w(1)\int_0^1 w(r)dr\}^2]$

b) $\Phi_2 \rightarrow (2/3)(\sigma^2/\sigma_u^2)[H' + (C_3/2D)(w(1)^2 - \sigma_u^2/\sigma^2) + (1/8D)(w(1)^2 - \sigma_u^2/\sigma^2)^2]$



c) $\Phi_3 \rightarrow (\sigma^2/\sigma_u^2)[H + (C_3/2D)(w(1)^2 - \sigma_u^2/\sigma^2) + (1/8D)(w(1)^2 - \sigma_u^2/\sigma^2)^2]$

where $H = (\tfrac{1}{2})[-w(1)^2 + 2w(1)C_1/D + C_2(w(1) - 2\int_0^1 w(r)dr)/D$
$\quad -(1/D^2)\{C_1^2 + C_2^2/12 + C_3^2 \int_0^1 w(r)^2 dr + 2C_1C_3 \int_0^1 w(r)dr$
$\quad + 2C_2C_3(\int_0^1 rw(r)dr - (1/2)\int_0^1 w(r)dr)\}]$

$H' = H + (\tfrac{1}{2})w(1)^2$

$C_1 = w(1)[\int_0^1 w(r)^2 dr - 12(\int_0^1 rw(r)dr)^2 + 18\int_0^1 w(r)dr\int_0^1 rw(r)dr - 6(\int_0^1 w(r)dr)^2]$
$\quad + \int_0^1 w(r)dr[6(\int_0^1 w(r)dr)^2 - 12\int_0^1 rw(r)dr \int_0^1 w(r)dr]$

$C_2 = 12w(1)[\tfrac{1}{2}\int_0^1 w(r)^2 dr + \int_0^1 w(r)dr \int_0^1 rw(r)dr - (\int_0^1 w(r)dr)^2]$
$\quad + 12\int_0^1 w(r)dr[(\int_0^1 w(r)dr)^2 - \int_0^1 w(r)^2 dr]$

$C_3 = 12(\int_0^1 rw(r)dr - \tfrac{1}{2}\int_0^1 w(r)dr)(\int_0^1 w(r)dr - w(1)/2) - w(1)\int_0^1 w(r)dr$

$D = \int_0^1 w(r)^2 dr - 12(\int_0^1 rw(r)dr)^2 + 12\int_0^1 w(r)dr\int_0^1 rw(r)dr - 4(\int_0^1 w(r)dr)^2$

Theorem 3.1 shows that the limiting distribution of the statistics considered have the same general form under a wide variety of possible innovation structure $\{u_t\}$. In the case the errors are i.i.d. we have $\sigma^2 = \sigma_u^2$ and the functionals defined in theorem 3.1 are independent of any nuisance parameters. The percentage points of their distribution can therefore be tabulated (See Dickey and Fuller (1981) - Tables IV, V and VI, page 1063. Their approach, however, is different).

Note that the i.i.d. condition is sufficient but not necessary to obtain $\sigma^2 = \sigma_u^2$. The equivalence holds, for example, when the innovations are martingale differences under mild additional conditions (e.g. moment condition (b) of assumption 2.1). Therefore unmodified versions of the Dickey-Fuller tests are valid asymptotically in the presence of some heterogeneity in the innovation sequence. In general, however, these tests will not have the correct asymptotic size when the innovations are such that $\sigma^2 \neq \sigma_u^2$. We must therefore consider some transformations of the $\Phi$-statistics.

## 4. THE LIMITING DISTRIBUTION OF TRANSFORMED STATISTICS

In the general case where $\sigma^2 \neq \sigma_u^2$ we need consistent estimators of $\sigma^2$ and $\sigma_u^2$ since the limiting distributions derived in Section 3 depend on these parameters. $S_0^2$, $\hat{S}^2$ and $\tilde{S}^2$ are consistent estimates of $\sigma_u^2$ in model (3). When there is a non-zero drift in the model, as in (3.a') both $\hat{S}^2$ and $\tilde{S}^2$ are consistent. Since we want to allow for a non-zero drift in regressions such as (2) we shall use $\tilde{S}^2$ as a consistent estimator of $\sigma_u^2$ when the null hypothesis is $H_0^3$. For the consistent estimation of $\sigma^2 = \lim_{T\to\infty} E[T^{-1} S_T^2]$ define

$$S_{T\ell}^2 = T^{-1}\Sigma_1^T u_t^2 + 2T^{-1} \Sigma_{\tau=1}^{\ell} \Sigma_{t=\tau+1}^{T} u_t u_{t-\tau}$$

where, under the null hypothesis, $u_t = y_t - y_{t-1}$. Alternatively, one may use the following estimator proposed by Newey and West (1985) which has the property of yielding a positive value by construction.

$$\sigma_{T\ell}^2 = T^{-1}\Sigma_1^T u_t^2 + 2T^{-1} \Sigma_{\tau=1}^{\ell} w(\tau,\ell) \Sigma_{t=\tau+1}^{T} u_t u_{t-\tau}$$

where $w(\tau,\ell) = 1-[\tau/(\ell + 1)]$. ($\sigma_{T\ell}^2$ is $2\Pi$ time the spectral density estimator of $\sigma^2$ at frequency zero, where a triangular lag window is used to smooth the estimated spectrum). $\sigma_{T\ell}^2$ is consistent for $\sigma^2$ under the same conditions as for $S_{T\ell}^2$ which, however, are stronger than assumption (2.1). We need (for a proof see Phililips (1986b), theorem 4.2):



<u>Assumption 4.1</u>:

(a) $\{u_t\}_1^\infty$ satisfies assumption 2.1(a), (c) and (d); part (b) of assumption 2.1 is replaced by the stronger moment condition

$$\sup_t E|u_t|^{2\beta} < \infty, \text{ for some } \beta > 2;$$

(b) $\ell \to \infty$ as $T \to \infty$ such that $\ell = o(T^{1/4})$.

Rather than using first differences $u_t = y_t - y_{t-1}$ in the construction of $S_{T\ell}^2$ and $\sigma_{T\ell}^2$ we could use the residuals from the regression equations (1) and (2). Since the coefficients in these regressions are consistent it is easy to show that these versions of $S_{T\ell}^2$ and $\sigma_{T\ell}^2$ (denoted $\hat{S}_{T\ell}^2$, $\tilde{S}_{T\ell}^2$, $\hat{\sigma}_{T\ell}^2$ and $\tilde{\sigma}_{T\ell}^2$) are also consistent estimates of $\sigma^2$ under the same conditions as those of Assumption 4.1 when the data are generated by model (3). Note that $\tilde{S}_{T\ell}^2$ and $\tilde{\sigma}_{T\ell}^2$ will be the preferred estimators when we wish to allow for a non-zero drift as in (3.a').

The proposed transformed statistics are defined as follows:

$$(4) \quad Z\Phi_1 = (\hat{S}^2/\hat{\sigma}_{T\ell}^2)\Phi_1 -(1/2\hat{\sigma}_{T\ell}^2)(\hat{\sigma}_{T\ell}^2-\hat{S}^2)[T(\hat{\alpha}-1)-(1/4)(\hat{\sigma}_{T\ell}^2-\hat{S}^2)\{T^{-2}\Sigma(y_{t-1}-\bar{Y}_{-1})^2\}^{-1}]$$

$$(5) \quad Z\Phi_2 = (\hat{S}^2/\hat{\sigma}_{T\ell}^2)\Phi_2-(1/3\ \hat{\sigma}_{T\ell}^2)(\hat{\sigma}_{T\ell}^2-\hat{S}^2)[T(\tilde{\alpha}-1)-(T^6/48D_x)(\hat{\sigma}_{T\ell}^2-\hat{S}^2)]$$

$$(6) \quad Z\Phi_3 = (\tilde{S}^2/\tilde{\sigma}_{T\ell}^2)\Phi_3-(1/2\ \tilde{\sigma}_{T\ell}^2)(\tilde{\sigma}_{T\ell}^2-\tilde{S}^2)[T(\tilde{\alpha}-1)-(T^6/48D_x)(\tilde{\sigma}_{T\ell}^2-\tilde{S}^2)]$$

where $D_x = (T^2(T^2-1)/12)\,\Sigma y_{t-1}^2 - T(\Sigma t y_{t-1})^2 + T(T+1)\Sigma t y_{t-1}\Sigma y_{t-1} - (T(T+1)(2T+1)/6)(\Sigma y_{t-1})^2$

is the determinant of the matrix (X'X) with

$X = (\underline{e}, \underline{t}, \underline{y}_{-1})$, $\underline{e}' = (1,\ldots,1)$, $\underline{t}' = (1 - T/2, 2-T/2, \ldots, T-T/2)$ and $\underline{y}'_{-1} = (y_0, y_1, \ldots, y_{T-1})$. It will be noted that in the definition of $D_x$ asymptotically negligeable terms are included. While it makes no difference for the asymptotic distribution, these extra terms are necessary to ensure a positive value of $D_x$ and make it invariant with respect to $y_0$, the initial observation.

The limiting distribution of these transformed statistics is derived in the next theorem.

Theorem 4.2: If $\{y_t\}_0^T$ is a sequence of random variables generated according to model (3) with the sequence $\{u_t\}$ satisfying assumption 4.1, then as $T \uparrow \infty$:

a) $Z\Phi_1 \rightarrow (\frac{1}{2})[w(1)^2+\{\int_0^1 w(r)^2 dr-(\int_0^1 w(r)dr)^2\}^{-1}\{(1/2)(w(1)^2-1)-w(1)\int_0^1 w(r)dr\}^2]$

b) $Z\Phi_2 \rightarrow (2/3)\ [H' + (C_3/2D)\ (w(1)^2 -1) + (1/8D)\ (w(1)^2 -1)^2]$

c) $Z\Phi_3 \rightarrow H + (C_3/2D)(w(1)^2 - 1) + (1/8D)(w(1)^2 - 1)^2$

Theorem 4.2 shows that the limiting distribution of the transformed statistics is the same for a very general class of innovation sequence $\{u_t\}$ which permit some degree of temporal dependence and heterogeneity. Furthermore, it is clear from the results of Theorem 3.1 that the limiting distribution of the transformed statistics is, under the conditions of theorem 4.2, the same as the limiting distribution of the original regression F-type statistics under the assumption of i.i.d. errors $\{u_t\}$. This implies that the critical values tabulated by Dickey and Fuller (1981) can be used as the appropriate critical values for the transformed statistics. This procedure yields a class of operational statistics for joint hypothesis testing in time series regression with a unit root. These are appropriate under a wide class of possible data generating mechanisms and require only a simple transformation of the original regression F-tests, defined by (4), (5) and (6).



## 5. LOCAL ASYMPTOTIC POWER

We may develop asymptotic power functions for these tests by considering a sequence of local alternatives to (3.b) given by

(7) $\alpha = e^{c/T} \sim 1 + c/T$

When $c = 0$, (7) reduces to the null hypothesis $\alpha = 1$. The idea of developing a noncentral asymptotic distribution using the specification (7) was first explored by Phillips (1986b) who proved the following lemma which extends the asymptotic distribution theory of sample moments given in lemma 2.4 to this more general framework.

Lemma 5.1: If the time series $\{y_t\}_0^T$ is generated by model (3) with (7) instead of (b) then as $T \uparrow \infty$:

(a) $T^{-3/2}\ \Sigma_1^T\ y_t \rightarrow \sigma\int_0^1 J_c(r)\ dr$

(b) $T^{-2}\ \Sigma_1^T\ y_t^2 \rightarrow \sigma^2\int_0^1 J_c(r)^2 dr$

(c) $T^{-5/2}\ \Sigma_1^T\ ty_t \rightarrow \sigma\int_0^1 rJ_c(r)\ dr$

(e) $T^{-1}\ \Sigma_1^T\ y_{t-1}\ u_t \rightarrow \sigma^2 \int_0^1 J_c(r)\ dw(r) + (\frac{1}{2})(\sigma^2 - \sigma_u^2)$

where $J_c(r) = \int_0^r e^{(r-s)c}\ dw(s)$

Using these results for sample moments we may derive the asymptotic distribution under local alternatives of the statistics considered in previous sections.

Theorem 5.2: If the time series $\{y_t\}_0^T$ is generated by model (3) with (7) instead of (b), then as $T \uparrow \infty$:

(a) $\Phi_1 \to (\sigma^2/2\sigma_u^2)\{J_c(1)^2 + c^2 Z_c + 2cX_c + X_c^2/Z_c + (c + X_c)(1-\sigma_u^2/\sigma^2) + (1/4)(1 - \sigma_u^2/\sigma^2)^2\}$

(b) $\Phi_2 \to (2\sigma^2/3\sigma_u^2)\{H_c' + (1/2D_c)(1-\sigma_u^2/\sigma^2)[cD_c + C_{3c} + (1/4)(1-\sigma_u^2/\sigma^2)]\}$

(c) $\Phi_3 \to (\sigma^2/\sigma_u^2)\{H_c + (1/2D_c)(1-\sigma_u^2/\sigma^2)[cD_c + C_{3c} + (1/4)(1 - \sigma_u^2/\sigma^2)]\}$

where $Z_c = \int_0^1 J_c(r)^2\, dr - (\int_0^1 J_c(r)\, dr)^2$

$X_c = \int_0^1 J_c(r)\, dw(r) - w(1) \int_0^1 J_c(r)\, dr$

$C_{3c} = 6\int_0^1 rJ_c r(r)\, dr\, w(1) - 4\int_0^1 J_c(r)\, dr\, w(1) - 12\int_0^1 r dw(r)\{\int_0^1 rJ_c(r)\, dr - (\tfrac{1}{2})\int_0^1 J_c(r) dr\} + \int_0^1 J_c(r) dw(r)$

$D_c = \int_0^1 J_c(r)^2\, dr - 12(\int_0^1 rJ_c(r)\, dr)^2 + 12\int_0^1 rJ_c(r) dr \int_0^1 J_c(r)\, dr - 4(\int_0^1 J_c(r) dr)^2$

$H_c' = H_c + (\tfrac{1}{2})\, J_c(1)$

and $H_c$ is such that when evaluated at $c = 0$ it is given by

$$H_{c=0} = H + (C_3/2D)(w(1)^2-1) + (1/8D)(w(1)^2-1)^2$$

Remark: The actual expression for $H_c$ is lengthy. Since its exact formulation is not needed in order to derive the next theorem we do not report it here.

Theorem 5.2 generalizes theorem 3.1 to near-integrated processes. Note in particular that, where evaluated at c=0, the functionals in theorem 5.2 reduces to those in theorem 3.1. The next theorem characterizes the limiting distribution of the transformed statistics to near-integrated processes.

Theorem 5.3: If the time series $\{y_t\}_0^T$ is generated by model (3) with (7) instead of (b) and assumption 4.1 instead of (d), then as $T \uparrow \infty$:

(a) $Z\Phi_1 \to (\tfrac{1}{2})[J_c(1) + c^2 Z_c + 2cX_c + (X_c^2/Z_c)]$



(b) $Z\Phi_2 \to (2/3)\, H_c'$

(c) $Z\Phi_3 \to H_c$

One can analyse the asymptotic power of the tests under a sequence of local alternatives by comparing the results of theorems 5.2 and 5.3. Theorem 5.2 gives the asymptotic power of the Dickey-Fuller test statistics when the functionals are evaluated at $\sigma^2 = \sigma_u^2$ corresponding to the case with i.i.d. errors. However, note that when evaluated at $\sigma_u^2 = \sigma^2$ the functionals of theorem 5.2 are identical to those of theorem 5.3. The latter gives the asymptotic power of the transformed statistics for general error structures. The combination of these two theorems shows that the transformed statistics have the same asymptotic local power under a general structure for the innovation sequence as do the Dickey-Fuller statistics under i.i.d. errors. In particular nothing is lost, under the power criterion considered here, by using the transformed statistics (the Z's) when the original Dickey-Fuller statistics are appropriate, i.e. under i.i.d. errors.

## 6. EXAMPLES

Box and Jenkins (1976, p. 525) list 197 concentration readings from a chemical process. The authors conclude that an ARMA(1,1) or an ARIMA(0,1,1) model should be fitted to the data. We obtained the following regression equation

$$y_t = \underset{(1.002)}{7.300} + \underset{(0.059)}{0.572}\, y_{t-1} + \hat{u}_t$$

Since we want to test for an ARIMA (0,1,1) against an ARMA(1,1), the errors are a first-order moving-average and the appropriate value of $\ell$, the truncation lag parameter in the estimator $\hat{\sigma}^2_{T\ell}$, should therefore be 1. The result obtained is $Z\Phi_1 =$ 34.38. Since the critical value at the 1 % level is 6.43 we clearly reject the null hypothesis of a unit root. The results are similar with larger values of $\ell$.

Box and Jenkins also list 369 observations of the price of an IBM stock (series B, p. 526). We fitted the following regression equation to this data set:

$$y_t = \underset{(2.185)}{0.170} + \underset{(0.005)}{0.999}\, y_{t-1} + \hat{u}_t$$

with the truncation lag parameter $\ell$ set to 1, $Z\Phi_1 = 0.287$. Since the critical value at the 5% level is 4.59, we do not reject the null hypothesis of a unit root without drift. Since one may suspect the presence of a deterministic trend, one may further check this conclusion by calculating the $Z\Phi_2$ and $Z\Phi_3$ statistics. For this purpose, we obtained the following regression equation:

$$y_t = \underset{(2.849)}{4.809} + \underset{(0.005)}{-0.0117}\,(t-T/2) + \underset{(0.006)}{0.989}\, y_{t-1} + \tilde{u}_t$$

with $\ell = 1$, $Z\Phi_2 = 2.213$, and $Z\Phi_3 = 3.063$. Since the critical values at the 5% level are 4.68 and 6.25 respectively, the same conclusion emerges. The results are similar using other values of $\ell$.

## MATHEMATICAL APPENDIX

In the following lemma we state the results derived in Phillips and Perron (1986) which will be used in the proofs of theorems (3.1) and (4.1). For a more elaborate treatment see Perron (1986). In what follows, $\Sigma$ signifies $\sum_{t=1}^{T}$.

<u>Lemma A.1</u> If $\{y_t\}_0^T$ is a sequence of random variables generated according to model (3), then as $T \to \infty$:



a) $T(\hat{\alpha}-1) \to \{\int_0^1 w(r)dr - (\int_0^1 w(r)dr)^2\}^{-1}\{(\tfrac{1}{2})(w(1)^2-\sigma_u^2/\sigma^2)-w(1)\int_0^1 w(r)dr\}$

b) $\hat{\sigma}^2 \underset{p}{\to} \sigma_u^2$

c) $T(\tilde{\alpha}-1) \to [C_3 + \tfrac{1}{2}(w(1)^2 - \sigma_u^2/\sigma^2)]/D$

d) $T^{3/2}\tilde{\beta} \to (\sigma/D)[C_2 + 6(w(1)^2 - \sigma_u^2/\sigma^2)(\tfrac{1}{2}\int_0^1 w(r)dr - \int_0^1 rw(r)dr)]$

e) $T^{1/2}\tilde{\mu} = (\sigma/D)[C_1 - (\tfrac{1}{2})\int_0^1 w(r)dr\,(w(1)^2 - \sigma_u^2/\sigma^2)]$

f) $\tilde{\sigma}^2 \underset{p}{\to} \sigma_u^2$ .

<u>Proof of Theorem 3.1:</u>

a) $\Phi_1 = (2\hat{S}^2)^{-1}[T\, S_0^2 - (T-2)\,\hat{S}^2]$

$$= (2\hat{S}^2)^{-1}[\Sigma(y_t - y_{t-1})^2 - \Sigma(y_t - \hat{\mu} - \hat{\alpha}y_{t-1})^2]$$

Using the fact that $\Sigma(y_t - y_{t-1})^2 = \Sigma u_t^2$ under the null hypothesis and $\hat{\mu} = \bar{Y}_{(0)} - \hat{\alpha}\bar{Y}_{(-1)}$

$$\Phi_1 = (2\hat{S}^2)^{-1}\{\Sigma u_t^2 - \Sigma[u_t - T^{-1}S_T - (\hat{\alpha}-1)(y_{t-1} - \bar{Y}_{-1})]^2\}$$

$$= (2\hat{S}^2)^{-1}\,[T^{-1}S_T^2 - T^2(\hat{\alpha}-1)^2\, T^{-2}\Sigma(y_{t-1} - \bar{Y}_{-1})^2 + 2T(\hat{\alpha}-1)\, T^{-1}\Sigma u_t(y_{t-1} - \bar{Y}_{-1})]$$

Applying lemma A.1 (a and b), lemmas 2.2-2.4, we obtain

$$\Phi_1 \to (2\sigma_u^2)^{-1}\{\sigma^2 w(1)^2$$
$$- \{\int_0^1 w(r)^2dr - (\int_0^1 w(r)dr)^2\}^{-2}\{(\tfrac{1}{2})(w(1)^2-\sigma_u^2/\sigma^2)-w(1)\int_0^1 w(r)dr\}^2$$
$$\cdot\, [\sigma^2 \int_0^1 w(r)^2dr - \sigma^2(\int_0^1 w(r)dr)^2]$$
$$+ 2[\{\int_0^1 w(r)^2dr - (\int_0^1 w(r)dr)^2\}^{-1}\{(\tfrac{1}{2})(w(1)^2 - \sigma_u^2/\sigma^2)-w(1)\int_0^1 w(r)dr\}]$$

$$\cdot\ [(\sigma^2/2)(w(1)^2 - \sigma_u^2/\sigma^2) - \sigma^2 w(1) \int_0^1 w(r)dr]\}$$

which yields the desired representation upon simplification.

We shall now consider the limiting distribution of $\Phi_3$. Results for the statistic $\Phi_2$ will follow as very simple modifications.

<u>Proof of part (c):</u>

$$\Phi_3 = (2\tilde{S}^2)^{-1}[T\{S_0^2 - (\bar{Y}_{(0)} - \bar{Y}_{(-1)})^2\} - (T-3)\tilde{S}^2]$$

$$= (2\tilde{S}^2)^{-1}[\Sigma(y_t - y_{t-1})^2 - T^{-1}(y_T - y_0)^2 - \Sigma(y_t - \tilde{\mu} - \tilde{\beta}(t-T/2) - \tilde{\alpha}\, y_{t-1})^2]$$

Again under the null hypothesis of model (3) we have $y_t - y_{t-1} = u_t$ and $y_T - y_0 = S_T$ so that

$$\Phi_3 = (2\tilde{S}^2)^{-1}[\Sigma u_t^2 - T^{-1}S_T^2 - \Sigma\{u_t - \tilde{\mu} - \tilde{\beta}(t - T/2) - (\tilde{\alpha} - 1)y_{t-1}\}^2]$$
$$= (2\tilde{S}^2)^{-1}[-T^{-1}S_T^2 + 2\Sigma u_t\{\tilde{\mu} + \tilde{\beta}(t - T/2) + (\tilde{\alpha} - 1)y_{t-1}\}$$
$$- \Sigma\{\tilde{\mu} + \tilde{\beta}(t - T/2) + (\tilde{\alpha} - 1)y_{t-1}\}^2]$$

We analyse each term of the expression in bracket successively.

<u>Term # 1:</u> $K_1 = T^{-1} S_T^2 \to \sigma^2 w(1)^2 \equiv A_1$ by lemma 2.2

<u>Term # 2:</u> $K_2 = 2\Sigma u_t(\tilde{\mu} + \tilde{\beta}(t - T/2) + (\tilde{\alpha} - 1)y_{t-1})$

$$= 2T^{1/2}\,\tilde{\mu}\,T^{-1/2}\,\Sigma u_t + 2T^{3/2}\tilde{\beta}(T^{-3/2}\Sigma t u_t - (1/2)T^{-1/2}\Sigma u_t)$$
$$+ 2T(\tilde{\alpha} - 1)\,T^{-1}\,\Sigma u_t y_{t-1}$$

Now $\Sigma u_t = S_T$, $T^{-3/2}\Sigma t u_t \to \sigma(w(1) - \int_0^1 w(r)dr)$ (see Phillips and Perron (1986), Lemma 2.3.c). Using lemma A.1 (c, d, e) and lemma 2.4(e) we obtain:

$$K_2 \to (2\sigma^2/D)\ w(1)[C_1 - (1/2)\int_0^1 w(r)dr(w(1)^2 - \sigma_u^2/\sigma^2)]$$
$$+ (2\sigma^2/D)((\tfrac{1}{2})w(1) - \int_0^1 w(r)dr)[C_2 + 6(w(1)^2 - \sigma_u^2/\sigma^2)((1/2)\int_0^1 w(r)dr - \int_0^1 rw(r)dr)]$$



$$+ (2\sigma^2/D)\ [(\tfrac{1}{2})(w(1)^2 - \sigma_u^2/\sigma^2)][C_3 + (\tfrac{1}{2})(w(1)^2 - \sigma_u^2/\sigma^2)]$$
$$= (2\sigma^2/D)\ [w(1)\ C_1 + ((\tfrac{1}{2})w(1) - \int_0^1 w(r)dr)\ C_2]$$
$$+ (2\sigma^2/D)(w(1)^2 - \sigma_u^2/\sigma^2)[-(\tfrac{1}{2})w(1)\int_0^1 w(r)dr + 6((\tfrac{1}{2})w(1) - \int_0^1 w(r)dr)$$
$$\cdot\ ((\tfrac{1}{2})\int_0^1 w(r)dr - \int_0^1 rw(r)dr) + C_3/2 + (1/4)(w(1)^2 - \sigma_u^2/\sigma^2)]$$

$$= A_2 + (2\sigma^2/D)\ (w(1)^2 - \sigma_u^2/\sigma^2)[C_3 + (1/4)(w(1)^2 - \sigma_u^2/\sigma^2)]$$

where $A_2 = (2\sigma^2/D)[w(1)\ C_1 + ((\tfrac{1}{2})w(1) - \int_0^1 w(r)dr)\ C_2]$

<u>Term # 3:</u> $K_3 = \Sigma[\tilde{\mu} + \tilde{\beta}(t - T/2) + (\tilde{\alpha} - 1)y_{t-1}]^2$

$$= \Sigma[\tilde{\mu}^2 + \tilde{\beta}^2(t^2 - tT + T^2/4) + (\tilde{\alpha} - 1)^2 y_{t-1}^2 + 2\tilde{\mu}\tilde{\beta}(t - T/2)$$
$$+ 2\tilde{\mu}(\tilde{\alpha} - 1)y_{t-1} + 2\tilde{\beta}(\tilde{\alpha} - 1)(t - T/2)y_{t-1}]$$
$$= T\tilde{\mu}^2 + T^3\tilde{\beta}^2(T^{-3}\Sigma t^2 - T^{-2}\,\Sigma t + 1/4) + T^2(\tilde{\alpha} - 1)T^{-2}\Sigma y_{t-1}^2$$
$$+ 2T^{1/2}\,\tilde{\mu}\,T^{3/2}\,\tilde{\beta}(T^{-2}\Sigma t - (\tfrac{1}{2})) + 2T^{1/2}\,\tilde{\mu}\,T(\tilde{\alpha} - 1)\,T^{-3/2}\,\Sigma y_{t-1}$$
$$+ 2T^{3/2}\,\tilde{\beta}\,T(\tilde{\alpha} - 1)(T^{-5/2}\,\Sigma t y_{t-1} - (\tfrac{1}{2})T^{-3/2}\,\Sigma y_{t-1})$$

Using lemma A.1 (c, d, e), and lemma 2.4 (a, b, c) we obtain

$$K_3 \to (2\sigma^2/D)\ [C_1^2/2 - (\tfrac{1}{2})C_1 \int_0^1 w(r)dr\ (w(1)^2 - \sigma_u^2/\sigma^2)$$
$$+ (1/8)\ (\int_0^1 w(r)dr)^2(w(1)^2 - \sigma_u^2/\sigma^2)^2 + C_2^2/24$$
$$+ (\tfrac{1}{2})C_2\ (w(1)^2 - \sigma_u^2/\sigma^2)((\tfrac{1}{2})\int_0^1 w(r)dr - \int_0^1 rw(r)dr)$$
$$+ (3/2)(w(1)^2 - \sigma_u^2/\sigma^2)^2((1/2)\int_0^1 w(r)dr - \int_0^1 rw(r)dr)^2$$
$$+ (\tfrac{1}{2})C_3^2 \int_0^1 w(r)^2 dr + (\tfrac{1}{2})C_3 \int_0^1 w(r)^2 dr\ (w(1)^2 - \sigma_u^2/\sigma^2)$$
$$+ (1/8)\int_0^1 w(r)^2 dr\ (w(1)^2 - \sigma_u^2/\sigma^2)^2 + C_1 C_3 \int_0^1 w(r)dr$$

$$+ (\tfrac{1}{2})C_1\int_0^1 w(r)dr(w(1)^2 - \sigma_u^2/\sigma^2) - (\tfrac{1}{2})C_3\,(\int_0^1 w(r)dr)^2(w(1)^2 - \sigma_u^2/\sigma^2)$$

$$- (1/4)(\int_0^1 w(r)dr)^2(w(1)^2 - \sigma_u^2/\sigma^2)^2 + C_3C_2[\int_0^1 rw(r)dr - (\tfrac{1}{2})\int_0^1 w(r)dr]$$

$$- 6C_3[-\int_0^1 rw(r)dr + (\tfrac{1}{2})\int_0^1 w(r)dr]^2\,(w(1)^2 - \sigma_u^2/\sigma^2)$$

$$- (\tfrac{1}{2})C_2\,[(1/2)\int_0^1 w(r)dr - \int_0^1 rw(r)dr]\,(w(1)^2 - \sigma_u^2/\sigma^2)$$

$$- 3[(\tfrac{1}{2})\int_0^1 w(r)dr - \int_0^1 rw(r)dr]^2\,(w(1)^2 - \sigma_u^2/\sigma^2)^2\,\Big]$$

A) Collecting terms not involving $(w(1)^2 - \sigma_u^2/\sigma^2)$

$$A_{31} = (2\sigma^2/D^2)[C_1^2/2 + C_2^2/24 + (\tfrac{1}{2})C_3^2\int_0^1 w(r)^2dr + C_1C_3\int_0^1 w(r)dr$$
$$+ C_3C_2(\int_0^1 rw(r)dr - (1/2)\int_0^1 w(r)dr)]$$

B) Collecting terms involving $(w(1)^2 - \sigma_u^2/\sigma^2)$ but not $(w(1)^2 - \sigma_u^2/\sigma^2)^2$

$$A_{32} = (2\sigma^2/D^2)(w(1)^2 - \sigma_u^2/\sigma^2)\Big[-(\tfrac{1}{2})C_1\int_0^1 w(r)dr + (\tfrac{1}{2})C_2((\tfrac{1}{2})\int_0^1 w(r)dr - \int_0^1 rw(r)dr)$$
$$+ (\tfrac{1}{2})C_3\int_0^1 w(r)^2dr + (\tfrac{1}{2})C_1\int_0^1 w(r)dr - (\tfrac{1}{2})C_3\,(\int_0^1 w(r)dr)^2$$
$$-6\,C_3[(\tfrac{1}{2})\int_0^1 w(r)dr - \int_0^1 rw(r)dr]^2 - (\tfrac{1}{2})C_2((\tfrac{1}{2})\int_0^1 w(r)dr - \int_0^1 rw(r)dr)\Big]$$
$$= (2\sigma^2/D^2)(w(1)^2 - \sigma_u^2/\sigma^2)\,C_3\{(\tfrac{1}{2})\int_0^1 w(r)^2dr - (1/2)(\int_0^1 w(r)dr)^2$$
$$- 6[(\tfrac{1}{2})\int_0^1 w(r)dr - \int_0^1 rw(r)dr]^2\}$$

Now the term in bracket is simply D/2 and

$$A_{32} = (\sigma^2/D)\,(w(1)^2 - \sigma_u^2/\sigma^2)\,C_3$$

c) Terms involving $(w(1)^2 - \sigma_u^2/\sigma^2)^2$

$$A_{33} = (2\sigma^2/D^2)(w(1)^2 - \sigma_u^2/\sigma^2)^2\,\Big[(1/8)(\int_0^1 w(r)dr)^2$$



$$+ (3/2)((\tfrac{1}{2})\int_0^1 w(r)dr - \int_0^1 rw(r)dr)^2$$
$$+ (1/8)\int_0^1 w(r)^2dr - (1/4)(\int_0^1 w(r)dr)^2 - 3((\tfrac{1}{2})\int_0^1 w(r)dr - \int_0^1 rw(r)dr)^2\Big]$$
$$= (\sigma^2/4D)(w(1)^2 - \sigma_u^2/\sigma^2)^2$$ since the term in bracket is simply D/8.

Therefore $K_3 \rightarrow A_{31} + A_{32} + A_{33}$

$$= A_{31} + (\sigma^2/D)(w(1)^2 - \sigma_u^2/\sigma^2)\,C_3 + (\sigma^2/4D)(w(1)^2 - \sigma_u^2/\sigma^2)^2$$

Finally $\Phi_3 = (2\tilde{S}^2)^{-1}(-K_1 + K_2 - K_3)$

$$\rightarrow (2\sigma_u^2)^{-1}[-A_1 + A_{21} - A_{31}]$$
$$+ (\sigma^2/D\sigma_u^2)[C_3 + (1/4)\,(w(1)^2 - \sigma_u^2/\sigma^2)]\,(w(1)^2 - \sigma_u^2/\sigma^2)$$
$$+ (\sigma^2/2D\sigma_u^2)\,(w(1)^2 - \sigma_u^2/\sigma^2)\,C_3 + (\sigma^2/4D)(w(1)^2 - \sigma_u^2/\sigma^2)^2$$
$$= (\sigma^2/\sigma_u^2)[H + (C_3/2D)(w(1)^2 - \sigma_u^2/\sigma^2) + (1/8D)(w(1)^2 - \sigma_u^2/\sigma^2)^2]$$

since $(\sigma^2/2)^{-1}\,[-A_1 + A_{21} - A_{31}] = H$.

<u>Proof of (b)</u>

$$\Phi_2 = (3\tilde{\sigma}^2)^{-1}\,[T\sigma_0^2 - (T-3)\,\tilde{\sigma}^2]$$

Note that the only differences with $\Phi_3$ are that the term in bracket is divided by 3 instead of 2 and that this expression has one term less $-T(\bar{Y}_{(0)} - \bar{Y}_{(-1)})^2 = -T^{-1}S_T^2$ under the null hypothesis, which converges to $-\sigma^2 w(1)$. We therefore correct the H term accordingly in the functional representation of the limiting distribution of $\Phi_3$ and multiply this expression by 2/3. Note that these differences do not affect the terms involving $(w(1)^2 - \sigma_u^2/\sigma^2)$.

**Proof of theorem 4.2:**

a) $Z\Phi_1 = (\hat{S}^2/\hat{\sigma}^2_{T\ell})\,\Phi_1$

$$- (1/2\hat{\sigma}^2_{T\ell})(\hat{\sigma}^2_{T\ell} - \hat{S}^2)[T(\hat{\alpha}-1) - (1/4)(\hat{\sigma}^2_{T\ell} - \hat{S}^2)\{T^{-2}\Sigma(y_{t-1} - \bar{Y}_{-1})^2\}^{-1}]$$

$$\to (\tfrac{1}{2})[w(1)^2 + \{\textstyle\int_0^1 w(r)^2 dr - (\int_0^1 w(r)dr)^2\}^{-1}[(\tfrac{1}{2})(w(1)^2 - \sigma_u^2/\sigma^2) - w(1)\int_0^1 w(r)dr]^2]$$

$$- (1/2\sigma^2)(\sigma^2 - \sigma_u^2)\Big[\{\textstyle\int_0^1 w(r)^2 dr - (\int_0^1 w(r)dr)^2\}^{-1} [(1/2)(w(1)^2 - \sigma_u^2/\sigma^2) - w(1)\int_0^1 w(r)dr]$$

$$- (1/4)(\sigma^2 - \sigma_u^2)\{\sigma[\textstyle\int_0^1 w(r)^2 dr - (\int_0^1 w(r)dr)^2]\}^{-1}\Big]$$

$$= (\tfrac{1}{2})\, w(1)^2 + (1/2B)[(1/4)w(1)^4 - (\tfrac{1}{2})w(1)^2\sigma_u^2/\sigma^2 + (1/4)\sigma_u^4/\sigma^4$$

$$- (w(1)^2 - \sigma_u^2/\sigma^2)\, w(1) \textstyle\int_0^1 w(r)dr + w(1)^2(\int_0^1 w(r)dr)^2]$$

$$- (1/2B)(1-\sigma_u^2/\sigma^2)[(\tfrac{1}{2})(w(1)^2 - \sigma_u^2/\sigma^2) - w(1)\textstyle\int_0^1 w(r)dr - (1/4)(1 - \sigma_u^2/\sigma^2)]$$

$$\text{where } B = \textstyle\int_0^1 w(r)^2 dr - (\int_0^1 w(r)dr)^2$$

$$= (\tfrac{1}{2})\, w(1)^2 + (1/2B)\,[(1/4)\, w(1)^4 - (\tfrac{1}{2})w(1)^2 + 1/4$$

$$- (w(1)^2 - 1)\, w(1) \textstyle\int_0^1 w(r)dr + w(1)^2(\int_0^1 w(r)dr)^2]$$

$$= (\tfrac{1}{2})\,\{w(1)^2 + (1/2B)\,[(\tfrac{1}{2})(w(1)^2 - 1) - w(1)\textstyle\int_0^1 w(r)dr]^2\}$$

and the result follows upon substitution for B.

c) $Z\Phi_3 = (\tilde{S}^2/\tilde{\sigma}^2_{T\ell})\,\Phi_3 - (1/2\tilde{\sigma}^2_{T\ell})(\tilde{\sigma}^2_{T\ell} - \tilde{S}^2)[T(\tilde{\alpha}-1) - (T^6/48D_x)(\tilde{\sigma}^2_{T\ell} - \tilde{S}^2)]$

Now $T^{-6}D_x \to \sigma^2 D/12$, see Phillips and Perron (1986) and using lemma A.1c, and theorem 3.1



$$Z\Phi_3 \to H + (C_3/2D)(w(1)^2 - \sigma_u^2/\sigma^2) + (1/8D)\,(w(1)^2 - \sigma_u^2/\sigma^2)^2$$

$$- (1/2D)(1 - \sigma_u^2/\sigma^2)[C_3 + (\tfrac{1}{2})(w(1)^2 - \sigma_u^2/\sigma^2) - (1/4)(1 - \sigma_u^2/\sigma^2)]$$

$$= H + (C_3/2D)(w(1)^2 - 1) + (1/8D)(w(1)^2 - 1)^2$$

The proof of part (b) is similar.

**Proof of theorem 5.2:**

a) We write $\Phi_1$ as

$$\Phi_1 = (2\hat{S}^2)^{-1}[T^{-1}(y_T - Y_0)^2 - T^2(\hat{\alpha}-1)^2 T^{-2}\Sigma(y_{t-1} - \bar{Y}_{-1})^2 + 2T(\hat{\alpha}-1)\, T^{-1}\Sigma(y_t - y_{t-1})(y_{t-1} - \bar{Y}_{-1})]$$

From theorem 6.2 of Phillips and Perron (1986)

$$T(\hat{\alpha}-1) \to c + [X_c + (\tfrac{1}{2})(1 - \sigma_u^2/\sigma^2)]/Z_c$$

$$T^{-2}\Sigma(y_{t-1} - \bar{Y}_{-1})^2 \to \sigma^2 Z_c$$

$$T^{-1}\Sigma(y_t - y_{t-1})(y_{t-1} - \bar{Y}_{-1}) = T^{-1}\Sigma((c/T)\, y_{t-1} + u_t)(y_{t-1} - \bar{Y}_{-1})$$

$$\to c\sigma^2 Z_c + \sigma^2 X_c + (\tfrac{1}{2})(\sigma^2 - \sigma_u^2)$$

Therefore

$$\Phi_1 \to (\sigma^2/2\sigma_u^2)\{J_c(1)^2 - [c + X_c + (\tfrac{1}{2})\,(1 - \sigma_u^2/\sigma^2)]^2/Z_c$$

$$+ 2\{c + [X_c + \tfrac{1}{2}(1 - \sigma_u^2/\sigma^2)][cZ_c + X_c + \tfrac{1}{2}(1 - \sigma_u^2/\sigma^2)]\}\}$$

$$= (\sigma^2/2\sigma_u^2)\{J_c(1)^2 + c^2 Z_c + 2cX_c + X_c^2/Z_c$$

$$+ (c + X_c)(1 - \sigma_u^2/\sigma^2) + (1/4)(1 - \sigma_u^2/\sigma^2)\}$$

as required.

c) We write $\Phi_3$ as

$$\Phi_3 = (2\tilde{S}^2)^{-1}[-T^{-1}[\Sigma(y_t - y_{t-1})]^2$$
$$+ 2\Sigma(y_t - y_{t-1})(\tilde{\mu} + \tilde{\beta}(t - T/2) + (\tilde{\alpha} - 1)\, y_{t-1})$$
$$- \Sigma(\tilde{\mu} + \tilde{\beta}(t - T/2) + (\tilde{\alpha} - 1)y_{t-1})^2]$$

From theorem 6.3 of Phillips and Perron (1986)

$$T(\tilde{\alpha} - 1) \to c + \{C_{3c} + (\tfrac{1}{2})(1 - \sigma_u^2/\sigma^2)\}/D_c$$

It is relatively easy to show that

$$T^{\frac{1}{2}}\tilde{\mu} \to (\sigma/D_c)\Big[w(1)[\textstyle\int_0^1 J_c(r)^2 dr - 12(\int_0^1 rJ_c(r)dr)^2$$
$$+ 18 \textstyle\int_0^1 rJ_c(r)dr \int_0^1 J_c(r)dr - 6(\int_0^1 J_c(r)dr)^2]$$
$$+ \textstyle\int_0^1 w(r)dr\, [6(\int_0^1 J_c(r)dr)^2 - 12\int_0^1 J_c(r)dr \int_0^1 rJ_c(r)dr]$$
$$+ 12\, c\textstyle\int_0^1 J_c(r)dr[(1/12)\int_0^1 J_c(r)^2 dr - (\int_0^1 rJ_c(r)dr)^2$$
$$+ (3/2)\textstyle\int_0^1 rJ_c(r)dr \int_0^1 J_c(r)dr - (\frac{1}{2})(\int_0^1 J_c(r)dr)^2]$$
$$- \textstyle\int_0^1 J_c(r)dr\int_0^1 J_c(r)dw(r) - \frac{1}{2}\int_0^1 J_c(r)dr\,(1 - \sigma_u^2/\sigma^2)\Big]$$

and

$$T^{3/2}\tilde{\beta} \to (12\,\sigma/D_c)\Big[w(1)[(\tfrac{1}{2})\textstyle\int_0^1 J_c(r)^2 dr + \int_0^1 J_c(r)dr\int_0^1 rJ_c(r)dr$$
$$- (\textstyle\int_0^1 J_c(r)dr)^2] + [c\int_0^1 rJ_c(r)dr\int_0^1 w(r)dr][(\int_0^1 J_c(r)dr)^2 - \int_0^1 J_c(r)^2 dr]$$
$$+ c\textstyle\int_0^1 J_c(r)dr[\int_0^1 J_c(r)dr \int_0^1 rJ_c(r)dr - (\frac{1}{2}) \int_0^1 J_c(r)^2 dr]$$
$$+ [\tfrac{1}{2}\textstyle\int_0^1 J_c(r)dr - \int_0^1 rJ_c(r)dr][c\int_0^1 J_c(r)^2 dr + \int_0^1 J_c(r)dw(r) + \frac{1}{2}(1 - \sigma_u^2/\sigma^2)]\Big]$$



The result follows, upon tedious but straightforward algebra, by using these results and lemma 5.1. The pattern of the proof follows exactly the one of theorem 3.1(c). The proof of part (b) is similar.

<u>Proof of theorem 5.3:</u>

a) The proof follows by substracting form the functional in 5.2.a, the limit of the correction term in $Z\Phi_1$ given by

$$(\tfrac{1}{2})(1 - \sigma_u^2/\sigma^2)[c + X_c/Z_c + (\tfrac{1}{2})(1 - \sigma_u^2/\sigma^2)/Z_c - (1/4)(1 - \sigma_u^2/\sigma^2)^2/Z_c]$$

The proof of parts (b) and (c) are similar.

## REFERENCES


BILLINGSLEY, P. (1968), Convergence of Probability Measures, John Whiley: New York.

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.

DICKEY, D.A. and W.A. FULLER (1981), "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", Econometrica, 49, 1057-1072.

HERRNDORF, N. (1984), "A Functional Central Limit Theorem for Weakly Dependent Sequences of Random Variables", Annals of Probability, 12, 141-153.

NEWEY, N.K. and K.D. WEST (1985), "A Simple Positive Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", Princeton Discussion Paper no. 92.

PERRON, Pierre (1986), "Hypothesis Testing in Time Series Regression With a Unit Root", Unpublished Ph.D. Dissertation, Yale University.

PHILLIPS, P.C.B. (1986a), "Time Series Regression with Unit Roots", Cowles Foundation Discussion Paper No. 740-R, Yale University.

PHILLIPS, P.C.B. (1986b), "Towards a Unified Asymptotic Theory for Autoregression", Cowles Foundation Discussion Paper No. 782, Yale University.

PHILLIPS, P.C.B. and P. PERRON (1986), "Testing for a Unit Root in Time Series Regression", mimeo.

SAID, S.E. and D.A. DICKEY (1984), "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order", Biometrika, 71, 599-608.