COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale University
New Haven, Connecticut 06520

Cowles Foundation Discussion Paper No. 979

Note: Cowles Foundation Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. Requests for single copies of a Paper will be filled by the Cowles Foundation within the limits of the supply. References in publications to Discussion Papers (other than acknowledgment that a writer had access to such unpublished material) should be cleared with the author to protect the tentative character of these papers.

TESTING THE NULL HYPOTHESIS OF STATIONARITY AGAINST THE ALTERNATIVE OF A UNIT ROOT: HOW SURE ARE WE THAT ECONOMIC TIME SERIES HAVE A UNIT ROOT?

by

Denis Kwiatkowski, Peter C. B. Phillips and Peter Schmidt

May 1991

TESTING THE NULL HYPOTHESIS OF STATIONARITY AGAINST THE ALTERNATIVE OF A UNIT ROOT: HOW SURE ARE WE THAT ECONOMIC TIME SERIES HAVE A UNIT ROOT?

Denis Kwiatkowski

Central Michigan University

Peter C.B. Phillips

Yale University

Peter Schmidt

Michigan State University

January, 1990

Revised November, 1990

The second and third authors gratefully acknowledge the support of the National Science Foundation.

1. INTRODUCTION

It is a well-established empirical fact that standard unit root tests fail to reject the null hypothesis of a unit root for many economic time series. This was first argued systematically in the influential article of Nelson and Plosser (1982), who applied Dickey-Fuller type tests (Dickey (1976), Fuller (1976), Dickey and Fuller (1979)) to 14 annual U.S. time series and failed to reject the hypothesis of a unit root in all but one of the series. These results are not changed by allowing for error autocorrelation using the augmented tests of Said and Dickey (1984) or the corrected test statistics of Phillips (1987) and Phillips and Perron (1988). Similar results are obtained for many other macroeconomic time series. A partial listing of empirical studies yielding these findings can be found in DeJong et al. (1989).

The standard conclusion that is drawn from this empirical evidence is that many or most aggregate economic time series contain a unit root. However, it is important to note that in this empirical work the unit root is set up as the null hypothesis to be tested, and the way in which classical hypothesis testing is carried out ensures that the null hypothesis is accepted unless there is strong evidence against it. Therefore, an alternative explanation for the common failure to reject a unit root is simply that most economic time series are not very informative about whether or not there is a unit root; or, equivalently, that standard unit root tests are not very powerful against relevant alternatives. Several more recent studies have argued that this is indeed the case. For example, DeJong et al. (1989) provide evidence that the Dickey-Fuller tests have low power against stable autoregressive alternatives with roots near unity, and Diebold and Rudebusch (1990) show that they also have low power against fractionally integrated alternatives.

Bayesian analysis offers an alternative means of evaluating how informative the data are regarding the presence of a unit root, by providing direct posterior evidence in support of stationarity and nonstationarity. Working from flat priors, DeJong and Whiteman (1990) found only two of the Nelson-Plosser series to have stochastic trends using this approach. Phillips (1990) used objective ignorance priors in extracting posteriors and found support for stochastic trends in five of the series.

These studies suggest that, in trying to decide by classical methods whether economic data are stationary or integrated, it would be useful to have available tests of the null hypothesis of stationarity as well as tests of the null hypothesis of a unit root. Specifically, if we test both the null hypothesis of a unit root and the null hypothesis of stationarity, there are four possible outcomes, all but one of which seem easily interpretable. If we accept the unit root and reject stationarity, we have found a unit root. If we reject a unit root and accept stationarity, we have found stationarity. If we accept both the unit root and stationarity, we conclude that we can't be very sure whether or not there is a unit root. Finally, if we reject both the unit root and stationarity, it is not clear what should be concluded.

This paper provides a straightforward test of the null hypothesis of stationarity against the alternative of a unit root. There have been relatively few previous attempts to test the null hypothesis of stationarity. Park and Choi (1988) consider a test statistic which is essentially the F statistic for "superfluous" deterministic trend variables; this statistic should be close to zero under the stationary null but not under the alternative of a unit root. Rudebusch (1990) considers the Dickey-Fuller test statistics, but estimates both trend stationary and difference stationary models and then uses the bootstrap to evaluate the distribution of these statistics under each model. Using the Nelson and Plosser data, he often cannot reject either the trend stationary model or the difference stationary model. DeJong et al. (1989) consider the Dickey-Fuller regression

$$(1) \qquad y_t = \alpha + \delta t + \rho y_{t-1} + \epsilon_t$$

in which the unit root corresponds to $\rho = 1$, but they also test the stationary null hypothesis $\rho = .85$. For most of the series used by Nelson and Plosser, they can reject neither $\rho = 1$ nor $\rho = .85$. Furthermore, this failure to reject both hypotheses is shown to be reasonable in terms of the powers of the tests, which they explore through Monte Carlo experimentation.

These are reasonable first attempts to test stationarity, but they all suffer from the lack of a plausible model in which the null of stationarity is naturally framed as a parametric restriction. Only DeJong et al. test a parametric restriction that implies stationarity, and their choice of ρ=.85 to represent stationarity (as opposed to ρ=.70 or .95 or whatever) is obviously arbitrary. Clearly stationarity is a composite null hypothesis in models like (1) above.

As a general statement, it is certainly not obvious that a parameterization (like the one underlying the Dickey-Fuller tests) that is natural and useful to test the null of a unit root should also be natural and useful to test the null of stationarity. In this paper we will use a parameterization which provides a plausible representation of both stationary and nonstationary variables and which leads naturally to a test of the hypothesis of stationarity. Specifically, we choose a components representation in which the time series under study is written as the sum of a deterministic trend, a random walk, and a stationary error. The null hypothesis of trend stationarity corresponds to the hypothesis that the variance of the random walk equals zero. Under the additional assumptions that the random walk is normal and that stationary error is normal white noise, the LM statistic for the trend stationarity hypothesis is the same as the locally best invariant (LBI) test statistic, and follows from Nabeya and Tanaka (1988). However, the assumption that the stationary error is white noise is not credible in many empirical applications, since it implies that under the null hypothesis the variable in question should have iid deviations from trend. We therefore proceed in the spirit of Phillips (1987) and Phillips and Perron (1988) by deriving the asymptotic distribution of the LM statistic under rather general conditions on the stationary error, and we propose a modified version of the LM statistic that is valid asymptotically under these more general conditions. The asymptotic distribution is non-standard, involving so-called higher order Brownian bridges.

When we apply this test to the Nelson-Plosser data, our results depend on the way that the deterministic trend is accommodated. For almost all series we can reject the hypothesis of level stationarity, but for many of the series we cannot reject the hypothesis of trend stationarity. The latter result is in broad agreement with the results of DeJong et al. (1989) and Rudebusch (1990), and with the aforementioned Bayesian analyses of DeJong and Whiteman (1990) and Phillips (1990). It suggests that for many series the existence of a unit root is in doubt, despite the failure of Dickey-Fuller tests (and other related tests) to reject the unit root hypothesis.

2. THE LM STATISTIC FOR THE STATIONARITY HYPOTHESIS

Let y_t, $t = 1, 2, \ldots, T$ be the observed series for which we wish to test stationarity. We assume that we can decompose the series into the sum of deterministic trend, a random walk and a stationary error:

(2) $$y_t = \xi t + r_t + \epsilon_t \ .$$

Here r_t is a random walk:

(3) $$r_t = r_{t-1} + u_t$$

where the u_t are iid $(0,\sigma_u^2)$. The initial value r_0 is treated as fixed and serves the role of an intercept. The stationarity hypothesis is simply $\sigma_u^2 = 0$ (or equivalently $\sigma_r^2 = 0$). Since ϵ_t is assumed to be stationary, under the null hypothesis y_t is trend stationary. We will also consider the special case of the model (2) in which we set $\xi = 0$, in which case under the null hypothesis y_t is stationary around a level (r_0) rather than around a trend.

The statistic we will use is both the LM statistic and the LBI test statistic for the hypothesis $\sigma_u^2 = 0$, under the stronger assumptions that the u_t are normal and that the ϵ_t are iid N(0, σ_ϵ^2). Nyblom (1986) considers a model equivalent to our model above and gives the LBI test statistic, but a more convenient expression follows from deriving the statistic as a special case of the statistic developed by Nabeya and Tanaka (1988) to test for random coefficients. (Other relevant references include Tanaka (1983), Nyblom and Makelainen (1983), Franzini and Harvey (1983), and Leybourne and McCabe (1989), and a general discussion can be found in Harvey (1989).) Nabeya and Tanaka consider the regression model

(4) $$y_t = x_t\beta_t + z_t'\gamma + \epsilon_t \quad ,$$

in which the scalar β_t is a normal random walk ($\beta_t = \beta_{t-1} + u_t$, with the u_t iid) and the errors ϵ_t are iid N(0, σ_ϵ^2). They test the hypothesis $\sigma_u^2 = 0$, so that they test the null hypothesis of constancy of regression coefficients against the alternative of random walk coefficients. Our model (2) is obviously the special case of their model in which $x_t = 1$ for all t, $z_t = t$, and their β_t is our r_t. If we set $\xi = 0$ in (2), so as to test the hypothesis of level stationarity, this corresponds to eliminating

z_t from their model, in which case we have the simpler model of Nyblom and Makelainen (1986) and Tanaka (1983).

Appendix A gives the details of the simplifications of the Nabeya and Tanaka statistic that apply in our model. The end result is very simple. Let e_t, t = 1, 2, ..., T be the residuals from the regression of y on an intercept and time trend. Let $\hat{\sigma}_\epsilon^2$ be the usual estimate of the error variance from this regression (the sum of squared residuals, divided by T). Define the partial sum process of the residuals:

$$S_t = \sum_{i=1}^{t} e_i \quad , \quad t = 1, 2, \ldots, T. \tag{5}$$

Then the LM (and LBI) statistic is

$$LM = \sum_{t=1}^{T} S_t^2 \,/\, \hat{\sigma}_\epsilon^2 \tag{6}$$

Furthermore, in the event that we wish to test the null hypothesis of level stationarity instead of trend stationarity, we simply define e_t as the residual from the regression of y on an intercept only (that is, $e_t = y_t - \bar{y}$) instead of as above, and the rest of the construction of the test statistic is unaltered.

The test is an upper tail test. Critical values that are valid asymptotically will be supplied in the next section.

Note that the LM statistic (6) can be interpreted as a functional of the partial sum process S_t given in (5). As such, the statistic is also related to the generic class of tests for a unit root developed recently by Stock (1990). However, unlike the tests considered in Stock's paper, our statistic is intended as a test for stationarity and not as a test of a unit root null. In consequence, the asymptotic theory of our tests statistic is quite different in certain respects from that of Stock, as we shall indicate below.

The statistic (6) also may arise in other contexts. Saikkonen and Luukkonen (1990) derive (6) as the locally best unbiased invariant test of the hypothesis $\theta = 1$ in the model $\Delta y_t = \epsilon_t - \theta\epsilon_{t-1}$, with $E(y_0)$ unknown and playing the role of intercept in our model, and with the ϵ_t iid normal. Since $\theta = 1$ is a stationarity hypothesis, as our hypothesis $\sigma_u^2 = 0$ also is, this is not really a surprising result.

3. ASYMPTOTIC THEORY

In this section we consider the asymptotic distribution of the LM statistic given in (6) above. The LM statistic was derived under the assumption that the errors ϵ_t were iid $N(0, \sigma_\epsilon^2)$. (It was also assumed that the random walk component r_t was normal, but this assumption is obviously irrelevant to the distribution of the test statistic under the null hypothesis, since the null eliminates the random walk component.) However, in this section we will consider the asymptotic distribution of the statistic under weaker assumptions about the errors. As argued in the Introduction, this is important because the series to which the stationarity test will be applied are typically highly dependent over time, and so the iid error assumption under the null is unrealistic. To allow for quite general forms of temporal dependence we may assume that the ϵ_t satisfy the (strong mixing) regularity conditions of Phillips and Perron (1988, p. 336) or the linear process conditions of Phillips and Solo (1989, theorems 3.3 and 3.14). These conditions put some limits on the degree of heterogeneity and autocorrelation allowed in the ϵ sequence but are otherwise fairly general. On the one hand, they are weaker conditions than stationarity, because some heterogeneity is allowed; on the other hand, they are stronger than stationarity alone, because there are limits on the degree of allowable autocorrelation. The Phillips-Perron regularity conditions have been used extensively by subsequent authors, including Leybourne and McCabe (1989). The Phillips-Solo conditions are especially useful because they conveniently allow for all ARMA processes, with either homogeneous or heterogeneous innovations.

Nabeya and Tanaka provide the asymptotic distribution of our test statistics for the case where the ϵ process is iid, and our results are therefore an extension of theirs. Also, their results are given in terms of limiting characteristic functions, while ours involve simple functionals of Brownian motion, which lead to more compact expressions. Finally, some of our results are a special case of results in McCabe and Leybourne (1988). (See also Leybourne and McCabe (1989).)

We define the "long run variance"

$$\text{(7)} \qquad \sigma^2 = \lim_{T\to\infty} T^{-1}E(S_T^2)$$

which will enter into the asymptotic distribution of the test statistic. A consistent estimator of σ^2, say $s^2(\ell)$, can be constructed from the residuals e_t, as in Phillips (1987) or Phillips and Perron (1988); specifically, we will use an estimator of the form

(8) $$s^2(\ell) = T^{-1} \sum_{t=1}^{T} e_t^2 + 2T^{-1} \sum_{s=1}^{\ell} w(s,\ell) \sum_{t=s+1}^{T} e_t e_{t-s}$$

Here $w(s,\ell)$ is an optional lag window (or weighting function) that corresponds to the choice of a spectral window. We will use the Bartlett window $w(s,\ell) = 1 - s/(\ell+1)$ as in Newey and West (1987), which guarantees the non-negativity of $s^2(\ell)$. For consistency of $s^2(\ell)$, it is necessary that the lag truncation parameter $\ell \to \infty$ as $T \to \infty$. The rate $\ell = o(T^{1/2})$ will usually be satisfactory under both the null (e.g., Andrews (1989)) and under the alternative (see Section 4 below).

For the tests of both the level stationary and trend stationary hypotheses, the denominator of the LM statistic in (6) is $\hat{\sigma}_\epsilon^2$, which converges in probability to σ_ϵ^2. However, when the errors are not iid, the appropriate denominator of the test statistic is an estimate of σ^2 instead of σ_ϵ^2. To establish this, we consider the numerator of the test statistic, normalized by division by T^2:

(9) $$\eta = T^{-2} \Sigma S_t^2 \quad .$$

We will show that η has an asymptotic distribution equal to σ^2 times a functional of a Brownian bridge, so that division by $s^2(\ell)$ (or by any other consistent estimate of σ^2) gives a statistic with an asymptotic distribution free of nuisance parameters. For clarity, we will consider separately the tests of the level stationarity and trend stationarity hypotheses.

3.a Level Stationary Hypothesis

The model is as in equation (2) with ξ set to zero, so that the residuals e_t are from a regression of y on intercept only; that is, $e_t = y_t - \bar{y}$. S_t is then the partial sum process of the residuals e, as in equation (5). Let η_μ be as defined in (9), with the subscript "μ" indicating that we have extracted a mean but not a trend from y. It is well known that the partial sums of deviations from means of a process satisfying the assumptions of Phillips and Perron (1988) converge to a Brownian bridge, and this implies that

$$(10) \qquad \eta_\mu \rightarrow \sigma^2 \int_0^1 V(r)^2 dr \quad .$$

Here V(r) is a standard Brownian bridge: $V(r) = W(r) - rW(1)$, where W(r) is a Wiener process (Brownian motion). The symbol "$\rightarrow$" in (10) signifies weak convergence of the associated probability measures. The limit (10) is a special case of a result obtained previously by McCabe and Leybourne (1988) in the context of tests for random regression coefficients.

As noted above, we now divide η by a consistent estimate of σ^2 to get the test statistic that we will actually use. We will indicate this division with a "^", so that the test statistic is

$$(11) \qquad \hat{\eta}_\mu = \eta_\mu \,/\, s^2(\ell) = T^{-2} \Sigma S_t^2 \,/\, s^2(\ell).$$

It follows immediately from (10) and from the consistency of $s^2(\ell)$ that

$$(12) \qquad \hat{\eta}_\mu \rightarrow \int_0^1 V(r)^2\, dr \quad .$$

Table 1 gives upper tail critical values of $\int V(r)^2\, dr$, calculated via a direct simulation, using a sample size of 2000, 50,000 replications, and the random number generator GASDEV/RAN3 of Press, Flannery, Teukolsky and Vetterling (1986). These critical values agree quite closely with those given by MacNeill (1978, Table 2, p. 431), Nyblom and Makelainen (1983, Table 1, p. 859), McCabe and Leybourne (1988, Table 3), and Nabeya and Tanaka (1988, Table 1, p. 232).

3.b Trend Stationary Hypothesis

The analysis of the trend stationary case is very similar to that of the level stationary case. The model is now exactly as in equation (2). Let e_t be the residuals from a regression of y_t on intercept and trend, and let S_t be the partial sum process of the e_t as in (5). Furthermore let η_τ be as defined in (9), where the subscript "τ" indicates that we have extracted a mean and a trend from y, and serves to distinguish the trend stationary case from the level stationary case.

The partial sum process of residuals from a regression of a process satisfying the assumptions of Phillips and Perron (1988) on intercept and trend converges to a so-called second-level Brownian bridge, as given by MacNeill (1978) or Schmidt and Phillips (1989), Appendix 3. Thus we have

$$(13) \qquad \eta_\tau \rightarrow \sigma^2 \int_0^1 V_2(r)\, dr \quad ,$$

where the second-level Brownian bridge $V_2(r)$ is given by

$$(14) \qquad V_2(r) = W(r) + (2r - 3r^2)\, W(1) + (-6r + 6r^2) \int_0^1 W(r)\, dr \; .$$

(Our V_2 is MacNeill's B_1, p. 426.) As previously, we use a "^" to indicate that the test statistic has been divided by a consistent estimate of σ^2, and in this notation the test statistic is

$$(15) \qquad \hat{\eta}_\tau = \eta_\tau \; / \; s^2(\ell) = T^{-2} \, \Sigma \, S_t^2 \; / \; s^2(\ell) \; .$$

Its asymptotic distribution is given by

$$(16) \qquad \hat{\eta}_\tau \rightarrow \int_0^1 V_2(r)^2 \, dr \quad .$$

The upper tail critical values of $\int V_2(r)^2 \, dr$ are also given in Table 1. They agree quite closely with the critical values given by MacNeill (1974, Table 2, p. 431) and Nabeya and Tanaka (1988, Table 2, p. 233).

Note that the limit theory given by (12) and (16) under the null of stationarity is different than that of Stock (1990). Under the null that the time series has a unit root, Stock works with functionals of the detrended series itself and shows (e.g., in his Theorem 2) that those converge weakly to functionals of a detrended Brownian motion. In our case, a partial sum process is constructed from the residuals of a regression with a time series that is stationary under our null, and in consequence the limit process is a Brownian bridge or higher level Brownian bridge, corresponding to the degree of the extracted trend. Thus, even though the functionals that appear in (12) and (16) are related to some of those employed by Stock (as in his extended Sargan-Barghava tests), the limit processes are different. As we shall see in the following Section, so too is the limit theory under the alternative.

4. CONSISTENCY OF THE TEST

In this section we consider the asymptotic distributions of the $\hat{\eta}_\mu$ and $\hat{\eta}_\tau$ tests under the alternative hypothesis that $\sigma_u^2 \neq 0$, so that y is an integrated process. Specifically, our interest is in showing that the tests are consistent. This is non-trivial because, under the alternative hypothesis, both the numerator $[T^{-2} \Sigma S_t^2]$ and the denominator $[s^2(\ell)]$ of the test statistics diverge. Basically, we show that the numerator is of order in probability T^2 [denoted $O_p(T^2)$] while the denominator is $O_p(\ell T)$, so that the test statistic is $O_p(T/\ell)$. Since $T/\ell \to \infty$ as $T \to \infty$, the tests are consistent.

We establish this result first for the level-stationary case. We start with the numerator of the statistic, and we first observe that, since the u_t are iid,

$$(17) \qquad T^{-1/2} r_{[bT]} = T^{-1/2} \sum_{j=1}^{[bT]} u_j \to \sigma_u W(b)$$

where $b \in [0,1]$ and [bT] is the integer part of bT. Then

$$(18) \qquad T^{-3/2} S_{[aT]} = T^{-3/2} \sum_{j=1}^{[aT]} (r_j - \bar{r}) + T^{-3/2} \sum_{j=1}^{[aT]} (\epsilon_j - \bar{\epsilon})$$

$$= T^{-3/2} \sum_{j=1}^{[aT]} (r_j - \bar{r}) + o_p(1)$$

$$= T^{-1} \sum_{j=1}^{[aT]} T^{-1/2} r_j - ([aT]/T)\ T^{-1/2}\ \bar{r}$$

$$\rightarrow \sigma_u \int_0^a W(b)db - a\sigma_u \int_0^1 W(b)db = \sigma_u \int_0^a \underline{W}(s)ds$$

where $\underline{W}(s)$ is the demeaned Wiener process

(19) $$\underline{W}(s) = W(s) - \int_0^1 W(b)db$$

Therefore

(20) $$T^{-4} \sum_{j=1}^{T} S_t^2 = T^{-1} \sum_{j=1}^{T} (T^{-3/2} S_t)^2 \rightarrow \sigma_u^2 \int_0^1 \left[\int_0^a \underline{W}(s)ds \right]^2 da$$

so that $T^{-2} \Sigma S_t^2$ is indeed $O_p(T^2)$ as claimed in the preceeding paragraph.

The argument for the denominator of the test statistic, $s^2(\ell)$, is more straightforward. From Phillips (1991, unnumbered equation between (A10) and (A11)) we have that

(21) $$(\ell T)^{-1} s^2(\ell) \rightarrow K\sigma_u^2 \int_0^1 \underline{W}(s)^2 ds$$

provided $T^{-1/2}\ell \rightarrow 0$ as $T \rightarrow \infty$. The constant K is defined by

(22) $$K = \int_{-1}^{1} k(s)ds$$

where k(s) represents the weighting function used in $s(\ell)$; in our case, $w(s,\ell) = k(s/\ell)$ in the notation of equation (8) above. For the Newey-West estimator, $k(s) = 1 - |s|$ and therefore K = 1. Obviously (21) implies that $s^2(\ell)$ is $O_p(\ell T)$.

Since $T^{-2} \Sigma S_t^2$ is $O_p(T^2)$ and $s^2(\ell)$ is $O_p(\ell T)$, we deduce that $\hat{\eta}_\mu$ is $O_p(T/\ell)$. Given that ℓ grows less quickly than T, the test is consistent. However, we have in fact established more than just the order in probability of the test statistic. Under the alternative hypothesis, (20) and (21) imply that

(23) $$(\ell/T)\hat{\eta}_\mu \rightarrow \int_0^1 \left(\int_0^a \underline{W}(s)ds \right)^2 da \; / \; K \int_0^1 \underline{W}(s)^2 ds$$

Note that this limit is nuisance parameter free because the scale effect from the variance $\sigma_u^2 \neq 0$ in the numerator and denominator of the limit cancels.

The analysis for the trend stationarity test statistic $\hat{\eta}_\tau$ is only slightly more complicated. We just need to replace the demeaned Wiener process $\underline{W}(s)$ above with the demeaned and detrended Wiener process $W^*(s)$:

(24) $$W^*(s) = W(s) + (6b - 4) \int_0^1 W(r)df + (-12b + 6) \int_0^1 rW(r)dr$$

This is given by Park and Phillips (1988, equation (16), p. 474), who prove the equivalent of our (18) above, when S_t is the partial sum process of the residuals of a random walk on intercept and time trend. (Actually, they allow for general integrated processes, but it is part of our maintained hypothesis that r_t is a random walk.) The rest of our analysis then follows without further change.

Again we note the distinction between our limit theory and that of Stock (1990). Under the alterantive of a unit root, our LM statistics $\hat{\eta}_\mu$ and $\hat{\eta}_\tau$ diverge, as does our estimate $s^2(\ell)$ of the long run variance of the stationary component in the model. In Stock's model, and under the null of a unit root, his statistics converge weakly to functionals of Brownian motion. Moreover, his estimate of the long-run variance converges to a non-zero constant, for both the unit root null and trend stationary alternatives. (See his Lemma 1 for the latter.)

5. APPLICATION TO THE NELSON-PLOSSER DATA

In this section we apply our tests for stationarity to the data analyzed by Nelson and Plosser (1982). These are U.S. annual data covering from 62 to 111 years, and ending in 1970. These data have been analyzed subsequently by many others, including Perron (1988) and DeJong et al. (1989). A rough assessment of their findings is as follows. For 12 of the 14 series, we clearly cannot reject the null hypothesis of a unit root. The unit root hypothesis is rejected at about the 5% level for the unemployment rate series, and it is rejected at about the 10% level for the industrial production series. The conventional wisdom is that these results indicate the presence of a unit root in most of the Nelson-Plosser series. We wish to check whether our approach to testing stationarity corroborates this reading of the data. In particular, we propose to test whether the data will reject a null hypothesis of stationarity, rather than just not reject a null of a unit root.

In Table 2 we first present the $\hat{\eta}_\mu$ test statistic which we use to test the null hypothesis of stationarity around a level. We consider values of the lag truncation parameter ℓ (used in the estimation of the long run variance) from zero to eight. The values of the test statistics are fairly sensitive to the choice of ℓ, and in fact for every series the value of the test statistic decreases as ℓ increases. This occurs because $s^2(\ell)$

increases as ℓ increases, and is a reflection of large and persistent positive autocorrelations in the series. Nevertheless, the outcome of the tests is not in very much doubt: for all series except the unemployment rate and the interest rate, we can reject the hypothesis of level stationarity.

The ability to reject the hypothesis of level stationarity is not very surprising in light of the obvious deterministic trends present in these series. We therefore proceed to test the null hypothesis of stationarity around a deterministic linear trend, for which $\hat{\eta}_\tau$ is the appropriate statistic. Once again the test statistics decline monotonically as ℓ increases, and in this case the choice of ℓ is important to the conclusions. If we did not correct for error autocorrelation at all, which corresponds to picking $\ell = 0$, we would reject the null hypothesis of trend stationarity for every series. As argued above, for temporally dependent series such as the ones under consideration, iid errors are not plausible under the null hypothesis, and our empirical results show the importance of allowing for error autocorrelation. Although the test statistics decline monotonically as ℓ increases, for most of the series the value of the long run variance estimate has settled down reasonably by the time we reach $\ell = 8$, and so the value of the test statistic has also settled down. Using the results for $\ell = 8$, we find that we can reject the hypothesis of trend stationarity at the 5% level for five series: industrial production, consumer prices, real wages, velocity and stock prices. For three other series (real GNP, nominal GNP and interest rate) we can reject the hypothesis of trend stationarity at the 10% level. We cannot reject the null hypothesis of trend stationarity at usual critical levels for six series: real per capita GNP, employment, unemployment rate, GNP deflator, wages and money. These empirical results seem to be very much in accord with the Bayesian posterior analysis in Phillips (1990).

Combining the results of our tests of the trend stationarity hypothesis with the results of the Dickey-Fuller tests of the unit root hypothesis, the following picture emerges. The unemployment series appears to be stationary, since we can reject the unit root hypothesis and cannot reject the trend stationarity hypothesis. Four series (consumer prices, real wages, velocity and stock prices) appear to have unit roots, since we can reject the trend stationarity hypothesis and cannot reject the unit root hypothesis. Three more series (real GNP, nominal GNP and the interest rate) probably have unit roots; we cannot reject the unit root hypothesis, and the evidence against the trend stationarity hypothesis is marginally significant. For six series (real per capita GNP, employment, unemployment rate, GNP deflator, wages and money) we cannot reject either the unit root hypothesis or the trend stationarity hypothesis, and the appropriate inference is that the data are not sufficiently informative to tell whether or not there is a unit root. Finally, for the industrial production series, there is evidence against both the unit root hypothesis and the trend stationary hypothesis, and thus it is not clear what to conclude. Presumably other alternatives, such as explosive roots, fractional integration or stationarity around a non-linear trend, would have to be considered.

5. CONCLUDING REMARKS

We have presented statistical tests of the hypothesis of stationarity, either around a level or around a deterministic linear trend. These tests could be extended to allow for non-linear trends, along the same lines as Schmidt and Phillips (1989, section 5). The tests are intended to complement unit root tests, such as the Dickey-Fuller tests. By testing both the unit root hypothesis and the stationarity hypothesis, we can distinguish series that appear to be stationary, series that appear to have a unit root, and series for which the data (or the tests) are not sufficiently informative to be sure whether they are stationary or integrated.

The main technical innovation of this paper is the allowance made for error autocorrelation. Correspondingly, the main practical difficulty in performing the tests is the estimation of the long run variance. The types of series that one would suspect might have a unit root are likely to be highly autocorrelated, even at fairly long lags. Thus the choice of the lag truncation parameter (ℓ) to use in estimating the long run variance is a difficult problem. In our empirical work it is clear that allowing for autocorrelation matters, and the difficulty in choosing ℓ is reflected in the dependence of the results on the choice of ℓ. In this regard it should be noted that we estimate the long run variance from residuals from a fit of the model with the stationarity hypothesis imposed (i.e., with $\sigma_u^2 = 0$ imposed), and so if the null hypothesis is not true we should expect $s^2(\ell)$ to diverge as ℓ increases, as indeed it does. An important topic for further research is to find an estimate of the long run variance σ^2 that is consistent under the null and that increases the rate of divergence of the LM statistic under the alternative.

APPENDIX A

DERIVATION OF THE LM STATISTIC

Equation (4) of the main text is equation (1.1) of Nabeya and Tanaka (1988, p. 218), and uses their notation. For this model the LM statistic for the hypothesis $\sigma_u^2 = 0$ is given by their equation (2.5), p. 219, as follows:

(A.1) $$LM = y'MD_xA_TD_xMy \;/\; y'My \quad .$$

Here M is the projection matrix onto the space orthogonal to (x,Z), so that (My) is the vector of residuals from the regression of y on (x,Z). In our model (2), (x,Z) corresponds to intercept and time trend, so (My) is the vector of residuals from a regression of y on intercept and time trend. (These residuals were called e_t, t = 1, 2, ..., T, in the main text.)

The denominator of the statistic in (A.1), y'My, is just the sum of squared residuals from this regression, and equals $T\hat{\sigma}_\epsilon^2$ in the notation of the text. Apart from a factor of T, which is inessential, this is the same as the denominator of the statistic in equation (6).

The matrix D_x in (A.1) equals identity when x corresponds to intercept, as it does in our case, and can therefore be ignored. Therefore the numerator of the test statistic in (A.1) equals $e'A_Te$, where e is the vector of residuals described in the last paragraph. The matrix A_T has $(t,s)^{th}$ element equal to min(t,s), so that

$$(A.2) \qquad A_T = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & T \end{bmatrix}$$

This matrix creates reverse partial sums. That is, the numerator of the test statistic equals

$$(A.3) \qquad e'A_T e = \sum_{t=1}^{T} R_t^2 \quad , \quad R_t = \sum_{i=t}^{T} e_i \quad .$$

This appears to differ from the numerator of the statistic in equation (6) of the main text, which relies on the forward partial sums S_t defined in (5). However, the two expressions are in fact equal. Because the sum of the residuals is zero, we have $R_1 = S_T = 0$, $S_t = -R_{T-t}$ $(t = 1, 2, \ldots, T)$, and the sum of squares of the S_t equals the sum of squares of the R_t.

TABLE 1

Upper tail critical values for $\hat{\eta}_\mu$

(Upper tail percentiles of the distribution of $\int_0^1 V(r)^2\,dr$)

Critical level:	.10	.05	.025	.01
Critical value:	.347	.463	.574	.739

Upper tail critical values for $\hat{\eta}_\tau$

(Upper tail percentiles of the distribution of $\int_0^1 V_2(r)^2\,dr$)

Critical level:	.10	.05	.025	.01
Critical value:	.119	.146	.176	.216

TABLE 2

STATIONARITY TESTS APPLIED TO NELSON-PLOSSER DATA

$\hat{\eta}_\mu$ Test for Level-Stationarity
(5% critical value is .463)

Series	Lag truncation parameter (ℓ) 0	1	2	3	4	5	6	7	8
Real GNP	5.96	3.06	2.08	1.59	1.30	1.11	0.97	0.86	0.78
Nominal GNP	5.81	2.98	2.04	1.56	1.28	1.09	0.95	0.85	0.77
Real per capital GNP	5.54	2.84	1.94	1.50	1.22	1.05	0.92	0.82	0.75
Industrial production	10.79	5.48	3.70	2.81	2.27	1.92	1.66	1.47	1.32
Employment	7.57	3.87	2.63	2.01	1.64	1.39	1.21	1.08	0.98
Unemployment rate	0.31	0.18	0.14	0.11	0.10	0.10	0.09	0.09	0.09
GNP deflator	7.51	3.82	2.59	1.97	1.60	1.35	1.18	1.04	0.94
Consumer prices	7.90	4.02	2.73	2.08	1.69	1.43	1.24	1.10	0.99
Wages	6.72	3.43	2.33	1.78	1.45	1.23	1.07	0.95	0.86
Real wages	6.96	3.55	2.40	1.83	1.48	1.26	1.09	0.97	0.88
Money	8.01	4.08	2.76	2.10	1.70	1.44	1.25	1.11	1.00
Velocity	8.40	4.29	2.90	2.21	1.80	1.52	1.32	1.17	1.05
Interest rate	0.78	0.42	0.30	0.24	0.20	0.17	0.16	0.14	0.13
Stock prices	8.01	4.10	2.79	2.13	1.74	1.48	1.29	1.15	1.04

$\hat{\eta}_\tau$ Test for Trend-Stationarity
(5% critical value is .146)

Series	Lag truncation parameter (ℓ) 0	1	2	3	4	5	6	7	8
Real GNP	.630	.337	.242	.198	.173	.158	.148	.141	.137
Nominal GNP	.755	.392	.273	.215	.181	.159	.143	.132	.124
Real per capital GNP	.528	.283	.204	.167	.147	.134	.126	.121	.118
Industrial production	.822	.446	.320	.257	.220	.196	.179	.166	.155
Employment	.526	.278	.198	.158	.136	.122	.112	.105	.101
Unemployment rate	.216	.124	.094	.079	.071	.066	.063	.061	.061
GNP deflator	.492	.256	.178	.140	.117	.103	.093	.086	.081
Consumer prices	1.85	.943	.641	.491	.401	.342	.301	.270	.246
Wages	.612	.317	.220	.173	.145	.128	.115	.107	.101
Real wages	.956	.511	.365	.293	.252	.226	.208	.194	.184
Money	.445	.228	.158	.124	.104	.092	.084	.079	.075
Velocity	1.78	.932	.647	.504	.418	.360	.319	.287	.262
Interest rate	.845	.457	.323	.255	.214	.186	.166	.151	.140
Stock prices	1.23	.646	.454	.359	.302	.264	.237	.216	.199

REFERENCES

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

DeJong, D.N. and C.H. Whiteman (1990), "Trends and Random Walks in Macroeconomic Time Series: A Reconsideration Based on the Likelihood Principle," Journal of Monetary Economics, forthcoming.

DeJong, D.N., J.C. Nankervis, N.E. Savin and C.H. Whiteman (1989), "Integration Versus Trend Stationarity in Macroeconomic Time Series," Working Paper No. 89-99, 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.

Diebold, F.X. and G.D. Rudebusch (1990), "On the Power of Dickey-Fuller Tests against Fractional Alternatives," Finance and Economics Discussion Series Paper 119, Board of Governors of the Federal Reserve System.

Franzini, L. and A.C. Harvey (1983), "Testing for Deterministic Trend and Seasonal Components in Time Series Models," Biometrika, 70, 673-682.

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

Harvey, A.C. (1989), Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge: Cambridge University Press.

Leybourne, S.J. and B.P.M. McCabe (1989), "On the Distribution of Some Test Statistics for Parameter Constancy," Biometrika, 76, 169-177.

MacNeill, I. (1978), "Properties of Sequences of Partial Sums of Polynomial Regression Residuals with Applications to Tests for Change of Regression at Unknown Times," Annals of Statistics, 6, 422-433.

McCabe, B.P.M. and S.J. Leybourne (1988), "Testing for Variation in Multivariate Random Coefficient Regression Models," unpublished manuscript, University of Sydney.

Nabeya, S. and K. Tanaka (1988), "Asymptotic Theory of a Test for the Constancy of Regression Coefficients against the Random Walk Alternative," Annals of Statistics, 16, 218-235.

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.

Newey, W.K. and K.D. West (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," Econometrica, 55, 703-708.

Nyblom, J. (1986), "Testing for Deterministic Linear Trend in Time Series," Journal of the American Statistical Association, 81, 545-549.

Nyblom, J. and T. Makelainen (1983), "Comparisons of Tests for the Presence of Random Walk Coefficients in a Simple Linear Model," Journal of the American Statistical Association, 78, 856-864.

Park, J.Y. and B. Choi (1988), "A New Approach to Testing for a Unit Root," Working Paper No. 88-23, Center for Analytical Economics, Cornell University.

Park, J.Y. and P.C.B. Phillips (1988), "Statistical Inference in Regressions with Integrated Processes: Part I," Econometric Theory, 4, 468-498.

Perron, P. (1988), "Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach," Journal of Economic Dynamics and Control, 12, 297-332.

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

Phillips, P.C.B. (1991), "Spectral Regression for Cointegrated Time Series," in W. Barnett (ed.), Nonparametric and Semiparametric Methods in Economics and Statistics, Cambridge University Press, forthcoming.

Phillips, P.C.B. (1990), "To Criticize the Critics: An Objective Bayesian Analysis of Stochastic Trends," Cowles Foundation Discussion Paper No. 950, Yale University.

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. and V. Solo (1989), "Asymptotics for Linear Processes," Cowles Foundation Discussion Paper No. 932, Yale University.

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.

Rudebusch, G.D. (1990), "Trends and Random Walks in Macroeconomic Time Series: A Reexamination," Working Paper Series, Number 105, Economic Activity Section, Board of Governors of the Federal Reserve System.

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

Schmidt, P. and P.C.B. Phillips (1989), "Testing for a Unit Root in the Presence of Deterministic Trends," unpublished manuscript, Michigan State University.

Saikkonen, P. and R. Luukkonen, "Testing for a Moving Average Unit Root," unpublished manuscript, University of Helsinki.

Stock, J.H. (1990), "A Class of Tests for Integration and Cointegration," unpublished manuscript, Harvard University.

Tanaka, K. (1983), "Non-Normality of the Lagrange Multiplier Statistic for Testing the Constancy of Regression Coefficients," Econometrica, 51, 1577-1582.