Econometrica, Vol. 66, No. 1 (January, 1998), 123–148

TREND FUNCTION HYPOTHESIS TESTING IN THE PRESENCE OF SERIAL CORRELATION

BY TIMOTHY J. VOGELSANG[1]

In this paper test statistics are proposed that can be used to test hypotheses about the parameters of the deterministic trend function of a univariate time series. The tests are valid in the presence of general forms of serial correlation in the errors and can be used without having to estimate the serial correlation parameters either parametrically or nonparametrically. The tests are valid for $I(0)$ and $I(1)$ errors. Trend functions that are permitted include general linear polynomial trend functions that may have breaks at either known or unknown locations. Asymptotic distributions are derived, and consistency of the tests is established. The general results are applied to a model with a simple linear trend. A local asymptotic analysis is used to compute asymptotic size and power of the tests for this example. Size is well controlled and is relatively unaffected by the variance of the initial condition. Asymptotic power curves are computed for the simple linear trend model and are compared to existing tests. It is shown that the new tests have nontrivial asymptotic power. A simulation study shows that the asymptotic approximations are adequate for sample sizes typically used in economics. The tests are used to construct confidence intervals for average GNP growth rates for eight industrialized countries using post-war data.

KEYWORDS: Wald test, hypothesis test, partial sum, unit root, structural change, conservative test.

1. INTRODUCTION

IN THIS PAPER STATISTICS ARE PROPOSED to test general linear hypotheses regarding parameters of the deterministic trend function of a univariate time series. The framework is general enough to include most deterministic trend functions that are linear in parameters including polynomial trend functions that may have breaks at known or unknown dates. The innovations of the time series may be serially correlated and have up to one unit root. A priori knowledge as to whether the innovations are $I(0)$ or $I(1)$ is not required. When the innovations are modeled as local to a unit root in a model with a simple linear trend, the tests can be carried out without knowledge of the local to unity parameter or knowledge of the variance of the initial condition. This result is important as the local to unity parameter and the variance of the initial condition cannot be consistently estimated.

[1] I thank Serena Ng, Pierre Perron, and Mark Watson for helpful comments and seminar participants at Cornell University, University of Rochester, Princeton University, and the 1996 Winter Meetings of the Econometrics Society. I am grateful to a co-editor and two anonymous referees for helpful comments and suggestions that helped to improve the paper. I also thank the Center for Analytic Economics at Cornell University.

Another useful property of the statistics from an applied perspective is that the statistics are asymptotically invariant to all the serial correlation parameters. Estimates of the serial correlation parameters, either parametric or nonparametric, are not needed, and the sometimes subjective finite sample choices such as lag length, information criteria, kernel or truncation lag can be completely avoided. All that is required is that a functional central limit theorem hold for the partial sums of the innovations.

Since the focus of this paper is trend function hypothesis testing, the serial correlation, i.e. the dynamics, are viewed as nuisance parameters. This would be the appropriate framework from an economic standpoint if the goal were, for example, forming confidence intervals on growth rates of GNP or testing for and identifying structural shifts in growth of GNP. Sometimes, however, dynamics are of interest, usually in a multivariate framework, e.g. business cycles, and individual series are often transformed to remove nonstationarities. These transformations commonly involve detrending and removing unit roots. Obviously, detrending requires a well specified trend function. From the work of Perron (1988, 1989, 1990) it is well known that misspecification of the trend function can result in highly misleading unit root tests. So, even if the trend function is not of direct interest, correct specification of the trend function is required for additional modeling.

The remainder of the paper is organized as follows. In order to motivate the usefulness of the statistics developed later in the paper, the next section contains a simple Monte Carlo experiment which illustrates some of the problems that arise when the form of serial correlation is not known. It is shown that OLS-based Wald statistics suffer from substantial finite sample size distortions. If the standard Wald statistic is normalized by the sample size, a test with correct size is obtained. This statistic has good power when the errors are $I(1)$ but lacks power when the errors are $I(0)$. New statistics are proposed in Section 3 that have good size and are powerful when the errors are $I(0)$. These statistics along with the normalized Wald statistic comprise a class of tests with good size and complementary power. Asymptotic results are presented including limiting null distributions for $I(0)$ and $I(1)$ errors and conditions under which consistency holds. In Section 4 the general results are applied to a model with a simple linear trend function. Limiting distributions are tabulated, and the statistics are evaluated by examining asymptotic size and power. In Section 5 results from finite sample size and power simulations for the simple linear trend model are given. As long as the errors are not $I(1)$ with an MA component with a root close to but not equal to one, the new tests have finite sample size close to the nominal level. Section 6 contains an empirical application. Confidence intervals for the average growth rate of real per capita GNP are constructed using quarterly post war data for eight industrialized countries. Confidence intervals for pre-1973 and post-1973 samples indicate that growth rates have slowed in many cases. Section 7 contains concluding remarks, and all proofs are given in an Appendix.

2. BACKGROUND AND MOTIVATION

The literature on trend function hypothesis testing is quite large and no attempt is made to summarize it here. However, it is useful to review some classic and some recent results so that the statistics proposed in this paper can be placed in context. To facilitate the discussion consider a very simple time series model for $t = 1, 2, \ldots, T$,

$$y_t = \beta_1 + \beta_2 t + u_t, \tag{1}$$

where $\{u_t\}$ is a mean zero error process. If $\{y_t\}$ is the logarithm of output, then β_2 is the average growth rate of output. Suppose one were interested in inference about β_2. When $\{u_t\}$ is $I(0)$, OLS estimates of (1) are efficient since they are asymptotically normal with variance equivalent to GLS. This follows directly from the classic results of Grenander and Rosenblatt (1957). Since the asymptotic variance in model (1) is proportional to the spectral density of $\{u_t\}$ at frequency zero, asymptotically valid inference can be carried out using any consistent estimate of the spectral density of $\{u_t\}$. If, on the other hand, $\{u_t\}$ is $I(1)$, OLS applied to (1) no longer has optimality properties. But upon first differencing, the model $\Delta y_t = \beta_2 + \Delta u_t$ is obtained and OLS again has optimality properties.

In practice there are two considerations that can make inference more difficult to carry out than suggested by the results of Grenander and Rosenblatt (1957). First, it is often unknown whether errors are $I(0)$ or $I(1)$, and in series with highly persistent errors often neither $I(0)$ nor $I(1)$ asymptotics provide good approximations to finite sample distributions. Second, sampling variability in spectral density estimates can lead to poor approximations of asymptotic distributions in finite samples. Furthermore, choices such as kernel and truncation lag (in nonparametric estimation) and lag length (in parametric estimation) can lead to conflicting results in practice.

The first consideration is addressed by Canjels and Watson (1997) where model (1) is analyzed, and the errors are modeled as local to a unit root, or nearly $I(1)$. They consider several feasible GLS estimators and find that asymptotic distributions depend on a local to unity parameter and the variance of the initial condition. Since these parameters cannot be consistently estimated from the data, they proposed conservative tests based on Bonferoni confidence intervals. They establish that nearly $I(1)$ asymptotics provide a good approximation to finite sample distributions for parameter values often found in economic data. Ultimately, they show that the Prais and Winsten (1954) feasible GLS estimator performs the best in practice. A drawback of the Canjels and Watson (1997) approach in a more general framework is that construction of the Bonferoni confidence intervals can be quite demanding, although for model (1) computing confidence intervals is fairly simple.

Regarding the second consideration, sampling variability is clearly unavoidable when estimating the spectrum and may not be an issue if the sample size is large enough. With regards to choices such as kernel and truncation lag in nonparametric estimation, much progress has been made recently on data dependent methods for making such choices. Robinson (1991) proposed and established the consistency of a data dependent method for choosing the truncation lag using a cross-validation procedure. Andrews (1991) proposed consistent data dependent methods for choosing the truncation lag using the "plug-in" method. Andrews (1991) also established the optimality of the quadratic spectral kernel in certain models. Andrews and Monahan (1992) found that prewhitening can improve the estimators considered by Andrews (1991). For parametric models of serial correlation Hall (1994) and Ng and Perron (1995) have shown that data dependent methods for choosing lag lengths in autoregressive approximations work well in unit root tests.

To illustrate how serial correlation in the errors can affect inference regarding β_2 in model (1) a simple Monte Carlo experiment was conducted. The null hypothesis was $\beta_2 = 0$ and data were generated using the following ARMA(1, 1) model for $\{u_t\}$:

$$(2) \qquad u_t = \alpha u_{t-1} + \eta_t + \theta\eta_{t-1},$$

with $\{\eta_t\}$ i.i.d. $N(0,1)$ random deviates and $\eta_0 = 0$. The $\{\eta_t\}$ were generated using the ran1( ) subroutine of Press et al. (1992) using $-T$ as the initial seed. Without loss of generality, $\beta_1 = 0$. In all cases 1,000 replications were used, and the nominal level was 5%. Results are given for $\alpha = 0.8$, 0.9, 0.95, 1.0, and $\theta = -1.0, -0.8, -0.4, 0.0, 0.4, 0.8$.

Three Wald statistics were used to test the null hypothesis. The first is a Wald test based on the OLS estimate of β_2 from (1). The asymptotic variance of OLS is proportional to the spectral density at frequency zero of $\{u_t\}$ which was estimated using the quadratic spectral kernel with the truncation lag chosen according to the automatic procedure of Andrews and Monahan (1992) using AR(1) prewhitening. This test is labeled OLS. The limiting distribution of this statistic is χ_1^2 when errors are $I(0)$. The second test is based on the Prais-Winsten (1954) feasible GLS estimate of (1) using the AR(1) transformation as suggested by Canjels and Watson (1997). Additional correlation in the model was estimated using an autoregressive spectral estimate with the lag length of the autoregressive approximation chosen using a data dependent method suggested by Ng and Perron (1995). The conservative Bonferoni test based on the local to $I(1)$ asymptotic approximation as recommended by Canjels and Watson (1997) was used. The limiting distribution of this statistic is χ_1^2. The third test is the standard Wald test (using the OLS estimate of the variance) normalized by T^{-1} which is labelled $T^{-1}W_T$. The 5% asymptotic critical value for $T^{-1}W_T$ was taken from Table II(ii) with details given in Section 3. Phillips and Durlauf (1988) derived the limiting distribution of $T^{-1}W_T$ under $I(1)$ errors, but they did not tabulate critical values nor suggest using $T^{-1}W_T$ as a test statistic. One of the contributions of this paper is to tabulate critical values for $T^{-1}W_T$ and to determine its properties as a test statistic.

Table I reports null rejection probabilities. Table I also reports results for statistics that are proposed later in the paper. The first result given by the table is that the OLS test has severe size distortions. As α approaches one, rejection probabilities become very large and worsen as T increases. This is not surprising since the OLS statistic diverges to ∞ when $\alpha = 1.0$. There are also distortions when the errors are clearly $I(0)$. In unreported simulations it was found that these distortions disappear if the OLS variance is assumed to be known, suggesting that sampling variability of the estimate of the spectrum is a source of size distortion. The GLS statistic has much better size with rejection probabilities near or below 0.05 except when $\alpha = 1.0$ and $\theta = -0.8$ in which case rejection probabilities exceed 0.15. If this range is excluded from the parameter space, GLS has good size. The $T^{-1}W_T$ statistic has size close to 0.05 since rejection probabilities are always near or below 0.05. GLS and $T^{-1}W_T$ are conservative when $\alpha < 1$ as rejection probabilities are close to zero.

If controlling size were the only concern, then GLS and $T^{-1}W_T$ would be good candidate statistics. Since these statistics are based on $I(1)$ asymptotics, power of these tests should be good for α close to one. This was shown by Canjels and Watson (1997) for GLS. When α is not close to one, GLS and $T^{-1}W_T$ are conservative and will lack power. In the next section statistics are proposed which maintain the good size properties of GLS and $T^{-1}W_T$ but are designed to be powerful when the errors are $I(0)$.

3. THE GENERAL MODEL AND ASYMPTOTIC RESULTS

3.1 *The General Model*

Consider a univariate time series process $\{y_t\}$, $t = 1, 2, \ldots, T$, generated by

$$y_t = f(t)'\beta + u_t, \tag{3}$$

where $f(t) = [f_1(t), f_2(t), \ldots, f_k(t)]'$ is a $(k \times 1)$ vector of trends, $\beta = (\beta_1, \beta_2, \ldots, \beta_k)'$ is a $(k \times 1)$ vector of trend parameters and $\{u_t\}$ is an error process satisfying

$$(1 - L\alpha_T)u_t = v_t, \quad t = 2, 3, \ldots, T, \quad u_1 = \sum_{i=0}^{[\kappa T]} \alpha_T^i v_{1-i}, \tag{4}$$

$$v_t = d(L)e_t, \quad d(L) = \sum_{i=0}^{\infty} d_i L^i, \quad \sum_{i=0}^{\infty} i|d_i| < \infty, \quad \text{and} \quad d(1)^2 > 0, \tag{5}$$

where $\{e_t\}$ is a martingale difference sequence with $E(e_t^2 | e_{t-1}, e_{t-2}, \ldots) = 1$ and $\sup_t Ee_t^4 < \infty$, L is the lag operator, and $[x]$ is the integer part of x. Assume there exists a $(k \times k)$ diagonal normalization matrix τ_T and a $(k \times 1)$ vector of functions $F(r)$ defined on $[0, 1]$ such that

$$\tau_T f(t) = F(t/T) + o(1), \quad \int_0^1 F(r)\, dr < \infty, \quad \text{and} \tag{6}$$

$$\det\left[\int_0^1 F(r)F(r)'\, dr\right] > 0.$$

TABLE I

FINITE SAMPLE NULL HYPOTHESIS REJECTION PROBABILITIES, ARMA(1, 1) ERRORS USING 5% ASYMPTOTIC CRITICAL VALUES

MODEL: $y_t = \beta_1 + \beta_2 t + u_t$, $z_t = \beta_1 t + \beta_2[\frac{1}{2}(t^2 + t)] + S_t$, $u_t = \alpha u_{t-1} + v_t + \theta v_{t-1}$; H_0: $\beta_2 = 0$; 1,000 REPLICATIONS

		$T = 100$					$T = 250$					$T = 500$				
α	θ	OLS	GLS	$T^{-1}W_T$	PS_T^1	PS_T^2	OLS	GLS	$T^{-1}W_T$	PS_T^1	PS_T^2	OLS	GLS	$T^{-1}W_T$	PS_T^1	PS_T^2
0.8	−1.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	−0.8	0.060	0.000	0.000	0.033	0.042	0.066	0.000	0.000	0.047	0.050	0.055	0.000	0.000	0.058	0.060
	−0.4	0.207	0.011	0.000	0.037	0.041	0.211	0.000	0.000	0.049	0.046	0.201	0.000	0.000	0.059	0.057
	0.0	0.103	0.004	0.000	0.024	0.036	0.082	0.001	0.000	0.037	0.045	0.059	0.000	0.000	0.045	0.058
	0.4	0.032	0.000	0.000	0.021	0.034	0.016	0.001	0.000	0.035	0.045	0.010	0.000	0.000	0.042	0.058
	0.8	0.021	0.000	0.000	0.020	0.036	0.007	0.000	0.000	0.033	0.045	0.004	0.000	0.000	0.042	0.058
0.9	−1.0	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	−0.8	0.222	0.000	0.000	0.075	0.061	0.247	0.000	0.000	0.067	0.051	0.260	0.000	0.000	0.075	0.058
	−0.4	0.332	0.023	0.000	0.033	0.041	0.308	0.005	0.000	0.035	0.041	0.293	0.000	0.000	0.049	0.054
	0.0	0.144	0.006	0.000	0.023	0.037	0.095	0.004	0.000	0.029	0.042	0.075	0.000	0.000	0.045	0.053
	0.4	0.045	0.003	0.000	0.019	0.036	0.027	0.002	0.000	0.029	0.041	0.010	0.000	0.000	0.043	0.053
	0.8	0.031	0.000	0.000	0.018	0.036	0.012	0.000	0.000	0.029	0.041	0.005	0.000	0.000	0.043	0.053
0.95	−1.0	0.004	0.000	0.000	0.002	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	−0.8	0.394	0.025	0.000	0.125	0.074	0.446	0.003	0.000	0.092	0.044	0.462	0.000	0.000	0.087	0.049
	−0.4	0.423	0.023	0.000	0.036	0.034	0.409	0.006	0.000	0.032	0.036	0.370	0.001	0.001	0.044	0.047
	0.0	0.212	0.007	0.001	0.018	0.031	0.122	0.002	0.000	0.023	0.036	0.106	0.001	0.001	0.038	0.047
	0.4	0.084	0.008	0.001	0.012	0.027	0.044	0.000	0.000	0.018	0.036	0.014	0.000	0.000	0.034	0.047
	0.8	0.048	0.001	0.001	0.012	0.027	0.029	0.000	0.000	0.018	0.036	0.007	0.000	0.000	0.034	0.046
1.0	−1.0	0.066	0.000	0.000	0.035	0.043	0.062	0.000	0.000	0.050	0.056	0.045	0.000	0.000	0.046	0.049
	−0.8	0.681	0.211	0.000	0.364	0.166	0.819	0.149	0.001	0.303	0.074	0.895	0.185	0.005	0.242	0.060
	−0.4	0.669	0.077	0.026	0.113	0.066	0.717	0.039	0.043	0.075	0.049	0.772	0.047	0.053	0.065	0.053
	0.0	0.487	0.045	0.056	0.054	0.058	0.436	0.037	0.052	0.048	0.043	0.454	0.043	0.060	0.055	0.053
	0.4	0.324	0.051	0.060	0.038	0.056	0.275	0.039	0.056	0.040	0.043	0.283	0.035	0.061	0.052	0.052
	0.8	0.261	0.058	0.060	0.035	0.055	0.223	0.048	0.057	0.038	0.043	0.240	0.040	0.062	0.052	0.052

Assumptions (4) and (5) are the same assumptions used by Canjels and Watson (1997). A functional central limit theorem applies to the partial sums of $\{e_t\}$ so that $T^{-1/2}\sum_{t=1}^{[rT]} e_t \Rightarrow w(r)$ where $w(r)$ is a standard Wiener process and $\Rightarrow$ denotes weak convergence in distribution. It directly follows that $T^{-1/2}\sum_{t=1}^{[rT]} v_t \Rightarrow d(1)w(r)$. The restriction $d(1)^2 > 0$ is equivalent to restricting the spectral density at frequency zero of $\{v_t\}$ to be positive. This rules out nondegenerate cases. Other conditions on $\{u_t\}$, such as mixing conditions popularized by Phillips (1987a), could be considered without changing any of the results that follow provided that the partial sums of $\{u_t\}$, $S_t = \sum_{j=1}^{t} u_j$, satisfy a functional central limit theorem. Assumptions (4) and (5) include $I(0)$ and $I(1)$ errors. When $\alpha_T = \alpha$ and $|\alpha| < 1$, $\{u_t\}$ is $I(0)$. When $\alpha_T = (1 - \bar{\alpha}/T)$, $\{u_t\}$ is a nearly $I(1)$ process[2] (a pure $I(1)$ process when $\bar{\alpha} = 0$). Throughout the paper $I(1)$ denotes errors with $\alpha_T = (1 - \bar{\alpha}/T)$. Under these assumptions a functional central limit theorem applies to $\{S_t\}$. When $\{u_t\}$ is $I(0)$, $T^{-1/2}S_{[rT]} \Rightarrow \sigma w(r)$ where $\sigma^2 = d(1)^2/(1-\alpha)^2$, when $\{u_t\}$ is $I(1)$, $T^{-1/2}u_{[rT]} \Rightarrow d(1)V_{\bar{\alpha}}(r)$ where $V_{\bar{\alpha}}(r) = w_{\bar{\alpha}}(r) + \exp(-r\bar{\alpha})\tilde{w}_{\bar{\alpha}}(\kappa)$, $w_{\bar{\alpha}}(r) = \int_0^r \exp(-\bar{\alpha}(r-s))\,dw(s)$, $\tilde{w}_{\bar{\alpha}}(\kappa) = \int_0^\kappa \exp(-\bar{\alpha}(\kappa - s))\,d\tilde{w}(s)$, and $\tilde{w}(r)$ is a standard Wiener process independent with $w(r)$. Assumption (4) incorporates the effects of the initial condition into the asymptotic distribution theory. The κ parameter governs the variance of the initial condition. When $\kappa = 0$, u_1 is an $O_p(1)$ random variable. When $\kappa > 0$, u_1 is $O_p(1)$ when $\{u_t\}$ is $I(0)$ but is $O_p(T^{1/2})$ when $\{u_t\}$ is $I(1)$.

Assumption (6) simplifies the asymptotic representation of the statistics and rules out ill behaved trend functions like $f_i(t) = 1/t$ and trends that are asymptotically collinear. Assumption (6) is general enough to permit polynomial trends possibly with a finite number of structural changes. More general assumptions on the trend function such as the conditions given in Grenander and Rosenblatt (1957) could be used, but the assumptions above suffice for most models of interest.

By forming partial sums of $\{y_t\}$, (3) can be transformed as

$$(7) \qquad z_t = g(t)'\beta + S_t,$$

where $g(t) = \sum_{j=1}^{t} f(j)$ and $z_t = \sum_{j=1}^{t} y_j$. When $\{u_t\}$ is $I(0)$, $\{z_t\}$ has $I(1)$ innovations, while when $\{u_t\}$ is $I(1)$, $\{z_t\}$ has $I(2)$ innovations. Define $G(r) = \int_0^r F(s)\,ds$. The properties of $g(t)$ and $G(r)$ follow from (6) as $T^{-1}\tau_T g(t) = G(t/T) + o(1)$, $\int_0^1 G(r)\,dr < \infty$ and $\det[\int_0^1 G(r)G(r)'\,dr] > 0$. It is convenient to write models (3) and (7) as

$$(8) \qquad Y = X_1\beta + u,$$

$$(9) \qquad Z = X_2\beta + S,$$

[2] Local to unity asymptotics is becoming standard in time series analysis involving integrated data. Theoretical background can be found in Bobkoski (1983), Chan and Wei (1987), Phillips (1987b), and Nabeya and Perron (1994).

where $Y=\{y_t\}$, $Z=\{z_t\}$, $u=\{u_t\}$, and $S=\{S_t\}$ are $(T\times 1)$ vectors and $X_1=\{f(t)'\}$ and $X_2=\{g(t)'\}$ are $(T\times k)$ matrices. Let $\hat{\beta}$ denote the OLS estimate of β from (8) and let β^* denote the OLS estimate of β from (9). Define the OLS estimates of the error variance in (8) and (9) as $s_y^2=T^{-1}Y'M_{x_1}Y$ and $s_z^2=T^{-1}Z'M_{x_2}Z$ where $M_x=I-x(x'x)^{-1}x'$.

3.2 *The Test Statistics*

In order to define the test statistics some preliminary developments are needed. Let t^{j-1} be the highest order polynomial of t in $f(t)$. Consider the regression models,

$$(10)\qquad y_t=f(t)'\beta+\sum_{i=j}^{m}\gamma_i t^i+u_t,$$

$$(11)\qquad z_t=g(t)'\beta+\sum_{i=j+1}^{m}\gamma_i t^i+S_t.$$

Let $J_T^1(m)$ denote the standard OLS Wald statistic normalized by T^{-1} for testing the joint hypothesis $\gamma_j=\gamma_{j+1}=\cdots=\gamma_m=0$ in (10) and let $J_T^2(m)$ denote the standard OLS Wald statistic normalized by T^{-1} for testing the joint hypothesis $\gamma_{j+1}=\gamma_{j+2}=\cdots=\gamma_m=0$ in (11). $J_T^1(m)$ is a unit root statistic proposed by Park and Choi (1988) and Park (1990). When the errors are $I(0)$, $J_T^1(m)$ converges to zero. When the errors are $I(1)$, $J_T^1(m)$ has a nondegenerate limiting distribution. As a unit root statistic, $J_T^1(m)$ is a left tailed test. $J_T^2(m)$ is not a unit root statistic and has nondegenerate limiting distributions for $I(0)$ and $I(1)$ errors with the $I(1)$ limiting distribution skewed much farther to the right than the $I(0)$ limiting distribution. $J_T^1(m)$ and $J_T^2(m)$ are used to smooth discontinuities in the limiting distributions of the statistics of interest as the errors go from $I(0)$ to $I(1)$.

Consider testing hypotheses regarding β of the form

$$H_0\colon R\beta=r,\qquad H_1\colon R\beta\neq r,$$

where R is a $(q\times k)$ matrix of constants and r is a $(q\times 1)$ matrix of constants. Using the OLS estimate of β from (8), the $T^{-1}W_T$ statistic is defined as

$$T^{-1}W_T=T^{-1}\left(R\hat{\beta}-r\right)'\left[R(X_1'X_1)^{-1}R'\right]^{-1}\left(R\hat{\beta}-r\right)/s_y^2.$$

A statistic similar to $T^{-1}W_T$ constructed from the OLS estimate of β from (9) is defined as

$$PS_T^i=T^{-1}(R\beta^*-r)'\left[R(X_2'X_2)^{-1}R'\right]^{-1}(R\beta^*-r)/\left(s_z^2\exp(bJ_T^i(m)\right)\qquad (i=1,2),$$

where b is a constant. A third statistic which is the standard OLS Wald statistic with s_y^2 replaced by another sample moment is

$$PSW_T^i = \frac{(R\hat{\beta} - r)'\left[R(X_1'X_1)^{-1}R'\right]^{-1}(R\hat{\beta} - r)}{\left[T^{-1}100s_z^2\exp(bJ_T^i(m))\right]} \qquad (i = 1, 2).$$

The 100 is included in the denominator of PSW_T^i to normalize the critical values and becomes important computationally in some applications including tests for structural change (see Vogelsang (1997b)). This normalization does not affect the size or power of PSW_T^i. Different values for b are used for each statistic. When $q = 1$, t statistics can be defined which are useful for one sided tests and constructing confidence intervals.

The PS_T^i and PSW_T^i statistics are designed to have power when the errors are $I(0)$ but remain robust to $I(1)$ errors in terms of size. Suppose that $b = 0$ so that the $J_T^i(m)$ statistics disappear from the definitions of PS_T^i and PSW_T^i. In this case PS_T^i and PSW_T^i have nondegenerate limiting distributions for both $I(0)$ and $I(1)$ errors, but the distributions are skewed much farther to the right when the errors are $I(1)$ compared to $I(0)$. This would make asymptotically valid inference possible only if it were known whether the errors were $I(0)$ or $I(1)$. When $b > 0$, the $J_T^i(m)$ statistics smooth the discontinuities of PS_T^i and PSW_T^i as the errors go from $I(0)$ to $I(1)$ by taking on large values for $I(1)$ errors and small values for $I(0)$ errors. For example, when the errors are $I(0)$, $J_T^1(m)$ is zero asymptotically and has no effect on the distribution of PS_T^1. When the errors are $I(1)$, the distribution of PS_T^1 is skewed to the right and at the same time $J_T^1(m)$ becomes nondegenerate and $\exp(bJ_T^1(m))$ takes on large values, reducing the critical values of PS_T^1. Therefore, the b's can be chosen to bring the $I(0)$ and $I(1)$ distributions of PS_T^i and PSW_T^i close together. Given a percentage point, b can be chosen so the $I(0)$ and $I(1)$ critical values are the same, providing a test that has the correct asymptotic size simultaneously for $I(0)$ and $I(1)$ errors. The $J_T^1(m)$ statistic has the nice property that it does not affect the asymptotic power of PS_T^1 and PSW_T^1 when the errors are $I(0)$. While $J_T^2(m)$ does affect power in the $I(0)$ case, PS_T^2 and PSW_T^2 have better finite sample size in certain cases and, unlike PS_T^1 and PSW_T^1, can be configured to be robust to fractionally integrated errors. Results with fractionally integrated errors are available upon request.

3.3 *Asymptotic Results*

In this section limiting distributions of the statistics under the null hypothesis are given and are expressed as functionals of standard Brownian motions. The limiting distributions of PS_T^i and PSW_T^i depend on the trends in the model since partial summing of the data places a unit root in the model by construction. Likewise, when the errors are $I(1)$, the limiting distribution of $T^{-1}W_T$ depends on the trends in the model. The asymptotic results obtained in this section are new and are used by Vogelsang (1994, 1997b) to construct tests for structural change in the trend function.

Because estimators of coefficients on some trends converge at different rates, additional notation is needed. Let $\mu_i \leq 0$ be the largest nonpositive power of T appearing in the nonzero elements in the ith row of $R\tau_T$. Define a $(q \times q)$ diagonal matrix A such that $A_{jj} = T^{\mu_j}$, and let $R^* = \lim_{T\to\infty} A^{-1} R\tau_T$. The R^* matrix pulls from R the elements of the constraint that converge the slowest and dominate the limiting distribution.

Limiting distributions are expressed in terms of the following functionals. Let $\nu(r)$ be a scalar function and let $B_1(r)$ and $B_2(r)$ be $(n_1 \times 1)$ and $(n_2 \times 1)$ vectors of functions all that lie on the space $[0,1]$. Let C be a $(q \times n_1)$ matrix of constants. Define the functionals,

$$K(\nu(r), B_1(r), C) \equiv \left[C\left(\int_0^1 B_1(r) B_1(r)' \, dr \right)^{-1} \int_0^1 B_1(r) \, d\nu(r) \right]'$$

$$\times \left[C\left(\int_0^1 B_1(r) B_1(r)' \, dr \right)^{-1} C' \right]^{-1}$$

$$\times C\left(\int_0^1 B_1(r) B_1(r)' \, dr \right)^{-1} \int_0^1 B_1(r) \, d\nu(r),$$

$$L(\nu(r), B_1(r), C) \equiv \left[C\left(\int_0^1 B_1(r) B_1(r)' \, dr \right)^{-1} \int_0^1 B_1(r) \nu(r) \, dr \right]'$$

$$\times \left[C\left(\int_0^1 B_1(r) B_1(r)' \, dr \right)^{-1} C' \right]^{-1}$$

$$\times C\left(\int_0^1 B_1(r) B_1(r)' \, dr \right)^{-1} \int_0^1 B_1(r) \nu(r) \, dr,$$

$$M(\nu(r), B_1(r)) \equiv \int_0^1 \nu(r)^2 \, dr - \int_0^1 B_1(r)' \nu(r) \, dr$$

$$\times \left(\int_0^1 B_1(r) B_1(r)' \, dr \right)^{-1} \int_0^1 B_1(r) \nu(r) \, dr,$$

$$J(\nu(r), B_1(r), B_2(r))$$

$$\equiv [M(\nu(r), B_1(r)) - M(\nu(r), B_2(r))] / M(\nu(r), B_2(r)).$$

In addition define $Q_1(r) = (F(r)', r^j, r^{j+1}, \ldots, r^m)'$ and $Q_2(r) = (G(r)', r^{j+1}, r^{j+2}, \ldots, r^m)'$. The limiting null distributions of PS_T^i, PSW_T^i, and $T^{-1} W_T$ depend on whether the errors are $I(0)$ or $I(1)$ and are summarized in the following theorems.

THEOREM 1: *Suppose* $\{u_t\}$ *follows* (4) *and* (5) *and* $\alpha_T = \alpha, |\alpha| < 1$ *so that* $\{u_t\}$ *is* $I(0)$. *If* (6) *holds for model* (3), *then under* H_0 *as* $T \to \infty$,

$$T^{-1} W_T \Rightarrow 0,$$

$$PS_T^1 \Rightarrow L(w(r), G(r), R^*)/M(w(r), G(r)),$$

$$PSW_T^1 \Rightarrow K(w(r), F(r), R^*)/[100M(w(r), G(r))],$$

$$PS_T^2 \Rightarrow L(w(r), G(r), R^*)/[M(w(r), G(r)) \times \exp\{bJ(w(r), G(r), Q_2(r))\}],$$

$$PSW_T^2 \Rightarrow K(w(r), F(r), R^*)/[100M(w(r), G(r)) \times \exp\{bJ(w(r), G(r), Q_2(r))\}].$$

THEOREM 2: *Suppose* $\{u_t\}$ *follows* (4) *and* (5) *and* $\alpha_T = (1 - \bar{\alpha}/T)$ *so* $\{u_t\}$ *is* $I(1)$. *If* (6) *holds for model* (3), *then under* H_0 *as* $T \to \infty$,

$$T^{-1} W_T \Rightarrow L(V_{\bar{\alpha}}(r), F(r), R^*)/M(V_{\bar{\alpha}}(r), F(r)),$$

$$PS_T^1 \Rightarrow L\left(\int_0^r V_{\bar{\alpha}}(s)\, ds, G(r), R^*\right) \Big/ \left[M\left(\int_0^r V_{\bar{\alpha}}(s)\, ds, G(r)\right) \exp\{bJ(V_{\bar{\alpha}}(r), F(r), Q_1(r))\}\right],$$

$$PSW_T^1 \Rightarrow L(V_{\bar{\alpha}}(r), F(r), R^*) \Big/ \left[100M\left(\int_0^r V_{\bar{\alpha}}(s)\, ds, G(r)\right) \exp\{bJ(V_{\bar{\alpha}}(r), F(r), Q_1(r))\}\right],$$

$$PS_T^2 \Rightarrow L\left(\int_0^r V_{\bar{\alpha}}(s)\, ds, G(r), R^*\right) \Big/ \left[M\left(\int_0^r V_{\bar{\alpha}}(s)\, ds, G(r)\right) \exp\left\{bJ\left(\int_0^r V_{\bar{\alpha}}(s)\, ds, G(r), Q_2(r)\right)\right\}\right],$$

$$PSW_T^2 \Rightarrow L(V_{\bar{\alpha}}(r), F(r), R^*) \Big/ \left[100M\left(\int_0^r V_{\bar{\alpha}}(s)\, ds, G(r)\right) \times \exp\left\{bJ\left(\int_0^r V_{\bar{\alpha}}(s)\, ds, G(r), Q_2(r)\right)\right\}\right].$$

The proofs are given in the Appendix. When $q = 1$, t statistics can be considered and their limiting null distributions can be expressed using similar notation. The distributions in the theorems are nonstandard and depend on the trends in the model, but critical values can be easily simulated case by case. With the exception of $\bar{\alpha}$ and κ, the distributions are asymptotically free of nuisance parameters, and the tests can be implemented without requiring estimates of serial correlation parameters (σ^2 or $d(1)$). For PS_T^i and PSW_T^i suppose the b's are chosen so that the $I(0)$ critical values and the $\bar{\alpha} = \kappa = 0$ $I(1)$ critical values are the same, and for $T^{-1}W_T$ suppose the $\bar{\alpha} = \kappa = 0$ $I(1)$ critical values are used. When implemented in this way for model (1), it is shown in the next section that practically speaking the statistics are conservative with respect to $\bar{\alpha}$ and κ and the tests can be used without knowledge of $\bar{\alpha}$ and κ. This robustness to $\bar{\alpha}$ and κ holds in many other models as well; see Vogelsang (1994, 1997b). Finally, to implement the PS_T^i and PSW_T^i statistics m must be chosen. The choice of m can be justified asymptotically by examining the size/power trade off when the errors are $I(1)$ as is done for model (1) in the next section.

3.4 *Consistency*

Conditions for consistency of the tests are established under the fixed alternative

$$H_1\colon R\beta = r + \delta, \tag{12}$$

where δ is a $(q \times 1)$ vector of constants. The results are given by the following theorems.

THEOREM 3: *Suppose* $\{u_t\}$ *follows* (4) *and* (5) *with* $\alpha_T = \alpha, |\alpha| < 1$ *so that* $\{u_t\}$ *is* $I(0)$, *and* (6) *and* (12) *hold for model* (3). *Then as* $T \to \infty, PS_T^i, PSW_T^i$ *diverge to* ∞. *If there exists a* j *such that* $\delta_j \neq 0$ *and* $A_{jj}^{-1} > T^{\gamma}$ *for* $\gamma > 0$, *then as* $T \to \infty, T^{-1}W_T$ *diverges to* ∞.

THEOREM 4: *Suppose* $\{u_t\}$ *follows* (4) *and* (5) *with* $\alpha_T = (1 - \bar{\alpha}/T)$ *so* $\{u_t\}$ *is* $I(1)$, *and* (6) *and* (12) *hold for model* (3). *If there exists a* j *such that* $\delta_j \neq 0$ *and* $A_{jj}^{-1} > T^{1/2}$, *then as* $T \to \infty$, $T^{-1}W_T$, PS_T^i, *and* PSW_T^i *diverge to* ∞. *If* $A_{jj}^{-1} \leq T^{1/2}$ *for all* i *such that* $\delta_j \neq 0$, *then as* $T \to \infty$, $T^{-1}W_T$, PS_T^i *and* PSW_T^i *are* $O_p(1)$.

The proofs are given in the Appendix. When the errors are $I(0)$, Theorem 3 establishes that PS_T^i and PSW_T^i are consistent while $T^{-1}W_T$ is consistent unless departures from the null involve trends that are not growing over time. For example, $T^{-1}W_T$ cannot be used to consistently test hypotheses about the constant. When the errors are $I(1)$, the tests are consistent provided that departures from the null do not exclusively involve parameters on regressors that are growing slower than $T^{1/2}$.

4. MODEL WITH SIMPLE LINEAR TREND

Using the theorems in Section 3, results are easily obtained for specific models. In this section results are presented for model (1). Transforming (1) by partial summing gives

$$z_t = \beta_1 t + \beta_2\left[\tfrac{1}{2}(t^2 + t)\right] + S_t.$$

In the notation of Section 3 we have $f(t) = [1, t]'$, $\tau_T = \mathrm{diag}(1, T^{-1})$, $F(r) = [1, r]'$, $g(t) = [t, \frac{1}{2}(t^2 + t)]'$, $G(r) = [r, \frac{1}{2}r^2]'$, $Q_1(r) = (1, r, r^2, \ldots, r^m)'$, and $Q_2(r) = (r, \frac{1}{2}r^2, r^3, \ldots, r^m)'$. The hypothesis of interest is H_0: $\beta_2 = \beta_0$ or H_0: $R\beta = r$ with $R = [0, 1]$, $r = \beta_0$ giving $A = [T^{-1}]$ and $R^* = [0, 1]$. The limiting null distributions follow directly from Theorems 1 and 2. Since $q = 1$, t-statistic versions of the tests can be used and are denoted by $t\text{-}PS_T^i$, $t\text{-}PSW_T^i$, and $T^{-1/2}t\text{-}W_T$.

The limiting null distributions of PS_T^i, PSW_T^i, $t\text{-}PS_T^i$, and $t\text{-}PSW_T^i$ in the $I(0)$ case are presented in Table II(i). The critical values were simulated using $N(0, 1)$ i.i.d. random deviates to approximate the Wiener processes in Theorem 1. The integrals were approximated by normalized sums of 1,000 steps using 10,000 replications. The random number generator used was ran1( ) taken from Press, et al. (1992) with initial seed of -1000. Only right tail critical values are reported as they are of interest for hypothesis testing. Left tailed tests can be carried out using the t statistics since they have symmetric distributions. The numbers in parentheses under the critical values are the b's that result in the $\bar{\alpha} = \kappa = 0$ $I(1)$ critical values with $m = 9$ being the same as the $I(0)$ critical values for that percentage point. The b's were obtained using simulations. The limiting null distributions of $T^{-1}W_T$ and $T^{-1/2}t\text{-}W_T$ for $\bar{\alpha} = \kappa = 0$ $I(1)$ errors are given in Table II(ii) with the critical values simulated using the same methods as for Table II(i). The limiting distribution of $T^{-1}W_T$ in this case was obtained by Phillips and Durlauf (1988), but they did not tabulate critical values. For the remainder of this subsection the focus is on the t statistic version of the tests as analogous results hold for the Wald statistics. Results are not reported for the $t\text{-}PSW_T^i$ statistics given the similarity to results for the $t\text{-}PS_T^i$ statistics.

TABLE II(i)

ASYMPTOTIC DISTRIBUTIONS: PS_T^i, PSW_T^i, $t\text{-}PS_T^i$, AND $t\text{-}PSW_T^i$, $I(0)$ ERRORS
MODEL: $y_t = \beta_1 + \beta_2 t + u_t$, $z_t = \beta_1 t + \beta_2[\frac{1}{2}(t^2 + t)] + S_t$;
H_0: $\beta_2 = \beta_0$; b'S GIVEN IN PARENTHESES

	90.0%	95.0%	97.5%	99.0%
PS_T^1	3.017	4.537	6.121	8.759
	(1.451)	(1.966)	(2.685)	(3.946)
PSW_T^1	0.674	1.013	1.436	1.986
	(1.535)	(2.085)	(2.848)	(4.227)
$t\text{-}PS_T^1$	1.331	1.720	2.152	2.647
	(0.494)	(0.716)	(0.995)	(1.501)
$t\text{-}PSW_T^1$	0.601	0.818	1.015	1.245
	(0.539)	(0.750)	(1.036)	(1.603)
PS_T^2	2.027	2.784	3.322	3.949
	(0.183)	(0.286)	(0.455)	(0.753)
PSW_T^2	0.448	0.612	0.785	0.887
	(0.191)	(0.305)	(0.468)	(0.783)
$t\text{-}PS_T^2$	1.152	1.392	1.677	1.849
	(0.050)	(0.095)	(0.147)	(0.265)
$t\text{-}PSW_T^2$	0.520	0.652	0.793	0.890
	(0.056)	(0.099)	(0.156)	(0.267)

TABLE II(ii)

ASYMPTOTIC DISTRIBUTIONS: $T^{-1}W_T$ AND $T^{-1/2}t\text{-}W_T$, $I(1)$ ERRORS WITH $\bar{\alpha}=0$, $\kappa=0$
MODEL: $y_t=\beta_1+\beta_2 t+u_t$, $z_t=\beta_1 t+\beta_2[\frac{1}{2}(t^2+t)]+S_t$; H_0: $\beta_2=\beta_0$

	90.0%	95.0%	97.5%	99.0%
$T^{-1}W_T$	5.161	7.727	11.004	15.370
$T^{-1/2}t\text{-}W_T$	1.724	2.298	2.835	3.479

In order to justify using $\bar{\alpha}=\kappa=0$ b's for $t\text{-}PS_T^i$ and $\bar{\alpha}=\kappa=0$ critical values for $T^{-1/2}t\text{-}W_T$, asymptotic rejection probabilities were simulated for $\bar{\alpha}=-0.6,-0.4,\ldots,12.0$, and $\kappa=0,0.2,\ldots,1.0$, with the nominal level 5% using 10,000 replications. To place the values of $\bar{\alpha}$ in perspective, $\bar{\alpha}=4$ corresponds to $\alpha=0.96$ in a sample with 100 observations. Results with $\kappa=0$ are reported in Table III, panel (a) for $t\text{-}PS_T^1$ $(m=2,3,\ldots,9)$, $t\text{-}PS_T^2$ $(m=9)$, and $T^{-1/2}t\text{-}W_T$. Rejection probabilities for $t\text{-}PS_T^1$ are at or below 0.05 regardless of m except when $\bar{\alpha}=-2$ (explosive errors) in which case rejection probabilities are slightly higher than 0.05. Similar results hold for $t\text{-}PS_T^2$. Therefore, asymptotic size of $t\text{-}PS_T^1$ and $t\text{-}PS_T^2$ does not depend on m. As long as $\bar{\alpha}\geq 0$, $T^{-1/2}t\text{-}W_T$ has rejection probabilities at or well below 0.05. Simulations reported in the working paper Vogelsang (1996) showed that rejection probabilities are not sensitive to κ and that asymptotic size remains at 0.05 for $\kappa>0$. Since, when the errors are $I(0)$, rejection probabilities are 0.05 by construction for $t\text{-}PS_T^1$ and $t\text{-}PS_T^2$ and zero by construction for $T^{-1/2}t\text{-}W_T$, the tests have asymptotic size of 0.05 whether the errors are $I(0)$ or are $I(1)$ provided explosive errors are ruled out.

Since asymptotic size of $t\text{-}PS_T^1$ and $t\text{-}PS_T^2$ does not depend on m, a natural way to choose m is to maximize asymptotic power. There does not appear to be an analytic method of maximizing power with respect to m, but heuristic evidence can be provided. Using the local alternatives H_1: $\beta_2=\beta_0+T^{-3/2}c$ for $I(0)$ errors and H_1: $\beta_2=\beta_0+T^{-1/2}c$ for $I(1)$ errors, local asymptotic distributions were obtained and used to compute asymptotic power functions. The local asymptotic distributions were simulated using similar methods to those used to obtain limiting null critical values. Asymptotic power depends on c/σ in the $I(0)$ case and on $c/d(1)$ in the $I(1)$ case. Since the expressions of these limiting distributions are not particularly useful themselves, they are not reported but can be found for $t\text{-}PS_T^1$ and $t\text{-}PSW_T^1$ in the working paper Vogelsang (1996).

Asymptotic power of $t\text{-}PS_T^1$ does not depend on m in the $I(0)$ case. Therefore, the choice of m was made on the basis of power in the $I(1)$ case which is given in Panel (b) of Table III for the same range of m as in panel (a). The range of local alternatives (relative to $d(1)$) is $5,10,\ldots 25$. Power is clearly increasing in m given $c/d(1)$. In unreported simulations it was found that power does not increase substantially for $m>9$ and flattens off. Thus, $m=9$ results in tests with approximately the highest power. Similar results were obtained for $t\text{-}PS_T^2$. In unreported simulations it was also found that when the errors are $I(0)$, power

TABLE III

ASYMPTOTIC SIZE AND POWER OF $t\text{-}PS_T^1$, $t\text{-}PS_T^2$, AND $T^{-1/2}t\text{-}W_T$; $I(1)$ ERRORS
NOMINAL SIZE = 0.05, USING $\bar{\alpha} = 0$, $\kappa = 0$ CRITICAL VALUES
MODEL: $y_t = \beta_1 + \beta_2 t + u_t$, $z_t = \beta_1 t + \beta_2[\frac{1}{2}(t^2 + t)] + S_t$, $u_t = (1 - \bar{\alpha}/T)u_{t-1} + v_t$;
H_0: $\beta_2 \leq \beta_0$, H_1: $\beta_2 = \beta_0 + T^{-1/2}c$, $\kappa = 0$

Panel (a): Null Rejection Probabilities

		$t\text{-}PS_T^1$								$t\text{-}PS_T^2$	
$\bar{\alpha}$	$c/d(1)$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$	$m=9$	$m=9$	$T^{-1/2}t\text{-}W_T$
−6	0.0	0.000	0.001	0.000	0.001	0.000	0.001	0.000	0.000	0.001	0.000
−4	0.0	0.003	0.004	0.003	0.003	0.003	0.002	0.003	0.003	0.004	0.070
−2	0.0	0.049	0.064	0.061	0.064	0.068	0.066	0.068	0.072	0.063	0.275
0	0.0	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050
2	0.0	0.036	0.033	0.029	0.028	0.027	0.026	0.026	0.026	0.033	0.033
4	0.0	0.037	0.031	0.028	0.023	0.025	0.024	0.023	0.023	0.028	0.000
6	0.0	0.035	0.030	0.028	0.025	0.025	0.024	0.023	0.024	0.030	0.000
8	0.0	0.033	0.030	0.029	0.027	0.027	0.024	0.024	0.024	0.033	0.000
10	0.0	0.033	0.029	0.029	0.029	0.028	0.026	0.026	0.025	0.034	0.000
12	0.0	0.031	0.030	0.029	0.030	0.029	0.027	0.026	0.027	0.037	0.000

Panel (b): Power

		$t\text{-}PS_T^1$							
$\bar{\alpha}$	$c/d(1)$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$	$m=9$
0	5	0.238	0.315	0.338	0.363	0.380	0.384	0.397	0.417
	10	0.269	0.384	0.429	0.456	0.485	0.492	0.511	0.540
	15	0.285	0.420	0.473	0.505	0.533	0.547	0.568	0.598
	20	0.296	0.445	0.501	0.534	0.570	0.582	0.605	0.633
	25	0.305	0.463	0.524	0.556	0.594	0.609	0.629	0.659
5	5	0.343	0.486	0.525	0.552	0.578	0.586	0.600	0.630
	10	0.386	0.565	0.623	0.652	0.690	0.698	0.717	0.745
	15	0.407	0.603	0.673	0.706	0.738	0.751	0.769	0.797
	20	0.420	0.630	0.700	0.739	0.773	0.784	0.802	0.828
	25	0.430	0.650	0.719	0.763	0.793	0.804	0.823	0.849
10	5	0.453	0.685	0.750	0.786	0.812	0.822	0.835	0.855
	10	0.499	0.769	0.832	0.863	0.886	0.896	0.909	0.927
	15	0.522	0.805	0.870	0.893	0.914	0.924	0.938	0.951
	20	0.537	0.826	0.890	0.909	0.931	0.941	0.950	0.963
	25	0.547	0.841	0.900	0.920	0.943	0.950	0.959	0.970

of $t\text{-}PS_T^2$ follows a similar pattern. Therefore, in all cases the asymptotic power results suggest the use of $m = 9$ in practice. In what follows, $m = 9$ is always used.

Since asymptotic size of the statistics is excellent whether the errors are $I(0)$ or $I(1)$, it is useful to compare asymptotic power of the statistics. In all cases that follow the nominal level was 5%. Consider the case where the errors are $I(0)$. Under the $I(0)$ local alternative, asymptotic power of $T^{-1/2}t\text{-}W_T$ is equal to size since the limiting distributions is the same as under the null hypothesis.

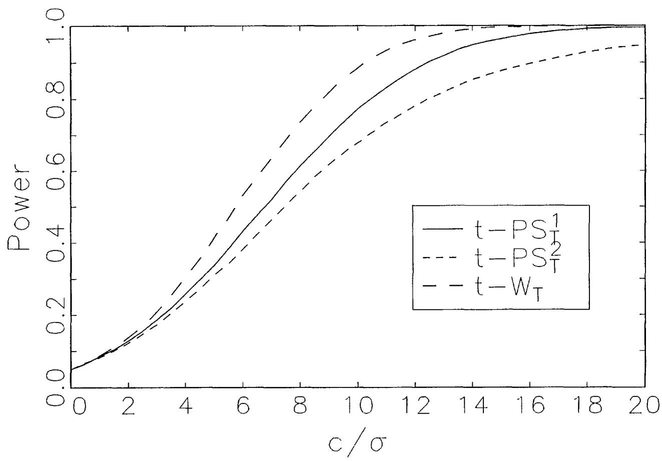

FIGURE 1.—Asymptotic power in model (1), $I(0)$ errors; H_0: $\beta_2 \le \beta_0$, H_1: $\beta_2 = \beta_0 + cT^{-3/2}$.

Asymptotic power of $t\text{-}PS_T^1$ and $t\text{-}PS_T^2$ is not degenerate and is plotted in Figure 1. As a baseline for comparison, asymptotic power of $t\text{-}W_T$, the unnormalized t statistic, is also plotted. When the errors are $I(0)$, $t\text{-}W_T$ can have optimality properties. Asymptotic power of the $t\text{-}W_T$ statistic was computed analytically since its limiting distribution is normal under both the null and local alternative. As expected, $t\text{-}W_T$ is the most powerful statistic, but $t\text{-}PS_T^1$ and $t\text{-}PS_T^2$ have nontrivial asymptotic power with $t\text{-}PS_T^1$ slightly more powerful than $t\text{-}PS_T^2$.

Now consider power when the errors are $I(1)$. Under the $I(1)$ local alternative all of the statistics including $T^{-1/2}t\text{-}W_T$ have nondegenerate limiting distributions. Asymptotic power is plotted for $\bar{\alpha} = 0, 5, 10$ in Figures 2–4. For a baseline

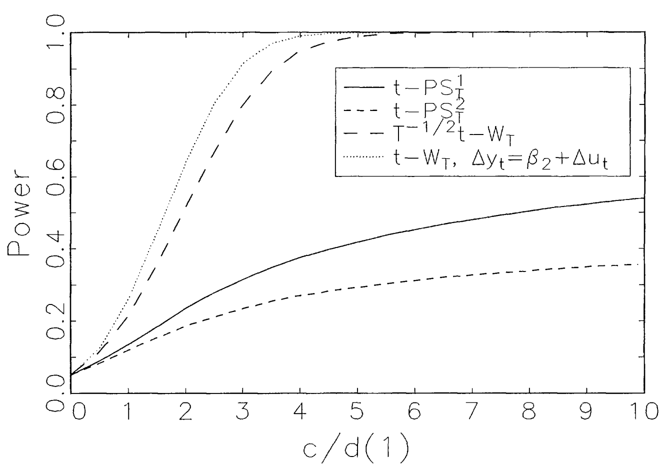

FIGURE 2.—Asymptotic power in model (1), $I(1)$ errors, $\bar{\alpha} = 0$; H_0: $\beta_2 \le \beta_0$, H_1: $\beta_2 = \beta_0 + cT^{-1/2}$.

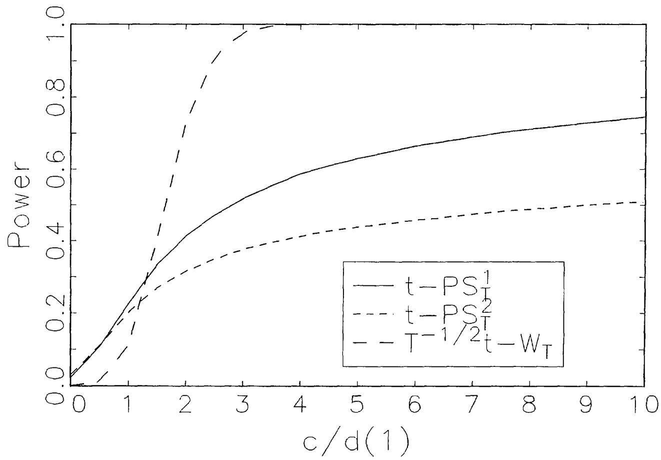

FIGURE 3.—Asymptotic power in model (1), $I(1)$ errors, $\bar{\alpha}=5$; H_0: $\beta_2 \leq \beta_0$, H_1: $\beta_2=\beta_0+cT^{-1/2}$.

of comparison, Figure 2 contains a plot of the asymptotic power of $t\text{-}W_T$ based on the model $\Delta y_t=\beta_2+\Delta u_t$ which is an optimal statistic when $\bar{\alpha}=0$ and $\{v_t\}$ are i.i.d. and normal. As can be seen from Figure 2, when $\bar{\alpha}=0$, $T^{-1/2}t\text{-}W_T$ is much more powerful than $t\text{-}PS_T^1$ and $t\text{-}PS_T^2$ and has comparable, but lower, power to the optimal $t\text{-}W_T$ statistic. As $\bar{\alpha}$ increases, the rankings of the statistics change with $t\text{-}PS_T^1$ and $t\text{-}PS_T^2$ having higher power for alternatives close to the null and $T^{-1/2}t\text{-}W_T$ having higher power for alternatives far from the null (see

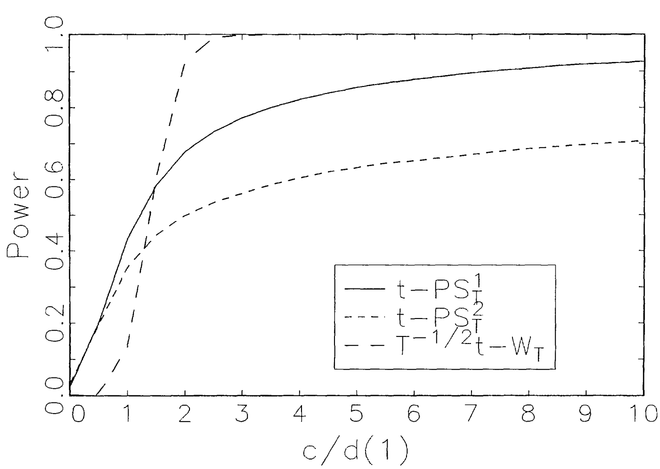

FIGURE 4.—Asymptotic power in model (1), $I(1)$ errors, $\bar{\alpha}=10$; H_0: $\beta_2 \leq \beta_0$, H_1: $\beta_2=\beta_0+cT^{-1/2}$.

Figures 3 and 4). These results suggest that the statistics will have complementary power in finite samples as the errors go from $I(0)$ to $I(1)$.

5. FINITE SAMPLE SIMULATIONS FOR LINEAR TREND MODEL

This section provides evidence on the finite sample size and power of the new tests in model (1). Results are given for the Wald versions of the tests with similar results holding for the t statistics. Results are only reported for PS_T^i as results for PSW_T^i are similar. Without loss of generality, consider testing the hypothesis H_0: $\beta_2 = 0$. In Table I rejection null probabilities are reported for PS_T^1 and PS_T^2 using $m = 9$ and 5% asymptotic critical values. Recall that $u_0 = 0$ which is equivalent to setting $\kappa = 0$. Unreported simulations (available upon request) using $u_0 \neq 0$ were also performed, and it was found that the finite sample behavior of the statistics does not depend in any practical way on u_0. When $T = 100$, rejection probabilities of PS_T^1 and PS_T^2 are close to or below 0.05 unless $\theta = -0.8$ and α is close to 1. If MA errors with a root close to but not equal to one cannot be ruled out, the tests are oversized. As T increases the size distortions disappear for PS_T^2. Recall that GLS is also oversized in the same region of the nuisance parameter space. If there is the possibility of $I(1)$ errors with an MA component with a root near but not equal to one, GLS and PS_T^1 should be used with caution.

Finite sample power with $\beta_2 = 0.0$, 0.05, 0.1, and 0.2 for $T = 100$ is reported in Table IV and was obtained using simulations with 1,000 replications and $\{u_t\}$ generated using (2) with $\alpha = 0.9$, 0.95, and 1.0 and $\theta = -0.4$, 0.0, and 0.4. These are parameter values where exact size is close to nominal size and power comparisons are meaningful. To mimic how the tests are used in practice, 5% asymptotic critical values were used so power is not size adjusted. There is no statistic that uniformly dominates the other statistics. When $\beta_2 = 0.2$, $T^{-1}W_T$ often has the highest power. When $\alpha = 1.0$, $T^{-1}W_T$ and GLS have the highest power. When $\alpha = 0.9$ and $\beta_2 \leq 0.1$, PS_T^1 and PS_T^2 have the highest power with PS_T^1 more powerful when $\theta < 0$ and PS_T^2 more powerful when $\theta > 0$. When $\alpha = 0.95$, either PS_T^1 or PS_T^2 has the most power for $\beta_2 \leq 0.1$, while for $\beta_2 = 0.2$, $T^{-1}W_T$ is the most powerful. These power results are consistent with the local asymptotic predictions.

6. EMPIRICAL APPLICATION

The t statistics from model (1) were used to construct confidence intervals for estimates of post WWII real GNP quarterly growth rates for the U.S., Canada, France, Germany, Italy, Japan, and the U.K. The sources of the data can be found in Banerjee, Lumsdaine, and Stock (1992). Perron (1991), Banerjee, Lumsdaine, and Stock (1992), and Vogelsang (1997a) concluded that there is considerable evidence of a shift in slope for many of the series. Jointly for France, Germany, and Italy, Bai, Lumsdaine, and Stock (1994) constructed a 90% confidence interval of (1972, 1975) for the estimated break date which

TABLE IV

FINITE SAMPLE POWER, ARMA(1,1) ERRORS USING 5% ASYMPTOTIC CRITICAL VALUES
MODEL: $y_t = \beta_1 + \beta_2 t + u_t$, $z_t = \beta_1 t + \beta_2[\frac{1}{2}(t^2 + t)] + S_t$, $u_t + \alpha u_{t-1} + v_t + \theta v_{t-1}$;
H_0: $\beta_2 = 0$; $T = 100$, 1,000 REPLICATIONS; POWER IS NOT SIZE ADJUSTED

α	θ	β_2	GLS	$T^{-1}W_T$	PS_T^1	PS_T^2
0.9	−0.4	0.0	0.023	0.000	0.033	0.041
		0.05	0.206	0.001	0.383	0.254
		0.10	0.325	0.178	0.702	0.390
		0.20	0.878	0.995	0.871	0.500
0.9	0.0	0.0	0.006	0.000	0.023	0.037
		0.05	0.024	0.001	0.116	0.131
		0.10	0.084	0.023	0.265	0.254
		0.20	0.316	0.734	0.502	0.378
0.9	0.4	0.0	0.003	0.000	0.019	0.036
		0.05	0.014	0.001	0.050	0.089
		0.10	0.049	0.007	0.147	0.187
		0.20	0.150	0.305	0.319	0.318
0.95	−0.4	0.0	0.023	0.000	0.036	0.034
		0.05	0.098	0.002	0.202	0.128
		0.10	0.197	0.159	0.411	0.231
		0.20	0.826	0.938	0.642	0.320
0.95	0.0	0.0	0.007	0.001	0.018	0.031
		0.05	0.014	0.002	0.054	0.067
		0.10	0.044	0.028	0.122	0.134
		0.20	0.302	0.514	0.251	0.230
0.95	0.4	0.0	0.008	0.001	0.012	0.027
		0.05	0.012	0.002	0.034	0.047
		0.10	0.032	0.007	0.066	0.094
		0.20	0.094	0.207	0.150	0.172
1.0	−0.4	0.0	0.077	0.026	0.113	0.066
		0.05	0.106	0.089	0.138	0.075
		0.10	0.262	0.250	0.228	0.118
		0.20	0.719	0.740	0.407	0.194
1.0	0.0	0.0	0.045	0.056	0.054	0.058
		0.05	0.076	0.080	0.059	0.058
		0.10	0.143	0.152	0.073	0.073
		0.20	0.414	0.425	0.131	0.116
1.0	0.4	0.0	0.051	0.060	0.038	0.056
		0.05	0.059	0.073	0.046	0.056
		0.10	0.088	0.112	0.050	0.060
		0.20	0.218	0.257	0.078	0.088

Note: The underlined entries are the highest power for that simulation.

includes 1973, the date often associated with slowdowns in growth. None of these studies reported confidence intervals for the estimated growth rates of GNP.

What can be said about the magnitude of GNP growth rates before 1973 compared to after 1973? Using asymptotic critical values, 90% confidence intervals for the full period as well as for periods before and after 1973 were constructed by inverting the $t\text{-}PS_T^i$ $(i = 1,2)$, $T^{-1/2}t - W_T$, and GLS statistics. The GLS statistic was computed using an autoregressive spectral density estimator with the order of the autoregression chosen using the general to specific data dependent method analyzed by Ng and Perron (1995) where the last included lag is checked for significance using a two-tailed 10% test. This data dependent method requires a maximal autoregressive lag length, *kmax*. Results are reported for $kmax = 5$ and 10. The selected lag length, k, is reported along with the estimate of the sum of autoregressive parameters, $\hat{\alpha}$. The results are given in Table V. In all cases the point estimates of β_2 are higher pre-1973 than post-1973, indicating that growth rates have slowed. For France, Germany, Italy, and Japan, the pre-1973 and post-1973 confidence intervals often do not overlap, indicating the drop in growth rates is statistically significant. Growth rates fell by as much as 50% and many confidence intervals indicate that point estimates are fairly precise.

This empirical application illustrates some interesting properties of the statistics. First, GLS confidence intervals are often sensitive to the choice of *kmax*. This is particularly true for post-1973 Canada and Germany where increasing *kmax* from 5 to 10 results in different choices of lag length and vastly different confidence intervals. This shows how different estimates of serial correlation parameters can have large effects on empirical results. These problems are avoided by the other statistics as estimates of serial correlation parameters are not required. Second, tighter confidence intervals are often obtained with $T^{-1/2}t\text{-}W_T$ compared to $t\text{-}PS_T^i$ when $\hat{\alpha}$ is close to one. This is not surprising as $T^{-1/2}t\text{-}W_T$ is often more powerful when errors are $I(1)$ as shown above. Third, confidence intervals for the full period are often wide when using PS_T^i. In many cases the confidence interval end points exceed ± 99. Such wide confidence intervals should simply be interpreted as noninformative; that is, the statistics cannot be used to reject any reasonable null hypothesis regarding β_2. This is equivalent to saying that the tests have low power. The asymptotic and finite sample results suggest that PS_T^i has low power when the errors are $I(1)$. This is confirmed by Table V as wide confidence intervals occur when $\hat{\alpha}$ is close to 1. When $\hat{\alpha}$ is not close to 1, the confidence intervals using PS_T^i are often much tighter as predicted by the theory.

There is another potential source of wide confidence intervals when using the $t\text{-}PS_T^i$ statistics for the full samples. In model (1) it is assumed that the slope is not changing over time, but many of the series have a slope shift around 1973. In a simulation experiment reported in the working paper, Vogelsang (1996), it was shown that a shift in slope causes the PS_T^i confidence intervals based on model (1) to be wide. As the shift in slope increases, confidence intervals

TABLE V

90% CONFIDENCE INTERVALS FOR GROWTH RATES OF POST WAR REAL GNP

			$T^{-1/2}t\text{-}W_T$			$t\text{-}PS_T^1$			$t\text{-}PS_T^2$			$GLS, kmax = 5$					$GLS, kmax = 10$[b]				
Country	T	Period	$\beta_{2\,\min}$	$\hat{\beta}_2$	$\beta_{2\,\max}$	$\beta_{2\,\min}$	β_2^*	$\beta_{2\,\max}$	$\beta_{2\,\min}$	β_2^*	$\beta_{2\,\max}$	$\hat{\alpha}$	k	$\beta_{2\,\min}$	$\hat{\beta}_{2pw}$	$\beta_{2\,\max}$	$\hat{\alpha}$	k	$\beta_{2\,\min}$	$\hat{\beta}_{2pw}$	$\beta_{2\,\max}$
Canada	166	48:1 − 89:2	0.88	1.11	1.34	−1.92	1.15	4.23	−13.6	1.15	15.9	0.97	1	0.92	1.11	1.29	0.97	10	0.92	1.11	1.29
	104	48:1 − 73:4	1.00	1.20	1.40	1.06	1.19	1.32	1.05	1.19	1.33	0.87	1	0.96	1.22	1.47	0.84	10	0.86	1.22	1.58
	62	74:1 − 89:2	0.48	0.80	1.12	−1.78	0.77	3.32	0.05	0.77	1.48	0.90	3	0.43	0.83	1.23	0.86	10	−16.3	0.83	18.0
France	106	6:1 − 89:2	0.38	0.82	1.26	< −99	0.89	> 99	< −99	0.89	> 99	0.97	3	0.75	0.89	1.03					
	44	63:1 − 73:4	1.04	0.31	1.58	1.23	1.31	1.39	1.18	1.31	1.44	0.06	1	1.27	1.31	1.35					
	62	74:1 − 89:2	0.35	0.53	0.71	−4.94	0.53	6.00	< −99	0.53	> 99	0.92	1	0.40	0.54	0.69					
France[a]	105	63:1 − 89:2	0.38	0.82	1.25	< −99	0.88	> 99	< −99	0.88	> 99	0.97	1	0.75	0.88	1.01	0.98	7	0.76	0.88	1.01
	43	63:1 − 73:4	1.16	1.31	1.46	1.19	1.32	1.45	1.27	1.32	1.37	0.38	1	1.31	1.32	1.34	0.42	6	0.97	1.32	1.67
	62	74:1 − 89:2	0.35	0.53	0.71	−4.94	0.53	6.00	< −99	0.53	> 99	0.92	1	0.40	0.54	0.69					
Germany	158	50:1 − 89:2	0.39	0.98	1.57	< −99	1.07	> 99	< −99	1.07	> 99	0.97	1	0.86	1.09	1.31	0.97	9	0.93	1.09	1.25
	96	50:1 − 73:4	0.87	1.37	1.88	< −99	1.44	> 99	< −99	1.44	> 99	0.93	1	1.14	1.46	1.78					
	62	74:1 − 89:2	0.24	0.50	0.76	−1.80	0.49	2.78	−3.98	0.49	4.96	0.78	5	0.07	0.51	0.95	0.80	10	−6.43	0.51	7.45
Italy	124	52:1 − 82:4	0.75	1.14	1.54	< −99	1.22	> 99	< −99	1.22	> 99	1.01	1	0.87	1.10	1.33					
	88	52:1 − 73:4	1.15	1.34	1.53	1.05	1.36	1.67	0.85	1.36	1.87	0.83	1	1.09	1.34	1.59	0.98	8	0.86	1.34	1.70
	36	74:1 − 82:4	0.07	0.62	1.17	−95.8	0.68	97.2	−22.3	0.68	23.7	0.75	4	−0.61	0.50	1.61	1.07	7	0.27	0.50	0.72
Japan	150	52:1 − 89:2	0.99	1.74	2.50	< −99	1.88	> 99	< −99	1.88	> 99	1.00	5	1.41	1.69	1.97	1.00	9	1.39	1.69	1.98
	88	52:1 − 73:4	1.93	2.30	2.67	0.63	2.29	3.96	< −99	2.30	> 99	0.89	5	1.90	2.26	2.61	0.84	8	1.69	2.26	2.82
	62	74:1 − 89:2	0.90	1.04	1.19	0.47	1.04	1.62	−0.96	1.04	3.05	0.77	5	0.86	1.04	1.22					
U.K.	118	60:1 − 89:2	0.34	0.56	0.78	−0.41	0.57	1.54	< −99	0.57	> 99	0.91	2	0.37	0.58	0.79	0.88	8	0.21	0.58	0.95
	56	60:1 − 73:4	0.50	0.73	0.96	0.61	0.72	0.84	0.51	0.72	0.93	0.54	1	0.63	0.73	0.82	0.23	8	0.44	0.73	1.01
	62	74:1 − 89:2	0.11	0.49	0.87	−33.8	0.42	34.6	< −99	0.42	> 99	0.95	2	0.26	0.51	0.76					
U.S.	170	47:1 − 89:2	0.58	0.78	0.98	−0.02	0.81	1.63	−2.91	0.81	4.52	0.94	3	0.51	0.80	1.09					
	108	47:1 − 73:4	0.63	0.88	1.13	0.16	0.87	1.59	−7.13	0.87	8.88	0.91	3	0.49	0.89	1.29					
	62	74:1 − 89:2	0.33	0.67	1.02	−7.97	0.64	9.24	−2.56	0.64	3.84	0.89	2	0.27	0.67	1.07					

[a] This series has the strike of 1968:2 interpolated out.
[b] Results are only reported for $kmax = 10$ when a different lag length is chosen compared to using $kmax = 5$.

become wider. In practice a wide PS_T^i confidence interval may be an indication of structural change in the trend function.

7. CONCLUSION

New tests of hypotheses regarding parameters of the trend function of a univariate time series were proposed. The tests can be applied to trend functions that are linear in parameters. The tests are asymptotically valid in the presence of general serial correlation and do not require estimation of the serial correlation parameters, either parametrically or nonparametrically. The tests are asymptotically valid whether the errors are $I(0)$ or $I(1)$, and are, for all practical purposes, invariant to the variance of the initial condition. The statistics should prove very useful in practice as they remove the need to specify the form of serial correlation and are relatively easy to compute. While separate critical values are needed for each specification of the trend function, Theorems 1 and 2 along with modern computing power make simulation of asymptotic critical values easy in most cases.

Dept. of Economics, Cornell University, Uris Hall, Ithaca, NY 14853-7601, U.S.A.; tjv2@cornell.edu

Manuscript received January, 1996; final revision received February, 1997.

APPENDIX

This appendix contains proofs of the theorems contained in the text. The proofs are simplified by first establishing some preliminary lemmas. The lemmas are stated without proof as they follow from standard results.

LEMMA A1: *Suppose* $\{u_t\}$ *follows* (4) *and* (5) *and* $\alpha_T = \alpha, |\alpha| < 1$ *so that* $\{u_t\}$ *is* $I(0)$. *Then if* (6) *holds for model* (3), *as* $T \to \infty$,

(a) $T^{-1/2}\tau_T X_1' u \Rightarrow \sigma \int_0^1 F(r)\, dw(r)$,
(b) $T^{-1}\tau_T X_1' X_1 \tau_T \Rightarrow \int_0^1 F(r)F(r)'\, dr$,
(c) $T^{-5/2}\tau_T X_2' S \Rightarrow \sigma \int_0^1 G(r) w(r)\, dr$,
(d) $T^{-3}\tau_T X_2' X_2 \tau_T \Rightarrow \int_0^1 G(r)G(r)'\, dr$.

LEMMA A2: *Suppose* $\{u_t\}$ *follows* (4) *and* (5) *and* $\alpha_T = (1 - \bar{\alpha}/T)$ *so* $\{u_t\}$ *is* $I(1)$. *Then if* (6) *holds for model* (3), *as* $T \to \infty$,

(a) $T^{-3/2}\tau_T X_1' u \Rightarrow d(1) \int_0^1 F(r) V_{\bar{\alpha}}(r)\, dr$,
(b) $T^{-7/2}\tau_T X_2' S \Rightarrow d(1) \int_0^1 G(r)[\int_0^r V_{\bar{\alpha}}(s)\, ds]\, dr$,
(c) $T^{-4} S'S \Rightarrow d(1)^2 \int_0^1 [\int_0^r V_{\bar{\alpha}}(s)\, ds]^2\, dr$.

LEMMA A3: *Suppose* $\{u_t\}$ *follows* (4) *and* (5) *and* $\alpha_T = \alpha, |\alpha| < 1$ *so that* $\{u_t\}$ *is* $I(0)$. *Then if* (6) *holds for model* (3), *as* $T \to \infty$,

(a) $T^{1/2}\tau_T^{-1}(\hat{\beta} - \beta) \Rightarrow \sigma [\int_0^1 F(r)F(r)'\, dr]^{-1} \int_0^1 F(r)\, dw(r)$,
(b) $s_y^2 \Rightarrow \sigma_u^2 = \lim_{T \to \infty} E(T^{-1} \sum_{t=1}^T u_t^2)$,
(c) $T^{1/2}\tau_T^{-1}(\beta^* - \beta) \Rightarrow \sigma [\int_0^1 G(r)G(r)'\, dr]^{-1} \int_0^1 G(r) w(r)\, dr$,
(d) $T^{-1} s_z^2 \Rightarrow \sigma^2 M(w(r), G(r))$.

LEMMA A4: *Suppose* $\{u_t\}$ *follows* (4) *and* (5) *and* $\alpha_T = (1 - \bar{\alpha}/T)$ *so* $\{u_t\}$ *is* $I(1)$. *Then if* (6) *holds for model* (3), *as* $T \to \infty$,

(a) $T^{-1/2}\tau_T^{-1}(\hat{\beta} - \beta) \Rightarrow d(1)[\int_0^1 F(r)F(r)'\,dr]^{-1}\int_0^1 F(r)V_{\bar{\alpha}}(r)\,dr$,

(b) $T^{-1}s_y^2 \Rightarrow d(1)^2 M(V_{\bar{\alpha}}(r), F(r))$,

(c) $T^{-1/2}\tau_T^{-1}(\beta^* - \beta) \Rightarrow d(1)[\int_0^1 G(r)G(r)'\,dr]^{-1}\int_0^1 G(r)[\int_0^r V_{\bar{\alpha}}(s)\,ds]\,dr$,

(d) $T^{-3}s_z^2 \Rightarrow d(1)^2 M(\int_0^r V_{\bar{\alpha}}(s)\,ds, G(r))$.

The $J_T^i(m)$ statistics can be expressed in terms of moments like s_y^2 and $T^{-1}s_z^2$. Their limiting distributions follows from straightforward arguments and details are omitted. Theorems 1–4 are proven for $T^{-1}W_T$ and PS_T^i. Similar arguments for PSW_T^i are omitted.

PROOF OF THEOREM 1: Using Lemma A1 (a), (d), Lemma A3 (a)–(d), $J_T^1(m) \Rightarrow 0$, and $J_T^2(m) \Rightarrow J(w(r), G(r), Q_2(r))$, we have

$$\begin{aligned}
T^{-1}W_T &= T^{-1}(R\hat{\beta} - r)'\left[R(X_1'X_1)^{-1}R'\right]^{-1}(R\hat{\beta} - r)/s_y^2 \\
&= T^{-1}[A^{-1}R\tau_T T^{1/2}\tau_T^{-1}(\hat{\beta} - \beta)]' \\
&\quad \times \left[A^{-1}R\tau_T(T^{-1}\tau_T X_1'X_1\tau_T)^{-1}\tau_T R'A^{-1}\right]^{-1}[A^{-1}R\tau_T T^{1/2}\tau_T^{-1}(\hat{\beta} - \beta)]/s_y^2 \\
&= T^{-1}O_p(1) \Rightarrow 0,
\end{aligned}$$

$$\begin{aligned}
PS_T^i &= T^{-1}(R\beta^* - r)'\left[R(X_2'X_2)^{-1}R'\right]^{-1}(R\beta^* - r)/(s_z^2\exp(bJ_T^i(m))) \\
&= [A^{-1}R\tau_T T^{1/2}\tau_T^{-1}(\beta^* - \beta)]'\left[A^{-1}R\tau_T(T^{-3}\tau_T X_2'X_2\tau_T)^{-1}\tau_T R'A^{-1}\right]^{-1} \\
&\quad \times [A^{-1}R\tau_T T^{1/2}\tau_T^{-1}(\beta^* - \beta)]/(T^{-1}s_z^2\exp(bJ_T^i(m))) \\
&\Rightarrow L(w(r), G(r), R^*)/M(w(r), G(r)), \qquad i = 1, \\
&\Rightarrow L(w(r), G(r), R^*)/[M(w(r), G(r))\exp\{bJ(w(r), G(r), Q_2(r))\}], \qquad i = 2.
\end{aligned}$$

PROOF OF THEOREM 2: Using Lemma A1 (b), (d), Lemma A4 (a)–(d), $J_T^1(m) \Rightarrow J(V_{\bar{\alpha}}(r), F(r), Q_1(r))$, and $J_T^2(m) \Rightarrow J(\int_0^r V_{\bar{\alpha}}(s)\,ds, G(r), Q_2(r))$, we have

$$\begin{aligned}
T^{-1}W_T &= T^{-1}(R\hat{\beta} - r)'\left[R(X_1'X_1)^{-1}R'\right]^{-1}(R\hat{\beta} - r)/s_y^2 \\
&= [A^{-1}R\tau_T T^{-1/2}\tau_T^{-1}(\hat{\beta} - \beta)]'\left[A^{-1}R\tau_T(T^{-1}\tau_T X_1'X_1\tau_T)^{-1}\tau_T R'A^{-1}\right]^{-1} \\
&\quad \times [A^{-1}R\tau_T T^{-1/2}\tau_T^{-1}(\hat{\beta} - \beta)]/(T^{-1}s_y^2) \\
&\Rightarrow L(V_{\bar{\alpha}}(r), F(r), R^*)/M(V_{\bar{\alpha}}(r), F(r)).
\end{aligned}$$

$$\begin{aligned}
PS_T^i &= T^{-1}(R\beta^* - r)'\left[R(X_2'X_2)^{-1}R'\right]^{-1}(R\beta^* - r)/(s_z^2\exp(bJ_T^i(m))) \\
&= [A^{-1}R\tau_T T^{-1/2}\tau_T^{-1}(\beta^* - \beta)]'\left[A^{-1}R\tau_T(T^{-3}\tau_T X_2'X_2\tau_T)^{-1}\tau_T R'A^{-1}\right]^{-1} \\
&\quad \times [A^{-1}R\tau_T T^{-1/2}\tau_T^{-1}(\beta^* - \beta)]/(T^{-3}s_z^2\exp(bJ_T^i(m))) \\
&\Rightarrow L\left(\int_0^r V_{\bar{\alpha}}(s)\,ds, G(r), R^*\right)\Big/ \\
&\quad \left[M\left(\int_0^r V_{\bar{\alpha}}(s)\,ds, G(r)\right)\exp\{bJ(V_{\bar{\alpha}}(r), F(r), Q_1(r))\}\right], \qquad i = 1, \\
&\Rightarrow L\left(\int_0^r V_{\bar{\alpha}}(s)\,ds, G(r), R^*\right)\Big/ \\
&\quad \left[M\left(\int_0^r V_{\bar{\alpha}}(s)\,ds, G(r)\right)\exp\left\{bJ\left(\int_0^r V_{\bar{\alpha}}(s)\,ds, G(r), Q_2(r)\right)\right\}\right], \qquad i = 2.
\end{aligned}$$

As a preliminary to the proofs of Theorems 3 and 4, note that under the alternative (12) s_y^2 and s_z^2 are exactly invariant to δ as the models are estimated under the alterative. Therefore s_y^2 and s_z^2 have the same limiting behavior as under the null hypothesis. Similarly, the limiting behavior of the $J_T^i(m)$ statistics are exactly invariant to δ and have the same limiting distribution under the null and alternative hypotheses. Recall that A is a diagonal matrix with elements that are nonpositive powers of T. Therefore, A^{-1} is a diagonal matrix with elements that are nonnegative powers of T and $TA^{-1} \to \infty$.

PROOF OF THEOREM 3: Rewriting the alternative (12) as $r = R\beta - \delta$, we have

$$\begin{aligned}
T^{-1}W_T &= T^{-1}(R\hat{\beta} - R\beta + \delta)'\left[R(X_1'X_1)^{-1}R'\right]^{-1}(R\hat{\beta} - R\beta + \delta)/s_y^2 \\
&= T^{-1}[A^{-1}R\tau_T T^{1/2}\tau_T^{-1}(\hat{\beta} - \beta) + T^{1/2}A^{-1}\delta]' \\
&\quad \times \left[A^{-1}R\tau_T(T^{-1}\tau_T X_1'X_1\tau_T)^{-1}\tau_T R'A^{-1}\right]^{-1} \\
&\quad \times [A^{-1}R\tau_T T^{1/2}\tau_T^{-1}(\hat{\beta} - \beta) + T^{1/2}A^{-1}\delta]/s_y^2 \\
&= A^{-1}\delta'\left[\sigma_u^2 R^* \left\{\int_0^1 F(r)F(r)'\,dr\right\}^{-1} R^*\right]^{-1} \delta A^{-1} + o_p(1).
\end{aligned}$$

If $A^{-1} \to \infty$, then $T^{-1}W_T \Rightarrow \infty$ since $\delta'[\sigma_u^2 R^*\{\int_0^1 F(r)F(r)'\,dr\}^{-1}R^{*\prime}]^{-1}\delta > 0$. If $A^{-1} = O_p(1)$, then $T^{-1}W_T = O_p(1)$. For PS_T^1 we have

$$\begin{aligned}
PS_T^1 &= T^{-1}(R\beta^* - R\beta + \delta)'\left[R(X_2'X_2)^{-1}R'\right]^{-1}(R\beta^* - R\beta + \delta)/(s_z^2\exp(bJ_T^1(m))) \\
&= [A^{-1}R\tau_T T^{1/2}\tau_T^{-1}(\beta^* - \beta) + T^{1/2}A^{-1}\delta]' \\
&\quad \times \left[A^{-1}R\tau_T(T^{-3}\tau_T X_2'X_2\tau_T)^{-1}\tau_T R'A^{-1}\right]^{-1} \\
&\quad \times [A^{-1}R\tau_T T^{1/2}\tau_T^{-1}(\beta^* - \beta) + T^{1/2}A^{-1}\delta]/(T^{-1}s_z^2\exp(bJ_T^1(m))) \\
&= TA^{-1}\delta'\left[M(w(r), G(r))R^*\left\{\int_0^1 G(r)G(r)'\,dr\right\}^{-1} R^{*\prime}\right]^{-1} \delta A^{-1} + o_p(1) \Rightarrow \infty,
\end{aligned}$$

since $\delta'[M(w(r), G(r))R^*\{\int_0^1 G(r)G(r)'\,dr\}^{-1}R^{*\prime}]^{-1}\delta > 0$ and $TA^{-1} \to \infty$. Similar arguments hold for PS_T^2 and are omitted.

PROOF OF THEOREM 4: Rewriting the alternative (12) as $r = R\beta - \delta$ we have

$$\begin{aligned}
T^{-1}W_T &= T^{-1}(R\hat{\beta} - R\beta + \delta)'\left[R(X_1'X_1)^{-1}R'\right]^{-1}(R\hat{\beta} - R\beta + \delta)/s_y^2 \\
&= [A^{-1}R\tau_T T^{1/2}\tau_T^{-1}(\hat{\beta} - \beta) + T^{-1/2}A^{-1}\delta]' \\
&\quad \times \left[A^{-1}R\tau_T(T^{-1}\tau_T X_1'X_1\tau_T)^{-1}\tau_T R'A^{-1}\right]^{-1} \\
&\quad \times [A^{-1}R\tau_T T^{1/2}\tau_T^{-1}(\hat{\beta} - \beta) + T^{-1/2}A^{-1}\delta]/(T^{-1}s_y^2).
\end{aligned} \tag{13}$$

$$\begin{aligned}
PS_T^i &= T^{-1}(R\beta^* - R\beta + \delta)'\left[R(X_2'X_2)^{-1}R'\right]^{-1} \\
&\quad \times (R\beta^* - R\beta + \delta)/(s_z^2\exp(bJ_T^i(m))) \\
&= [A^{-1}R\tau_T T^{1/2}\tau_T^{-1}(\beta^* - \beta) + T^{-1/2}A^{-1}\delta]' \\
&\quad \times \left[A^{-1}R\tau_T(T^{-3}\tau_T X_2'X_2\tau_T)^{-1}\tau_T R'A^{-1}\right]^{-1} \\
&\quad \times [A^{-1}R\tau_T T^{1/2}\tau_T^{-1}(\beta^* - \beta) + T^{-1/2}A^{-1}\delta]/(T^{-1}s_z^2\exp(bJ_T^i(m))).
\end{aligned} \tag{14}$$

The terms in (13) and (14) are $O_p(1)$ and nonzero with the possible exception of $T^{-1/2}A^{-1}\delta$. If there exists a j such that $\delta_j > 0$ and $a_{jj}^{-1} > T^{1/2}$, then $T^{-1/2}A^{-1}\delta \to \infty$ and $T^{-1}W_T \Rightarrow \infty$ and $PS_T^i \Rightarrow \infty$. If $a_{jj}^{-1} \leq T^{1/2}$ for all j, then $T^{-1/2}A^{-1}\delta = O_p(1)$ and $T^{-1}W_T = O_p(1)$ and $PS_T^i = O_p(1)$.

REFERENCES

ANDREWS, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

ANDREWS, D. W. K., AND J. C. MONAHAN (1992): "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica*, 60, 953–966.

BAI, J., R. L. LUMSDAINE, AND J. H. STOCK (1994): "Testing for and Dating Breaks in Stationary and Nonstationary Multivariate Time Series," forthcoming in *Review of Economic Studies*.

BANERJEE, A., R. L. LUMSDAINE, AND J. H. STOCK (1992): "Recursive and Sequential Tests of the Unit Root and Trend-Break Hypotheses: Theory and International Evidence," *Journal of Business and Economic Statistics*, 10, 271–288.

BOBKOSKI, M. J. (1983): "Hypothesis Testing in Nonstationary Time Series," Ph.D. Dissertation, Department of Statistics, University of Wisconsin.

CANJELS, E., AND M. W. WATSON (1997): "Estimating Deterministic Trends in the Presence of Serially Correlated Errors," *Review of Economics and Statistics*, 79, 184–200.

CHAN, C. H., AND C. Z. WEI (1987): "Asymptotic Inference for Nearly Stationary AR(1) Processes," *The Annals of Statistics*, 15, 1050–1063.

GRENANDER, U., AND M. ROSENBLATT (1957): *Statistical Analysis of Stationary Time Series*. New York: Chelsea Publishing Company.

HALL, A. (1994): "Testing for a Unit Root in a Time Series With Pretest Data Based Model Selection," *Journal of Business and Economic Statistics*, 12, 461–470.

NABEYA, S., AND P. PERRON (1994): "Local Asymptotic Distributions Related to the AR(1) Model with Dependent Errors," *Journal of Econometrics*, 62, 229–264.

NG, S., AND P. PERRON (1995): "Unit Root Tests in ARMA Models With Data Dependent Methods for the Truncation Lag," *Journal of the American Statistical Association*, 90, 268–281.

PARK, J. Y. (1990): "Testing for Unit Roots and Cointegration by Variable Addition," in *Advances in Econometrics: Cointegration, Spurious Regressions and Unit Roots*, ed. by T. Fomby and F. Rhodes. London: Jai Press, pp. 107–134.

PARK, J. Y., AND B. CHOI (1988): "A New Approach to Testing for a Unit Root," CAE Working Paper #88-23, Cornell University.

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.

——— (1989): "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Econometrica*, 57, 1361–1401.

——— (1990): "Testing for a Unit Root in a Time Series with a Changing Mean," *Journal of Business and Economic Statistics*, 8, 153–162.

——— (1991): "A Test for Changes in a Polynomial Trend Function for a Dynamic Time Series," Mimeo, Princeton University.

PHILLIPS, P. C. B. (1987a): "Time Series Regression with Unit Roots," *Econometrica*, 55, 277–302.

——— (1987b): "Towards a Unified Asymptotic Theory for Autoregression," *Biometrika*, 74, 535–547.

PHILLIPS, P. C. B., AND S. DURLAUF (1988): "Trends Versus Random Walks in Time Series Analysis," *Econometrica*, 56, 1333–1354.

PRAIS, S. J., AND C. B. WINSTEN (1954): "Trend Estimators and Serial Correlation," Cowles Foundation Discussion Paper 383.

PRESS, W. H., B. P. FLANERY, S. A. TEUKOLSKY, AND W. T. VETTERLING (1992): *Numerical Recipes: The Art of Scientific Computing*. Cambridge, U.K.: Cambridge University Press.

ROBINSON, P. M. (1991): "Automatic Frequency Domain Inference on Semiparametric and Nonparametric Models," *Econometrica*, 59, 1329–1363.

VOGELSANG, T. J. (1994): "Testing for a Shift in Mean Without Having to Estimate Serial Correlation Parameters," CAE Working Paper #94-24, Cornell University, forthcoming in *Journal of Business and Economic Statistics*.

——— (1996): "Trend Function Hypothesis Testing in the Presence of Serial Correlation," CAE Working Paper #96-11, Cornell University.

——— (1997a): "Wald-type Tests for Detecting Shifts in the Trend Function of a Dynamic Time Series," forthcoming in *Econometric Theory*, 13, No. 6.

——— (1997b): "Testing for a Shift in Trend When Serial Correlation is of Unknown Form," CAE Working Paper 97-11, Cornell University.

LINKED CITATIONS

You have printed the following article:

Trend Function Hypothesis Testing in the Presence of Serial Correlation
Timothy J. Vogelsang
Econometrica, Vol. 66, No. 1. (Jan., 1998), pp. 123-148.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199801%2966%3A1%3C123%3ATFHTIT%3E2.0.CO%3B2-P

[Footnotes]

[2] **Asymptotic Inference for Nearly Nonstationary AR(1) Processes**
N. H. Chan; C. Z. Wei
The Annals of Statistics, Vol. 15, No. 3. (Sep., 1987), pp. 1050-1063.
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28198709%2915%3A3%3C1050%3AAIFNNA%3E2.0.CO%3B2-5

[2] **Towards a Unified Asymptotic Theory for Autoregression**
P. C. B. Phillips
Biometrika, Vol. 74, No. 3. (Sep., 1987), pp. 535-547.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28198709%2974%3A3%3C535%3ATAUATF%3E2.0.CO%3B2-S

References

Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation
Donald W. K. Andrews
Econometrica, Vol. 59, No. 3. (May, 1991), pp. 817-858.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199105%2959%3A3%3C817%3AHAACCM%3E2.0.CO%3B2-P

LINKED CITATIONS

An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator
Donald W. K. Andrews; J. Christopher Monahan
Econometrica, Vol. 60, No. 4. (Jul., 1992), pp. 953-966.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199207%2960%3A4%3C953%3AAIHAAC%3E2.0.CO%3B2-Q

Recursive and Sequential Tests of the Unit-Root and Trend-Break Hypotheses: Theory and International Evidence
Anindya Banerjee; Robin L. Lumsdaine; James H. Stock
Journal of Business & Economic Statistics, Vol. 10, No. 3. (Jul., 1992), pp. 271-287.
Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28199207%2910%3A3%3C271%3ARASTOT%3E2.0.CO%3B2-C

Estimating Deterministic Trends in the Presence of Serially Correlated Errors
Eugene Canjels; Mark W. Watson
The Review of Economics and Statistics, Vol. 79, No. 2. (May, 1997), pp. 184-200.
Stable URL:
http://links.jstor.org/sici?sici=0034-6535%28199705%2979%3A2%3C184%3AEDTITP%3E2.0.CO%3B2-A

Asymptotic Inference for Nearly Nonstationary AR(1) Processes
N. H. Chan; C. Z. Wei
The Annals of Statistics, Vol. 15, No. 3. (Sep., 1987), pp. 1050-1063.
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28198709%2915%3A3%3C1050%3AAIFNNA%3E2.0.CO%3B2-5

Testing for a Unit Root in Time Series with Pretest Data-Based Model Selection
Alastair Hall
Journal of Business & Economic Statistics, Vol. 12, No. 4. (Oct., 1994), pp. 461-470.
Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28199410%2912%3A4%3C461%3ATFAURI%3E2.0.CO%3B2-F

Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag
Serena Ng; Pierre Perron
Journal of the American Statistical Association, Vol. 90, No. 429. (Mar., 1995), pp. 268-281.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28199503%2990%3A429%3C268%3AURTIAM%3E2.0.CO%3B2-M

LINKED CITATIONS

The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis
Pierre Perron
Econometrica, Vol. 57, No. 6. (Nov., 1989), pp. 1361-1401.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198911%2957%3A6%3C1361%3ATGCTOP%3E2.0.CO%3B2-W

Testing for a Unit Root in a Time Series with a Changing Mean
Pierre Perron
Journal of Business & Economic Statistics, Vol. 8, No. 2. (Apr., 1990), pp. 153-162.
Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28199004%298%3A2%3C153%3ATFAURI%3E2.0.CO%3B2-0

Time Series Regression with a Unit Root
P. C. B. Phillips
Econometrica, Vol. 55, No. 2. (Mar., 1987), pp. 277-301.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198703%2955%3A2%3C277%3ATSRWAU%3E2.0.CO%3B2-R

Towards a Unified Asymptotic Theory for Autoregression
P. C. B. Phillips
Biometrika, Vol. 74, No. 3. (Sep., 1987), pp. 535-547.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28198709%2974%3A3%3C535%3ATAUATF%3E2.0.CO%3B2-S

Automatic Frequency Domain Inference on Semiparametric and Nonparametric Models
P. M. Robinson
Econometrica, Vol. 59, No. 5. (Sep., 1991), pp. 1329-1363.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199109%2959%3A5%3C1329%3AAFDIOS%3E2.0.CO%3B2-4