Econometric Theory, **13**, 1997, 818–849. Printed in the United States of America.

WALD-TYPE TESTS FOR DETECTING BREAKS IN THE TREND FUNCTION OF A DYNAMIC TIME SERIES

TIMOTHY J. VOGELSANG
Cornell University

In this paper, test statistics for detecting a break at an unknown date in the trend function of a dynamic univariate time series are proposed. The tests are based on the mean and exponential statistics of Andrews and Ploberger (1994, *Econometrica* 62, 1383–1414) and the supremum statistic of Andrews (1993, *Econometrica* 61, 821–856). Their results are extended to allow trending and unit root regressors. Asymptotic results are derived for both $I(0)$ and $I(1)$ errors. When the errors are highly persistent and it is not known which asymptotic theory ($I(0)$ or $I(1)$) provides a better approximation, a conservative approach based on nearly integrated asymptotics is provided. Power of the mean statistic is shown to be nonmonotonic with respect to the break magnitude and is dominated by the exponential and supremum statistics. Versions of the tests applicable to first differences of the data are also proposed. The tests are applied to some macroeconomic time series, and the null hypothesis of a stable trend function is rejected in many cases.

1. INTRODUCTION

Inherent in statistically modeling economic time series is the problem of specifying the deterministic trend function. Incorrect specification of the trend function can be problematic because estimates of the parameters governing the dynamic behavior of the model may be inconsistent. For example, Nelson and Kang (1981) demonstrated that inappropriate detrending of a random walk process can lead to spurious estimates of dynamic parameters. In addition, inference is often adversely affected when the trend function is misspecified. In the context of testing for a unit root, Perron (1988, 1989, 1990) showed that failure to include a time trend (when the series is trending) or failure to account for possible breaks in the trend function can result in highly misleading inference.

In practice, it is often assumed that parameters in the deterministic trend function do not vary over time. However, for many economic time series, even casual

I thank Pierre Perron for many helpful suggestions and guidance and Gregory Chow, Rene Garcia, and Robin Lumsdaine for helpful comments on an earlier draft. I thank Peter Phillips and two referees for constructive comments that helped to substantially improve the paper. Any remaining errors are of course my own. This paper is based on Chapter 4 of my dissertation and was supported in part by the International Finance Section at Princeton University and the Center for Analytic Economics at Cornell University. Robin Lumsdaine kindly provided some of the data used. Address correspondence to: Timothy J. Vogelsang, Department of Economics, Uris Hall, Cornell University, Ithaca, NY 14850-7601, USA; e-mail: tjv2@cornell.edu.

  0266-4666/97 $9.00 + .10

observation of the data can suggest the possibility of an unstable trend function over time. The longer the time span being investigated, the greater the chance that some form of structural change has occurred in the trend parameters. It is important to detect potential structural change in the trend function to arrive at a reasonably specified model.

In the time series econometrics literature, a great deal of attention has recently been given to the subject of detecting structural change. Earlier work on structural change has confined its focus to detecting parameter breaks in a univariate context under restrictive assumptions such as independent and identically distributed (i.i.d.) data, nontrending data, and/or stationary data, that is, no unit roots. More recent work has successfully relaxed many of these restrictions. Andrews (1993) relaxed the i.i.d. assumption by developing Wald, Lagrange multiplier (LM), and likelihood ratio (LR) tests in a general regression framework that allows dependent and heterogeneously distributed data, although trends are not permitted. Hansen (1990) developed an LM test in a similar framework. Kramer, Ploberger, and Alt (1988) developed a CUSUM test valid in the presence of serial correlation. The case of trending data with stationary errors was considered by Kim and Siegmund (1989) and by Chu (1989) and Chu and White (1992), with the latter two studies allowing for serial correlation in the errors. For the case of trending data with a unit root, Banerjee, Lumsdaine, and Stock (1992) (hereafter BLS, 1992) proposed a test for detecting breaks in the slope of the trend function in the presence of a unit root. Perron (1991) proposed tests of structural change in the polynomial trend of a univariate dynamic time series. He derived results for both stationary and unit root errors and suggested a conservative test when the order of integration is unknown. Results within a Bayesian framework were provided by Zivot and Phillips (1994). Tests for multiple structural changes were proposed by Bai and Perron (1997). Finally, recent tests valid within a multivariate framework, including cointegration, have been explored by Hansen (1992a) and Bai, Lumsdaine, and Stock (1997).

The purpose of this paper is to add to the literature and provide a procedure that can be used to test for structural change in the trend function of a univariate time series that allows serial correlation in the errors. The trend function is modeled as a polynomial in time. Asymptotic results are obtained for both $I(0)$ and $I(1)$ errors. The alternative hypothesis is a single break in the trend function at an unknown date. Because the date of the break is only identified under the alternative, test statistics are constructed using the methodology of Andrews (1993) and Andrews and Ploberger (1994). The tests involve computing Wald statistics for a break in trend over a range of possible break dates and taking the supremum and exponential averages of the statistics. Because the results of Andrews (1993) and Andrews and Ploberger (1994) do not permit trending regressors or unit root errors, their results are extended to apply in the present case. When the errors are highly persistent and it is not known a priori whether the errors are better characterized as being $I(0)$ or $I(1)$, a conservative approach is suggested.

To justify the conservative approach, asymptotic results are provided modeling the errors as local to unity following Phillips (1987), Chan and Wei (1987), and others. Simulating the asymptotic distributions, it is shown that the distribution functions of the statistics are monotonically decreasing with respect to the local to unity parameter. Therefore, a conservative test can be constructed using the unit root critical values. The size and power of the conservative tests are explored using local asymptotic analysis and finite sample simulations. One particularly interesting result obtained is that in certain empirically relevant cases power can be nonmonotonic with respect to the magnitude of the break. This result compares to those of Perron (1991), where nonmonotonic power was found for dynamic extensions of the statistics of Gardner (1969) and MacNeill (1978).

The remainder of the paper is organized as follows. In Section 2, the model and statistics are presented. In Section 3, the limiting distributions under the null hypothesis are derived and critical values are tabulated. Local asymptotic size is also examined in this section. In Section 4, local asymptotic power is explored for the special case of a model with a simple linear trend. Finite sample results using simulation experiments are provided in Section 5. Striking power results are obtained for the case of two breaks in the trend function. In Section 6, versions of the tests applicable to first differences of the data are discussed. The size and power of the first-difference tests are compared and contrasted with the original statistics. An empirical application is presented in Section 7 using the international postwar GNP/GDP data used by BLS (1992), the international historic real GDP/GNP series considered by Kormendi and Meguire (1990), and the data used by Nelson and Plosser (1982). The results indicate that many macroeconomic time series have trend functions with parameters that are not constant over time. Section 8 has concluding comments, and proofs of the theorems in the text are given in the Appendix.

2. DETECTING TREND BREAKS: THE MODEL AND TEST STATISTICS

Consider the following data-generating process (DGP) for a univariate time series process, $\{y_t\}_1^T$, with a break in trend at unknown time T_b^c,

$$y_t = f(t)\theta + g(t,T_b^c)\gamma + v_t, \tag{1}$$

$$A(L)v_t = e_t, \tag{2}$$

where $f(t) = (1,t,t^2,\dots,t^p)$, $g(t,T_b^c) = 1(t > T_b^c)\{1,t - T_b^c,(t - T_b^c)^2,\dots,(t - T_b^c)^p\}$, $\theta = (\theta_0,\theta_1,\dots,\theta_p)'$, $\gamma = (\gamma_0,\gamma_1,\dots,\gamma_p)'$, $A(L) = 1 - a_1 L - \cdots - a_{k+1} L^{k+1}$, and $1(\cdot)$ is the indicator function. The autoregressive polynomial $A(z)$ is assumed to have at most one real valued root on the unit circle and all others strictly outside the unit circle, and the error process $\{e_t\}$ is assumed to be i.i.d. $(0,\sigma_e^2)$ with finite fourth moment. Under (1) and (2), $\{y_t\}$ is an autoregressive, stationary *or* unit root, process around a pth-order deterministic time trend with a break at date T_b^c. The null hypothesis of a stable trend function is given by

$$H_0: \gamma = 0. \tag{3}$$

Under the alternative, at least one of the trend polynomials has a break,

$$H_1: \gamma_i \neq 0 \quad \text{for at least one } i = 0,1,\dots,p. \tag{4}$$

For example, if $p = 1$, then for $t \leq T_b^c$ the intercept and the growth rate of $\{y_t\}$ are θ_0 and θ_1, respectively, whereas for $t > T_b^c$ the intercept and growth rate are $(\theta_0 + \gamma_0)$ and $(\theta_1 + \gamma_1)$. For the asymptotic analysis, it is assumed that the ratio of the true break date, T_b^c, to the sample size, T, remains a fixed proportion, λ_c, as T increases, that is, $T_b^c = \lambda_c$T.

It is convenient to factor the polynomial $A(L)$ according to the augmented Dickey–Fuller (ADF) procedure as $A(L) = (1 - \alpha L) - C(L)(1 - L)$, where $C(L) = \sum_{i=1}^{k} c_i L^i$, $c_i = -\sum_{j=i+1}^{k+1} a_j$, and $\alpha = \sum_{j=1}^{k+1} a_j$. Applying this factorization to v_t and defining $\pi = \alpha - 1$ gives

$$\Delta v_t = \pi v_{t-1} + \sum_{i=1}^{k} c_i \Delta v_{t-i} + e_t. \tag{5}$$

If α is modeled local to unity as $\alpha = 1 - \bar{\alpha}/T$ where $\bar{\alpha}$ is the local to unity parameter, then by standard results $T^{-1/2} v_{[rT]} \Rightarrow \sigma w_{\bar{\alpha}}(r)$, where $\sigma^2 = \sigma_e^2/(1 - C(1))^2$, $w_{\bar{\alpha}}(r) = \int_0^r \exp(-\bar{\alpha}(r - s))\, dw(s)$, where $w(r)$ is a standard Wiener process, $[rT]$ is the integer part of rT, and $\Rightarrow$ denotes weak convergence. Note that $1 - C(1) \neq 0$ because $\{v_t\}$ has at most one unit root. When $\{v_t\}$ is $I(0)$, then $T^{-1/2} \sum_{t=1}^{[rT]} v_t \Rightarrow \sigma_e^2 A(1)^{-2} w(r)$.

Using $A(L)$ and the ADF factorization, (1) can be rewritten as

$$\Delta y_t = f(t)\beta + g(t,T_b^c)\delta + d(t,T_b^c)\eta + \pi y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-i} + e_t,$$

where $d(t,T_b^c) = \{1(t = T_b^c + 1), 1(t = T_b^c + 2), \ldots, 1(t = T_b^c + k)\}$, $\eta = (\eta_1, \eta_2, \ldots, \eta_k)'$, and β, δ, and η are implicitly defined by $f(t)\beta = A(L)f(t)\theta$ and $g(t,T_b^c)\delta + d(t,T_b^c)\eta = A(L)g(t,T_b^c)\gamma$. Because the one-time dummy variables $d(t,T_b^c)$ are asymptotically negligible, it is convenient to drop them from the model and consider

$$\Delta y_t = f(t)\beta + g(t,T_b^c)\delta + \pi y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-i} + e_t. \tag{6}$$

Under the null hypothesis of no structural change, $\gamma = 0$, and it directly follows that $\delta = 0$. Therefore, test statistics can be constructed by estimating (6) and testing the hypothesis that $\delta = 0$. Writing the model in the form given by (6) is useful because serial correlation in the errors is handled by including enough lags of Δy_t. Because (6) is routinely estimated in unit root testing, tests based on (6) can be routinely computed in practice.

Because λ_c is a parameter that is present only under the alternative hypothesis, this testing problem falls within the class of tests proposed by Andrews and Ploberger (1994). Andrews and Ploberger (1994) derived optimal test statistics under quite general conditions that apply to models in which some parameters are present only under the alternative. These tests are of average exponential form over all possible values of the parameters that are present only under the alternative.

Suppose that (6) is estimated by ordinary least squares (OLS) using the break date $T_b = [\lambda T]$, where $\lambda \in [\lambda^*, 1 - \lambda^*] \subset (0,1)$. Note that the break date used in the estimation, T_b, may differ from the true break date, T_b^c. Let $W_T^p(\lambda)$ denote the

Wald statistic for testing $\delta = 0$. Define the discrete set of possible break dates to be $\Lambda = (T_b^*, T_b^* + 1, \ldots, T - T_b^*)$, where $T_b^* = [\lambda * T]$. The parameter λ^* is often called the amount of trimming. Trimming, or the requirement that the set of break dates maps into a closed subset of (0,1), is necessary for the asymptotic results to be nondegenerate. Two statistics from the class of statistics proposed by Andrews and Ploberger (1994) are

$$\text{Mean } W_T^p = T^{-1} \sum_{T_b \in \Lambda} W_T^p(T_b/T), \tag{7}$$

$$\text{Exp } W_T^p = \log\left(T^{-1} \sum_{T_b \in \Lambda} \exp\left(\frac{1}{2} W_T^p(T_b/T)\right)\right). \tag{8}$$

These statistics will be in the class of optimal statistics provided the regression has stationary and nontrending regressors. Therefore, when $p = 0$ and $\{v_t\}$ is $I(0)$, the mean and exponential tests are optimal, but when $p \geq 1$ and/or $\{v_t\}$ is $I(1)$, the optimality results do not apply. It is of interest to note that the Exp W_T^p statistic is designed to have power in detecting alternatives distant from the null, large breaks, whereas the Mean W_T^p statistic is designed to have power in detecting alternatives close to the null, small breaks.

A third related statistic proposed originally by Quandt (1960) and generalized by Andrews (1993) is the supremum statistic defined as

$$\text{Sup } W_T^p = \sup_{T_b \in \Lambda} W_T^p(T_b/T). \tag{9}$$

The Sup W_T^p statistic is not a member of the class of optimal statistics proposed by Andrews and Ploberger (1994) but is useful because it provides an estimate of the true break date ratio λ_c. See Bai (1993) for details on setting confidence intervals for estimates of λ_c using the supremum statistic in regression models.

The Mean W_T^p, Exp W_T^p, and Sup W_T^p statistics have received some attention in recent studies. Bai et al. (1997) considered all three statistics in testing for a mean break in multivariate models. Hansen (1992a) examined the Mean W_T^p and Sup W_T^p statistics in testing for structural change in cointegration models, whereas BLS (1992) used the Sup W_T^p statistic to test for trend breaks in models with unit root errors.

3. THE LIMITING DISTRIBUTIONS UNDER THE NULL HYPOTHESIS

In this section, the limiting behavior of the statistics under the null hypothesis of no break in the trend function is investigated. The asymptotic results depend on whether $\{v_t\}$ is $I(0)$ or $I(1)$, and separate theorems are given for the two cases. To facilitate the presentation of the results, some additional notation is needed. Define $F(r) = (1, r, r^2, \ldots, r^p)$ and $G(r,\lambda) = 1(r > \lambda)(1, r - \lambda, (r - \lambda)^2, \ldots, (r - \lambda)^p) \equiv 1(r > \lambda)F(r - \lambda)$, where $r \in [0,1]$. Note that $F(r)$ and $G(r,\lambda)$ are $(p + 1)$ row vectors of functions defined on the unit interval and that $\tau_1 f(t) = F(t/T)$ and $\tau_1 g(t, T_b) = G(t/T, \lambda)$, where τ_1 is a $(p + 1) \times (p + 1)$

diagonal matrix with diagonal elements $1, T^{-1}, T^{2}, \ldots, T^{-p}$. Let $\tilde{G}(r,\lambda)$ and $\tilde{w}_{\bar{\alpha}}(r)$ denote, respectively, the residuals from the projections of $G(r,\lambda)$ and $w_{\bar{\alpha}}(r)$ onto the space spanned by $F(r)$, for example, $\tilde{G}(r,\lambda) = G(r,\lambda) - F(r)(\int_0^1 F(r)'F(r)\,dr)^{-1}\int_0^1 F(r)'G(r,\lambda)\,dr$. Using this notation, the limiting results are compactly summarized in the following theorems.

THEOREM 1. *Let* $\{y_t\}$ *be the stochastic process described by* (1)–(3). *If the process* $\{v_t\}$ *is* $I(0)$, *that is*, $|\alpha| < 1$, *as* $T \to \infty$:

$$W_T^p(T_b/T) \Rightarrow \int_0^1 \tilde{G}(r,\lambda)\,dw(r)\left(\int_0^1 \tilde{G}(r,\lambda)'\tilde{G}(r,\lambda)\,dr\right)^{-1}\int_0^1 \tilde{G}(r,\lambda)'\,dw(r)$$
$$\equiv W_s^p(\lambda)$$

uniformly in λ,

$$\text{Mean } W_T^p \Rightarrow \int_{\lambda^*}^{1-\lambda^*} W_s^p(\lambda)\,d\lambda \equiv \text{Mean } W_s^p,$$

$$\text{Exp } W_T^p \Rightarrow \log\left\{\int_{\lambda^*}^{1-\lambda^*}\exp\left(\frac{1}{2}W_s^p(\lambda)\right)d\lambda\right\} \equiv \text{Exp } W_s^p,$$

$$\text{Sup } W_T^p \Rightarrow \operatorname*{Sup}_{\lambda\in[\lambda^*,1-\lambda^*]} W_s^p(\lambda)\,d\lambda \equiv \text{Sup } W_s^p.$$

THEOREM 2. *Let* $\{y_t\}$ *be the stochastic process described by* (1)–(3). *If the process* $\{v_t\}$ *is nearly* $I(1)$, *that is*, $\alpha = 1 - \bar{\alpha}/T$, *as* $T \to \infty$:

$$W_T^p(T_b/T) \Rightarrow H_1^p(\lambda)'H_2^p(\lambda)^{-1}H_1^p(\lambda) \equiv W_u^p(\lambda)$$

uniformly in λ,

$$\text{Mean } W_T^p \Rightarrow \int_{\lambda^*}^{1-\lambda^*} W_u^p(\lambda)\,d\lambda \equiv \text{Mean } W_u^p,$$

$$\text{Exp } W_T^p \Rightarrow \log\left\{\int_{\lambda^*}^{1-\lambda^*}\exp\left(\frac{1}{2}W_u^p(\lambda)\right)d\lambda\right\} \equiv \text{Exp } W_u^p,$$

$$\text{Sup } W_T^p \Rightarrow \operatorname*{Sup}_{\lambda\in[\lambda^*,1-\lambda^*]} W_u^p(\lambda)\,d\lambda \equiv \text{Sup } W_u^p,$$

where

$$H_1^p(\lambda) = \int_0^1 \tilde{G}(r,\lambda)'\,dw(r) - \left(\int_0^1 \tilde{G}(r,\lambda)'\tilde{w}_{\bar{\alpha}}(r)\,dr\int_0^1 \tilde{w}_{\bar{\alpha}}(r)\,dw(r)\right)\Big/\int_0^1 \tilde{w}_{\bar{\alpha}}(r)^2\,dr,$$

$$H_2^p(\lambda) = \int_0^1 \tilde{G}(r,\lambda)'\tilde{G}(r,\lambda)\,dr - \left(\int_0^1 \tilde{G}(r,\lambda)'\tilde{w}_{\bar{\alpha}}(r)\,dr\int_0^1 \tilde{G}(r,\lambda)\tilde{w}_{\bar{\alpha}}(r)\,dr\right)\Big/\int_0^1 \tilde{w}_{\bar{\alpha}}(r)^2\,dr.$$

Note from the theorems that the limiting distributions of the statistics are nonstandard. When $p = 0$, Theorem 1 is a special case of Andrews and Ploberger (1994), Andrews (1993), and Bai et al. (1997). Critical values from those studies are available for a range of values for λ^*. For the other cases, the limiting results are new, and critical values were obtained using simulation methods. When the errors are $I(0)$ and purely $I(1)(\bar{\alpha} = 0)$, the distributions are free of nuisance parameters and depend only on p and λ^*. Asymptotic critical values for $p = 0,1,2$ and $\lambda^* = 0.01, 0.15$ are tabulated in Table 1 for the $I(0)$ case and Table 2 for the $I(1)(\bar{\alpha} = 0)$ case. The critical values were calculated by simulation methods using $N(0,1)$ i.i.d. random deviates to approximate the Wiener processes implicit in $W_s^p(\lambda)$ and $W_u^p(\lambda)$. The integrals were approximated by the normalized sums of 1,000 steps using 10,000 replications.

Several observations can be made based on the tabulated critical values. First, as p increases, the distributions become skewed farther to the right. Second, as the trimming is increased, the critical values become larger; however, the critical values of the supremum statistic do not depend heavily on the amount of trim-

TABLE 1. Asymptotic distributions of Mean W_T^p, Exp W_T^p, and Sup W_T^p

	Mean W_s^p			Exp W_s^p			Sup W_s^p		
%	$p=0$	$p=1$	$p=2$	$p=0$	$p=1$	$p=2$	$p=0$	$p=1$	$p=2$
				Stationary Case, $\lambda^* = 0.01$					
.01	0.18	0.54	1.01	0.08	0.34	0.67	1.72	4.10	4.76
.025	0.22	0.65	1.14	0.11	0.40	0.78	2.01	4.55	6.32
.05	0.26	0.73	1.30	0.14	0.47	0.91	2.35	5.07	6.93
.10	0.32	0.86	1.49	0.17	0.57	1.06	2.77	6.69	7.64
.50	0.73	1.63	2.56	0.49	1.20	1.92	5.03	8.63	10.97
.90	2.00	3.49	4.74	1.59	2.76	3.70	9.24	13.62	16.06
.95	2.66	4.42	5.65	2.20	3.52	4.41	10.85	15.44	17.89
.975	3.34	5.36	6.69	2.80	4.18	5.22	12.46	17.26	19.57
.99	4.21	6.64	8.14	3.63	5.24	6.24	14.49	19.90	21.65
				Stationary Case, $\lambda^* = 0.15$					
.01	0.08	0.29	0.56	−0.30	−0.12	0.13	0.76	2.40	4.04
.025	0.10	0.35	0.67	−0.28	−0.06	0.24	0.92	2.74	4.55
.05	0.12	0.42	0.77	−0.26	0.01	0.34	1.14	3.16	5.03
.10	0.16	0.50	0.92	−0.23	0.09	0.50	1.41	3.66	5.63
.50	0.47	1.10	1.76	0.06	0.70	1.35	3.22	6.38	8.84
.90	1.58	2.70	3.58	1.23	2.33	3.18	7.32	11.25	13.96
.95	2.20	3.50	4.41	1.89	3.13	3.98	9.00	13.29	15.84
.975	2.85	4.35	5.25	2.53	3.88	4.68	10.69	15.12	17.61
.99	3.70	5.55	6.47	3.46	5.05	5.78	13.02	17.51	19.90

Note: The critical values were calculated via simulation methods using $N(0,1)$ i.i.d. random deviates to approximate the Wiener processes defined in the distributions given by Theorem 1. The integrals were approximated by the normalized sums of 1,000 steps using 10,000 replications.

TABLE 2. Asymptotic distributions of Mean W_T^p, Exp W_T^p, and Sup W_T^p

	Mean W_u^p			Exp W_u^p			Sup W_u^p		
%	$p=0$	$p=1$	$p=2$	$p=0$	$p=1$	$p=2$	$p=0$	$p=1$	$p=2$
				Unit root case, $\lambda^* = 0.01$					
.01	0.54	1.50	2.55	0.33	1.15	2.12	3.35	7.39	10.54
.025	0.65	1.72	2.89	0.42	1.37	2.44	4.01	8.18	11.54
.05	0.77	1.95	3.24	0.53	1.60	2.74	4.66	8.99	12.44
.10	0.95	2.24	3.68	0.69	1.91	3.17	5.57	10.09	13.65
.50	1.91	3.96	6.00	1.82	3.69	5.50	9.74	15.00	19.21
.90	3.32	7.14	10.18	4.02	6.98	9.58	16.14	22.60	28.11
.95	3.91	8.22	11.74	4.84	8.18	11.09	18.20	25.27	31.35
.975	4.53	9.29	13.17	5.68	9.27	12.50	20.23	27.76	34.45
.99	5.35	10.54	14.80	6.69	10.56	14.42	22.64	30.44	38.43
				Unit root case, $\lambda^* = 0.15$					
.01	0.28	1.02	1.78	−0.12	0.70	1.69	2.16	6.36	9.82
.025	0.35	1.17	2.10	−0.05	0.94	2.08	2.73	7.15	10.75
.05	0.44	1.38	2.40	0.05	1.19	2.39	3.32	8.05	11.69
.10	0.59	1.63	2.79	0.23	1.53	2.86	4.22	9.11	12.91
.50	1.51	3.17	4.86	1.52	3.49	5.34	9.02	14.38	18.88
.90	2.87	6.12	8.65	3.87	6.90	9.54	15.78	22.29	27.99
.95	3.43	7.19	10.00	4.71	8.12	11.07	17.88	25.10	31.29
.975	3.99	8.07	11.32	5.57	9.24	12.47	20.08	27.56	34.39
.99	4.65	9.17	13.02	6.60	10.54	14.34	22.48	30.36	38.35

Note: The critical values were calculated via simulation methods using $N(0,1)$ i.i.d. random deviates to approximate the Wiener processes defined in the distributions given by Theorem 2. The integers were approximated by the normalized sums of 1,000 steps using 10,000 replications.

ming. Third, the limiting distributions are different in the $I(0)$ case as compared with the $I(1)$ case. If it is known a priori that the errors are $I(0)$, then the critical values from Table 1 should be used. If it is known a priori that the errors are purely $I(1)$, then the critical values from Table 2 should be used.

In practice, it is often the case that the errors are highly persistent (α between 0.8 and 1), in which case it may not be obvious which distribution theory ($I(0)$ or $I(1)$) provides a better finite sample approximation. In this situation, the local to unity asymptotic results of Theorem 2 can provide an approximation. Given a value of α in a finite sample, the corresponding value of $\bar{\alpha}$ can be obtained by $\bar{\alpha} = T(1 - \alpha)$ and asymptotic critical values obtained from Theorem 2. The practical limitation of this approach is that when α is not known, $\bar{\alpha}$ cannot be consistently estimated from the data. One way around this problem is to use a conservative approach. Given a nominal significance level, a simple conservative test can be constructed by using the largest critical value across values of $\bar{\alpha}$. This test will have asymptotic size equal to the nominal level by construction. This approach is further simplified because, given a nominal level, the critical values of the statistics are monotonically decreasing in $\bar{\alpha}$. Therefore, a conservative test can be constructed by using the $\bar{\alpha} = 0$ critical values that are given in Table 2.

To demonstrate that the critical values are monotonically decreasing in $\bar{\alpha}$, limiting distributions for $\bar{\alpha} = 0,4,8,\ldots,16,20$ were simulated using the same techniques as those used for Tables 1 and 2. Plots of the distribution functions for the mean statistic with $p = 1$ are given in Figure 1. Plots for other values of p and the other two statistics are qualitatively similar and are not reported. For all nominal levels, the critical values are the largest for $\bar{\alpha} = 0$.

An attractive feature of this conservative approach is that the size of the test will be asymptotically correct. A drawback of this conservative approach is that power will be penalized when $\bar{\alpha} > 0$ (when the errors are not exactly $I(1)$). One way to investigate the potential power loss is to examine the rejection probabilities under the null hypothesis for particular values of $\bar{\alpha} > 0$. Conditional on $\bar{\alpha}$, these rejection probabilities can be interpreted as asymptotic size. Using the simulated asymptotic distributions, asymptotic size of the statistics for $\bar{\alpha} = 0,2,4,\ldots,18,20$ was computed and is plotted in Figure 2. Results are only reported for $p = 1$ and are qualitatively similar for other values of p. The nominal size was 0.05. From the figure, asymptotic size is 0.05 when $\bar{\alpha} = 0$, which is true by construction. As $\bar{\alpha}$ increases, asymptotic size steadily decreases as expected. Thus, power of the conservative tests will be less than power of the tests had $\bar{\alpha}$ been known, and this loss in power increases as $\bar{\alpha}$ grows. A detailed examination of the power properties of the tests is given in the next section.

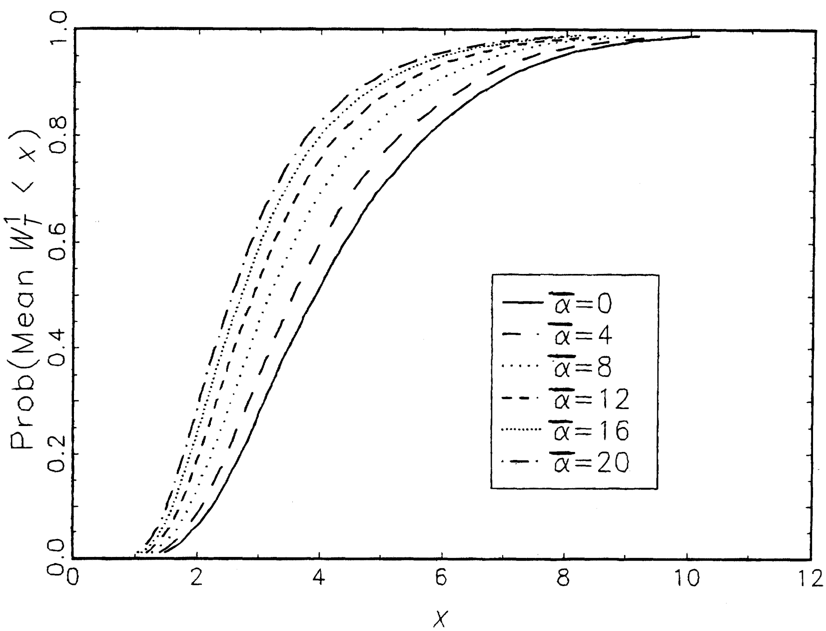


FIGURE 1. Cumulative distribution function of Mean W_T^1, nearly $I(1)$ errors, $p = 1$.

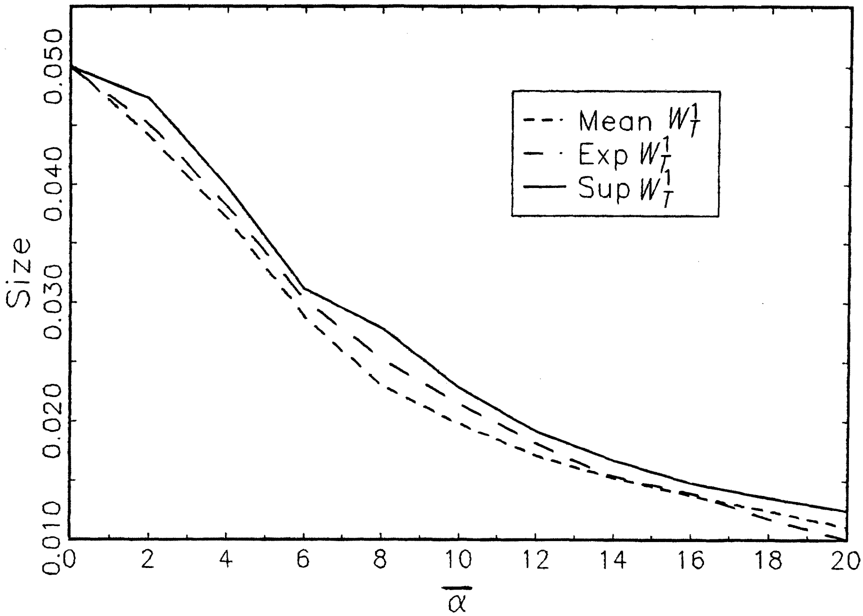


FIGURE 2. Conservative test using $\bar{\alpha} = 0$ critical values. Asymptotic size, 5% nominal level, $p = 1$.

4. LOCAL ASYMPTOTIC POWER: $p = 1$

This section explores the local asymptotic power properties of the statistics for the case of $p = 1$. Attention is focused on $p = 1$ to keep the exposition as simple as possible. Results are only reported for the case where the errors are nearly $I(1)$ and the alternative is a shift in slope ($\gamma_0 = 0$). This parameterization is appropriate for many macroeconomic time series. A local asymptotic analysis for $I(0)$ errors can be found in the working paper version of this paper (Vogelsang, 1994).

Consider the local alternative $\gamma_1 = T^{-1/2}\bar{\gamma}_1$, which leads to the model

$$y_t = \theta_0 + \theta_1 t + T^{-1/2}\bar{\gamma}_1 1(t > T_b^c)(t - T_b^c) + v_t. \tag{10}$$

Note that the change in slope converges to the null value of zero at the usual rate of $T^{-1/2}$. Transforming (10) using $A(L)$ yields the following model:

$$\Delta y_t = \beta_0 + \beta_1 t + T^{-1/2}\delta_0 1(t > T_b^c) + T^{-1/2}\delta_1 1(t > T_b^c)(t - T_b^c) + \pi y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-i} + \epsilon_t, \tag{11}$$

where $\beta_0 = (1 - \alpha)\theta_0 + (\alpha - C(1))\theta_1, \beta_1 = (1 - \alpha)\theta_1, \delta_0 = (\alpha - C(1))\bar{\gamma}_1, \delta_1 = (1 - \alpha)\bar{\gamma}_1$, and $\epsilon_t = e_t + T^{-1/2}\bar{\gamma}_1 \sum_{i=1}^{k-1} 1(t = T_b^c + i + 1)\sum_{j=i+1}^{k} c_j$. Using the notation from the previous section, $G(r,\lambda) = 1(r > \lambda)(1, r - \lambda)$ and let $G_1(r,\lambda) = 1(r > \lambda)(r - \lambda)$.

The following theorem provides the limiting distributions of the statistics under model (10) with nearly integrated errors.

THEOREM 3. *Let* $\{y_t\}$ *be the stochastic process described by* (1), (2), *and* (10). *If the process* $\{v_t\}$ *is nearly* $I(1)$, *that is,* $\alpha = 1 - \bar{\alpha}/T$, *as* $T \to \infty$:

$$W_T^1(T_b/T) \Rightarrow [\nu' L_2(\lambda,\lambda_c)' + L_1(\lambda)']L_2(\lambda,\lambda)^{-1}[L_2(\lambda,\lambda_c)\nu + L_1(\lambda)] \equiv L(\lambda)$$

uniformly in λ,

$$\text{Mean } W_T^1 \Rightarrow \int_{\lambda^*}^{1-\lambda^*} L(\lambda)\, d\lambda,$$

$$\text{Exp } W_T^1 \Rightarrow \log\left\{\int_{\lambda^*}^{1-\lambda^*} \exp\left(\frac{1}{2} L(\lambda)\right) d\lambda\right\},$$

$$\text{Sup } W_T^1 \Rightarrow \sup_{\lambda\in[\lambda^*,1-\lambda^*]} L(\lambda)\, d\lambda,$$

where

$$L_1(\lambda) = \int_0^1 \tilde{G}(r,\lambda)'\, dw(r) - \left(\int_0^1 \tilde{G}(r,\lambda)'[\kappa\tilde{G}_1(r,\lambda_c) + \tilde{w}_{\bar{\alpha}}(r)]\, dr \times \frac{\int_0^1 [\kappa\tilde{G}_1(r,\lambda_c) + \tilde{w}_{\bar{\alpha}}(r)]\, dw(r)}{\int_0^1 [\kappa\tilde{G}_1(r,\lambda_c) + \tilde{w}_{\bar{\alpha}}(r)]^2\, dr}\right),$$

$$L_2(\lambda_1,\lambda_2) = \int_0^1 \tilde{G}(r,\lambda_1)'\tilde{G}(r,\lambda_2)\, dr - \left(\int_0^1 \tilde{G}(r,\lambda_1)'[\kappa\tilde{G}_1(r,\lambda_c) + \tilde{w}_{\bar{\alpha}}(r)]\, dr \times \frac{\int_0^1 \tilde{G}(r,\lambda_2)[\kappa\tilde{G}_1(r,\lambda_c) + \tilde{w}_{\bar{\alpha}}(r)]\, dr}{\int_0^1 [\kappa\tilde{G}_1(r,\lambda_c) + \tilde{w}_{\bar{\alpha}}(r)]^2\, dr}\right),$$

$$\kappa = \bar{\gamma}_1/\sigma, \quad \nu = \bar{\delta}/\sigma_e, \quad \textit{and} \quad \bar{\delta} = \bar{\gamma}_1(1 - C(1),\bar{\alpha})'.$$

The local asymptotic distributions depend on $\bar{\gamma}_1$, $\bar{\alpha}$, $C(1)$, σ^2, and σ_e^2. Therefore, local asymptotic power depends on these parameters. It is difficult to determine analytically how power depends on $\bar{\gamma}_1$ and $\bar{\alpha}$ because the distributions are complicated functions of these parameters. Therefore, local power functions were simulated using the same techniques as those used to generate the asymptotic critical values in Section 3. To control the effects of $C(1)$ and σ_e^2 on power, $C(1) = 0$ and $\sigma_e^2 = 1$ ($\sigma^2 = 1$) for all simulations. This parameterization corresponds to $\{v_t\}$ being a nearly integrated random walk with unit variance innovations. Results are reported for Mean W_T^1 and Exp W_T^1 only. Results for Sup W_T^1 are very similar to those of Exp W_T^1 and are available upon request. Local power is reported for $\bar{\gamma}_1 = 0,2,\ldots,18,20$ and $\bar{\alpha} = 0,4,\ldots,16,20$. The nominal level was 0.05 and $\lambda^* = 0.01$ in all cases. Results for Exp W_T^1 are given in Figure 3, and results for Mean W_T^1 are given in Figure 4.

First, consider power of the Exp W_T^1 statistic. As $\bar{\alpha}$ increases, power is higher provided $\bar{\gamma}_1$ is not too close to the null. If $\bar{\gamma}_1$ is close to the null, power is decreasing in $\bar{\alpha}$, and this occurs because of the conservative nature of the test. However, power is quite poor in general for $\bar{\gamma}_1 \leq 2$. Therefore, power increases as the errors become less persistent unless the break is very small in which case power is low in general.

The results for the Mean W_T^1 statistic are much more striking. As seen in Figure 4, the power of Mean W_T^1 is nonmonotonic in $\bar{\gamma}_1$ except when $\bar{\alpha} = 0$. The nonmonotonicities become more pronounced as $\bar{\alpha}$ increases. For small breaks, power is increasing in $\bar{\alpha}$ and for medium breaks power is decreasing in $\bar{\alpha}$, whereas

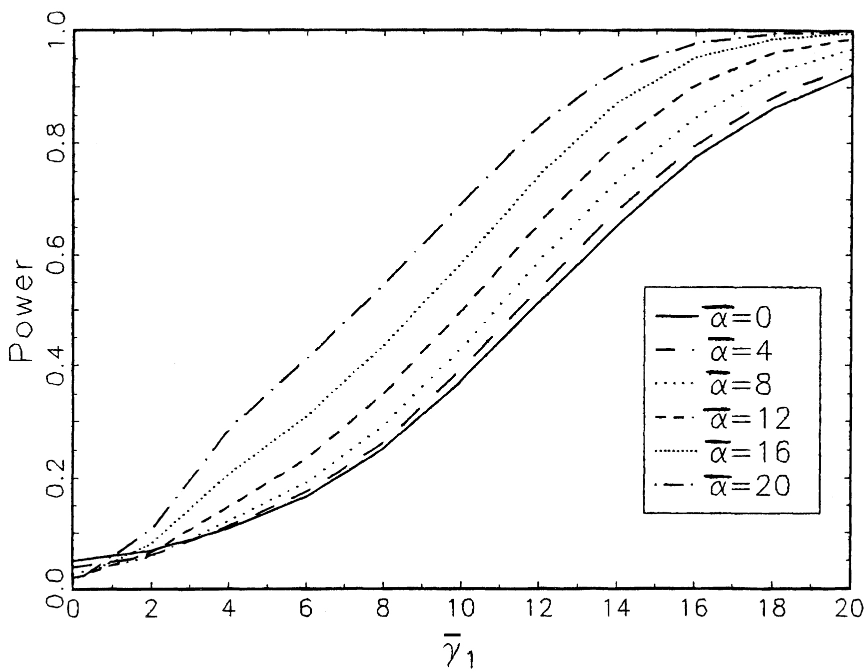


FIGURE 3. Asymptotic power Exp W_T^1, nearly $I(1)$ errors, $p = 1$.

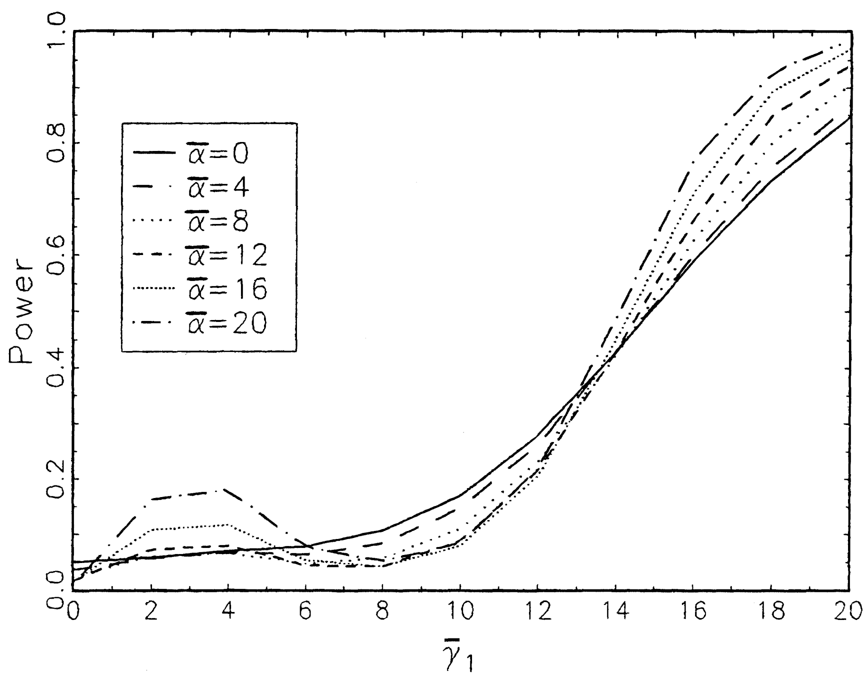


FIGURE 4. Asymptotic power Mean W_T^1, nearly $I(1)$ errors, $p = 1$.

for large breaks power is again increasing in $\bar{\alpha}$. This suggests that Mean W_T^1 will have problems detecting medium-sized shifts, and the problems become more pronounced the less persistent the errors.

Finally, by comparing Figures 3 and 4, it is quite apparent that Exp W_T^1 dominates Mean W_T^1 in terms of power except when the shift is small in magnitude. This is consistent with the justification for the statistics given by Andrews and Ploberger (1994).

5. POWER IN FINITE SAMPLES

In this section, simulation results are presented which explore the finite sample power of the statistics. Select results are reported for $p = 0$ and $p = 1$. Size simulations were also performed, and the results correspond very closely to the local size results of Section 3 and are omitted. The remainder of this section reports finite sample power for two cases: (1) a single break in trend and (2) two breaks in trend. Results for the Sup W_T statistic are not reported because they are very similar to the results for Exp W_T.

5.1 Finite Sample Power for a One-Time Break

To investigate the finite sample power of the tests against a one-time break in the trend function, series of size $T = 100$ were generated using the following DGP:

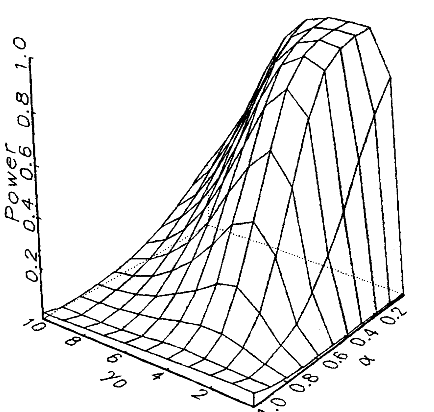
$$y_t = \gamma_0 1(t > \lambda_c T) + v_t, \qquad (p = 0) \tag{12}$$

$$y_t = \gamma_0 1(t > \lambda_c T) + \gamma_1 1(t > \lambda_c T)(t - \lambda_c T) + v_t, \qquad (p = 1) \tag{13}$$

$$v_t = \alpha v_{t-1} + e_t, \tag{14}$$

with $\{e_t\}$ i.i.d. $N(0,1)$ and $v_0 = 0$. The parameters θ_0 and θ_1 were set to zero because the statistics are exactly invariant to their values. Models (12,14) and (13,14) generate series having a break in the trend function at date $T_b^c = \lambda_c T$ with AR(1) errors. All power experiments had 2,000 replications with the test statistics calculated using regression (6) with $k = 0$. For model (12,14), regression (6) was estimated using $p = 0$; for model (13,14), regression (6) was estimated using $p = 1$. The values of 0.0,0.1,...,1.0 were used for α. The simulations were conducted for three values of λ_c, 0.25, 0.5, and 0.75. Results are only reported for $\lambda_c = 0.5$. Results for $\lambda_c = 0.25, 0.75$ can be found in Vogelsang (1994). The 5% conservative asymptotic critical values were used in all simulations. The power simulations were not size adjusted because the interest is in examining the power of the conservative procedure by using asymptotic critical values. Trimming in all cases was 1%.

Consider first the case of $p = 0$. This case was not covered by the local asymptotic analysis. Power was calculated for $\gamma_0 = 0,1,\dots,10$. Because the errors have standard deviation of one, γ_0 measures intercept shifts in units of standard deviations of the errors. Thus, $\gamma_0 = 1.0$ is a moderate break, whereas $\gamma_0 = 10$ is a very large break. The resulting power functions are plotted in Figure 5. Several observations can be made. As α decreases, power increases. Thus, as the errors become less persistent, power increases. If α is held fixed, power is increasing in γ_0 for Exp W_T^0 but is nonmonotonic in γ_0 for Mean W_T^0. Surprisingly, large breaks cannot be detected using Mean W_T^0 even if the errors are not highly persistent. The



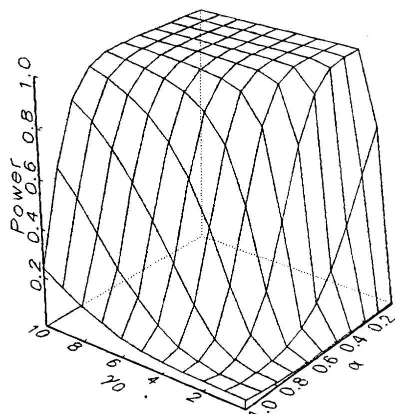


FIGURE 5. Power against a change in mean, $p = 0$, $T = 100$, 5% nominal size. Left: Mean W_T^0, $\lambda_c = 0.5$, 1% trimming. Right: Exp W_T^0, $\lambda_c = 0.5$, 1% trimming.

main conclusion to draw from these results is that Exp W_T^0 (and Sup W_T^0) dominate Mean W_T^0 in terms of power.

Now consider the case of trending data, $p = 1$. Power in detecting a shift in intercept, γ_0, is not reported because it is similar to the results for $p = 0$. Power for detecting a shift in slope is reported in Figure 6. Power is plotted for $\gamma_1 = 0.0, 0.3, \ldots, 3.0$, and $\gamma_0 = 0$ in all cases. The results are very similar to the results of the local asymptotic analysis from Section 4, and a detailed discussion is not required. Again, the main implication is that the Mean W_T^1 statistic has nonmonotonic power in γ_1 whereas the other statistics have monotonic power.

What is the reason behind the nonmonotonic power of the Mean W_T statistic? All three statistics can be viewed as weighted averages of Wald statistics across a set of possible break dates. It is the weights that determine whether or not power is monotonic. The Mean W_T statistic places equal weights across break dates. The Exp W_T statistic places higher weights on large values of the Wald statistics and smaller weights on small values of the Wald statistics because the exponential function is an increasing function. The Sup W_T statistic places a weight of one on the largest Wald statistic and a weight of zero on the other Wald statistics. Perron (1991) proposed dynamic versions of the statistics of Gardner (1969) and MacNeill (1978) and found that these statistics exhibit nonmonotonic power similar to that of Mean W_T. It is easy to show that the Perron statistic is simply a weighted average of LM statistics across all break dates based on regression (6) for testing $\gamma_0 = 0$ with weights proportional to $\lambda(1 - \lambda)$. Thus, for $p = 0$, the Perron and Mean W_T^0 statistics are very similar.

Nonmonotonic power can be explained in the following way. To keep ideas concrete, suppose $p = 0$ and that there is a shift in intercept, $\gamma_0 \neq 0$. If the Wald

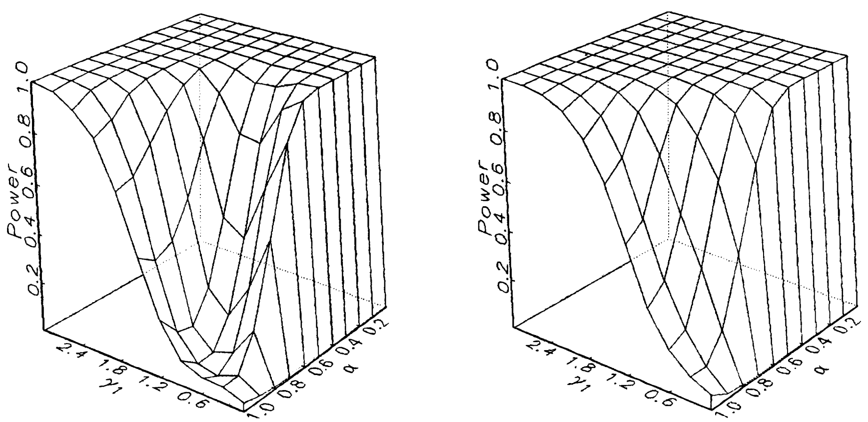


FIGURE 6. Power against a change in trend, $p = 1$, $T = 100$, 5% nominal size. Left: Mean W_T^1, $\lambda_c = 0.5$, 1% trimming. Right: Exp W_T^1, $\lambda_c = 0.5$, 1% trimming.

or LM statistic is computed using the true break date, T_b^c, the Wald or LM statistic will be large. Conversely, if the Wald or LM statistic is computed using a break date distant from the true break date, the regression will be misspecified. From the results of Perron (1990) it is well known that if a shift in mean is misspecified in a regression like (6) the estimate of π will be biased toward zero and the fitted model cannot be distinguished from a model with a unit root and no shift in trend. Therefore, the Wald or LM statistic will be relatively small. As the shift in mean increases, the Wald or LM statistics increase at the true break date but decrease at break dates far from the true break date. The average and, hence, power may fall as the magnitude of the break increases. The reason the Exp W_T^0 and Sup W_T^0 statistics exhibit monotonic power is that they place most or all of the weight on the Wald statistics near the correct break date. For a detailed explanation of why the Mean W_T^1 statistic has nonmonotonic power, see the simulation study of Vogelsang (1996).

5.2 Power with Two Breaks

It is interesting to examine the power of the tests in the presence of two breaks in the trend function at dates λ_1 and λ_2. For simplicity, assume that the magnitudes of the breaks are equal. Results are only reported for $p = 1$ and slope changes. Results for $p = 0$ and $p = 1$ with intercept shifts can be found in Vogelsang (1994). The following DGP was used:

$$y_t = \gamma_1[1(t > \lambda_1 T)(t - \lambda_1 T) + 1(t > \lambda_2 T)(t - \lambda_2 T)] + v_t, \tag{15}$$

with v_t still defined by (14). The values of 0.25 and 0.75 were used for λ_1 and λ_2, respectively. The number of replications was 2,000, and the sample size was kept at 100. Again, $\gamma_1 = 0.0, 0.3, \ldots, 3.0$ and $\alpha = 0.0, 0.1, \ldots, 1.0$.

The results are shown in Figure 7. Strikingly, it can be seen that both statistics now show nonmonotonic power functions. The Sup W_T^1 statistic also exhibits nonmonotonic power. Interestingly, power is now much lower as compared with the single break case, as a comparison to Figure 6 indicates. Therefore, if two breaks in the same direction have occurred, none of the statistics will be able to detect the shifts unless the shifts are relatively small and the errors are not too persistent. If two large breaks in the same direction have occurred, the statistics cannot detect them even though the breaks would be obvious in the data.

6. THE STATISTICS IN FIRST DIFFERENCES

In this section, versions of the tests that are applicable to first differences of the data are developed. It is useful to consider taking first differences because this will result in efficiency gains when $\bar{\alpha} = 0$. Results are only provided for $p = 1$ and are easily generalized. Details are kept to a minimum to simplify the exposition.

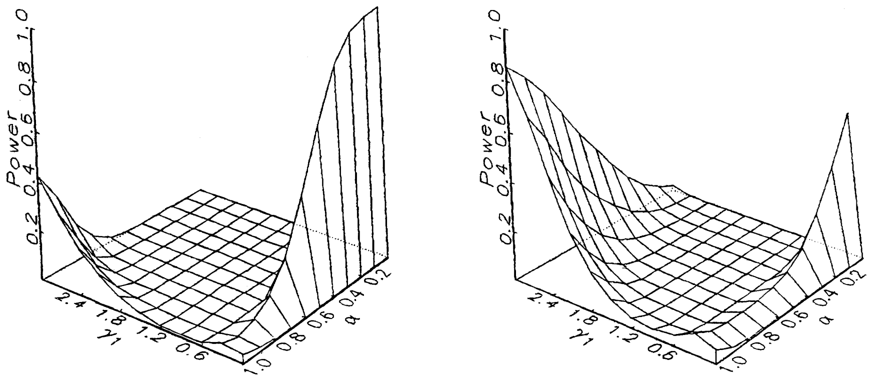


FIGURE 7. Power against two changes in trend, $p = 1$, $T = 100$, 5% nominal size. Left: Mean W_T^1, $\lambda_1 = 0.25$, $\lambda_2 = 0.75$, 1% trimming. Right: Exp W_T^1, $\lambda_1 = 0.25$, $\lambda_2 = 0.75$, 1% trimming.

The DGP under the null hypothesis of no break in the trend function is still governed by equations (1)–(3). First differencing and transforming (1) with $p = 1$, the regression of interest can be written as

$$\Delta y_t = \beta_1 + \delta_1 1(t > T_b^c) + \sum_{i=1}^{k+1} a_i \Delta y_{t-i} + \Delta e_t. \tag{16}$$

Note that β_0 and δ_0 vanish upon first differencing, and so the first-difference procedure cannot be used to detect breaks in the intercept of a series. Suppose that (16) is estimated by OLS. Let $WD_T(\lambda)$ denote the Wald test for testing $\delta_1 = 0$ in (16). Define the first-differenced statistics as

$$\text{Mean } WD_T^1 = T^{-1} \sum_{T_b \in \Lambda} WD_T(T_b/T), \tag{17}$$

$$\text{Exp } WD_T^1 = \log\left(T^{-1} \sum_{T_b \in \Lambda} \exp\left(\frac{1}{2} WD_T(T_b/T)\right)\right). \tag{18}$$

$$\text{Sup } WD_T^1 = \sup_{T_b \in \Lambda} WD_T(T_b/T). \tag{19}$$

Let $\hat{G}_0(r,\lambda) = 1(r > \lambda) - (1 - \lambda)$. Note that $\hat{G}_0(r,\lambda)$ is the residual from a projection of $1(r > \lambda)$ onto the space spanned by the identity function on [0,1]. The next two theorems give the asymptotic null distributions of (17)–(19).

THEOREM 4. *Let* $\{y_t\}$ *be the stochastic process described by* (1)–(3). *If the process* $\{v_t\}$ *is* $I(0)$, *that is,* $|\alpha| < 1$, *as* $T \to \infty$:

$WD_T^1(T_b/T) \Rightarrow 0$

uniformly in λ,

$$\text{Mean } WD_T^1 \Rightarrow 0, \qquad \text{Exp } WD_T^1 \Rightarrow 0, \qquad \textit{and} \qquad \text{Sup } WD_T^1 \Rightarrow 0.$$

THEOREM 5. *Let* $\{y_t\}$ *be the stochastic process described by* (1)–(3). *If the process* $\{v_t\}$ *is nearly* $I(1)$, *that is,* $\alpha = 1 - \bar{\alpha}/T$, *as* $T \to \infty$:

$$\begin{aligned} WD_T^1(T_b/T) &\Rightarrow \left(\int_0^1 \hat{G}_0(r,\lambda)\, dw(r) - \bar{\alpha}\sigma\sigma_e^{-1} \int_0^1 \hat{G}_0(r,\lambda) w_{\bar{\alpha}}(r)\, dr \right)^2 \\ &\qquad \times \left(\int_0^1 \hat{G}_0(r,\lambda)^2\, dr \right)^{-1} \\ &\equiv WD_u^1(\lambda) \end{aligned}$$

uniformly in λ,

$$\text{Mean } WD_T^1 \Rightarrow \int_{\lambda^*}^{1-\lambda^*} WD_u^1(\lambda)\, d\lambda \equiv \text{Mean } WD_u^1,$$

$$\text{Exp } WD_T^1 \Rightarrow \log\left\{ \int_{\lambda^*}^{1-\lambda^*} \exp\left(\frac{1}{2} WD_u^1(\lambda) \right) d\lambda \right\} \equiv \text{Exp } WD_u^1,$$

$$\text{Sup } WD_T^1 \Rightarrow \sup_{\lambda \in [\lambda^*, 1-\lambda^*]} WD_u^1(\lambda)\, d\lambda \equiv \text{Sup } WD_u^1.$$

In the stationary case, the statistics converge to zero. However, when the errors are nearly integrated, the statistics have nondegenerate distributions. When $\bar{\alpha} = 0$, the limiting distributions reduce to the expressions given by Theorem 1 with $p = 0$. If the data are highly persistent, then a conservative test can be constructed by using the value of $\bar{\alpha}$ that results in the largest critical values based on the limiting distributions given by Theorem 5. As is true for the level statistics, using $\bar{\alpha} = 0$ critical values gives a conservative test.

To demonstrate this fact, the distribution functions of the statistics were simulated using the same techniques in Section 3 for $\bar{\alpha} = 0,4,\ldots,16,20$. The distributions are only reported for the Mean WD_T^1 statistic and are qualitatively similar for the other two statistics. Figure 8 depicts plots of the distribution functions for Mean WD_T^1. As is clearly evident, the critical values are monotonically decreasing in $\bar{\alpha}$, and so the $\bar{\alpha} = 0$ critical values provide a conservative test.

To compare the power properties of the first-difference tests with the levels test, the local asymptotic distributions of the first-difference statistics were derived under the same local alternative, (10). The results are summarized in the following theorem.

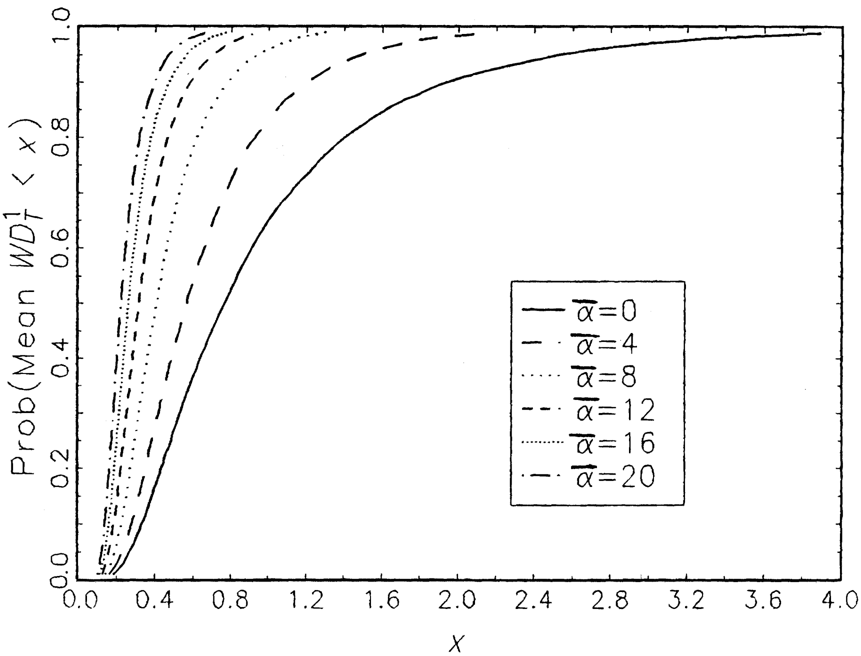


FIGURE 8. c.d.f. of Mean WD_T^1, nearly $I(1)$ errors, $p = 1$.

THEOREM 6. *Let* $\{y_t\}$ *be the stochastic process described by* (1), (2), *and* (10). *If the process* $\{v_t\}$ *is nearly* $I(1)$, *that is,* $\alpha = 1 - \bar{\alpha}/T \to \infty$:

$$\begin{aligned} WD_T^1(T_b/T) \Rightarrow & \left(\int_0^1 \hat{G}_0(r,\lambda)^2\,dr\right)^{-1} \\ & \times \left(\bar{\gamma}_1 \sigma_e^{-1} \int_0^1 \hat{G}_0(r,\lambda_c)\hat{G}_0(r,\lambda)\,dr + \int_0^1 \hat{G}_0(r,\lambda)\,dw(r)\right. \\ & \left. - \bar{\alpha}\sigma\sigma_e^{-1}\int_0^1 \hat{G}_0(r,\lambda) w_{\bar{\alpha}}(r)\,dr\right)^2 \\ \equiv & \; LD(\lambda) \end{aligned}$$

uniformly in λ,

$$\text{Mean } WD_T^1 \Rightarrow \int_{\lambda^*}^{1-\lambda^*} LD(\lambda)\,d\lambda,$$

$$\text{Exp } WD_T^1 \Rightarrow \log\left\{\int_{\lambda^*}^{1-\lambda^*} \exp\left(\frac{1}{2} LD(\lambda)\right) d\lambda\right\},$$

$$\text{Sup } WD_T^1 \Rightarrow \sup_{\lambda \in [\lambda^*, 1-\lambda^*]} LD(\lambda)\,d\lambda.$$

Using the results of Theorem 6, asymptotic power curves were simulated for the first-difference statistics. Results are reported only for the Exp WD_T statistic. The results for Mean WD_T and Sup WD_T are similar to Exp WD_T. Figure 9 depicts asymptotic power of Exp WD_T for $\bar{\gamma}_1 = 0,2,4,\ldots,12$ and for $\bar{\alpha} = 0,4,8,\ldots,16,20$. For small breaks, power is decreasing in $\bar{\alpha}$, but for large breaks power is increasing in $\bar{\alpha}$. Note that power is monotonic in $\bar{\gamma}_1$. If the plots in Figure 8 are compared with the plots in Figure 3, it can be determined under what conditions the first-difference tests will be preferred to the levels tests and vice versa. To facilitate such a comparison, differences of power in the two figures are plotted in Figure 10. Points above the zero axis are where the levels test has higher power, and points below the zero axis are where the fist-difference test has higher power. If $\bar{\alpha} = 0$, then the first-difference test delivers higher power. Therefore, the first-difference tests should perform better with highly persistent series. But as the errors become less persistent ($\bar{\alpha}$ increases), the levels tests are more powerful in detecting small breaks, whereas the first-difference tests remain more powerful in detecting large breaks.

To see if the first-difference tests also exhibit nonmonotonic power when there are two breaks, simulations were run using DGP (15) and the statistics were computed using regression (16) with $k = 1$. As before, 2,000 replications were used and the 5% asymptotic critical value with $\bar{\alpha} = 0$ was used. Power was

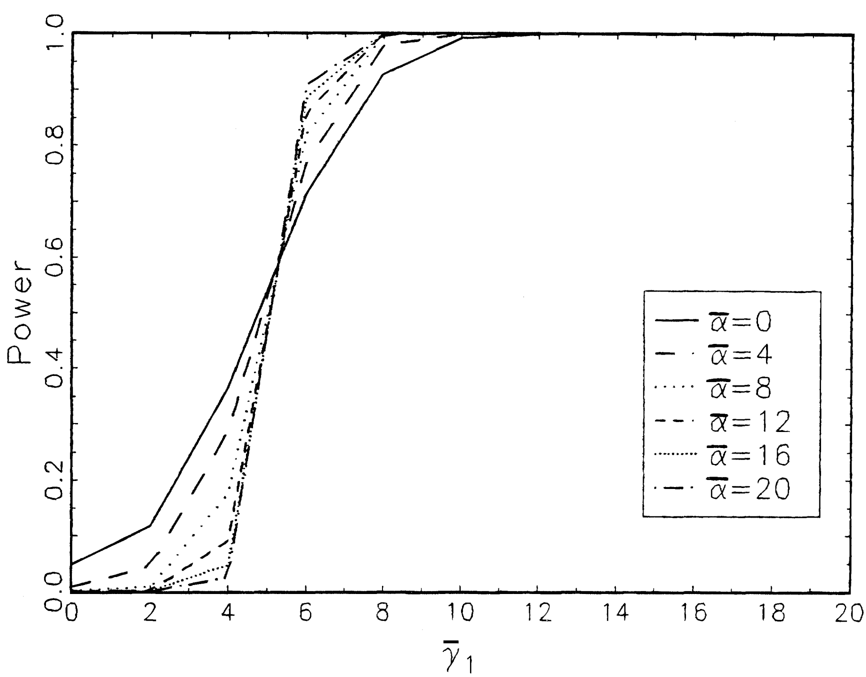


FIGURE 9. Asymptotic power Exp WD_T^1, nearly $I(1)$ errors, $p = 1$.

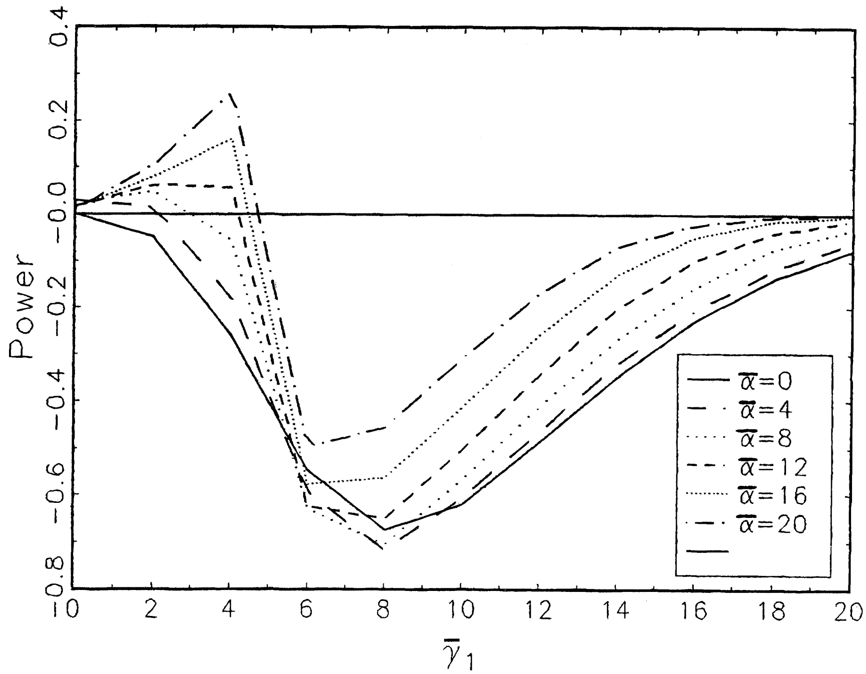


FIGURE 10. Difference in asymptotic power of Exp W_T^1 and Exp WD_T^1, nearly $I(1)$ errors, $p = 1$.

simulated for $\gamma_1 = 0.0, 0.1, \ldots, 1.0$, and the results are plotted in Figure 11. Notice that the power curves are monotonic with respect to γ_1. Except for small breaks, for instance, $\gamma_1 = 0.2$, power of the first-difference tests dominates the levels tests when there are two breaks in the same direction.

7. EMPIRICAL RESULTS

To illustrate the usefulness of the statistics in the $p = 1$ case, they were applied to three sets of data that recently have been analyzed in the macroeconomics literature. The first series is comprised of international postwar real GNP/GDP data that were used by BLS (1992). This set of data is very similar to the data analyzed by Perron (1991). The countries include Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States. All series in this set were quarterly. The sources can be found in BLS (1992). The second set of data is taken from Kormendi and Meguire (1990). The historical real GNP/GDP series for 12 countries including Australia, Canada, Denmark, Finland, France, Germany, Italy, the Netherlands, Norway, Sweden, the United Kingdom, and the United States are used. These series are annual and span more than 100 years (1870–1986). The sources can be found in Kormendi and Meguire (1990). In a recent study, Ben-

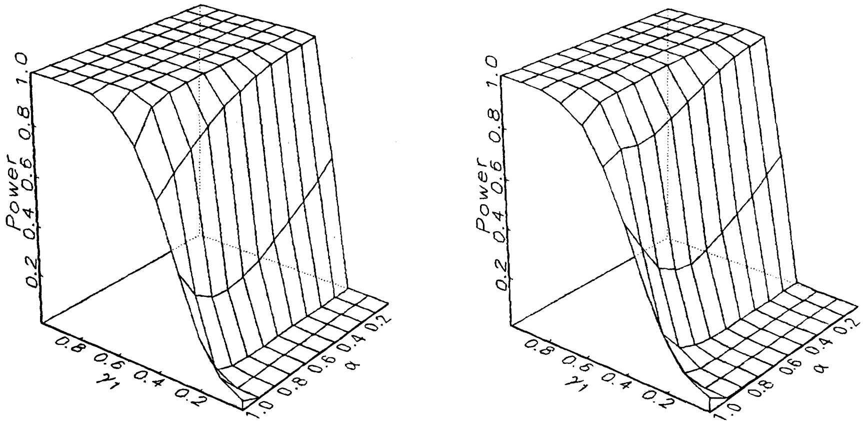


FIGURE 11. Power against two changes in trend, $p = 1$, $T = 100$, 5% nominal size. Left: Mean WD_T^1, $\lambda_1 = 0.25$, $\lambda_2 = 0.75$, 1% trimming. Right: Exp WD_T^1, $\lambda_1 = 0.25$, $\lambda_2 = 0.75$, 1% trimming.

David and Papell (1995) applied the supremum statistics from this paper to a broader set of series from the data set in Kormendi and Meguire (1990). The third set of data is the data used by Nelson and Plosser (1982). The Nelson–Plosser data were also used by Perron (1989) in the context of unit root tests allowing for a break in the trend function and by Chu and White (1992) who implemented tests for structural change in the trend function. These data are annual, spanning from the late 1800's to 1970. The exact sources can be found in Nelson and Plosser (1982). Note that all three sets of data comprise series that are trending over time.

For the tests based on the levels of the data, the statistics were calculated using estimates from (6) (with $p = 1$) and 1% trimming. Critical values were taken from Table 2. In practice, one must choose a value for k, the order of the estimated autoregression. An approach recommended by Perron and Vogelsang (1992) was used. One first estimates regression (6) by using a maximal value of 10 for k. One then tests the significance of the coefficient on the last included lag by using a 5% two-tailed t-test. Asymptotic normality of the t-test is used to carry out inference. Asymptotic normality holds whether the errors are stationary or have a unit root. If this coefficient is significant, the procedure is stopped. Otherwise, k is reduced by 1 and regression (6) is estimated by using $k = 9$. This continues until one either finds significance or until $k = 0$. When $k = 10$ and the coefficient on the 10th lag was significant, the maximal value of k was increased to 15. Asymptotically, this procedure yields the same distributions under the null hypothesis as when the order of the autoregressive is known. See Hall (1994) and Ng and Perron (1995) for theoretical details. The results are reported in Table 3 in columns 3–5 along with the break date chosen by the Sup W_T^1 statistic in column 6.

Using the Sup W_T^1 statistic, the null hypothesis is rejected for two of the postwar GNP series (France[1] and Japan), seven of the historical GNP series, and six

TABLE 3. Empirical results

Series	Sample	Levels statistics Mean	Exp	Sup	T_b	First-difference statistics Mean	Exp	Sup	T_b
International quarterly real GDP/GNP									
Canada	48:1–89:2	5.05	5.44	19.64	81:2	1.73	1.22	6.01	75:2
France	63:1–89:2	4.46	14.18[a]	39.46[a]	68:2	5.09[a]	6.17[a]	14.17[b]	74:1
France*	63:1–89:2	4.69	4.12	13.57	72:3	2.01[d]	3.53[b]	12.87[b]	66:2
Germany	50:1–89:2	2.13	1.45	7.19	80:1	4.08[b]	3.17[b]	9.16	60:3
Italy	52:1–82:4	5.36	4.81	17.74	69:4	5.25[a]	4.36[a]	15.18[a]	74:2
Japan	52:1–89:2	13.64[a]	20.77[a]	49.10[a]	68:3	3.95[b]	6.49[a]	20.98[a]	73:2
U.K.	60:1–89:2	2.58	5.81	20.25	79:4	0.59	0.25	1.63	73:4
U.S.	47:1–89:2	3.62	2.78	16.36	64:4	0.54	0.27	10.19[d]	51:3
Historical annual real GDP/GNP									
Australia	1870–1986	10.03[b]	9.09[c]	23.39[d]	1927	0.62	0.31	6.66	1946
Canada	1870–1986	4.74	13.18[a]	43.11[a]	1929	1.10	1.20	7.51	1933
Denmark	1870–1986	3.13	5.03	28.56[b]	1939	0.81	0.57	4.67	1943
Finland	1900–1986	3.58	5.64	20.06	1918	0.66	0.51	11.73[c]	1918
France	1870–1986	6.13	16.21[a]	56.40[a]	1942	0.87	1.25	12.66[b]	1947
Germany	1870–1986	5.55	6.01	19.65	1946	0.40	0.29	7.57	1946
Italy	1870–1986	5.11	8.04[d]	25.29[c]	1945	1.31	2.13[d]	16.76[a]	1945

Netherlands	1900–1986	3.93	7.77[d]	38.48[a]	1945	0.51	0.76	8.57	1944
Norway	1870–1986	2.81	3.90	17.40	1947	4.01[b]	3.53[b]	15.98[a]	1947
Sweden	1870–1986	3.46	2.67	9.22	1960	0.91	0.48	3.08	1918
U.K.	1870–1986	5.08	14.25[a]	38.00[a]	1918	0.73	0.43	5.13	1932
U.S.	1870–1986	2.08	2.22	11.54	1929	0.64	0.29	2.16	1944
			Nelson–Plosser data, annual						
CPI	1860–1970	5.32	4.88	13.48	1872	1.31	0.75	9.89[c]	1879
Employment	1890–1970	2.19	2.02	11.43	1929	0.21	−0.06	4.23	1906
GNP deflator	1889–1970	1.91	4.83	18.29	1920	0.46	0.13	5.39	1940
Industrial Productivity	1860–1970	1.93	3.15	26.98[c]	1929	0.62	0.14	1.86	1952
Interest rate	1900–1970	9.13[c]	7.77[d]	25.74[c]	1962	3.09[c]	6.53[a]	14.68[a]	1967
Money stock	1889–1970	1.94	2.01	10.89	1928	0.16	−0.08	3.02	1920
Nominal GNP	1909–1970	3.24	6.09	31.60[a]	1929	0.26	−0.03	7.87	1932
Per capita GNP	1909–1970	2.28	4.31	16.38	1938	0.26	−0.03	5.77	1921
Real GNP	1909–1970	2.95	4.27	21.34	1929	0.27	−0.03	5.76	1938
Real wages	1900–1970	3.03	5.38	23.37[d]	1940	0.51	0.16	8.48	1938
Stock prices	1871–1970	7.07	8.72[c]	24.96[d]	1936	0.54	0.21	5.89	1947
Unemployment rate	1890–1970	2.90	4.33	23.45[d]	1929	0.11	−0.11	2.42	1933
Velocity	1869–1970	4.58	3.41	11.66	1947	1.60	0.98	8.84	1949
Wages	1900–1970	2.67	4.00	17.00	1929	0.28	−0.01	3.99	1920

Note: *a*, *b*, *c*, and *d* denote significance at the 1%, 2.5%, 5%, and 10% levels, respectively, and * denotes the French GNP series with the strike of 1968:2 interpolated out. See BLS (1992) for details on the interpolation.

of the Nelson–Plosser data series. The Exp W_T^1 statistic allows rejection in almost as many cases as the Sup W_T^1 statistic but never rejects when the latter fails to do so. However, the Mean W_T^1 statistic rejects the null hypothesis in only three cases: the Japanese postwar real GDP series, the Australian historical GDP series, and the Nelson–Plosser interest rate.

The first-difference statistics were also applied to the series. The same values of k used in the levels tests were used for the first-difference tests. This was done to avoid potential problems caused by first differencing a stationary process and inducing noninvertible moving average errors with a unit root. The results are reported in Table 3 in columns 7–9, and the value of T_b chosen by Sup WD_T^1 is reported in the last column. The critical values were taken from Table 1 ($p = 0$ column). For the postwar GNP/GDP series, rejection of the null hypothesis is now possible for six of the eight series. The null hypothesis can be rejected for four of the series in which rejection was not possible by using the tests in levels. Interestingly, for France, Japan, and Italy, the Sup WD_T^1 statistic picked break dates near the time of the oil crisis (1973). For the other two sets of data, in only three instances could the null hypothesis be rejected when the tests in levels failed to do so. They were the historical GNP series for Finland and Norway and the Nelson–Plosser CPI series. The rejections for the Norway series were the strongest.

The results of this section clearly indicate that the assumption of a stable linear trend will not result in well-specified univariate models for many of the series examined. This is true whether one views these series as being stationary or as having a unit root. It is important to keep in mind that a rejection has little to say beyond whether a stable trend leads to a well-specified model. Even though the tests are designed to detect a single break in the trend function, a rejection does not indicate what specification should be used for the trend function because the tests will have power in detecting more general alternatives. Other techniques would be needed to guide further specification of the trend function.

8. CONCLUDING REMARKS

In this paper, test statistics for detecting a break at an unknown date in the trend function of a dynamic univariate time series were developed. Asymptotic results were derived for both $I(0)$ and $I(1)$ errors. When the errors are highly persistent and it is not known which asymptotic theory ($I(0)$ or $I(1)$) provides a better approximation, a conservative approach based on nearly integrated asymptotics was provided. The tests were based on the mean and exponential statistics of Andrews and Ploberger (1994) and the supremum statistic of Andrews (1993). Their results were extended to allow trending and unit root regressors. An analysis of the power of the statistics revealed that the mean statistic suffers from nonmonotonic power as the break magnitude increases, whereas the exponential and supremum statistics have monotonic power. Interestingly, all three statistics have nonmonotonic power when there are two breaks in the same direction. Versions of the tests applicable to first differences of the data were also proposed.

It was found that first differencing results in superior power when the errors are purely $I(1)$ or are highly persistent with a medium or large break. When the data are less persistent (but still near a unit root) and the break small, the levels tests have better power. Therefore, the levels and first-difference tests complement each other.

NOTE

1. The rejection of the France series is due to the labor strike of May 1968. For the series that interpolates out the strike observation, the null hypothesis cannot be rejected.

REFERENCES

Andrews, D.W.K. (1993) Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821–856.

Andrews, D.W.K. & W. Ploberger (1994) Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* 62, 1383–1414.

Bai, J. (1993) Estimation of Structural Change based on Wald-Type Statistics. Mimeo, Massachusetts Institute of Technology.

Bai, J., R.L. Lumsdaine, & J.H. Stock (1997) Testing for and dating breaks in integrated and cointegrated time series. *Review of Economic Studies*, forthcoming.

Bai, J. & P. Perron (1997) Testing for and estimation of multiple structural changes. *Econometrica*, forthcoming.

Banerjee, A., R.L. Lumsdaine, & J.H. Stock (1992) Recursive and sequential tests of the unit root and trend break hypothesis: Theory and international evidence. *Journal of Business and Economic Statistics* 10, 271–287.

Ben-David, D. & D.H. Papell (1995) The great wars, the great crash and steady state growth: Some new evidence about old stylized fact. *Journal of Monetary Economics* 36, 453–475.

Chan, C.H. & C.Z. Wei (1987) Asymptotic inference for nearly stationary AR(1) processes. *Annals of Statistics* 15, 1050–1063.

Chu, C.-S.J. (1989) New Tests for Parameter Constancy in Stationary and Nonstationary Regression Models. Mimeo, University of California at San Diego.

Chu, C.-S.J. & H. White (1992) A direct test for changing trend. *Journal of Business and Economic Statistics* 10, 289–300.

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

Gardner, L.A. (1969) On detecting changes in the mean of normal variates. *Annals of Mathematical Statistics* 40, 116–126.

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.

Hansen, B.E. (1990) Lagrange Multiplier Tests for Parameter Constancy in Non-Linear Models. Mimeo, Department of Economics, University of Rochester.

Hansen, B.E. (1992a) Tests for parameter instability in regressions with $I(1)$ processes. *Journal of Business and Economic Statistics* 10, 321–335.

Hansen, B.E. (1992b) Convergence to stochastic integrals for dependent heterogeneous processes. *Econometric Theory* 8, 489–500.

Kim, H.J. & D. Siegmund (1989) The likelihood ratio test for a change-point in simple linear regression. *Biometrika* 76, 409–423.

Kormendi, R.C. & P. Meguire (1990) A multicountry characterization of the nonstationarity of aggregate output. *Journal of Money, Credit and Banking* 22, 77–93.

Kramer, W., W. Ploberger, & R. Alt (1988) Testing for structural change in dynamic models. *Econometrica* 56, 1355–1369.

MacNeill, I.B. (1978) Properties of sequences of partial sums of polynomial regression residuals with applications to test for change of regression at unknown times. *Annals of Statistics* 6, 422–433.
Nelson, C.R. & H. Kang (1981) Spurious periodicity in inappropriately detrended time series. *Econometrica* 49, 741–751.
Nelson, C.R. & C.I. Plosser (1982) Trends and random walks in macroeconomic time series. *Journal of Monetary Economics* 10, 139–162.
Ng, S. & 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.
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.
Perron, P. (1989) The great crash, the oil price shock and the unit root hypothesis. *Econometrica* 57, 1361–1401.
Perron, P. (1990) Testing for a unit root in a time series with a changing mean. *Journal of Business and Economic Statistics* 8, 153–162.
Perron, P. (1991) A Test for Changes in a Polynomial Trend Function for a Dynamic Time Series. Mimeo, Université de Montréal.
Perron, P. & T.J. Vogelsang (1992) Nonstationarity and level shifts with an application to purchasing power parity. *Journal of Business and Economic Statistics* 10, 301–320.
Phillips, P.C.B. (1987) Towards a unified asymptotic theory for autoregression. *Biometrika* 74, 535–547.
Quandt, R. (1960) Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association* 55, 324–330.
Vogelsang, T.J. (1994) Wald-Type Tests for Detecting Shifts in the Trend Function of a Dynamic Time Series. CAE Working paper 94-12, Cornell University.
Vogelsang, T.J. (1996) Sources of Nonmonotonic Power when Testing for a Shift in the Trend Function of a Dynamic Time Series. Working paper, Department of Economics, Cornell University.
Zivot, E. & P.C.B. Phillips (1994) A Bayesian analysis of trend determination in economic time series. *Econometric Reviews* 13, 291–336.

APPENDIX

In this Appendix, proofs of the theorems are given. The approach is to establish convergence of the Wald statistic uniformly in λ for each theorem. Then, by the continuous mapping theorem, convergence of the mean, exponential, and supremum statistics directly follows. The Wald statistics are expressed as continuous functions of stochastic integrals defined with respect to λ, and the results of Hansen (1992b) are used to establish weak convergence uniformly in λ.

To simplify the presentation, it is convenient to write the model in matrix form as $\tilde{Y} = \tilde{g}(T_b^c)\gamma + \tilde{v}$, or $\Delta\tilde{Y} = \tilde{g}(T_b^c)\delta + \tilde{Z}\phi + \tilde{e}$, where Y, ΔY, v, and e are the $T \times 1$ vectors $\{y_t\}$, $\{\Delta y_t\}$, $\{v_t\}$, and $\{e_t\}$, $\tilde{g}(T_b^c)$ is the $T \times (p+1)$ matrix $\{g(t,\lambda_c)\}$, $Z = (Z_1, Z_2)$, with Z_1 the $T \times 1$ vector $\{y_{t-1}\}$, Z_2 the $T \times k$ matrix $\{\Delta y_{t-1}, \ldots, \Delta y_{t-k}\}$, and $\phi = (\pi, c')'$ with $c = (c_1, \ldots, c_k)'$. A tilde over a vector or matrix denotes the residuals from the projection onto the space spanned by f, which is the $T \times (p+1)$ matrix $\{f(t)\}$. Recall that a tilde over a function defined on the space $[0,1]$ denotes the residuals from the projection onto the space spanned by $F(r)$. Define $X = (X_1, X_2)$ where X_1 is the $T \times 1$ vector $\{v_{t-1}\}$ and X_2 is the $T \times k$ matrix $\{\Delta v_{t-1}, \ldots, \Delta v_{t-k}\}$. In addition to τ_1 defined previously, let τ_2 denote a $(k+1) \times$

$(k+1)$ diagonal matrix with diagonal elements $(T^{-1/2},1,\ldots,1)$. Unless otherwise noted, all sums run from 1 to T and given a matrix x, M_x denotes the matrix $I - x(x'x)^{-1}x'$.

The next three lemmas establish convergence of moments that appear in the definitions of the Wald statistics. Using the lemmas, the theorems are easy to establish.

LEMMA A1. *As $T \to \infty$, the following hold uniformly in λ,*

(a) $T^{-1}\tau_1 \tilde{g}(T_b)'\tilde{g}(T_b)\tau_1 \Rightarrow \int_0^1 \tilde{G}(r,\lambda)'\tilde{G}(r,\lambda)\,dr$,
(b) $T^{-1/2}\tau_1 \tilde{g}(T_b)'e \Rightarrow \sigma_e \int_0^1 \tilde{G}(r,\lambda)'dw(r)$.

Proof.

$$
\begin{aligned}
\text{(a) } T^{-1}\tau_1 \tilde{g}(T_b)'\tilde{g}(T_b)\tau_1 &= T^{-1}\tau_1 g(T_b)'g(T_b)\tau_1 \\
&\quad - T^{-1}\tau_1 g(T_b)'f\tau_1(T^{-1}\tau_1 f'f\tau_1)^{-1}T^{-1}\tau_1 f'g(T_b)\tau_1 \\
&= T^{-1}\Sigma G(t/T,\lambda)'G(t/T,\lambda) \\
&\quad - T^{-1}\Sigma G(t/T,\lambda)'F(t/T)[T^{-1}\Sigma F(t/T)'F(t/T)]^{-1} \\
&\quad \times T^{-1}\Sigma F(t/T)'G(t/T,\lambda) \\
&\Rightarrow \int_0^1 G(r,\lambda)'G(r,\lambda)\,dr \\
&\quad - \left(\int_0^1 G(r,\lambda)'F(r)\,dr\right. \\
&\quad \left.\times\left[\int_0^1 F(r)'F(r)\,dr\right]^{-1}\int_0^1 F(r)'G(r,\lambda)\,dr\right) \\
&= \int_0^1 \tilde{G}(r,\lambda)'\tilde{G}(r,\lambda)\,dr,
\end{aligned}
$$

and is trivially uniform in λ.

$$
\begin{aligned}
\text{(b) } T^{-1/2}\Sigma G(t/T,\lambda)'e_t &= T^{-1/2}\sum_{T_b+1}^{T} F(t/T-\lambda)'e_t \\
&\Rightarrow \sigma_e \int_\lambda^1 F(r-\lambda)'\,dw(r) = \sigma_e \int_0^1 G(r,\lambda)'\,dw(r),
\end{aligned}
$$

uniformly in λ by Theorem 2.1 of Hansen (1992b) because $\{e_t\}$ is a martingale difference sequence and $F(t/T-\lambda) \Rightarrow F(r-\lambda)$ uniformly in λ. Using this result, it directly follows that

$$
\begin{aligned}
T^{-1/2}\tau_1 \tilde{g}(T_b)'e &= T^{-1/2}\Sigma G(t/T,\lambda)'e_t \\
&\quad - (T^{-1}\Sigma G(t/T,\lambda)'F(t/T)[T^{-1}\Sigma F(t/T)'F(t/T)]^{-1} \\
&\quad \times T^{-1/2}\Sigma F(t/T)'e_t) \\
&\Rightarrow \sigma_e \int_0^1 \tilde{G}(r,\lambda)'\,dw(r)
\end{aligned}
$$

uniformly in λ. ■

LEMMA A2. *Let* $|\alpha| < 1$ *so that* $\{v_t\}$ *is* $I(0)$. *As* $T \to \infty$, *the following hold uniformly in* λ:

(a) $T^{-1/2}\widetilde{X}'\tilde{e} \Rightarrow \Psi$ *where* Ψ *is a* $(k+1) \times 1$ *multivariate normal random variable.*
(b) $T^{-1}\widetilde{X}'\widetilde{X} \Rightarrow \Omega$ *where* $\Omega = EX_t'X_t$ *with* $X_t = (v_t, \Delta v_{t-1}, \ldots, \Delta v_{t-k})$ *and* Ω^{-1} *exists.*
(c) $T^{-1}\tau_1 \tilde{g}(T_b)'\widetilde{X} = o_p(1)$.

Proof. These are standard results (see Fuller 1976) and trivially hold uniformly in λ. ■

LEMMA A3. *Let* $\alpha = 1 - \bar{\alpha}/T$ *so that* $\{v_t\}$ *is nearly* $I(1)$. *As* $T \to \infty$, *the following hold uniformly in* λ:

(a) $T^{-1/2}\tau_2\widetilde{X}'\tilde{e} = [T^{-1}\widetilde{X}_1'\tilde{e}, T^{-1/2}\widetilde{X}_2'\tilde{e}] \Rightarrow \left[\sigma\sigma_e \int_0^1 \bar{w}_{\bar{\alpha}}(r)\,dw(r), \Psi_2'\right]'$,

where Ψ_2 *is a* $k \times 1$ *multivariate normal random variable.*

(b) $$T^{-1}\tau_2\widetilde{X}'\widetilde{X}\tau_2 = \begin{bmatrix} T^{-2}\widetilde{X}_1'\widetilde{X}_1 & T^{-3/2}\widetilde{X}_1'\widetilde{X}_2 \\ T^{-3/2}\widetilde{X}_2'\widetilde{X}_1 & T^{-1}\widetilde{X}_2'\widetilde{X}_2 \end{bmatrix} = \begin{bmatrix} \sigma^2 \int_0^1 \bar{w}_{\bar{\alpha}}(r)^2\,dr & 0_{1\times k} \\ 0_{k\times 1} & \Omega_{22} \end{bmatrix},$$

where $\Omega_{22} = EX_{2t}'X_{2t}$, $X_{2t} = (\Delta v_{t-1}, \ldots, \Delta v_{t-k})$, *and* Ω_{22}^{-1} *exists.*

(c) $$T^{-1}\tau_1\tilde{g}(T_b)'\widetilde{X}\tau_2 = [T^{-3/2}\tilde{g}(T_b)'\widetilde{X}_1, T^{-1}\tilde{g}(T_b)'\widetilde{X}_2]$$
$$\Rightarrow \left[\sigma \int_0^1 \widetilde{G}(r,\lambda)'\bar{w}_{\bar{\alpha}}(r)\,dr, 0_{(p+1)\times k}\right].$$

Proof.

(a) $T^{-1}\widetilde{X}_1'\tilde{e} = T^{-1}\Sigma\tilde{v}_{t-1}e_t \Rightarrow \sigma\sigma_e\int_0^1 \bar{w}_{\bar{\alpha}}(r)\,dw(r)$ uniformly in λ from Theorem 4.4 of Hansen (1992b). Because $\{\Delta v_t\}$ is a stationary process with innovations $\{e_t\}$, it follows from standard results (e.g., Fuller, 1976) that $T^{-1/2}\widetilde{X}_2'\tilde{e} \Rightarrow \Psi_2'$ and is trivially uniform in λ.
(b) $T^{-2}\widetilde{X}_1'\widetilde{X}_1 = T^{-2}\Sigma\tilde{v}_{t-1}^2 \Rightarrow \sigma^2\int_0^1 \bar{w}_{\bar{\alpha}}(r)^2\,dr$. Because $\{\Delta v_t\}$ is stationary, it follows from standard results that $T^{-1}\widetilde{X}_2'\widetilde{X}_2 \Rightarrow \Omega_{22}$ and $T^{-3/2}\widetilde{X}_1'\widetilde{X}_2 = o_p(1)$. These trivially hold uniformly in λ.
(c) $T^{-3/2}\tilde{g}(T_b)'\widetilde{X}_1 = T^{-3/2}\Sigma\widetilde{G}(t/T,\lambda)'\tilde{v}_{t-1} \Rightarrow \sigma\int_0^1 \widetilde{G}(r,\lambda)'\bar{w}_{\bar{\alpha}}(r)$ uniformly in λ because $T^{-3/2}\Sigma G(t/T,\lambda)'v_{t-1} = T^{-3/2}\Sigma_{T_b+1}^{T} F(t/T - \lambda)'v_{t-1} \Rightarrow \sigma\int_\lambda^1 F(r-\lambda)'w_{\bar{\alpha}}(r) = \sigma\int_0^1 G(r,\lambda)'w_{\bar{\alpha}}(r)$ uniformly in λ. Because $\{\Delta v_t\}$ is stationary, it follows from standard results that $T^{-1}\tilde{g}(T_b)'\widetilde{X}_2 = o_p(1)$. ■

Proof of Theorem 1. Under the null hypothesis $\gamma = \delta = 0$, $\widetilde{Y} = \tilde{v}$ and $\Delta\widetilde{Y} = \widetilde{Z}\phi + \tilde{e}$. Direct calculation gives

$$W_T^p(T_b/T) = T^{-1/2}\tilde{e}'M_{\widetilde{X}}\tilde{g}(T_b)\tau_1[T^{-1}\tau_1\tilde{g}(T_b)'M_{\widetilde{X}}\tilde{g}(T_b)\tau_1]^{-1}$$
$$\times T^{-1/2}\tau_1\tilde{g}(T_b)'M_{\widetilde{X}}\tilde{e}/s^2.$$

$$T^{-1/2}\tau_1\tilde{g}(T_b)'M_{\widetilde{X}}\tilde{e} = T^{-1/2}\tau_1\tilde{g}(T_b)'e - T^{-1}\tau_1\tilde{g}(T_b)'\widetilde{X}[T^{-1}\widetilde{X}'\widetilde{X}]^{-1}T^{-1/2}\widetilde{X}'e$$
$$\Rightarrow \sigma_e\int_0^1 \widetilde{G}(r,\lambda)'\,dw(r)$$

uniformly in λ by Lemmas A1 and A2.

$$T^{-1}\tau_1 \tilde{g}(T_b)'M_{\tilde{X}}\tilde{g}(T_b)\tau_1 = T^{-1}\tau_1 \tilde{g}(T_b)'\tilde{g}(T_b)\tau_1 - T^{-1}\tau_1 \tilde{g}(T_b)'\tilde{X}[T^{-1}\tilde{X}'\tilde{X}]^{-1}T^{-1}\tilde{X}'\tilde{g}(T_b)\tau_1 \Rightarrow \int_0^1 \tilde{G}(r,\lambda)'\tilde{G}(r,\lambda)\,dr$$

uniformly in λ by Lemmas A1 and A2.

$s^2 = T^{-1}\tilde{e}'M_{\tilde{X}}\tilde{e} + o_p(1) = T^{-1}\tilde{e}'\tilde{e} + o_p(1) \Rightarrow \sigma_e^2.$

The theorem follows from simple algebraic manipulation. ■

Proof of Theorem 2. Under the null hypothesis $\gamma = \delta = 0$, $\tilde{Y} = \tilde{v}$, and $\Delta\tilde{Y} = \tilde{Z}\phi + \tilde{e}$. Direct calculation gives

$$W_T^p(T_b/T) = T^{-1/2}\tilde{e}'M_{\tilde{X}}\tilde{g}(T_b)\tau_1[T^{-1}\tau_1\tilde{g}(T_b)'M_{\tilde{X}}\tilde{g}(T_b)\tau_1]^{-1} \times T^{-1/2}\tau_1\tilde{g}(T_b)'M_{\tilde{X}}\tilde{e}/s^2.$$

$$T^{-1/2}\tau_1\tilde{g}(T_b)'M_{\tilde{X}}\tilde{e} = T^{-1/2}\tau_1\tilde{g}(T_b)'e - T^{-1}\tau_1\tilde{g}(T_b)'\tilde{X}\tau_2[T^{-1}\tau_2\tilde{X}'\tilde{X}\tau_2]^{-1}T^{-1/2}\tau_2\tilde{X}'e$$

$$\Rightarrow \sigma_e\int_0^1 \tilde{G}(r,\lambda)'\,dw(r) - \left(\sigma_e\int_0^1 \tilde{G}(r,\lambda)'\tilde{w}_{\bar{\alpha}}(r)\int_0^1 \tilde{w}_{\bar{\alpha}}(r)\,dw(r)\right)\bigg/\int_0^1 \tilde{w}_{\bar{\alpha}}(r)^2\,dr \equiv \sigma_e H_1^p(\lambda)$$

uniformly in λ by Lemmas A1 and A3.

$$T^{-1}\tau_1\tilde{g}(T_b)'M_{\tilde{X}}\tilde{g}(T_b)\tau_1 = T^{-1}\tau_1\tilde{g}(T_b)'\tilde{g}(T_b)\tau_1 - T^{-1}\tau_1\tilde{g}(T_b)'\tilde{X}\tau_2[T^{-1}\tau_2\tilde{X}'\tilde{X}\tau_2]^{-1}T^{-1}\tau_2\tilde{X}'\tilde{g}(T_b)\tau_1$$

$$\Rightarrow \int_0^1 \tilde{G}(r,\lambda)'\tilde{G}(r,\lambda)\,dr - \left(\int_0^1 \tilde{G}(r,\lambda)'\tilde{w}_{\bar{\alpha}}(r)\int_0^1 \tilde{G}(r,\lambda)\tilde{w}_{\bar{\alpha}}(r)\right)\bigg/\int_0^1 \tilde{w}_{\bar{\alpha}}(r)^2\,dr = H_2^p(\lambda)$$

uniformly in λ by Lemmas A1 and A3. Because $\lambda \in [\lambda^*, 1-\lambda^*]$, $H_2^p(\lambda)^{-1}$ exists.

$s^2 = T^{-1}\tilde{e}'M_{\tilde{X}}\tilde{e} + o_p(1) = T^{-1}\tilde{e}'\tilde{e} + o_p(1) \Rightarrow \sigma_e^2.$

The theorem follows from simple algebraic manipulation. ■

Proof of Theorem 3. Because $p = 1$, $\tau_1 = \text{diag}(1, T^{-1})$ and $\tau_2 = \text{diag}(T^{-1/2}, 1)$. Under the local alternative (10), $\tilde{Y} = T^{-1/2}\tilde{g}_1(T_b^c)\bar{\gamma}_1 + \tilde{v}$ and $\Delta\tilde{Y} = T^{-1/2}\tilde{g}(T_b^c)\delta + \tilde{Z}\phi + \tilde{e} + o_p(1)$, where $\tilde{Z} = [T^{-1/2}\tilde{g}_1(T_b^c)\bar{\gamma}_1 + \tilde{X}_1, \tilde{X}_2] + o_p(1)$. Recall that under (10) $\delta = \bar{\gamma}_1[\alpha - C(1), 1-\alpha] = \bar{\gamma}_1[1 - \bar{\alpha}/T - C(1), \bar{\alpha}/T]$. Therefore, $\tau_1^{-1}\delta = \bar{\gamma}_1[1 - \bar{\alpha}/T - C(1), \bar{\alpha}] \Rightarrow \bar{\gamma}_1[1 - C(1), \bar{\alpha}] = \bar{\delta}$. First consider convergence of the following moments:

$$T^{-1/2}\tau_2\tilde{Z}'e = [T^{-3/2}\tilde{g}_1(T_b^c)'e\bar{\gamma}_1 + T^{-1}\tilde{X}_1'e, T^{-1/2}\tilde{X}_2'e] + o_p(1) \Rightarrow \left\{\sigma_e\int_0^1 [\bar{\gamma}_1\tilde{G}_1(r,\lambda_c) + \sigma\tilde{w}_{\bar{\alpha}}(r)]\,dw(r), \Psi_2\right\}, \quad \textbf{(A.1)}$$

uniformly in λ by Lemmas A1 and A3.

$$T^{-1}\tau_2\tilde{Z}'\tilde{Z}\tau_2$$

$$= \begin{bmatrix} T^{-1}[T^{-1}\tilde{g}_1(T_b^c)\bar{\gamma}_1 + T^{-1/2}\tilde{X}_1]'[T^{-1}\tilde{g}_1(T_b^c)\bar{\gamma}_1 + T^{-1/2}\tilde{X}_1] & T^{-3/2}[T^{-1}\tilde{g}_1(T_b^c)\bar{\gamma}_1 + T^{-1/2}\tilde{X}_1]'\tilde{X}_2 \\ T^{-3/2}\tilde{X}_2'[T^{-1}\tilde{g}_1(T_b^c)\bar{\gamma}_1 + T^{-1/2}\tilde{X}_1] & T^{-1}\tilde{X}_2'\tilde{X}_2 \end{bmatrix}$$

$$+ o_p(1)$$

$$\Rightarrow \begin{bmatrix} \int_0^1 [\bar{\gamma}_1 G_1(r,\lambda_c) + \sigma\widetilde{w}_{\bar{\alpha}}(r)]^2\,dr & 0_{1\times k} \\ 0_{k\times 1} & \Omega_{22} \end{bmatrix}, \tag{A.2}$$

uniformly in λ by Lemmas A1 and A3.

$$T^{-1}\tau_1\tilde{g}(T_b)'\tilde{Z}\tau_2 = [T^{-1}\tau_1\tilde{g}(\lambda)'\tilde{g}_1(\lambda_c)T^{-1}\bar{\gamma}_1 + T^{-1}\tau_1\tilde{g}(\lambda)'\tilde{X}_1, T^{-1}\tau_1\tilde{g}(\lambda)'\tilde{X}_2]$$

$$+ o_p(1)$$

$$\Rightarrow \left[\bar{\gamma}_1\int_0^1 \tilde{G}(r,\lambda)'\tilde{G}_1(r,\lambda_c)\,dr + \sigma\int_0^1 \tilde{G}(r,\lambda)'\widetilde{w}_{\bar{\alpha}}(r)\,dr, 0_{2\times k}\right], \tag{A.3}$$

uniformly in λ by Lemmas A1 and A3. Straightforward but tedious calculations give

$$s^2 \Rightarrow \sigma_e^2. \tag{A.4}$$

Direct calculation gives

$$W_T^1(T_b/T) = [\delta'\tau_1^{-1}T^{-1}\tau_1\tilde{g}(T_b^c)'M_{\tilde{Z}}\tilde{g}(T_b)\tau_1 + T^{-1/2}\tilde{e}'M_{\tilde{Z}}\tilde{g}(T_b)\tau_1]$$

$$\times [T^{-1}\tau_1\tilde{g}(T_b)'M_{\tilde{Z}}\tilde{g}(T_b)\tau_1]^{-1}$$

$$\times [T^{-1}\tau_1\tilde{g}(T_b)'M_{\tilde{Z}}\tilde{g}(T_b^c)\tau_1\tau_1^{-1}\delta + T^{-1/2}\tau_1\tilde{g}(T_b)'M_{\tilde{Z}}\tilde{e}]/s^2$$

$$\Rightarrow [\sigma_e^{-1}\bar{\delta}'L_2(\lambda,\lambda_c)' + L_1(\lambda)']L_2(\lambda,\lambda)^{-1}[\sigma_e^{-1}L_2(\lambda,\lambda_c)\bar{\delta} + L_1(\lambda)],$$

uniformly in λ by (A.1)–(A.4), Lemma A1, and algebraic manipulation. ■

Proof of Theorem 4. Let a tilde over a vector or matrix denote demeaned values and a circumflex over a function on the space [0,1] denote residuals from a projection onto the space spanned by the constant function 1. Write the DGP as $\Delta\hat{Y} = \Delta\hat{v}$, $\Delta\hat{Y} = \hat{Z}^*a + \Delta\hat{e}$ where Z^* is the $T \times (k+1)$ matrix $\{\Delta y_{t-1}, \ldots, \Delta y_{t-k-1}\}$. Let X^* denote the $T \times (k+1)$ matrix $\{\Delta v_{t-1}, \ldots, \Delta v_{t-k-1}\}$. Direct calculation gives

$$WD_T^1(T_b/T) = T^{-1/2}\Delta\hat{e}'M_{\hat{X}^*}\hat{g}_0(T_b)[T^{-1}\hat{g}_0(T_b)'M_{\hat{X}^*}\hat{g}_0(T_b)]^{-1}T^{-1/2}\hat{g}_0(T_b)'M_{\hat{X}^*}\Delta\hat{e}/s^2.$$

$[T^{-1}\tau_1\hat{g}_0(T_b)'M_{\hat{X}^*}\hat{g}_0(T_b)]^{-1} = O_p(1)$ by Lemmas A1 and A2. $T^{-1/2}\hat{g}_0(T_b)'M_{\hat{X}^*}\Delta\hat{e} = T^{-1/2}\tau_1\hat{g}_0(T_b)'M_{\hat{X}^*}B(L)\hat{e}$ where $B(L) = 1 - L$. The limit of $T^{-1/2}\hat{g}_0(T_b)'M_{\hat{X}^*}B(L)\hat{e}$ is proportional to $B(1) = 0$ and is thus zero. Therefore, the numerator of WD_T^p is $o_p(1)$. The proof is complete by showing $s^2 \Rightarrow 2\sigma_e^2$:

$$s^2 = T^{-1}\Delta\hat{e}'M_{\hat{X}^*}\Delta\hat{e} + o_p(1) = T^{-1}\Delta\hat{e}'\Delta\hat{e} + o_p(1) = T^{-1}\Sigma(\hat{e}_t - \hat{e}_{t-1})^2 + o_p(1)$$

$$\Rightarrow 2\sigma_e^2.$$ ■

Proof of Theorem 5. It is convenient to write the DGP as $\Delta\hat{Y} = \hat{Z}_1\pi + \hat{Z}_2c + \hat{e}$, or $\Delta\hat{v} = \hat{X}_1\pi + \hat{X}_2c + \hat{e}$ and note that $M_{\hat{Z}^*}\hat{Z}_2 = 0$. Direct calculation gives

$$WD_T^1(T_b/T) = T^{-1/2}(\hat{X}_1\pi + \hat{e})'M_{\hat{X}^*}\hat{g}_0(T_b)[T^{-1}\hat{g}_0(T_b)'M_{\hat{X}^*}\hat{g}_0(T_b)]^{-1}$$
$$\times T^{-1/2}\hat{g}_0(T_b)'M_{\hat{X}^*}(\hat{X}_1\pi + \hat{e})/s^2.$$

Using $\pi = -\bar{\alpha}/T$ gives

$$\begin{aligned} T^{-1/2}\hat{g}_0(T_t)'M_{\hat{X}^*}(\hat{X}_1\pi + \hat{e}) &= -T^{-1/2}\hat{g}_0(T_b)'M_{\hat{X}^*}\hat{X}_1\bar{\alpha} + T^{-1/2}\hat{g}_0(T_b)'M_{\hat{X}^*}\hat{e} \\ &= -T^{-3/2}\hat{g}_0(T_b)'\hat{X}_1\bar{\alpha} + T^{-1/2}\hat{g}_0(T_b)'\hat{e} + o_p(1) \\ &\Rightarrow \sigma_e\int_0^1 \hat{G}_0(r,\lambda)\,dw(r) - \bar{\alpha}\sigma\int_0^1 \hat{G}_0(r,\lambda)\widetilde{w}_{\bar{\alpha}}(r)\,dr \end{aligned}$$

uniformly in λ by Lemmas A1 and A3.

$$T^{-1}\hat{g}_0(T_b)'M_{\hat{X}^*}\hat{g}_0(T_b) \Rightarrow \int_0^1 \hat{G}_0(r,\lambda)^2\,dr$$

uniformly in λ by Lemmas A1 and A3.

$$\begin{aligned} s^2 &= T^{-1}(\hat{X}_1\pi + \hat{e})'M_{\hat{X}^*}(\hat{X}_1\pi + \hat{e}) + o_p(1) = T^{-1}\hat{e}'M_{\hat{X}^*}\hat{e} + o_p(1) \\ &= T^{-1}\hat{e}'\hat{e} + o_p(1) \Rightarrow \sigma_e^2. \end{aligned}$$

The theorem follows by simple algebra. ■

Proof of Theorem 6. Under the local alternative (10), $\Delta\hat{Y} = T^{-1/2}g_0(T_b^c)\bar{\gamma}_1 + \Delta\hat{v} + o_p(1)$, which implies that $Z^* = X^* + o_p(1)$. Direct calculation gives

$$\begin{aligned} WD_T^1(T_b/T) = {} & T^{-1/2}(T^{-1/2}g_0(T_b^c)\bar{\gamma}_1 + \Delta\hat{v})'M_{\hat{Z}^*}\hat{g}_0(T_b) \\ & \times [T^{-1}\hat{g}_0(T_b)'M_{\hat{Z}^*}\hat{g}_0(T_b)]^{-1} \\ & \times T^{-1/2}\hat{g}_0(T_b)'M_{\hat{Z}^*}(T^{-1/2}g_0(T_b^c)\bar{\gamma}_1 + \Delta\hat{v})/s^2. \end{aligned}$$

$$T^{-1}\hat{g}_0(T_b)'M_{\hat{Z}^*}\hat{g}_0(T_b) \Rightarrow \int_0^1 \hat{G}_0(r,\lambda)^2\,dr$$

uniformly in λ by Lemmas A1 and A3.

$$\begin{aligned} T^{-1/2}\hat{g}_0(T_b)'M_{\hat{Z}^*}(T^{-1/2}g_0(T_b^c)\bar{\gamma}_1 + \Delta\hat{v}) &= T^{-1}\hat{g}_0(T_b)'M_{\hat{Z}^*}g_0(T_b^c)\bar{\gamma}_1 \\ &\quad + T^{-1/2}\hat{g}_0(T_b)'M_{\hat{Z}^*}(\hat{X}_1\pi + \hat{e}) + o_p(1) \\ &\Rightarrow \bar{\gamma}_1\int_0^1 \hat{G}_0(r,\lambda)\hat{G}_0(r,\lambda_c)\,dr \\ &\quad + \sigma_e\int_0^1 \hat{G}_0(r,\lambda)\,dw(r) \\ &\quad - \bar{\alpha}\sigma\int_0^1 \hat{G}_0(r,\lambda)\widetilde{w}_{\bar{\alpha}}(r)\,dr, \end{aligned}$$

uniformly in λ by Lemmas A1 and A3. Straightforward algebra gives $s^2 \Rightarrow \sigma_e^2$, and algebraic manipulation completes the proof. ■

LINKED CITATIONS

You have printed the following article:

Wald-Type Tests for Detecting Breaks in the Trend Function of a Dynamic Time Series
Timothy J. Vogelsang
Econometric Theory, Vol. 13, No. 6. (Dec., 1997), pp. 818-849.
Stable URL:
http://links.jstor.org/sici?sici=0266-4666%28199712%2913%3A6%3C818%3AWTFDBI%3E2.0.CO%3B2-U

References

Tests for Parameter Instability and Structural Change With Unknown Change Point
Donald W. K. Andrews
Econometrica, Vol. 61, No. 4. (Jul., 1993), pp. 821-856.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199307%2961%3A4%3C821%3ATFPIAS%3E2.0.CO%3B2-I

Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative
Donald W. K. Andrews; Werner Ploberger
Econometrica, Vol. 62, No. 6. (Nov., 1994), pp. 1383-1414.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199411%2962%3A6%3C1383%3AOTWANP%3E2.0.CO%3B2-7

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

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

LINKED CITATIONS

A Direct Test for Changing Trend
Chia-Shang James Chu; Halbert White
Journal of Business & Economic Statistics, Vol. 10, No. 3. (Jul., 1992), pp. 289-299.
Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28199207%2910%3A3%3C289%3AADTFCT%3E2.0.CO%3B2-P

On Detecting Changes in the Mean of Normal Variates
L. A. Gardner, Jr.
The Annals of Mathematical Statistics, Vol. 40, No. 1. (Feb., 1969), pp. 116-126.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196902%2940%3A1%3C116%3AODCITM%3E2.0.CO%3B2-3

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

Tests for Parameter Instability in Regressions with I(1) Processes
Bruce E. Hansen
Journal of Business & Economic Statistics, Vol. 10, No. 3. (Jul., 1992), pp. 321-335.
Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28199207%2910%3A3%3C321%3ATFPIIR%3E2.0.CO%3B2-K

Convergence to Stochastic Integrals for Dependent Heterogeneous Processes
Bruce E. Hansen
Econometric Theory, Vol. 8, No. 4. (Dec., 1992), pp. 489-500.
Stable URL:
http://links.jstor.org/sici?sici=0266-4666%28199212%298%3A4%3C489%3ACTSIFD%3E2.0.CO%3B2-%23

The Likelihood Ratio Test for a Change-Point in Simple Linear Regression
Hyune-Ju Kim; David Siegmund
Biometrika, Vol. 76, No. 3. (Sep., 1989), pp. 409-423.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28198909%2976%3A3%3C409%3ATLRTFA%3E2.0.CO%3B2-5

LINKED CITATIONS

A Multicountry Characterization of the Nonstationarity of Aggregate Output
Roger C. Kormendi; Philip Meguire
Journal of Money, Credit and Banking, Vol. 22, No. 1. (Feb., 1990), pp. 77-93.
Stable URL:
http://links.jstor.org/sici?sici=0022-2879%28199002%2922%3A1%3C77%3AAMCOTN%3E2.0.CO%3B2-S

Testing for Structural Change in Dynamic Models
Walter Krämer; Werner Ploberger; Raimund Alt
Econometrica, Vol. 56, No. 6. (Nov., 1988), pp. 1355-1369.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198811%2956%3A6%3C1355%3ATFSCID%3E2.0.CO%3B2-4

Properties of Sequences of Partial Sums of Polynomial Regression Residuals with Applications to Tests for Change of Regression at Unknown Times
Ian B. MacNeill
The Annals of Statistics, Vol. 6, No. 2. (Mar., 1978), pp. 422-433.
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28197803%296%3A2%3C422%3APOSOPS%3E2.0.CO%3B2-9

Spurious Periodicity in Inappropriately Detrended Time Series
Charles R. Nelson; Heejoon Kang
Econometrica, Vol. 49, No. 3. (May, 1981), pp. 741-751.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198105%2949%3A3%3C741%3ASPIIDT%3E2.0.CO%3B2-Y

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

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

LINKED CITATIONS

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

Nonstationarity and Level Shifts with an Application to Purchasing Power Parity
Pierre Perron; Timothy J. Vogelsang
Journal of Business & Economic Statistics, Vol. 10, No. 3. (Jul., 1992), pp. 301-320.
Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28199207%2910%3A3%3C301%3ANALSWA%3E2.0.CO%3B2-L

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

Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes
Richard E. Quandt
Journal of the American Statistical Association, Vol. 55, No. 290. (Jun., 1960), pp. 324-330.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28196006%2955%3A290%3C324%3ATOTHTA%3E2.0.CO%3B2-%23