Inference on Unit Roots and Trend Breaks in Macroeconomic Time Series

by

Christian J. Murray
University of Houston

and

Eric Zivot
University of Washington

This version:

September 3, 1998

Correspondence:

Christian J. Murray
Department of Economics
University of Houston
Houston, TX 77204-5882

Tel: (713) 743-3835
Fax: (713) 743-3798
Email: cjmurray@uh.edu

1. Introduction

The most damaging criticism of the hypothesis advanced by Nelson and Plosser (1982), that U.S. output contains a unit root, has come through the allowance of structural change under the alternative hypothesis of trend stationarity. This was originally due to Perron (1989) and Rappoport and Reichlin (1989) who argued that Nelson and Plosser had overstated the frequency of permanent shocks by failing to allow for the possibility of a one time structural change. Perron showed that the real GNP series used by Nelson and Plosser is no longer consistent with the unit root hypothesis if a change in level, occurring at 1929, is considered. Perron's conclusion is that from 1909 to 1970, there is only one permanent shock, a negative one, and the rest of the variation in output is transitory around a time trend.

In Perron (1989), the date of the trend break, 1929, was assumed to be known *a priori*. This drew criticism originally from Christiano (1992) who suggested that Perron's results may be tainted by the assumption that the break date was known. Using a bootstrap procedure, he demonstrated that if the break date is allowed to be data dependent, then the critical values are much larger (in absolute value) than those tabulated by Perron. Zivot and Andrews (1992) and Banerjee et al. (1992) derived the limiting distribution of the unit root statistic when the break date is endogenized. Zivot and Andrews (1992) demonstrate that Perron's conclusion that U.S. GDP is stationary around a broken time trend still holds once critical values are adjusted to reflect estimation of the break date.

Since Perron, the literature has been flooded by papers which study the asymptotic distribution of unit root and/or trend break statistics under various methods for selecting the break date. This paper adds to the literature by deriving the asymptotic distribution of statistics on structural change coefficients, as well as statistics testing the joint null hypothesis of a unit root and no structural change. The latter potentially offer an increase in power over statistics which just test the unit root null. We then apply our results to the Maddison (1995) annual U.S. real GDP series, and post-war quarterly chained U.S. real GDP.

This paper is organized as follows. Section 2 reviews the literature on testing for unit roots and trend breaks. Section 3 presents the test statistics and derives their asymptotic distributions. Section 4 analyzes finite sample size and power. Section 5 applies our results to U.S. GDP. Section 6 summarizes and offers concluding remarks.

2. Testing for Unit Roots and Trend Breaks: A Brief Review of the Literature

Scattered throughout the literature is a plethora of results on the asymptotic distribution of unit root and structural change statistics when the break date is endogenized. In this section, we review these results for models which allow for (at most) one break in trend, and point out what has yet to be done. We divide the cases into trending and non-trending data.

2.1 Non-trending data

For non-trending data, the null hypothesis is a driftless unit root process with or without break, and the alternative is a stationary process with a one time change in mean. There are two methods of modeling trend breaks in the literature. The additive outlier (AO) approach models the break as an abrupt change, while the innovational outlier (IO) approach allows the break to occur gradually. Since most of the empirical work has used the (IO) approach, we concentrate on this method. For a detailed discussion of modeling innovational and additive outliers, the reader is referred to Vogelsang and Perron (1994). In general, all statistics for non-trending data are asymptotically invariant to a mean shift under the null hypothesis. We thus present the following unit root null hypothesis without a break in level:

$$\mathrm{H}_0 : \; y_t = y_{t-1} + u_t \tag{1}$$

where $u_t = \psi^*(L)e_t$; $\psi(L) = (1-\rho L)\psi^*(L)$; $\psi^*(L) = A^*(L)^{-1}B(L)$; $e_t \sim iid(0,\sigma^2)$ and $A^*(L)$ and $B(L)$ are p^{th} and q^{th} order lag polynomials with roots strictly outside the unit circle. Under the null hypothesis, $\rho = 1$ and $\theta = 0$. The alternative hypothesis allows for a one time change in mean of a stationary process ($\rho < 1$) and is as follows:

$$\mathrm{H}_1 : \; y_t = a + \psi(L)(\theta DU(T_B)_t + \mathrm{e}_t), \tag{2}$$

where $DU(T_B)_t = 1$ if $t > T_B$ and 0 otherwise; $DU(T_B)_t$ is the "step dummy" capturing a level shift at time T_B (the break date). Under the alternative θ represents the immediate change in mean and $\psi(1)\theta$ represents the long run change in mean. For this model, the test regression is:

$$y_t = \hat{a} + \hat{\theta} DU(\hat{T}_B)_t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{c}_i \Delta y_{t-1} + \hat{e}_t , \tag{3}$$

where $\hat{T}_B$ is the estimated break date and $\hat{k}$ is the estimated lag length.

Following the theoretical treatment of Said and Dickey (1984), *k* lagged differences are included in the regression equation to account for serial correlation in the innovation sequence. Said and Dickey prove that if *k* diverges as *T* diverges, but at a slower rate, then the asymptotic distribution of the ADF test is unaffected. For unit root tests which allow for structural change at an unknown point, there does not exist a proof that such a result is valid. In subsequent theoretical derivations, we shall assume that the errors are iid which simplifies the presentation of the results. We follow Zivot and Andrews (1992) and conjecture that adding *k* lags to the regression will correct for serial correlation.

As another matter, the correct number of lagged terms to include in the regression equation is unknown and must be chosen by the researcher. Choosing *k* too small results in a size bias, while choosing *k* too large results in a loss of power. In practice, certain data dependent methods for selecting *k* lead to an increase in power over fixing *k* as in Said and Dickey (1984) (unless of course you happen to choose the correct value of *k*). For standard ADF regressions, Hall (1994) proves that a number of such data based procedures leave the asymptotic distribution of the unit root statistic unaffected when the error terms follow a pure AR(p) process. Ng and Perron (1995) extend Hall's results to the ARMA(p,q) case. Among the methods analyzed are a general to specific (GS) strategy and the Schwartz information criterion (SIC). As long as the maximum lag in the selection set is allowed to grow appropriately with the sample size, both methods are shown to have zero probability of underfitting as the sample size diverges. This implies that the asymptotic critical values, which assume that *k* is known, are valid under such data dependent methods for selecting *k*. While such a result is likely to hold for unit root tests with structural change at an unknown point, a proof is apt to be quite involved. Again we conjecture that such a result exists, and in the subsequent empirical application, we shall employ both GS and SIC.

It should also be noted that for a particular regression, the lag length and break date are determined simultaneously. This will influence the finite sample performance of the test statistics. The appropriate method used to choose T_B is context specific. If rejection of the unit root hypothesis is desired, then the break date should be that which minimizes the unit root statistic. However, if one is just concerned with the dating of structural change, then choosing the break date to maximize some function of $\hat{\theta}$ is appropriate.

Table 1 presents the relevant statistics from regression (3), and their origin. Blank spaces indicate what has yet to be done. Perron and Vogelsang (1992) derive the asymptotic distribution of the unit root statistic where the break date is chosen to minimize the unit root statistic. This is denoted as $\inf t_\rho$. They demonstrate that the AO approach is asymptotically equivalent to the IO approach. They also consider the distribution of the unit root statistic when the break is chosen to minimize the one sided t-test of no structural change. This statistic is denoted as $t_{\rho,\inf(\theta)}$. In general, when a dummy variable statistic is used to choose the break date, the asymptotic equivalence of unit root statistics between the AO and IO approaches does not hold.

Table 1. Non-trending data

$$y_t = \hat{a} + \hat{\theta} DU(\hat{T}_B)_t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{c}_i \Delta y_{t-1} + \hat{e}_t$$

Statistic	Origin
$\inf t_p$	PV: AO/IO
$t_{\rho,\inf(\theta)}$	PV: AO/IO
$Wald_{\theta,\inf(\rho)}$	
$\sup Wald_\theta$	V: AO - P: AI/IO
$\sup Wald_{\theta,\rho}$	

PV is Perron and Vogelsang (1992). V is Vogelsang (1997)

Incorporating *a priori* knowledge of the sign of the break date can lead to an increase in power. Perron and Vogelsang demonstrate that $t_{\rho,\inf(\theta)}$ has greater power than $\inf t_\rho$ when the break date is negative. A similar result holds for $t_{\rho,\sup(\theta)}$ when $\theta > 0$.

The literature also contains some distributional results for tests statistics concerning structural change coefficients. Perron and Vogelsang (1992) derive the asymptotic distribution of the mean shift statistic, but critical values are not reported. Vogelsang (1997), modeling the break as an additive outlier, derives the asymptotic distribution of the mean-Wald, exp-Wald, and sup-Wald tests of the no structural change null for *I(0)* and *I(1)* data. The mean-Wald and exp-Wald tests cannot be used to estimate the break date, whereas the sup-Wald test can. These are extensions of the tests considered by Andrews (1993), and optimal tests considered by Andrews and Ploberger (1994) for deterministically and stochastically trending data. However, the optimality properties do not carry over to trending or integrated data. Vogelsang just considers the 2-sided Wald test that $\theta = 0$. Also of interest are the 1-sided t-tests that $\theta = 0$, which may lead to higher power if the sign of the break date is known *a priori*.

Two statistics in this context have not yet been computed. The first concerns inference on θ when T_B is chosen to minimize t_ρ. Second is the Wald test of the joint null that $\rho = 1$ and $\theta = 0$. We denote this statistic as $\sup Wald_{\theta,\rho}$. This may offer an increase in power over the $\inf t_\rho$ and $\sup Wald_\theta$ statistics which do not explicitly test a subset of the null hypothesis.

2.2 Trending data

For trending data, three different alternative hypotheses have been considered. The first, labeled Model A by Perron (1989) allows for a change in level under the alternative hypothesis. Model B allows for a change in the growth rate under the alternative, and Model C allows for both types of structural change. In general, all statistics for trending data are asymptotically invariant to a level shift under the null, but not to a change in slope. Thus, statistics for Model B and Model C will have different limiting distributions depending on whether a change in growth is allowed under the null. However, as pointed out by Vogelsang and Perron (1994), for changes in growth of the size typically encountered in practice, the no break asymptotics provide a better approximation to the finite sample distribution of the unit root statistics. We will thus present the models without a change in level or growth under the null.

All three models have the common null hypothesis:

$$\mathrm{H}_0 : \; y_t = \mu + y_{t-1} + u_t \, , \tag{4}$$

where $\{u_t\}$ obeys the restrictions in (1). The three alternative hypotheses can be written as follows:

$$H_1^A : \; y_t = a + \mu t + \psi(L)(\theta DU(T_B)_t + e_t) \, , \tag{5}$$

$$H_1^B : \; y_t = a + \mu t + \psi(L)(\gamma DT(T_B)_t + e_t) \, , \tag{6}$$

and

$$H_1^C : \; y_t = a + \mu t + \psi(L)(\theta DU(T_B)_t + \gamma DT(T_B)_t + e_t) \, . \tag{7}$$

The "ramp" dummy $DT(T_B)_t$ is $t - T_B$ if $t > T_B$ and 0 otherwise, γ is the immediate change in growth allowed under the latter two alternatives, and $\psi(1)\gamma$ is the long run change. The corresponding test regressions are:

$$y_t = \hat{a} + \hat{\theta} DU(\hat{T}_B)_t + \hat{b}t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{c}_i \Delta y_{t-1} + \hat{e}_t \, , \tag{8}$$

$$y_t = \hat{a} + \hat{\gamma} DT(\hat{T}_B)_t + \hat{b}t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{c}_i \Delta y_{t-1} + \hat{e}_t \, , \tag{9}$$

and

$$y_t = \hat{a} + \hat{\theta} DU(\hat{T}_B)_t + \hat{\gamma} DT(\hat{T}_B)_t + \hat{b}t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{c}_i \Delta y_{t-1} + \hat{e}_t \, . \tag{10}$$

For the sake of clarity, we shall discuss the origin of the statistics for all three models separately. Table 2A contains a description of the Model A statistics.

Table 2A. Trending data - Model A

$$y_t = \hat{a} + \hat{\theta} DU(\hat{T}_B)_t + \hat{b}t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{c}_i \Delta y_{t-1} + \hat{e}_t$$

Statistic	Origin
$\inf t_\rho^A$	ZA: IO - BLS: IO VP: AO - P: IO
$t^A_{\rho,Wald(\theta)}$	BLS: IO VP: AO - P: IO
$Wald^A_{\theta,\inf(\rho)}$	
$\sup Wald^A_\theta$	BLS: IO
$\sup Wald^A_{\theta,\rho}$	

ZA is Zivot and Andrews (1992). BLS is Banerjee, Lumsdaine, and Stock (1992). VP is Vogelsang and Perron (1994). P is Perron (1997).

Choosing the break date to minimize the unit root statistic, Zivot and Andrews (1992), Banerjee et al. (1992), and Perron (1997) derive the distribution of unit root statistic, $\inf t_{\rho}$, under the no break null with innovational outliers. Modeling the break as an additive outlier, Vogelsang and Perron (1994) derive $\inf t_{\rho}^{A}$ with no break under the null. Vogelsang and Perron (1994) and Perron (1997) also derive the unit root statistic when T_B is chosen via a statistic on $\hat{\theta}$. We will generically refer to this as $t^{A}_{\rho,Wald(\theta)}$, even though the break date is usually chosen to maximize or minimize the 1-sided t-test that θ = 0.

Banerjee et al. (1992) derive the Wald test that $\theta = 0$, denoted $\sup Wald_{\theta}^{A}$. Also of interest are the 1-sided t-tests of the same hypothesis.

As in the case of non-trending data, neither $Wald^{A}_{\theta,\inf(\rho)}$ nor $\sup Wald^{A}_{\theta,\rho}$ have yet been considered. The former is appropriate when one performs the Zivot-Andrews unit root test, and then wishes to perform inference on θ, conditional on the chosen break date. As mentioned before, the latter may offer an increase in power over either $\inf t_{\rho}^{A}$ or $\sup Wald_{\theta}^{A}$.

Table 2B presents analogous results for Model B. Since they basically mirror Table 2A, we forgo a discussion.

Table 2B. Trending data - Model B

$$y_t = \hat{a} + \hat{\gamma} DT(\hat{T}_B)_t + \hat{b}t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{c}_i \Delta y_{t-1} + \hat{e}_t$$

Statistic	Origin
$\inf t_{\rho}^{B}$	ZA: IO - BLS: IO VP: AO - P: AO
$t^{B}_{\rho,Wald(\gamma)}$	BLS: IO VP: AO - P: AO
$Wald^{B}_{\gamma,\inf(\rho)}$	
$\sup Wald_{\gamma}^{B}$	BLS: IO
$Wald^{B}_{\gamma,\rho}$	

ZA is Zivot and Andrews (1992). BLS is Banerjee, Lumsdaine, and Stock (1992). VP is Vogelsang and Perron (1994). P is Perron (1997).

Model C results are presented in Table 3C. Given that there are 2 structural change coefficients, there are many more cases to consider. To conserve space, we shall primarily focus on what has not yet been done. Although Vogelsang derives the mean-Wald, exp-Wald, and sup-Wald tests of the hypothesis that Model C contains no level shift or a change in growth, again for both *I(0)* and *I(1)* data, the individual 1-sided and 2-sided tests are of interest. There is also the joint Wald test that $\theta = \gamma = 0$ when the break date minimizes the unit root statistic, $Wald^{C}_{\theta,\gamma,\inf(\rho)}$. Finally, there is the joint Wald test that $\rho = 1$ and $\theta = \gamma = 0$.

Table 2C. Trending data - Model C

$$y_t = \hat{a} + \hat{\theta} DU(\hat{T}_B)_t + \hat{\gamma} DT(\hat{T}_B)_t + \hat{b}t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{c}_i \Delta y_{t-1} + \hat{e}_t$$

Statistic	Origin
$\inf t^{C}_{p}$	ZA: IO VP: AO - P: IO
$t^{C}_{\rho,Wald(\gamma)}$	VP: AO - P: AO
$Wald^{C}_{\theta,\gamma,\inf(\rho)}$	
$\sup Wald^{C}_{\theta}$	
$\sup Wald^{C}_{\gamma}$	
$\sup Wald^{C}_{\theta,\gamma}$	V: AO
$\sup Wald^{C}_{\theta,\gamma,\rho}$	

In Section 3, we shall catalog the distributions of test statistics for no structural change, and derive the distributions of test statistics for the joint null hypothesis that there is a unit root without a break in trend.

3. Asymptotic Distribution of the Test Statistics

In this section, we derive the asymptotic distributions of structural change statistics, as well as the joint distributions of statistics concerning the largest autoregressive root and structural change coefficients. The latter potentially offer a gain in power over tests which do not explicitly test the unit root hypothesis. In the theorems to follow, we restrict the innovation sequence to be iid, but the results remain valid in the presence of

ARMA(p,q) errors as long as k lagged difference terms are included in the regression. We consider non-trending and trending data separately.

3.1 Non-trending data

Following Zivot and Andrews (1992) and Banerjee et al. (1992) we specify a no break null hypothesis and innovational outliers. Recall the null hypothesis and test regression:

$$\mathrm{H}_0 : y_t = y_{t-1} + u_t , \tag{1'}$$

and

$$y_t = \hat{a} + \hat{\theta} DU(\hat{T}_B)_t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{c}_i \Delta y_{t-1} + \hat{e}_t \ . \tag{3'}$$

Let $\lambda \equiv \frac{T_B}{T}$ be the break fraction. For all the results which follow, we assume that λ remains constant as $T \to \infty$.

We first consider 4 different statistics to test the null hypothesis that θ = 0. Let $\sup Wald_\theta$ and $\sup t_{|\theta|}$ be the 2-sided tests where λ is chosen to maximize the Wald statistic, and the absolute value of the t-statistic respectively. Also, let $\sup t_\theta$ and $\inf t_\theta$ be the 1-sided tests which maximize and minimize the t-statistic respectively. The latter should be used if one has *a priori* knowledge of the sign of θ.

Following Zivot and Andrews, we can characterize the asymptotic distributions of these statistics in terms of projection residuals. Let $DU^*(\lambda, r)$ be the projection residual from the continuous time regression:

$$DU(\lambda, r) = \hat{\alpha}_0 + \hat{\alpha}_1 W(r) + DU^*(\lambda, r);$$

where $DU(\lambda, r) = 1$ if $r > \lambda$ and 0 otherwise, and $W(r)$ is standard Browning motion. That is, $\hat{\alpha}_0$ and $\hat{\alpha}_1$ solve

$$\min_{\alpha} \int_0^1 \left| DU(\lambda, r) - \alpha_0 - \alpha_1 W(r) \right|^2 dr \ .$$

Since we are considering 4 test statistics of the null that θ = 0, it is helpful to introduce some simplifying notation. Let $g_1(x) = x^2$, $g_2(x) = |x|$, and $g_3(x) = x$. Also, let $t_\theta(\lambda)$ denote the t-test for $\theta = 0$ as a function of the break fraction λ. For example $g_1[t_\theta(\lambda)]$ corresponds to the Wald test that $\theta = 0$. We then have the following theorem.

Theorem 3.1.A. Let $\{y_t\}$ be generated under the null hypothesis (1) and let $\{u_t\}$ be iid, mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of (0,1). Then,

$$\sup_{\lambda \in \Lambda} g_i\left[t_\theta(\lambda)\right] \Rightarrow \sup_{\lambda \in \Lambda} g_i\left[\left(\int_0^1 DU^*(\lambda,r)^2\,dr\right)^{-1/2}\left(\int_0^1 DU^*(\lambda,r)dW(r)\right)\right],$$

for i=1,2,3, and

$$\inf_{\lambda \in \Lambda} t_\theta(\lambda) \Rightarrow \inf_{\lambda \in \Lambda} \left(\int_0^1 DU^*(\lambda,r)^2\,dr\right)^{-1/2}\left(\int_0^1 DU^*(\lambda,r)dW(r)\right)$$

as $T \to \infty$, where $\Rightarrow$ denotes weak convergence in distribution in the sense of Billingsley (1968). The proof of this theorem proceeds along the lines of Zivot and Andrews (1992) and is therefore omitted.

We also consider the distribution of the step dummy t-statistic when the break date is chosen to minimize the unit root statistic. This is useful in circumstances where the unit root statistic is calculated as in Zivot and Andrews (1992), and then one wants to perform inference on θ. Following is the distribution of the Wald test for $\theta = 0$, choosing the break date to minimize t_ρ. We denote this statistic as $Wald_{\theta,\inf(\rho)}$.

Theorem 3.1.B. Let $\{y_t\}$ be generated under the null hypothesis (1) and let $\{u_t\}$ be iid, mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of (0,1). Then,

$$Wald_{\theta,\inf(\rho)}(\lambda) \Rightarrow \left(\int_0^1 DU^*(\lambda^*,r)^2\,dr\right)^{-1}\left(\int_0^1 DU^*(\lambda^*,r)dW(r)\right)^2$$

where

$$\lambda^* = \underset{\lambda \in \Lambda}{\arg\min} \left(\int_0^1 W^*(\lambda, r)^2 \, dr \right)^{-1/2} \left(\int_0^1 W^*(\lambda, r) dW(r) \right)$$

and the last term is the Perron and Vogelsang (1992) unit root statistic.

We now turn to the distribution of the Wald test of the null that $\theta = 0$ and $\rho = 1$. Let $\sup Wald_{\theta,\rho}$ denote this test statistic. Let $X_1(\lambda, r)' = (DU(\lambda, r), W(r))$ and $X_2(r) = 1$. Then $X^*{}_1(\lambda, r)$ is the projection residual from the continuous time regression which minimizes:

$$\min_{\alpha_0} \int_0^1 \left\| X_1(\lambda, r) - \alpha_0 X_2(r) \right\|^2 dr \, .$$

We then have the following result:

Theorem 3.1.C. Let $\{y_t\}$ be generated under the null hypothesis (1) and let $\{u_t\}$ be iid, mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of (0,1). Then,

$$\sup_{\lambda \in \Lambda} Wald_{\theta,\rho}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 X_1^*(\lambda, r) dW(r) \right)' \left(\int_0^1 X_1^*(\lambda, r) X_1^*(\lambda, r)' dr \right)^{-1} \left(\int_0^1 X_1^*(\lambda, r) dW(r) \right).$$

The asymptotic critical values for the statistics in Theorems 3.1.A – 3.1.C are presented in Table 3. The first row corresponds to the sup-Wald test analyzed by Vogelsang (1997) when the data are integrated. To simulate the asymptotic critical values, we generated driftless random walks with N(0,1) errors using the GAUSS rndn function. We set the sample size at 1000 and calculated the finite sample versions of the terms in Theorem 3.1. This was repeated this 50,000 times. An upper bound on the standard errors of the critical values is 0.0022.

3.2 Trending data

We now turn to the analysis of trending data. Recall the null hypothesis and test regressions:

$$\mathrm{H}_0 : \; y_t = \mu + y_{t-1} + u_t \, , \tag{4'}$$

$$y_t = \hat{a} + \hat{\theta} DU(\hat{T}_B)_t + \hat{b}t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{c}_i \Delta y_{t-1} + \hat{e}_t \ , \tag{8'}$$

$$y_t = \hat{a} + \hat{\gamma} DT(\hat{T}_B)_t + \hat{b}t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{c}_i \Delta y_{t-1} + \hat{e}_t \ , \tag{9'}$$

and

$$y_t = \hat{a} + \hat{\theta} DU(\hat{T}_B)_t + \hat{\gamma} DT(\hat{T}_B)_t + \hat{b}t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{c}_i \Delta y_{t-1} + \hat{e}_t \ . \tag{10'}$$

We first consider individual 1-sided and 2-sided tests for in Models A, B, and C. Let $DU^A(\lambda, r)$, $DT^B(\lambda, r)$, $DU^C(\lambda, r)$, and $DT^C(\lambda, r)$ be the projection residuals from the following continuous time regressions:

$$DU(\lambda, r) = \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 W(r) + DU^A(\lambda, r) \ ,$$

$$DT(\lambda, r) = \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 W(r) + DT^B(\lambda, r) \ ,$$

$$DU(\lambda, r) = \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 DT(\lambda, r) + \hat{\alpha}_3 W(r) + DU^C(\lambda, r) \ ,$$

and

$$DT(\lambda, r) = \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 DU(\lambda, r) + \hat{\alpha}_3 W(r) + DT^C(\lambda, r)$$

respectively; where $DT(\lambda, r) = r - \lambda$ if $r > \lambda$ and 0 otherwise. Letting $t_\theta^A(\lambda)$, $t_\gamma^B(\lambda)$, $t_\theta^C(\lambda)$, and $t_\gamma^C(\lambda)$ denote 4 t-statistics under consideration, we have the following result.

Theorem 3.2.A. Let $\{y_t\}$ be generated under the null hypothesis (4) and let $\{u_t\}$ be iid, mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of (0,1). Then,

$$\sup_{\lambda \in \Lambda} \ g_i \left[t_\theta^A(\lambda) \right] \Rightarrow \sup_{\lambda \in \Lambda} \ g_i \left[\left(\int_0^1 DU^A(\lambda, r)^2 \, dr \right)^{-1/2} \left(\int_0^1 DU^A(\lambda, r) dW(r) \right) \right],$$

$$\inf_{\lambda \in \Lambda} t_\theta^A(\lambda) \Rightarrow \inf_{\lambda \in \Lambda} \left(\int_0^1 DU^A(\lambda, r)^2 \, dr \right)^{-1/2} \left(\int_0^1 DU^A(\lambda, r) dW(r) \right),$$

$$\sup_{\lambda \in \Lambda} g_i \left[t_\gamma^B(\lambda) \right] \Rightarrow \sup_{\lambda \in \Lambda} g_i \left[\left(\int_0^1 DT^B(\lambda, r)^2 \, dr \right)^{-1/2} \left(\int_0^1 DT^B(\lambda, r) dW(r) \right) \right],$$

$$\inf_{\lambda \in \Lambda} t_\gamma^B(\lambda) \Rightarrow \inf_{\lambda \in \Lambda} \left(\int_0^1 DT^B(\lambda, r)^2 \, dr \right)^{-1/2} \left(\int_0^1 DT^B(\lambda, r) dW(r) \right),$$

$$\sup_{\lambda \in \Lambda} g_i \left[t_\theta^C(\lambda) \right] \Rightarrow \sup_{\lambda \in \Lambda} g_i \left[\left(\int_0^1 DU^C(\lambda, r)^2 \, dr \right)^{-1/2} \left(\int_0^1 DU^C(\lambda, r) dW(r) \right) \right],$$

$$\inf_{\lambda \in \Lambda} t_\theta^C(\lambda) \Rightarrow \inf_{\lambda \in \Lambda} \left(\int_0^1 DU^C(\lambda, r)^2 \, dr \right)^{-1/2} \left(\int_0^1 DU^C(\lambda, r) dW(r) \right),$$

$$\sup_{\lambda \in \Lambda} g_i \left[t_\gamma^C(\lambda) \right] \Rightarrow \sup_{\lambda \in \Lambda} g_i \left[\left(\int_0^1 DT^C(\lambda, r)^2 \, dr \right)^{-1/2} \left(\int_0^1 DT^C(\lambda, r) dW(r) \right) \right],$$

and

$$\inf_{\lambda \in \Lambda} t_\gamma^C(\lambda) \Rightarrow \inf_{\lambda \in \Lambda} \left(\int_0^1 DT^C(\lambda, r)^2 \, dr \right)^{-1/2} \left(\int_0^1 DT^C(\lambda, r) dW(r) \right),$$

for $i = 1,2,3$ and $g_i(\cdot)$ as in Section 3.1.

Critical values for these statistics are presented in Tables 4, 5, and 6 for Models A, B, and C respectively. The first rows in Tables 4 and 5 correspond to the F-statistics computed by Banerjee et al. (1992).

We next consider the Wald test for the joint hypothesis that there is neither a level shift nor a change in growth in Model C ($\theta = \gamma = 0$). Denote this statistic as $\sup Wald_{\theta,\gamma}^{C}$.

Let $D^{C}(\lambda,r)' = (DU(\lambda,r), DT(\lambda,r))$ and $X_2^{C}(r)' = (1, r, W(r))$. Then $D^{C*}(\lambda,r)$ is the projection residual from the continuous time regression which minimizes:

$$\min_{\alpha} \int_0^1 \left\| D^{C}(\lambda,r) - \alpha X_2^{C}(r) \right\|^2 dr .$$

We then have the following result.

Theorem 3.2.D. Let $\{y_t\}$ be generated under the null hypothesis (4) and let $\{u_t\}$ be iid, mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of (0,1). Then,

$$\sup_{\lambda \in \Lambda} Wald_{\theta,\gamma}^{C}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 D^{C*}(\lambda,r)dW(r) \right)' \left(\int_0^1 D^{C*}(\lambda,r) D^{C*}(\lambda,r)' dr \right)^{-1} \left(\int_0^1 D^{C*}(\lambda,r)dW(r) \right).$$

Critical values for this statistic are in the 9th row of Table 6 and correspond to the sup-Wald test analyzed by Vogelsang (1997) for *I(1)* data.

Also reported in the 5th rows of Tables 4 and 5 and the 10th row of Table 6 are $\sup Wald_{\theta,\inf(\rho)}^{A}$, $\sup Wald_{\gamma,\inf(\rho)}^{B}$, and $\sup Wald_{\theta,\gamma,\inf(\rho)}^{C}$, the Wald tests of no structural change when the break is chosen to minimize the unit root statistic. These distributions are derived in the following theorem.

Theorem 3.2.E. Let $\{y_t\}$ be generated under the null hypothesis (4) and let $\{u_t\}$ be iid, mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of (0,1). Then,

$$Wald^{A}_{\theta,\inf(\rho)}(\lambda) \Rightarrow \left(\int_0^1 DU^{A}(\lambda^{A},r)^2\,dr\right)^{-1}\left(\int_0^1 DU^{A}(\lambda^{A},r)dW(r)\right)^2,$$

$$Wald^{B}_{\gamma,\inf(\rho)}(\lambda) \Rightarrow \left(\int_0^1 DT^{B}(\lambda^{B},r)^2\,dr\right)^{-1}\left(\int_0^1 DT^{B}(\lambda^{B},r)dW(r)\right)^2,$$

and

$$Wald^{C}_{\theta,\gamma,\inf(\rho)}(\lambda) \Rightarrow \left(\int_0^1 D^{C*}(\lambda^{C},r)^2\,dr\right)^{-1}\left(\int_0^1 D^{C*}(\lambda^{C},r)dW(r)\right)^2,$$

where

$$\lambda^{A} = \underset{\lambda\in\Lambda}{\arg\min}\left(\int_0^1 W^{A}(\lambda,r)^2\,dr\right)^{-1/2}\left(\int_0^1 W^{A}(\lambda,r)dW(r)\right),$$

$$\lambda^{B} = \underset{\lambda\in\Lambda}{\arg\min}\left(\int_0^1 W^{B}(\lambda,r)^2\,dr\right)^{-1/2}\left(\int_0^1 W^{B}(\lambda,r)dW(r)\right),$$

and

$$\lambda^{C} = \underset{\lambda\in\Lambda}{\arg\min}\left(\int_0^1 W^{C}(\lambda,r)^2\,dr\right)^{-1/2}\left(\int_0^1 W^{C}(\lambda,r)dW(r)\right),$$

where the last three terms are the Zivot-Andrews unit root statistics for Models A, B, and C respectively.

We conclude this section by deriving the limiting distributions for $\sup Wald^{A}_{\theta,\rho}$, $\sup Wald^{B}_{\gamma,\rho}$, and $\sup Wald^{C}_{\theta,\gamma,\rho}$; the tests of the joint null hypothesis of a unit root and no structural change. Let $X_1^{A}(\lambda,r)' = (DU(\lambda,r),W(r))$, $X_1^{B}(\lambda,r)' = (DT(\lambda,r),W(r))$, $X_1^{C}(\lambda,r)' = (DU(\lambda,r)DT(\lambda,r),W(r))$, and $X_2^{A}(r)' = X_2^{B}(r)' = X_2^{C}(r)' = (1,r)$. Let $X_1^{j*}(\lambda,r)$ denote the projection residual from the continuous time regression of $X_1^{j}(\lambda,r)$ on $X_2^{j}(r)$ for *j* = 1,2,3. We then have the following.

Theorem 3.2.F. Let $\{y_t\}$ be generated under the null hypothesis (4) and let $\{u_t\}$ be iid, mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of (0,1). Then,

$$\sup_{\lambda \in \Lambda} Wald^{A}_{\theta,\rho}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 X_1^{A*}(\lambda,r)dW(r) \right)' \left(\int_0^1 X_1^{A*}(\lambda,r)X_1^{A*}(\lambda,r)'dr \right)^{-1} \left(\int_0^1 X_1^{A*}(\lambda,r)dW(r) \right),$$

$$\sup_{\lambda \in \Lambda} Wald^{B}_{\gamma,\rho}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 X_1^{B*}(\lambda,r)dW(r) \right)' \left(\int_0^1 X_1^{B*}(\lambda,r)X_1^{B*}(\lambda,r)'dr \right)^{-1} \left(\int_0^1 X_1^{B*}(\lambda,r)dW(r) \right),$$

and

$$\sup_{\lambda \in \Lambda} Wald^{C}_{\theta,\gamma,\rho}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 X_1^{C*}(\lambda,r)dW(r) \right)' \left(\int_0^1 X_1^{C*}(\lambda,r)X_1^{C*}(\lambda,r)'dr \right)^{-1} \left(\int_0^1 X_1^{C*}(\lambda,r)dW(r) \right).$$

Critical values for these are reported in the final rows of Tables 4 - 6.

4. Finite Sample Size and Power

In this section, we ascertain the finite sample properties of the statistics presented in Section 3, in terms of size and power. Table 7 presents the empirical size of selected test statistics at the 5% and 10% nominal significance levels for sample sizes T=100 and 200. Since all statistics are asymptotically invariant to μ, we set it to zero. The number of iterations is 5,000 for all simulations to follow.

Generally, the asymptotic critical values provide a reasonable approximation to the finite sample distributions, for sample sizes as low as 100. Doubling the sample size only results in a slight mitigation of the size distortion.

We now turn to the relative power of the joint and individual test statistics. We consider two values for ρ under the alternative hypothesis; 0.9 and 0.7. The range of θ and γ considered are 0.5, 1, 2, 3. The last value corresponds to a break 3 times the size of the innovation standard deviation. We also consider negative values of θ and γ. The results are similar and are available upon request from the authors.

Table 6 presents the power for non-trending data at the 5% and 10% significance levels. A few points are in order. First, the 1-sided test for no structural change has higher power than the 2-sided test. This result corroborates the finding of Vogelsang and Perron (1994) and Perron (1997) that imposing a sign for the trend break leads to an increase in power. Second $\sup W_{\theta,\rho}$ statistic uniformly dominates $\sup W_{\theta}$ and $\sup t_{\theta}$ for all values of θ and ρ. However, all statistics are dominated by the 1-sided unit root test. This result is analogous to a finding of Dickey (1984). They demonstrate that in standard ADF tests, the t-statistic for the unit root null has more power than F-tests of the joint null hypothesis of a unit root and no time trend. In practice, if the researcher is interested in performing inference on θ, then for non-trending data, the $\sup Wald_{\theta,\rho}$ test should be implemented.

A different picture emerges for trending data. Tables 9-11 present size adjusted power for Models A, B, and C respectively. The 1-sided unit root statistic no longer uniformly dominates the others in term of power. For Model A, the test which has the highest power depends on ρ. For ρ = 0.9, and a level shift in the range of 1 to 2 innovation standard deviations, the $\sup t_{\theta}^{A}$ test is the clear winner. However, for ρ = 0.7, the $\sup Wald_{\theta,\rho}^{A}$ outperforms the $\sup t_{\theta}^{A}$ test.

Turning now to Model B, for changes in growth in the range of 0.5 to 1 innovation standard deviation, the one sided test , $\sup t_{\gamma}^{B}$, is the clear winner for both value of T and ρ. There does not appear to be a distinct advantage to performing the $\sup Wald_{\gamma,\rho}^{B}$ test over the 1-sided test for structural change or a unit root. For larger changes in growth, all tests perform remarkably well. But as we shall see in the next section, changes in growth this size do not occur is U.S. output.

5.   Application to U. S. GDP

As an empirical application, we reconsider the Zivot-Andrews unit root tests on annual and quarterly U.S. real GDP analyzed by Murray and Nelson (1998). Murray and Nelson perform the Model A unit root test on the Maddison (1995) annual GDP series

(1870-1994), and the Model B unit root test on post-war quarterly chained U.S. GDP (1947.1–1997.3). They demonstrate that whether the lag length is selected by the general to specific (GS) strategy or the Schwartz information criterion (SIC), the unit root null is rejected at the 5% for annual GDP, but not at the 10% level for quarterly GDP. These regressions are presented in Table 12. Since each test considered in this section chooses the same break date for annual and quarterly data (1929 and 1972.2), we present the results for each series and lag selection procedure as one regression. As in Perron (1989) and Zivot and Andrews (1992), the maximum lag length considered is 8 for annual data and 12 for quarterly data. For either frequency, GS chooses the maximum lag allowed, while SIC chooses only 1.

Using the critical values from the 5th row in Tables 4 and 5, we can assess whether or not $\hat{\theta}$ or $\hat{\gamma}$, the step and ramp dummy coefficients, are significant when the break date is chosen to minimize t_ρ. (These are the $Wald^A_{\theta,\inf(\rho)}$ and $Wald^B_{\gamma,\inf(\rho)}$ statistics). For the annual series (Model A), the level shift is significant at the 10% level for GS, but insignificant for SIC. While both methods of lag selection result in rejection of the unit root null, GS suggests stationarity around a broken trend, while SIC indicates stationarity around a constant trend.

To assess whether there has been structural change while not explicitly testing the unit root hypothesis, we perform the 1-sided $\inf t^A_\theta$ test. This statistic is significant at the 5% level for GS, but not significant at the 10% level for SIC.

We now turn our attention to the $\sup Wald^A_{\theta,\rho}$ statistic, which tests the joint null hypothesis. This statistic is significant at the 5% level for both methods of lag selection. It thus suggests that GDP is stationary around a broken trend. The disagreement between $\sup Wald^A_{\theta,\rho}$ and $\inf t^A_\theta$, for SIC, may be due to poor power properties of $\inf t^A_\theta$ when only a subset of the null is violated, *i.e.* $\rho < 1$ and $\theta \neq 0$.

In Section 4 we demonstrated that for level shifts of the size estimated for this series (1 to 2 innovation standard deviations), and a non-local autoregressive root (0.7), the power of the $\sup Wald^A_{\theta,\rho}$ statistic dominates the 1-sided tests for structural change, but is dominated by the 1-sided Zivot-Andrews unit root test. Given that none of the statistics have an appreciable finite sample size distortion, these results lead us to conclude that annual GDP is stationary around a broken trend.

Analogous statistics for quarterly GDP are also presented in Table 12. For this series, which appears to have a unit root, we find that the pre and post break growth rates, based on the $Wald^{B}_{\gamma,\inf(\rho)}$ statistic, are not statistically different under either method of lag selection. A different picture emerges if we compute the 1-sided $\inf t^{B}_{\gamma}$ test for a change in growth. Under GS, there is not a statistical difference in growth rates, but the SIC results in a rejection of the null at the 10% level.

Turning to the $\sup W^{B}_{\gamma,\rho}$, for both methods of lag selection the joint test corroborates the Zivot-Andrews unit root test. Neither is significant at the 10% level.

We demonstrated in Section 4 that for the small changes in growth (less that 1 innovation standard deviation) that this series appears to exhibit, the 1-sided tests for structural change uniformly dominate all other statistics in terms of power. Since both methods of lag selection lead to different outcomes for the $\inf t^{B}_{\gamma}$ test, we can conclude that there is a unit root, but we are uncertain as to whether the rate of growth has changed in the postwar period.

6. Summary and Concluding Remarks

The purpose of this paper has been to fill in the gaps in the literature concerning the asymptotic distributions of test statistics for a unit root and/or structural change. We derive 1 and 2-sided tests for the null of no structural change as well as joint tests of the hypothesis that a time series is integrated without structural change. The motivation for the latter is the potential increase in power over tests which do not explicitly test the unit root hypothesis.

For Model A, no clear winner emerges. For level shifts of the size estimated for U.S. real GDP, the joint test has higher power than individual tests for non-local autoregressive roots. However, the situation is reversed for a local root. For Model B, the 1-sided tests for structural change dominate the joint tests for small changes in growth, regardless of the size of the autoregressive root.

We apply the tests derived here to annual and quarterly U.S. real GDP. Almost all tests agree that the 1870-1994 annual GDP series is stationary around a broken time trend with a change in level occurring at 1929. While all tests indicate that the 1947.1-1997.3 quarterly GDP series has a unit root, there is not a consensus as to whether or not the growth rate began to slow in 1972.2.

References

Andrews, Donald W. K., 1993, Tests for parameter instability and structural change with unknown change point, *Econometrica*, 61, 821-856.

Andrews, Donald W. K. and Werner Ploberger, 1994, Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica*, 62, 1383-1414.

Banerjee, Anindya, Robin L. Lumsdaine, and James 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, 271-287.

Billingsley, Patrick, 1968, Convergence of probability measures, Wiley, New York.

Christiano, Lawrence J., 1992, Searching for a break in GNP, *Journal of Business and Economic Statistics*, 10, 237-250.

Dickey, David A., 1984, Powers of unit root tests, *Proceedings of the Business and Economics Statistics Section, American Statistical Association*, 489-493.

Hall, Alastair, 1994, Testing for a unit root in time series with pretest date-based model selection, *Journal of Business and Economic Statistics*, 12, 271-306.

Maddison, Angus, 1995, Monitoring the world economy 1820-1992, Paris: OECD.

Murray, Christian J. and Charles R. Nelson, 1998, The uncertain trend in U.S. GDP, working paper, University of Washington.

Nelson, Charles R. and Charles I. Plosser, 1982, Trends and random walks in macroeconomic time series: Some evidence and implications, *Journal of Monetary Economics*, 10, 139-162.

Ng, Serena and Pierre Perron, 1995, Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag, *Journal of the American Statistical Association*, 90, 268-281.

Perron, Pierre, 1989, The great crash, the oil price shock, and the unit root hypothesis, *Econometrica*, 57, 1361-1401.

Perron, Pierre, 1997, Further evidence on breaking trend functions in macroeconomic variables, *Journal of Econometrics*, 80, 355-385.

Perron, Pierre and Timothy Vogelsang, 1992, Nonstationarity and level shifts with an application to purchasing power parity, *Journal of Business and Economic Statistics*, 10, 301-320.

Rappoport, P. and L. Reichlin, 1989, Segmented trends and nonstationary time series, *The Economic Journal*, 99, 168-177.

Said, Said E. and David A. Dickey, 1984, Testing for unit root in autoregressive moving average models of unknown order, *Biometrika*, 71, 599-608.

Vogelsang, Timothy J., 1993, Essays on testing for nonstationarities and structural change in time series models, Unpublished Ph.D. Dissertation, Princeton University, Princeton, NJ.

Vogelsang, Timothy, J., 1997, Wald-type tests for detecting breaks in the trend function of a dynamic time series, *Econometric Theory*, 13, 818-849.

Vogelsang, Timothy, J. and Pierre Perron, 1994, Additional tests for a unit root allowing for a break in the trend function at an unknown time, working paper, Cornell University.

Zivot, Eric and Donald W. K. Andrews, 1992, Further evidence on the great crash, the oil price shock and the unit root hypothesis, *Journal of Business and Economic Statistics*, 10, 251-270.