NBER WORKING PAPERS SERIES

# RECURSIVE AND SEQUENTIAL TESTS OF THE UNIT ROOT AND TREND BREAK HYPOTHESES: THEORY AND INTERNATIONAL EVIDENCE

Anindya Banerjee

Robin L. Lumsdaine

James H. Stock

Working Paper No. 3510

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1990

The authors thank T.W. Anderson, Donald Andrews, Bruce Hansen, Lars Hansen, G.S. Maddala, Greg Mankiw, Alain Monfort, and Mark Watson for helpful comments and discussions. This research was initiated while Banerjee was Visiting Scholar, Kennedy School of Government, Harvard University. Stock thanks the Sloan Foundation and the National Science Foundation (grant no. SES-89-10601) for financial support. This paper is part of NBER's research program in Economic Fluctuations. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #3510
November 1990

# RECURSIVE AND SEQUENTIAL TESTS OF THE UNIT ROOT AND TREND BREAK DOWN HYPOTHESES: THEORY AND INTERNATIONAL EVIDENCE

## ABSTRACT

This paper investigates the possibility, raised by Perron (1989, 1990a), that aggregate economic time series can be characterized as being stationary around broken trend lines. Unlike Perron, we treat the break date as unknown *a priori*. Asymptotic distributions are developed for recursive, rolling, and sequential tests for unit roots and/or changing coefficients in time series regressions. The recursive and rolling tests are based on a time series of recursively estimated coefficients, computed using increasing subsamples of the data. The sequential statistics are computed using the full data set and a sequence of regressors indexed by a "break" date. When applied to data on real postwar output from seven OECD countries, these techniques fail to reject the unit root hypothesis for five countries (including the U.S.), but suggest stationarity around a shifted trend for Japan.

Anindya Banerjee
Wadham College
Oxford University
Oxford OX1 3PN
UNITED KINGDOM

Robin L. Lumsdaine
Dept. of Economics
Harvard University
Cambridge, MA 02138

James H. Stock
Dept. of Economics, Evans Hall
University of CA - Berkeley
Berkeley, CA 94720

## 1. Introduction

There is a large literature on the persistence exhibited by aggregate output, in particular whether output is well characterized as containing a unit autoregressive root. One alternative to the unit root (or "integrated") model, suggested by Perron (1989), is that log output ($y_t$) is stationary around a deterministic time trend that has one slope in an initial fraction of the sample and, later, a different slope. Using quarterly data for the postwar U.S., Perron (1989) presented evidence against the unit root null in favor of this trend-shift alternative when the trend shift is associated with the first oil price shock. Evans (1989) and Perron (1990a) suggested the related model in which there is a shift in the intercept, in Perron's case possibly in conjunction with a shift in the slope of the deterministic trend. These models and empirical findings are important, for four reasons. First, as Perron emphasizes, if the trend-shift/stationary model is correct, then studies such as Cochrane's (1988) and Cogley's (1990) have attributed too much persistence to innovations in GNP, and conventional unit root test statistics will incorrectly fail to reject the unit root null. Second, in the spirit of Harvey (1985) and Watson (1986), this provides a parsimonious model for a slowly changing trend component of output which might be useful as data description. Interpreted narrowly, the single-shift/deterministic-trend model has little economic appeal, but interpreted more broadly it can be thought of as a metaphor for there being a few large events that determine the growth path of output over a decade or two -- in the U.S., the Depression or, later, the productivity slowdown. Once these decade-shaping events are taken into account, output exhibits business cycle properties in the sense of mean-reversion over business cycle horizons. Third, current empirical research relies heavily on techniques built upon the integrated/stationary classification of time series: if series that are stationary with breaking trends are incorrectly classified as integrated, incorrect inferences are likely to follow. Fourth, if the breaking-trend/stationary model fits many time series better than the integrated model, then the empirical relevance of the growing literature in theoretical econometrics on unit roots and cointegration is brought into serious question.

The empirical focus of this paper is on international patterns of persistence and possible permanent shifts in growth trends. We are, however, persuaded by Christiano's (1988) argument that the date of the break ought not be treated as known -- Perron's approach -- but rather should be treated as unknown *a priori*. After all, the hypothesis that there might have been a break in the U.S. output process around the first oil shock has intuitive appeal precisely because we know, before performing formal tests for breaks, that this major event *was* followed by a period of slower growth. This paper therefore starts with the presumption that, if there is a break, its date is not known *a priori* but rather is gleaned from the data.

The literature on persistence of output includes several international comparisons (Campbell and Mankiw [1989], Clark [1989], Cogley [1990], and Kormendi and Meguire [1988]). These are based on full-sample techniques rather than procedures that explicitly allow for changing coefficients, and thus leave unanswered some intriguing questions. Once the break point is treated as unknown *a priori*, is there evidence of a break in the drift of output? Is output stationary around a changing deterministic trend? If so, is this pattern consistent across countries, or is it idiosyncratic to specific countries? In particular, if there are identified breaks, are they associated with the productivity slowdown of the mid 1970's, and do they have the same timing across countries?

Our two objectives are first to develop econometric techniques (and the associated distribution theory) appropriate for answering these questions, and second to apply these techniques to international data on output (real GNP or GDP) for seven OECD countries. Once the break point is treated as unknown, the usual distribution theory does not apply: the relevant statistic now is, say, the absolute maximum of an increasing number of unit root test statistics, one for each possible break date. Christiano (1988) and Evans (1989) recognized this issue and addressed it by numerical simulations.

The methodological contribution of this paper is to provide an asymptotic distribution theory for statistics pertaining to the shifting-root/shifting-trend hypotheses. Three classes of statistics are considered: recursive, rolling, and sequential. Recursive statistics are computed over subsamples $t=1,\ldots,k$, for $k=k_0,\ldots,T$, where $k_0$ is a startup value and T is the size of the full sample. "Rolling" statistics are computed using samples that are a constant fraction $\delta_0$ of the full sample, rolling through the sample. Statistics that we term "sequential" are computed using the full sample, sequentially incrementing the date of the hypothetical break (or shift). The term "recursive" derives from Brown, Durbin and Evans' (1975) treatment of recursive estimation. Some recursive techniques are currently implemented in Hendry's (1987) statistical package PC-GIVE; also see Dufour (1982). An example of a sequential statistic is Perron's (1989) unit root test with a trend shift at date k, computed sequentially for $k = k_0,\ldots, T-k_0$, where $k_0$ allows for trimming the initial and final parts of the sample. Another sequential statistic is Quandt's (1960) likelihood ratio statistic, which entails computing the sequence of likelihood ratio statistics for a break in all the coefficients and then taking the maximum. These statistics are shown to have natural representations as stochastic processes defined on the unit interval, and their limiting distributions are characterized by functionals of Wiener processes. These results extend related work by McCabe and Harrison (1980), Sen (1980, 1982), Dufour (1982), James, James and Siegmund (1987), Kramer, Ploberger and Alt (1988), and Ploberger, Kramer and Kontrus (1989). The primary extension is to the case of a unit root in the regressors and to recursive and sequential tests for unit roots.

These techniques are applied to data on postwar real output for Canada, France, Germany, Italy, Japan, the U.K., and the U.S. In only one case (Japan) is the unit root hypothesis rejected in favor of the trend-shift hypothesis. We also investigate the possibility that output has a unit root, but that its drift (or mean growth rate) shifted at an unknown date over this period. This broken drift could proxy for a permanent shift in the process of technological change or, as here, for a productivity slowdown. In four of the seven countries, the no-break/unit-root null is rejected in favor of the broken-drift/unit-root alternative, with the break in the early 1970's.

The recursive and sequential statistics are described, and their asymptotic properties studied, in Sections 2 and 3. Section 4 presents critical values and a Monte Carlo experiment. The empirical results are presented in Section 5, and conclusions are summarized in Section 6.[1]

## 2. Recursive and Rolling Test Statistics

The primary focus of this section is on two recursive tests for a unit root, the Dickey-Fuller t-statistic and a modification of the Sargan-Bhargava (1983) and Bhargava (1986) statistics, although the results are more general in that some specialize to stationary time series as well. These two tests examine the null hypothesis that the process has a unit root, against the alternative that the process has a root less than one, perhaps with a nonzero linear time trend. Thus the statistics involve some form of detrending. The motivation for considering recursive unit root tests is that the process might be well-approximated as having a unit root over part of the sample but not over another part.

*A. Recursive and Rolling Dickey-Fuller Statistics*

It is assumed that observations on $\{y_t\}$ are generated according to:

$$(2.1) \quad \text{Model I:} \quad y_t = \mu_0 + \mu_1 t + \alpha y_{t-1} + \beta(L)\Delta y_{t-1} + \epsilon_t, \quad t=1,\ldots,T$$

where $\beta(L)$ is a lag polynomial of known order p with the roots of $1-\beta(L)L$ outside the unit circle. Under the null hypothesis $\alpha=1$ and $\mu_1=0$. For estimation and testing, $\alpha$ and $\mu_1$ are unconstrained. The errors are assumed to satisfy:

*Assumption A.* $\epsilon_t$ is a martingale difference sequence with $E(\epsilon_t^2|\epsilon_{t-1},\ldots) = \sigma^2$, $E(|\epsilon_t|^i|\epsilon_{t-1},\ldots) = \kappa_i$, i=3, 4, and $\sup_t E(|\epsilon_t|^{4+\gamma}|\epsilon_{t-1},\ldots) = \bar{\kappa} < \infty$ for some $\gamma > 0$.

The t-statistic testing the hypothesis that $\alpha=1$ in Model I, computed over the full sample of T observations, is the standard Dickey-Fuller (1979) t-statistic for testing for a unit root, including a constant and a time trend in the regression. The extension here is to the time series of recursively computed estimators and t-statistics. Because of the unit root under the null hypothesis, it is convenient to define transformed regressors $Z_t$ and a transformed parameter vector $\theta$ so that under the null (2.1) can be rewritten, (2.2) $$y_t = \theta' Z_{t-1} + \epsilon_t$$

where $Z_t = [Z_t^{1\prime}\ Z_t^2\ Z_t^3\ Z_t^4]'$ , where $Z_t^1 = (\Delta y_t - \bar{\mu}_0 \cdots \Delta y_{t-p+1} - \bar{\mu}_0)'$, $Z_t^2 = 1$, $Z_t^3 = (y_t - \bar{\mu}_0 t)$, and $Z_t^4 = t$, where $\bar{\mu}_0 = E\Delta y_t = \mu_0/(1-\beta(1))$, and where $\theta = (\theta_1'\ \theta_2\ \theta_3\ \theta_4)'$ with $\theta_1 = (\beta_1 \cdots \beta_p)'$, $\theta_2 = \mu_0 + \beta(1)\bar{\mu}_0 + \mu_1$, $\theta_3 = \alpha$, and $\theta_4 = \mu_1 + \alpha\bar{\mu}_0$. The transformed regressors $Z_t$ are linear combinations of the original regressors in (2.1), with the linear combination chosen to isolate the regressors with different stochastic properties; specifically, $Z_t^1$ are mean zero stationary regressors and $y_t - \bar{\mu}_0 t$ is an integrated process with no deterministic component. (This transformation is adopted from and discussed in Sims, Stock and Watson [1990].) Because the elements of $\theta$ converge at different rates, define the scaling matrix $T_T = \mathrm{diag}(T^{1/2} I_p,\ T^{1/2},\ T,\ T^{3/2})$, partitioned conformably with $Z_t$ and $\theta$. Let $\Omega_p$ denote the covariance matrix of $\Delta y_t, \ldots, \Delta y_{t-p+1}$, so $EZ_t^1 Z_t^{1\prime} = \Omega_p$. Also, take as initial conditions $Z_s = 0$, $s \le 0$, so that sums can be written starting at t=1.

The recursive OLS estimator of the coefficient vector is

(2.3) $$\hat{\theta}(\delta) = (\textstyle\sum_{t=1}^{[T\delta]} Z_{t-1} Z_{t-1}')^{-1} (\sum_{t=1}^{[T\delta]} Z_{t-1} y_t),\quad 0<\delta_0 \le \delta \le 1 .$$

Thus

(2.4) $$T_T(\hat{\theta}(\delta) - \theta) = V_T(\delta)^{-1} \phi_T(\delta)$$

where $V_T(\delta) = T_T^{-1} \sum_{t=1}^{[T\delta]} Z_{t-1} Z_{t-1}' T_T^{-1}$ and $\phi_T(\delta) = T_T^{-1} \sum_{t=1}^{[T\delta]} Z_{t-1} \epsilon_t$. Thus (2.3) and (2.4) respectively provide representations of the recursive least squares estimator and its scaled deviation from $\theta$ as random elements of $D[0,1]$.

There are analogous expressions for a general recursively computed Wald statistic and for the Dickey-Fuller t-statistic testing the hypothesis that $\alpha=1$. Suppose that the Wald statistic tests the q hypotheses $R\theta=r$, where without loss of generality the hypotheses are ordered so that R is upper block triangular when partitioned conformably with $\theta$, that is, the first restrictions involve coefficients on $Z_t^1$ (and perhaps $Z_t^2$, $Z_t^3$ and $Z_t^4$), the next restrictions involve coefficients on $Z_t^2$ (and perhaps $Z_t^3$ and $Z_t^4$), and so forth. The test statistics are

(2.5) $$F_T(\delta) = (R\hat{\theta}(\delta)-r)'[R(\sum_{t=1}^{[T\delta]} Z_{t-1}Z_{t-1}')^{-1}R']^{-1}(R\hat{\theta}(\delta)-r)/q\hat{\sigma}^2(\delta), \quad \delta_0\le\delta\le 1$$

(2.6) $$t_{DF}(\delta) = T(\hat{\theta}_3(\delta)-1)/[V_T(\delta)^{33}\hat{\sigma}^2(\delta)]^{½}, \quad \delta_0\le\delta\le 1,$$

where $\hat{\sigma}^2(\delta) = [T\delta]^{-1}\sum_{t=1}^{[T\delta]}(y_t-\hat{\theta}(\delta)'Z_{t-1})^2$ and $V_T(\delta)^{ij}$ denotes the (i,j) element of $V_T(\delta)^{-1}$. Algebraic manipulations have been used to rewrite $t_{DF}(\delta)$ from Model I in terms of the transformed regression (2.2). Finally, define $R^*$ (partitioned conformably with $\theta$) so that $R_{ii}^* = R_{ii}$, $i=1,\ldots,4$, $R_{12}^* = R_{12}$, and $R_{ij}^* = 0$ otherwise.

Let "=>" denote weak convergence on $D[0,1]$. The asymptotic behavior of the recursive estimators and test statistics is summarized in the following theorem.

*Theorem 1*. Suppose that $y_t$ is generated by Model I with $\mu_1=0$ and $\alpha=1$, and that Assumption A holds. Then, for $0<\delta_0\le\delta\le 1$,

a) $V_T(\cdot) \Rightarrow V(\cdot)$, $\phi_T(\cdot) \Rightarrow \phi(\cdot)$, and
$T_T(\hat{\theta}(\cdot)-\theta) \Rightarrow \theta^*(\cdot)$, where $\theta^*(\delta) = V(\delta)^{-1}\phi(\delta)$,
where $V(\delta)$ and $\phi(\delta)$ are partitioned conformably with $T_T$ and
$\phi(\delta) = \sigma[B(\delta),\ W(\delta),\ ½b\sigma(W(\delta)^2-\delta),\ \delta W(\delta)-\int_0^\delta W(\lambda)d\lambda]'$,

$V_{11} = \delta\Omega_p$, $V_{1j} = 0$, $j = 2,3,4$, $V_{22} = \delta$, $V_{23} = \sigma b\int_0^\delta W(\lambda)d\lambda$, $V_{24} = \frac{1}{2}\delta^2$, $V_{33} = \sigma^2 b^2\int_0^\delta W(\lambda)^2 d\lambda$, $V_{34} = \sigma b\int_0^\delta \lambda W(\lambda)d\lambda$, and $V_{44} = (1/3)\delta^3$, where $W(\delta)$ is standard Brownian motion on [0,1] and $B(\delta)$ is p-dimensional Brownian motion with covariance matrix $\Omega_p$, W and B are independent, and $b=(1-\beta(1))^{-1}$.

b) Under the null $R\theta=r$,

$$F_T(\cdot) \Rightarrow [R^*\theta^*(\cdot)]'[R^*V(\cdot)^{-1}R^{*\prime}]^{-1}[R^*\theta^*(\cdot)]/q\sigma^2 = F^*(\cdot).$$

c) Under the null $\alpha=1$,

$$\hat{t}_{DF}(\cdot) \Rightarrow [\sigma^2 V(\cdot)^{33}]^{-\frac{1}{2}}\theta_3^*(\cdot) = t_{DF}^*(\cdot).$$

Proofs of theorems are given in Appendix A. Some remarks serve to highlight different aspects of this result.

(a) Kramer, Ploberger and Alt (1988) and Ploberger, Kramer and Kontrus (1989) consider the case in which the regressors are stationary lagged dependent variables; our results provide explicit proofs in their case that the relevant "denominator" matrices are uniformly consistent.

(b) As in the $\delta$=1 case, $V(\cdot)$ is block diagonal. Thus the recursive estimation of the nuisance parameters $(\beta_1,\ldots,\beta_p)$ does not affect the asymptotic distribution of the recursive Dickey-Fuller statistic. The novel feature of these results is that they apply uniformly in $\delta$; the "marginals" at any fixed $\delta$ are those that would be obtained using conventional (fixed $\delta$) asymptotics. For example, $t_{DF}(\delta')$, evaluated at a fixed $\delta'$, has the "Dickey-Fuller" t-statistic distribution. Thus the limiting stochastic process $t_{DF}(\cdot)$ can be thought of as a Dickey-Fuller t-statistic process. Because $V(\cdot)$ is block diagonal, when the restrictions in $R^*$ involve only coefficients on $Z_{t-1}^1$, $F(\cdot)$ similarly can be thought of as a $\chi_q^2/q$ process. Also, the distribution of the recursive "demeaned" Dickey-Fuller statistic obtains as a special case by omitting t as a regressor in (2.1). The asymptotic representations apply for $0<\delta_0\leq\delta\leq1$, accounting for $[T\delta_0]$ startup observations.

(c) The transformation of the original regressors to $Z_t$ is used to obtain a nondegenerate joint limiting representation of the estimators. This is the device used by Fuller (1976) and Dickey and Fuller (1979). As in the $\delta=1$ case, because t is included as a regressor the distributions of the processes based on $\hat{\beta}(\delta)$ and $\hat{\alpha}(\delta)$ do not depend on the nuisance parameter $\mu_0$. For further discussion in the $\delta=1$ case, see Sims, Stock and Watson (1990).

(d) Although Theorem 1 is stated for the null model in which $\alpha=1$ and $\mu_1=0$, the results are sufficiently general to handle the case $|\alpha|<1$, $\mu_1\neq0$. This follows by redefining the variables in Model I. Specifically, let the left-hand variable be $\Delta y_t$ rather than $y_t$ and exclude $y_{t-1}$ from the regression. Then the regressors are $(Z_t^1,1,t)$, where $Z_t^1$ has mean zero and is stationary. Thus associating the I(0) regressors with $\Delta y_t$ in the notation of (2.1) and omitting the terms in $y_t$ in the statement of the theorem provides the limiting process for the recursive estimators for the case of a stationary autoregression when $\mu_1=0$. If $\mu_1$ is nonzero, an additional modification so that $Z_t^1$ remains mean zero and stationary (by subtracting $\mu_1 t$ from $\Delta y_t$) results in Theorem 1 applying to the case of a regression involving an AR(p) process that is stationary around a time trend. With these modifications, the result concerning $F_T(\cdot)$ (Theorem 1(b)) applies directly, although of course the result on $t_{DF}(\cdot)$ is no longer germane.

(e) Asymptotic representations for rolling estimators and test statistics obtain as a consequence of Theorem 1. Because a fixed fraction $\delta_0$ of the sample is used, unlike recursive estimators the sampling variability of rolling coefficient estimators is constant through the sample. Also, rolling statistics might better be able to detect some breaks, because for $\delta_0<½$ any break will be excluded from at least one part of the rolling sample and included in another. Let

$$\hat{\theta}(\delta;\delta_0) = (\sum_{t=[T(\delta-\delta_0)]+1}^{[T\delta]} Z_{t-1}Z'_{t-1})^{-1}(\sum_{t=[T(\delta-\delta_0)]+1}^{[T\delta]} Z_{t-1}y_t)$$

so that $T_T(\hat{\theta}(\delta;\delta_0)-\theta) = V_T(\delta;\delta_0)^{-1}\phi_T(\delta;\delta_0)$, where $V_T(\delta;\delta_0) = T_T^{-1}(\sum_{t=[T(\delta-\delta_0)]+1}^{[T\delta]} Z_{t-1}Z'_{t-1})T_T^{-1} = V_T(\delta)-V_T(\delta-\delta_0)$ and $\phi_T(\delta;\delta_0) = T_T^{-1}(\sum_{t=[T(\delta-\delta_0)]+1}^{[T\delta]} Z_{t-1}y_t) = \phi_T(\delta)-\phi_T(\delta-\delta_0)$. From Theorem 1(a), $V_T(\cdot;\delta_0) \Rightarrow V(\cdot;\delta_0)$, where $V(\delta;\delta_0) = V(\delta)-V(\delta-\delta_0)$, and $\phi_T(\cdot;\delta_0) \Rightarrow \phi(\cdot;\delta_0)$, where $\phi(\delta;\delta_0) = \phi(\delta)-\phi(\delta-\delta_0)$. Thus $T_T(\hat{\theta}(\cdot;\delta_0)-\theta) \Rightarrow \theta^*(\cdot;\delta_0)$, where $\theta^*(\delta;\delta_0) = V(\delta;\delta_0)^{-1}\phi(\delta;\delta_0)$. Representations for rolling F- and t-statistics follow accordingly.

*B. Recursive and Rolling Sargan-Bhargava statistics.*

Sargan and Bhargava (1983) derived the uniformly most powerful test of the null hypothesis that $y_t$ follows a Gaussian random walk against the stationary first-order alternative. Bhargava (1986) extended these statistics to the case in which the null is a Gaussian random walk with drift and the alternative is a stationary AR(1) around a potentially nonzero time trend, and he showed his test to be most powerful invariant. This discussion focuses on Bhargava's statistic, appropriate for the empirical application; the Sargan-Bhargava statistic is handled analogously. Bhargava's statistic is

(2.7) $$R_2 = \sum_{t=2}^{T}(\Delta y_t^B(T))^2/\sum_{t=2}^{T}(y_t^B(T))^2.$$

where $y_t^B(S) = y_t - ((t-1)/(S-1))y_S - ((S-t)/(S-1))y_1 - (\bar{y}(S)-½(y_1+y_S))$, where $1\le t\le S$ and $\bar{y}(S) = S^{-1}\sum_{r=1}^{S} y_r$.

Although these statistics have desirable optimality properties in the Gaussian AR(1) case, unfortunately they have the wrong size when they are applied to series in which the spectral density of $\Delta y_t$ at frequency zero ($s_{\Delta y}(0)$) differs from $\sigma^2/2\pi$. This difficulty can be addressed by modifying the Sargan-Bhargava statistic as suggested in Stock (1988), by replacing $T^{-1}\sum_{t=1}^{T}(\Delta y_t^B(T))^2$ by an estimator of $2\pi s_{\Delta y}(0)$. The full-sample version of this statistic was found to have good size and power. In its recursive form, this modified statistic (actually, a modification of $T^{-1}R_2^{-1}$) is

$$(2.8) \qquad \hat{R}(\delta) = [T\delta]^{-2}\sum_{t=2}^{[T\delta]}(y_t^B([T\delta])/(2\pi\hat{s}_{\Delta y}(0))^{\frac{1}{2}})^2$$

where $\hat{s}_{\Delta y}(0)$ is a consistent estimator of $s_{\Delta y}(0)$. It is assumed here that $\hat{s}_{\Delta y}(0)$ is computed using the full sample; only the numerator in the statistic is computed recursively.

The asymptotic representation of the recursive modified Sargan-Bhargava statistic (2.8) is summarized in the following Theorem.

*Theorem 2.* Assume that $\Delta y_t = \bar{\mu}_0 + c(L)\epsilon_t$, $t=1,\ldots,T$, where $\epsilon_t$ satisfies Assumption A, $c(1)\neq 0$, and $\sum_{j=0}^{\infty}|c_j|<\infty$, and assume that $\hat{s}_{\Delta y}(0) \xrightarrow{p} s_{\Delta y}(0)$. Then, on $0<\delta_0\leq\delta\leq 1$, $\hat{R}(\cdot) \Rightarrow R^*(\cdot)$, where $R^*(\delta) = \delta^{-2}\int_0^\delta W^B(\lambda,\delta)^2 d\lambda$ and $W^B(\lambda,\delta) = W(\lambda)-(\lambda/\delta-\frac{1}{2})W(\delta) - \int_0^\delta W(s)ds$ for $\lambda\leq\delta$.

This result assumes that $\hat{s}_{\Delta y}(0) \xrightarrow{p} s_{\Delta y}(0)$. Such estimators are discussed extensively in the literature on tests for unit roots. Note that not all consistent estimators under the null result in consistent tests; for discussions and recommendations, see Phillips and Ouliaris (1990), Stock (1988), and Stock and Watson (1988).

Although Theorem 2 is stated for the special case of the statistic $\hat{R}(\delta)$, this type of result extends directly to the wider class of unit root tests discussed in Stock (1988), which have the form $g(v_{[T\cdot]})$, where g is a continuous mapping from $D[0,1]$ to $\mathbb{R}^1$ and where $v_{[T\lambda]}$ is an appropriate element of $D[0,1]$ that obeys a FCLT. In this notation the recursive modified Sargan-Bhargava statistic is $g(f) = \delta^{-2}\int_0^\delta f^2(\lambda)d\lambda$, with $v_{[T\lambda]} = y^B_{[T\lambda]}([T\delta])/(2\pi\hat{s}_{\Delta y}(0))^{\frac{1}{2}}$.

Theorem 2 can be extended to rolling modified Sargan-Bhargava (and $g(\cdot)$-class) tests for unit roots, by computing the statistic on rolling overlapping subsamples of length $[\delta_0 T]$, with $\delta_0$ fixed. If, as in Theorem 2, $s_{\Delta y}(0)$ is estimated over the full sample, the extension to the rolling statistics is straightforward and is omitted here.

## 3. Sequential Tests for Changes in Coefficients

The statistics analyzed in this section are computed sequentially using the full sample. These allow for a single shift or break in a deterministic trend at an unknown date. The model considered is

$$(3.1)\quad \text{Model II:}\quad y_t = \mu_0 + \mu_1\tau_{1t}(k) + \mu_2 t + \alpha y_{t-1} + \beta(L)\Delta y_{t-1} + \omega' x_{t-1}(k) + \epsilon_t$$

for $t=1,\ldots,T$, where $\beta(L)$ is a lag polynomial of known order p. Unlike model I, model II allows for a m-vector of additional stationary regressors, $x_{t-1}(k)$. As in Section 2, it is convenient to transform the regressors to $Z_t = [Z_t^{1\prime}\ 1\ (y_t-\bar{\mu}_0 t)\ \tau_{1t+1}\ t]'$, where $Z_t^1 = (\Delta y_t-\bar{\mu}_0,\ldots,\ \Delta y_{t-p+1}-\bar{\mu}_0,\ x_t(k)')'$ and $\bar{\mu}_0 = E\Delta y_t$; to let $\theta = [\theta_1'\ \theta_2\ \theta_3\ \theta_4\ \theta_5]'$, where $\theta_1 = [\beta'\ \omega']'$, $\theta_2 = \mu_0+\beta(1)\bar{\mu}_0+\mu_2$, $\theta_3=\alpha$, $\theta_4=\mu_1$, and $\theta_5=\mu_2+\alpha\bar{\mu}_0$; and to set $Z_s=0$ for $s\le 0$.

The deterministic regressor $\tau_{1t}(k)$ captures the possibility of a shift or jump in the trend at period k. Following Perron (1989, 1990a), consider two cases:

(3.2) Case A (shift in trend): $\tau_{1t}(k) = (t-k)1(t>k)$

(3.3) Case B (jump in trend): $\tau_{1t}(k) = 1(t>k)$,

where $1(\cdot)$ is the indicator function. For Case A, the t-statistic testing $\mu_1=0$ provides information about whether there has been a shift (change in slope) in the trend; for Case B, this t-statistic provides information about whether there has been a break in the trend. Case B will also be referred to as the mean-shift model (cf. Perron [1990a]).

Let $\omega_0$ denote the value of $\omega$ under the null. It is assumed that those $x_t$ terms involving k do not enter under the null. Specifically, the disturbances and $\{x_t(k)\}$ are assumed to satisfy,

*Assumption B.*

(i) Let $M_T$ be the sigma field generated by $\{\epsilon_t, x_t(k), \epsilon_{t-1}, x_{t-1}(k), , \ldots\}$. Then $E(\epsilon_t|M_{t-1})=0$, $E(\epsilon_t^2|M_{t-1})=\sigma^2$, $E(|\epsilon_t|^i|M_{t-1})=\kappa_i$, i=3,4, and $E(|\epsilon_t|^{4+\gamma}|M_{t-1}) \leq \bar{\kappa} < \infty$ for some $\gamma>0$.

(ii) $\{x_t([T\delta])\}$ is such that: $Ex_t([T\delta])=0$, $T^{-1}\sum_{t=1}^{T} Z_t^1 Z_t^{1\prime} \xrightarrow{p} \Sigma(\delta)$, $T^{-3/2}\sum_{t=1}^{T} Z_t^1 y_t \Rightarrow 0$, and $(T^{-\frac{1}{2}}\sum_{t=1}^{[T\lambda]} \epsilon_t, T^{-\frac{1}{2}}\sum_{t=1}^{[T\lambda]} \omega_0' x_t(k)) \Rightarrow (\sigma W(\lambda), \pi H(\lambda))$, all uniformly in $\delta$, where W and H are standard 1-dimensional Brownian motions, W and H are not necessarily independent, and $\pi$ is a constant.

The leading case in which Assumption B is satisfied is when $\alpha=1$, $1-\beta(L)L$ has all its roots outside the unit circle, and $x_t$ is omitted so $Z_t^1$ consists of p lags of $\Delta y_{t-1}$. In this case, Model II accounts for a break in the deterministic trend in the Dickey-Fuller regression (2.1). The formulation (3.1) generalizes this leading case to include additional mean zero stationary regressors other than $\Delta y_{t-j}$ and certain regressors that depend on k. For example, setting $x_t(k) = (\Delta y_t - \bar{\mu}_0)1(t>k)$ permits testing the null hypothesis that the coefficient on $\Delta y_{t-1}$ in Model II is constant against the alternative that it changes once at an unknown date. It can be verified that this definition of $x_t(k)$ satisfies Assumption B. Also, note that $Ex_t=0$ is assumed without loss of generality, as long as a constant is included in the regression.

The estimators and test statistics are computed using the full T observations, for $k=k_0, k_0+1, \ldots, T-k_0$, where $k_0=[T\delta_0]$. The resulting statistics are thus sequential rather than recursive. Let R be a $q\times(m+p+4)$ matrix of restrictions on $\theta$. The stochastic processes constructed from the sequential estimators and Wald test statistic are, for $\delta_0\le\delta\le 1-\delta_0$,

$$(3.4)\qquad \hat{\theta}(\delta) = (\textstyle\sum_{t=1}^T z_{t-1}([T\delta])z_{t-1}([T\delta])')^{-1}(\sum_{t=1}^T z_{t-1}([T\delta])y_t)$$

$$(3.5)\qquad \Upsilon_T(\hat{\theta}(\delta)-\theta) = \Gamma_T(\delta)^{-1}\Psi_T(\delta)$$

$$(3.6)\quad F_T(\delta) = [R\hat{\theta}(\delta)-r]'[R(\textstyle\sum_{t=1}^T z_{t-1}([T\delta])z_{t-1}([T\delta])')^{-1}R']^{-1}[R\hat{\theta}(\delta)-r]/q\hat{\sigma}^2(\delta)$$

where $\hat{\sigma}^2(\delta)=T^{-1}\sum_{t=1}^T(y_t-\hat{\theta}(\delta)'z_{t-1}([T\delta]))^2$, $\Gamma_T(\delta)=\Upsilon_T^{-1}\sum_{t=1}^T z_{t-1}([T\delta])z_{t-1}([T\delta])'\Upsilon_T^{-1}$ and $\Psi_T(\delta) = \Upsilon_T^{-1}\sum_{t=1}^T z_{t-1}([T\delta])\epsilon_t$. Here, $\Upsilon_T=\Upsilon_{AT}$ in case A and $\Upsilon_T=\Upsilon_{BT}$ in case B, where $\Upsilon_{AT} = \mathrm{diag}(T^{1/2}I_{p+m}, T^{1/2}, T, T^{3/2}, T^{3/2})$ and $\Upsilon_{BT} = \mathrm{diag}(T^{1/2}I_{p+m}, T^{1/2}, T, T^{1/2}, T^{3/2})$.

The next theorem provides asymptotic representations for the standardized sequential coefficients.

*Theorem 3.* Suppose that $y_t$ is generated according to Model II with $\mu_1=\mu_2=0$ and $\alpha=1$ and that Assumption B holds. Then:

a) In case A, $T_{AT}(\hat{\theta}(\cdot)-\theta) \Rightarrow \Gamma(\cdot)^{-1}\Psi(\cdot)$, where

$$\Psi(\delta) = \sigma[(\Sigma(\delta)^{\frac{1}{2}}B(1))',\ W(1),\ \int_0^1 J(\lambda)dW(\lambda),$$
$$(1-\delta)W(1)-\int_\delta^1 W(\lambda)d\lambda,\ W(1)-\int_0^1 W(\lambda)d\lambda]',$$

$\Gamma_{11} = \Sigma(\delta)$, $\Gamma_{1j} = 0$, $j=2,\dots,5$, $\Gamma_{22} = 1$, $\Gamma_{23}=\int_0^1 J(\lambda)d\lambda$, $\Gamma_{24} = \frac{1}{2}(1-\delta)^2$, $\Gamma_{25} = \frac{1}{2}$, $\Gamma_{33} = \int_0^1 J(\lambda)^2 d\lambda$, $\Gamma_{34} = \int_\delta^1(\lambda-\delta)J(\lambda)d\lambda$, $\Gamma_{35} = \int_0^1 \lambda J(\lambda)d\lambda$, $\Gamma_{44} = (1-\delta)^3/3$, $\Gamma_{45} = (1-\delta^3)/3-\frac{1}{2}\delta(1-\delta^2)$, and $\Gamma_{55} = 1/3$, where $W(\delta)$ is a standard Brownian motion process, $b=(1-\beta(1))^{-1}$, $J(\lambda) = b\pi H(\lambda) + \sigma b W(\lambda)$, $B(\lambda)$ is (p+m)-dimensional standard Brownian motion, and B is independent of (W,H).

b) In case B, $T_{BT}(\hat{\theta}(\cdot)-\theta) \Rightarrow \Gamma(\cdot)^{-1}\Psi(\cdot)$, where $\Psi$ is as in (a) except that $\Psi_4(\delta) = \sigma(W(1)-W(\delta))$, and where $\Gamma$ is as in (a) except for $\Gamma_{24} = 1-\delta$, $\Gamma_{34} = \int_\delta^1 J(\lambda)d\lambda$, $\Gamma_{44} = 1-\delta$, and $\Gamma_{45} = \frac{1}{2}(1-\delta^2)$, where B, W, J, and b are as defined in (a).

Several remarks are in order.

(a) When $x_t(k)$ does not appear as a regressor and $\delta$ is fixed, this reduces to the model and results presented in Perron (1989). Theorem 3 generalizes this result to the case in which the estimator and test statistic processes are random elements of $D[0,1]$, indexed by $\delta$. Note, however, that Perron considered the case of unknown (possibly infinite) AR order p, whereas here p is assumed to be finite and known.

(b) This result applies for $0<\delta_0\le\delta\le(1-\delta_0)<1$. Thus the test for the change in the coefficients is constrained not to be at the ends of the sample. In practice, this requires choosing a "trimming" value $k_0=[T\delta_0]$ to evaluate the process $F_T(\delta)$.

(c) Formal representations for the $F_T(\delta)$ statistic, or for sequential Dickey-Fuller statistics, obtain using the "$R^*$" device used in Theorem 1(b). The limiting process for $T^{\frac{1}{2}}(\hat{\theta}_1(\cdot)-\theta)$ and $F_T$ statistics testing restrictions on $Z^1_{t-1}$ can be thought of respectively as Gaussian and $\chi^2_q/q$ statistic processes, with the "marginals" of each process (for fixed evaluation points $\delta=\delta'$) having their respective Gaussian or $\chi^2_q/q$ distribution. Unlike the limiting processes obtained in Theorem 1, these processes are not adapted to the sequence of sigma fields generated by $\{W(\lambda), B(\lambda)\}$.

(d) This result provides joint uniform convergence of all the estimators and test statistics. Thus in particular it provides the asymptotic representation of continuous functions of one or more of these processes. One example is a rule considered by Christiano (1988): compute the Dickey-Fuller t-statistic in Model II, $t_{DF}(\delta)$, for $k_0/T \le \delta \le 1$, and let $t_{DF}^{min*} = \min_{k_0\le k\le T} t_{DF}(k/T)$.

(e) A related sequential statistic is the Quandt (1960) likelihood ratio statistic, which tests for a break in any or all of the coefficients. This entails estimating $2(T-k_0)$ separate regressions over the subsamples $1,\ldots,[T\delta]$ and $[T\delta]+1,\ldots,T$. The likelihood ratio statistic is computed for each possible break point, and the Quandt LR statistic is the maximum of these. Because each statistic involves the full sample, the Quandt LR is sequential. However, because of its use of subsample regressions, its asymptotic representation is obtained using the results of Theorem 1 for model I.

In the notation of Section 2, Quandt (1960) considered $-2\ln\hat{\lambda}^* = \max_{k_0\le k\le T-k_0} -2\ln\lambda(k)$, where $\lambda(k) = \hat{\sigma}^k_{1,k}\hat{\sigma}^{T-k}_{k+1,T}/\hat{\sigma}^T_{1,T}$ and $\hat{\sigma}_{1,k}$ is the standard error of the regression using observations $1,\ldots,k$, etc. Because $\hat{\sigma}_{1,[T\delta]} \Rightarrow \sigma$ under the null (a consequence of Theorem 1), the LR statistic is asymptotically

$$(3.7)\quad -2\ln\lambda([T\delta]) = -\sigma^{-2}\sum_{t=1}^{[T\delta]}(y_t-\hat{\theta}(\delta)'Z_{t-1})^2 - \sigma^{-2}\sum_{t=[T\delta]+1}^{T}(y_t-\tilde{\theta}(\delta)'Z_{t-1})^2 + \sigma^{-2}\sum_{t=1}^{T}(y_t-\hat{\theta}(1)'Z_{t-1})^2$$

where $\tilde{\theta}(\delta) = (\sum_{t=[T\delta]+1}^{T}Z_{t-1}Z_{t-1}')^{-1}(\sum_{t=[T\delta]+1}^{T}Z_{t-1}y_t)$ and $\hat{\theta}(\delta)$ is given in (2.3). By algebraic manipulation,

$$(3.8)\quad -2\ln\lambda([T\delta]) = \sigma^{-2}\{T_T(\hat{\theta}(\delta)-\theta)\}'V_T(\delta)\{T_T(\hat{\theta}(\delta)-\theta)\} + \sigma^{-2}\{T_T(\tilde{\theta}(\delta)-\theta)\}'\tilde{V}_T(\delta)\{T_T(\tilde{\theta}(\delta)-\theta)\} - \sigma^{-2}\{T_T(\hat{\theta}(1)-\theta)\}'V_T(1)\{T_T(\hat{\theta}(1)-\theta)\}$$

where $\tilde{V}_T(\delta) = T_T^{-1}\sum_{t=[T\delta]+1}^{T}Z_{t-1}Z_{t-1}'T_T^{-1}$. Using the identity $T_T(\tilde{\theta}(\delta)-\theta) = \{V_T(1)-V_T(\delta)\}^{-1}\{\phi_T(1)-\phi_T(\delta)\}$ and the results from Theorem 1, one obtains

$$(3.9)\quad -2\ln\lambda([T\cdot]) \Rightarrow \sigma^{-2}\phi(\cdot)'\{V(\cdot)^{-1}+(V(1)-V(\cdot))^{-1}\}\phi(\cdot) - \sigma^{-2}\phi(1)'\{V(1)^{-1}-(V(1)-V(\cdot))^{-1}\}\phi(1) - 2\sigma^{-2}\phi(\cdot)'(V(1)-V(\cdot))^{-1}\phi(1).$$

This specializes to Chu's (1989, Section 2) result for the Quandt LR statistic in the special case that only stationary regressors and a constant are included. Then $V(\delta)=\delta V(1)$, and from Theorem 1 $\sigma^{-1}V(1)^{-\frac{1}{2}}\phi(\delta) = W_{p+1}(\delta)$ (i.e., $\sigma^{-1}V(1)^{-\frac{1}{2}}\phi(\delta)$ is distributed as a (p+1)-dimensional standard Brownian motion). Algebraic manipulation then yields the representation, $-2\delta(1-\delta)\ln\hat{\lambda}([T\delta]) \Rightarrow W_{p+1}^{*}(\delta)'W_{p+1}^{*}(\delta)$, where $W_{p+1}^{*}(\delta) = W_{p+1}(\delta)-\delta W_{p+1}(1)$ is a (p+1)-dimensional standard Brownian Bridge. Thus

$$(3.10)\quad -2\ln\hat{\lambda}^{*} \Rightarrow \sup_{\delta_0\le\delta\le 1-\delta_0}\{W_{p+1}^{*}(\delta)'W_{p+1}^{*}(\delta)/(\delta(1-\delta))\}.$$

Chu (1989) also provides critical values, a Monte Carlo analysis, alternative related test statistics, and relates the result (3.10) to earlier ones in the literature. Hansen (1990) recently proposed a related "mean Chow" statistic, the average of these likelihood ratios, the distribution of which is the multivariate generalization of the limiting distribution of the Anderson-Darling (1954) statistic, subject to trimming.

(f) A statistic that will be used in the empirical analysis is the sequential t-statistic testing the hypothesis that the coefficient on $\tau_{1t}(k)$ is zero in case B ($\tau_{1t}(k)=1(t>k)$), under the maintained hypothesis that $\alpha=1$ and $\mu_2=0$. This corresponds to a shift in the intercept in a p-th order autoregression for $\Delta y_t$. The distribution of this statistic ($t_{\tau_1}(\cdot)$) is obtained from Theorem 3. Because $\Gamma(\cdot)$ is block diagonal, the asymptotic distribution of $t_{\tau_1}(\cdot)$ does not depend on the nuisance parameters $(\beta_1,\ldots,\beta_p)$ and has the limiting representation $t_{\tau_1}(\cdot) \Rightarrow t^*_{\tau_1}(\cdot)$, where $t^*_{\tau_1}(\delta) = W^*_1(\delta)/(\delta(1-\delta))^{\frac{1}{2}}$ and $W^*_1(\delta) = W(\delta)-\delta W(1)$ is the 1-dimensional Brownian bridge.

## 4. Monte Carlo Results

This section reports asymptotic critical values and examines the size and power of selected recursive and sequential statistics. All statistics (except $t_{\tau_1}(\cdot)$ in the restricted ($\alpha=1$, $\mu_2=0$) model) include $(1,t)$ as regressors to allow for a possible time trend under the alternative.

The first five statistics examined are recursive tests for unit roots: the full-sample Dickey-Fuller statistic, $t_{DF}$ ($=t_{DF}(1)$ in the notation of (2.6)); the maximal Dickey-Fuller statistic, $t^{max}_{DF}=\max_{k_0\le k\le T} t_{DF}(k/T)$; the minimal Dickey-Fuller statistic, $t^{min}_{DF}=\min_{k_0\le k\le T} t_{DF}(k/T)$; $t^{diff}_{DF} = t^{max}_{DF}-t^{min}_{DF}$;

and the minimal value of the recursive Modified Sargan-Bhargava statistic, $\hat{R}^{min} = \min_{k_0 \leq k \leq T} \hat{R}(k/T)$, where $\hat{R}(\delta)$ is defined in (2.8). Asymptotic critical values, computed by Monte Carlo integration, are reported in Table 1 for $\delta_0=0.25$. Not surprisingly, the critical value for $t_{DF}^{min}$ is well below the full-sample Dickey-Fuller critical value.

An initial examination of the size of these recursive statistics is reported in Table 2. In this experiment, T=100, the true model is an AR(1) in first differences, and a single lag of $\Delta y_t$ is included in the regression (2.1). The first three statistics have sizes near their asymptotic levels, but the size for $t_{DF}^{diff}$ substantially exceeds its level. The size of $\hat{R}^{min}$ deteriorates sharply as the autocorrelation increases. This is surprising in light of the very good size performance of $\hat{R}(1)$ found in Stock (1988).

Table 3 examines the power of these recursive statistics against the hypothesis that the largest root shifts, being less than one in half the sample and equaling one in the other half. The recursive statistics $t_{DF}^{max}$ and $t_{DF}^{min}$ have somewhat better power than $t_{DF}$ when the root is initially less than one, but each has power less than its level against the alternative that the root is initially one, then less than one. In this latter case, $t_{DF}^{diff}$ often detects a shift, although it does not in the case that the other three statistics have power. One interpretation of these findings is that when the root is initially one, the recursive sample moments for k after the true break point are dominated by the initial, large sample moments. However, because of the coefficient shift, $\hat{\sigma}(\delta)$ tends to increase, driving $t_{DF}(\delta)$ towards zero. In any case, no single recursive statistic seems to provide a reliable test against this structural break alternative.

The second set of statistics consists of three sequential statistics: the maximum of the sequential F-statistics, $F_T^{max} = \max_{k_0 \leq k \leq T-k_0} F_T(k/T)$ testing the hypothesis that $\mu_1=0$; the sequential Dickey-Fuller statistic evaluated at the value of k ($\hat{k}$, equivalently $\hat{\delta}$) that maximizes $F_T(\delta)$, $t_{DF}(\hat{\delta})$; and $t_{DF}^{min*}=\min_{k_0\leq k\leq T-k_0} t_{DF}(k/T)$, the minimal Dickey-Fuller statistic over all the sequentially computed Dickey-Fuller statistics. These statistics were proposed by Christiano (1988) to extend Perron's analysis to the case in which k is unknown. We consider Case A (the trend-shift model (3.2)) in detail and present asymptotic critical values for Case B (the mean-shift model (3.3)).

Critical values for these statistics in Case A are reported in Table 4 for $\delta_0=.15$. The large critical value of the maximal F-statistic also was found by Christiano (1988). Table 5 examines the size of these procedures when the true model has no trend break and follows an AR(1) in first differences. In each case, the size is approximately the level of the test.

Table 6 reports the power of these sequential tests against some trend-shift/stationary alternatives. The trend shift is calibrated so that the change in the slope is large, 20% or 40% of the standard deviation of the innovation. These results suggest several generalizations. First, the maximal F-statistic detects the trend break with high probability, particularly if it occurs later in the sample. Second, both the unit root tests reject with high probability, more often if the break occurs early than late in the sample. Third, the break point is identified rather accurately, in the sense that a large fraction of the estimated break points $\hat{k}$ occur within $\pm$ .05T of the true break point. Fourth, the full-sample Dickey-Fuller statistic fails to detect stationarity around a shifting trend, particularly if the break is in the second half of the sample. This confirms Perron's (1989) results and interpretation: the permanent shift in the deterministic trend is mistaken for a persistent innovation to a stochastic trend.

Table 7 provides asymptotic critical values when there is a shift in the mean (Case B). The final column reports critical values for the maximal absolute t-statistic on $\tau_{1t}(k)$ for the restricted model with $\mu_2$=0 and $\alpha$=1.

## 5. Empirical Results

The previous results are used here to examine whether shifts or breaks in trends provide a suitable model for the apparent persistence in seasonally adjusted output in seven OECD countries: Canada, France, West Germany, Italy, Japan, the United Kingdom, and the United States. Log real output for each of the countries are graphed in Figure 1. The data sources are discussed in Appendix B. As in the previous section, the recursive statistics were computed using trimming of 25% ($\delta_0$=.25), and the sequential statistics used $\delta_0$=.15.

*Results.*

We first computed full-sample statistics for the seven countries, modeling each series as an AR(2) or AR(4) in first (or second) differences. The corresponding Dickey-Fuller statistics are tests of the null hypothesis of one (respectively, two) unit root; a constant and a time trend were included in the regression with first differences, a constant was included in the regression with second differences. For each of the seven countries, the hypothesis of one unit root could not be rejected, but in each case the hypothesis of two unit roots could be rejected. We therefore adopted the single unit root model (with nonzero drift) as the interesting null hypothesis for each of the seven countries and proceeded with the computation of other statistics using models (2.1) and (3.1), where the order of the lag polynomial $\beta$(L) is 4 and the null hypothesis is the existence of a single unit root. In what follows, (3.1) is specified with no additional regressors ($x_t$).

The recursive Dickey-Fuller t-statistic is plotted in Figure 2; the dashed lines are the 10% and 5% critical values for $t_{DF}^{min}$. The recursive statistics are summarized in Table 8. In no case is the standard non-recursive Dickey-Fuller statistic significant at the 25% level, and neither $t_{DF}^{min}$ nor $t_{DF}^{max}$ rejects the unit root null at the 15% (asymptotic) level. For all countries but Italy, the recursive statistics provide no evidence against the unit root null. For Italy the evidence is mixed: the minimal recursive and full-sample $t_{DF}$ statistics provide no evidence against the unit root null, but $\hat{R}^{min}$ is significant at the 10% (but not 5%) level. However, the poor size performance of $\hat{R}^{min}$ in Section 4 suggests caution in interpreting this result.

The trend shift hypothesis is examined in the first columns of Table 9, and the sequential Dickey-Fuller t-statistics are graphed in Figure 3. The dashed lines in Figure 3 are the 10% and 5% significance levels for the minimal sequential Dickey-Fuller t-statistic. The unit-root/no-break null can be rejected against the trend-shift/stationary alternative for only one country, Japan; the rejection is at the 5% significance level. The same conclusions obtain whether the minimal sequential Dickey-Fuller statistic or $t_{DF}(\hat{\delta})$ is used, where $t_{DF}(\hat{\delta})$ is the Dickey-Fuller statistic evaluated at the value of $\delta$ that maximizes the F-statistic on the break term. Although the break points vary across countries, for France, Italy and possibly Germany they are clustered in 1971-1973. There is some evidence against the unit root null in the case of Canada as well: the p-value for the full-sample $t_{DF}$ in Table 8 was .62, but the p-value for the sequential statistics is .12. In the other countries, the p-values do not change markedly from those in Table 8.

The mean-break model (model B), examined in the second part of Table 9, presents a somewhat different picture. The sequential Dickey-Fuller statistics for this model are plotted in Figure 4. Only one country, Canada, indicates a rejection of the unit root null at the 5% level, with the break in 1981:3.

The final statistic examines the possibility that output for these countries is I(1), but that there has been a single shift in the mean growth rate. This corresponds to Model B of Section 3, with the imposition of the unit root null ($\alpha=1$) and a zero time trend in first differences ($\mu_2=0$). This model is examined by computing the maximal t-statistic on $\tau_{1t}(k)=1(t>k)$ in Model B under the restrictions that $\alpha=1$ and $\mu_2=0$. This statistic and 5% critical values are plotted in Figure 5. (Movement outside the critical band indicates rejection at the 5% level using the maximal absolute t-statistic). In four cases -- France, Germany, Italy, and Japan -- the restriction of a constant drift is rejected in favor of the hypothesis of a shift in the drift.

*Sensitivity analysis*

We performed three additional analyses to investigate the sensitivity of the results to potential measurement errors or anomalies in the data. The results are briefly summarized here; details are available from the authors upon request. First, in 1968:2 France experienced a major strike, over which we have interpolated for the reported results. When the original data are used, the results do not change, except for the sequential mean-shift statistics; then $t_{DF}^{min*}$ rejects the constant-drift unit root null at the 5% level, with the break sharply identified as 68:2. We view this as an artifact: reversion to "trend" after the strike is best thought of rather prosaically as people returning to work, not as reflecting trend stationary behavior in the long-run factors driving French economic growth, such as technical progress and labor productivity.

Second, the results for Germany are for 1950:1-1989:2; for Japan, 1952:1-1989:2. Because these data start near the end of World War II, the earliest observations might have unusually large measurement error. But the conclusions about the significance of unit root tests do not change upon repeating the analysis over 1955:1-1989:2, although the evidence against the constant-drift/unit-root null in favor of the mean-shift/unit-root alternative (using $|t_{\tau_1}(\hat{\delta})|$) is less strong for Germany.

Third, the computations were repeated using 8 rather than 4 lags for the Dickey-Fuller regressions and for the autoregressive spectral density estimator used to construct $\hat{R}(\delta)$. The qualitative results are largely unchanged. The most notable differences are: for Canada, the unit root null is no longer rejected in favor of the mean-shift/stationary alternative (Case B) -- the p-value is now .14; for the U.K., the unit root null is rejected against Case B (p-value = .06).

*Summary*

These results suggest rather different characterizations of the long-run properties of output for these countries. In two countries, Canada and Japan, there is evidence against the unit root null using these statistics. For Canada, the unit root null is rejected against the mean-shift/stationary alternative, with the break point in 1981:3. Thus the recession of the early 1980's is represented as a permanent downward shift in trend growth; after the recovery, output again is stationary along its original growth path.

The results for Japan indicate that the unit root null is rejected against the trend-shift/stationary alternative, with the break in 1969:4. This shift is apparent in Figure 1: from 1952:1 to 1969:4, on average Japanese output grew at 9.2% per year; since then, it has grown at 4.4%.

The results for the remaining countries provide no evidence against the unit root hypothesis. Based on the restricted mean-shift model of Figure 5, however, output for France, Germany, and Italy seem better characterized as being integrated, but having suffered a permanent reduction in the rate of growth of output. For Italy and France, this slowdown appears around 1974, the time of the first oil shock. For Germany, the sequential t-statistic is less precise in identifying a specific break point, although the statistic is significantly negative just before 1974. For these countries, then, output is well characterized as being integrated, but with average growth that is slower over the period of the productivity slowdown.

The results for the U.K. are qualitatively different, providing no evidence against the unit root null. The t-statistics in Figure 5 do not indicate a statistically significant productivity slowdown; indeed, the growth rate increased in the 80's, although not significantly so using these procedures.

The results for the U.S. indicate no rejections of the unit root null against any of the various hypotheses. These results parallel Christiano's (1988) failure to reject the no-break/unit-root hypothesis using bootstrapped critical values. They accord with Banerjee, Dolado and Galbraith's (1990) failure to reject the trend-break alternative for the U.S. (using the uniform critical values tabulated in Section 4) for longer annual data series that include the Depression. They are also consistent with the findings in Zivot and Andrew's (1989) closely related paper, especially with their failure to reject the unit root null against a trend break alternative for real GNP using the longer Nelson-Plosser (1982) data when uniform ($t_{DF}^{min*}$) critical values are used. Although Perron (1989) finds evidence for a trend break -- and a rejection of the unit root -- in 1973, this conclusion is based on the assumption that the break point is known *a priori*; when the break point is treated as unknown, the evidence is much weaker.

*Comparison with previous literature and discussion*

Several recent papers extend Cochrane's (1988) study of persistence in U.S.

output to international data. Each study differs in its sample period and, to various degrees, in the statistical measure used. Although point estimates of persistence are not comparable across studies, relative rankings are.

Campbell and Mankiw (1989) examined the same seven countries considered here over 1957-1986 and measured persistence by the size of a (bias adjusted) variance ratio for long (5-10 year) differences. They concluded that persistence in the U.K. was less than in the U.S., but greater in each of the other countries. Cogley (1990) computed modified variance ratios over 1870-1985 for nine countries, including Canada, France, Italy, the U.K., and the U.S. Although his data set is much longer, his conclusions are similar to Campbell and Mankiw's: the U.S. exhibited the least persistence, followed by Canada; the largest variance ratios were for France and Italy. Kormendi and Meguire (1988) also used variance ratios to analyze long annual data on 12 countries and postwar data for 32 countries, including Canada, France, Germany, Italy, the U.K., and the U.S. Of these six, using bias-unadjusted measures they too found U.S. output to exhibit the least persistence (the smallest variance ratios), with French, German, and Italian output exhibiting the most persistence.

Clark (1989) used a different technique -- a stochastic trend-cycle decomposition of the form studied by Harvey (1985) and Watson (1986) -- to study the relative importance of "cyclical" components for the seven countries we consider, over approximately 1960-1986. A notable feature of his results is that an I(1) trend fit well for five of the countries; for France and especially Japan, however, the fit of the model was substantially improved when an I(2) trend (a stochastic trend with a random walk drift) was introduced. He interprets this as providing a flexible way to account for the slower growth in these two countries in the latter half of the sample.

The striking feature of the variance ratio results is that in each study the variance ratios are highest in the countries for which we identify deterministic breaks (with either I(1) or I(0) stochastic components) -- Japan, Germany, France, and Italy -- and lowest for the U.S., for which we find no evidence against the no-break unit root hypothesis. Also, Clark finds evidence of extreme persistence, in the form of an integrated drift, for France and Japan. These results have a consistent explanation. If the deterministic-break specification is valid, then this break is by definition highly persistent. If the break is not taken into account explicitly, then it will be misidentified by variance ratio statistics as a large but otherwise typical shock to output. In a model in which the drift is forced to be either constant or integrated (Clark's [1989] model), a sufficiently large permanent change will be modeled as an integrated drift. If the deterministic break view is correct, variance ratio statistics could give quite misleading views of persistence.

Interpreted more broadly, these results suggest that not all shocks to output are the same: the shocks associated with the oil crisis of 1974-5 were considerably more persistent than other shocks before and after, so much so that our procedures classify them as deterministic breaks rather than a large negative realization. This treats the single-break model as a simple way to separate massive, economy-changing shocks -- the Depression, World War II, the productivity slowdown in the 1970's -- from the other shocks to output that, while persistent, exhibit less permanence. To us, interpreting these broken trends as deterministic is unsatisfying, for this conditions on elements of aggregate activity, such as productivity growth rates or changes in fiscal or (possibly) monetary management rules, that are unpredictable. We instead prefer to interpret these rejections as metaphors for these countries having long-run trends that are smooth with occasional large shocks (see Perron (1989, 1990a)). This is compatible with Blanchard and Watson's (1986) "large shock/small shock" hypothesis, which they developed using U.S. data -- except that this better describes Japan and is more accurately termed the "persistent shock/less persistent shock" hypothesis. It is also compatible with Hamilton's (1989) model of random regime switches, although the regimes here last much longer than the business-cyle switches Hamilton identified for U.S. GNP. In Japan, the trend growth rate dropped around 1970. In contrast, the Canadian trend growth rate has stayed relatively constant, but the growth path shifted downward in the early 1980's.

## 6. Summary and Conclusions

The results in Sections 2 and 3 provide a framework for evaluating the asymptotic distributions of several recursive or sequential statistics. These results resolve some open questions, such as the distribution of the process of the recursive least squares estimators. There are, however, several theoretical questions that these results only begin to address. In particular, each of the test statistics discussed here has the flavor of general diagnostic tests against possible changes in coefficients, either autoregressive coefficients or coefficients describing the deterministic components of the process. This suggests the value of obtaining formal results on the power of these tests against various structural break alternatives.

For Germany, Italy, and France, these results provide a new characterization of the productivity slowdown as being a reduction in the rate of growth of the I(1) output process over this period; there is no evidence in favor of the trend-break/stationary hypothesis for any of these countries. This slowdown occurred at approximately the same time for each of these countries, 1974. The analysis also provides new insights into the trend growth of Japan and Canada. For Japan, the unit root null is rejected against the alternative that output is stationary around a trend that slowed significantly around 1970. For Canada, the unit root null is rejected against the alternative that the linear trend growth path shifted downward after the 1979-1982 recession, although this is sensitive to the number of autoregressive lags used. Finally, the empirical analysis provides little evidence against the unit root null for the U.K. and -- in contrast to Perron (1989) -- none for the U.S.

## Appendix A: Proofs of Theorems

Proof of Theorem 1

To simplify notation, it is assumed that $\bar{\mu}_0=0$. This is done without loss of generality, since $Z_t^1 = (\Delta y_t-\bar{\mu}_0,\ldots,\Delta y_{t-p+1}-\bar{\mu}_0)$ and $Z_t^3 = (y_t-\bar{\mu}_0 t)$. Also, throughout the notation $\phi_T(\delta)\Rightarrow\phi(\delta)$ or $\phi_T(\cdot)\Rightarrow\phi(\cdot)$ is used interchangeably.

(a) First consider $\phi_T$. Let $C(L) = (1-\beta(L)L)^{-1}$ so that $\Delta y_t = C(L)\epsilon_t$, let $b = C(1) = (1-\beta(1))^{-1}$, and let $s=[T\delta]$. The uniform convergence results $\phi_{2T}(\delta) = T^{-\frac{1}{2}}\sum_{t=1}^{s}\epsilon_t\Rightarrow\sigma W(\delta)$ and $\phi_{4T}(\delta) = T^{-3/2}\sum_{t=1}^{s}t\epsilon_t \Rightarrow \sigma[\delta W(\delta)-\int_0^\delta W(\lambda)d\lambda]$ are immediate consequences of Assumption A and the Functional Central Limit Theorem (FCLT) (Herrndorf [1984]; also see Hall and Heyde [1980]). Consider $\phi_{3T}$ and write $y_t=C(1)\xi_t+U_t$, where $\xi_t=\sum_{r=1}^{t}\epsilon_r$ and $U_t=C^*(L)\epsilon_t$, where $C^*(L) = (1-L)^{-1}[C(L)-C(1)]$. Then

$$\phi_{3T}(\delta) = T^{-1}\sum_{t=1}^{s}y_{t-1}\epsilon_t = C(1)T^{-1}\sum_{t=1}^{s}\xi_{t-1}\epsilon_t + T^{-1}\sum_{t=1}^{s}U_{t-1}\epsilon_t.$$

Now $T^{-1}\sum_{t=1}^{s}\xi_{t-1}\epsilon_t = \frac{1}{2}(T^{-1}\xi_s^2-T^{-1}\sum_{t=1}^{s}\epsilon_t^2)$. Because $\nu_t = \epsilon_t^2-\sigma^2$ is a martingale difference sequence (MDS) with $\sup_t E|\nu_t|^{2+\gamma}\leq\bar{\kappa}<\infty$ by Assumption A, $T^{-1}\sum_{t=1}^{s}\epsilon_t^2 = (s/T)\sigma^2 + T^{-1}\sum_{t=1}^{s}\nu_t \Rightarrow \delta\sigma^2$. Because $T^{-1}\xi_s^2 \Rightarrow \sigma^2W(\delta)^2$, $T^{-1}\sum_{t=1}^{s}\xi_{t-1}\epsilon_t \Rightarrow \frac{1}{2}\sigma^2(W(\delta)^2-\delta)$. Because $U_{t-1}\epsilon_t$ is a MDS with $\sup_t E|U_{t-1}\epsilon_t|^4<\infty$ (this from the moment assumptions on $\epsilon_t$ and from the 1-summability of C(L) [e.g. Stock (1987)] and thus the absolute summability of $C^*(L)$), $T^{-\frac{1}{2}}\sum_{t=1}^{s}U_{t-1}\epsilon_t$ obeys a FCLT. Thus $T^{-1}\sum_{t=1}^{s}U_{t-1}\epsilon_t \Rightarrow 0$, so $\phi_{3T}(\delta) \Rightarrow \frac{1}{2}b\sigma^2(W(\delta)^2-\delta)$. Finally, because $Z_{t-1}^1\epsilon_t$ is a MDS with $\sup_t E(Z_{it-1}^1\epsilon_t)^4 \leq \kappa_4^2(\sum_{j=0}^{\infty}|c_j|)^4 < \infty$, $\phi_{1T}(\delta) = T^{-\frac{1}{2}}\sum_{t=1}^{s}Z_{t-1}^1\epsilon_t \Rightarrow \sigma B(\delta)$, where $B(\delta)$ has covariance matrix $EZ_t^1Z_t^{1\prime} = \Omega_p$. It follows from Chan and Wei (1988, Theorem 2.2) that B and W are independent.

Next consider $V_T(\delta)$. The uniform convergence of each element of $V_T$, with the exception of $V_{11T}$ and $V_{13T}$, either obtains by direct calculation or is a consequence of $T^{-\frac{1}{2}}y_s \Rightarrow b\sigma W(\delta)$ and the continuous mapping theorem. For example, $V_{33T}(\delta) = T^{-2}\sum_{t=1}^{[T\delta]}y_t^2 \Rightarrow b^2\sigma^2\int_0^\delta W(r)^2dr$. Note however that this formally holds only for $\delta$ fixed; to show convergence of the process $V_{33T}(\cdot) \Rightarrow V_{33}(\cdot)$, it must further be shown that $g(\delta;f) = \int_0^\delta f(r)^2dr$ is a continuous mapping from $D[0,1]$ to $D[0,1]$. This argument is made in Zivot and Andrews (1989) and is not repeated here.

To demonstrate the convergence of $V_{11T}(\delta)$ it is sufficient to consider its (1,1) element; the argument for the other elements is similar. Now $(V_{11T})_{11} = T^{-1}\Sigma_{t=1}^{s}(\Delta y_t)^2$ (recall $\bar{\mu}_0=0$). Define $\gamma_0=E(\Delta y_t)^2$, $X_t=(\Delta y_t)^2-\gamma_0$, and $S_i=\Sigma_{t=1}^{i}X_t$. With these definitions,

$$T^{-1}\sum_{t=1}^{i}(\Delta y_t)^2 = (i/T)\gamma_0 + T^{-1}S_i, \quad i=1,\ldots,T.$$

Thus the desired result follows if $\Pr[\max_{i\leq T}|T^{-1}S_i| > \delta] \to 0$ for all $\delta>0$. This will be shown by first showing that $X_t$ is a mixingale and second applying the mixingale extension of Doob's inequality.

From Hall & Heyde (1980, p. 19), $X_t$ is a mixingale if there exists sequences of nonnegative constants $d_t$ and $\Psi_m$ such that $\Psi_m \to 0$ as $m \to \infty$ and

(i) $\|E(X_t|F_{t-m})\|_2 \leq \Psi_m d_t$

(ii) $\|X_t-E(X_t|F_{t+m})\|_2 \leq \Psi_{m+1}d_t$

for all $t\geq1$ and $m\geq0$, where $\|X\|_2=(EX^2)^{\frac{1}{2}}$. Condition (ii) is automatically satisfied by $\{X_t\}$ because $\{F_t\}$ is an increasing sequence of $\sigma$-fields and $X_t$ is an adapted stochastic process so that $E(X_t|F_{t+m})=X_t$; thus $\|X_t-E(X_t|F_{t+m})\|_2=0$.

Next turn to condition (i). Now

$$(\|E(X_t|F_{t-m})\|_2)^2 = E((E(X_t|F_{t-m}))^2)$$
$$= E((E[(C(L)\epsilon_t)^2 - \gamma_0|\epsilon_{t-m},\epsilon_{t-m-1},\ldots])^2)$$

$$= E([\sum_{j=0}^{m-1}c_j^2\sigma^2 + (\sum_{j=m}^{\infty}c_j\epsilon_{t-j})^2 - \sum_{j=0}^{\infty}c_j^2\sigma^2]^2)$$
$$= E([(\sum_{j=m}^{\infty}c_j\epsilon_{t-j})^2 - \sum_{j=m}^{\infty}c_j^2\sigma^2]^2)$$
$$= E[(\sum_{j=m}^{\infty}c_j\epsilon_{t-j})^4] - (\sum_{j=m}^{\infty}c_j^2\sigma^2)^2$$
$$= \sum_{j=m}^{\infty}\sum_{k=m}^{\infty}\sum_{\ell=m}^{\infty}\sum_{r=m}^{\infty}c_jc_kc_\ell c_r E(\epsilon_{t-j}\epsilon_{t-k}\epsilon_{t-\ell}\epsilon_{t-r}) - (\sum_{j=m}^{\infty}c_j^2\sigma^2)^2$$
$$\le \kappa_4(\sum_{j=m}^{\infty}|c_j|)^4$$

where the final inequality obtains by Assumption A. Now $\sum_{j=m}^{\infty}|c_j| \le K_1\sum_{j=m}^{\infty}\lambda^j = K_1\lambda^m/(1-\lambda)$, where $K_1$ is a constant and $\lambda$ is the absolute value of the largest root of $1-L\beta(L)$, which satisfies $|\lambda|<1$ by assumption. Thus $\|E(X_t|F_{t-m})\|_2 \le d_t\Psi_m$, where $d_t = [\kappa_4K_1^4/(1-\lambda)^4]^{1/2}$ and $\Psi_m = \lambda^{2m}$. Thus condition (i) is satisfied with $\Psi_m\to 0$ as $m\to\infty$ and $(\Delta y_t)^2$ is a mixingale.

Next, apply Chebyschev's inequality and the mixingale extension of Doob's inequality (Hall and Heyde [1980], Lemma 2.1) to show that $T^{-1}S_{[T\lambda]} \Rightarrow 0$. The condition of this lemma is that $\Psi_m$ be $O(m^{-1/2}(\log(m))^{-2})$, which is satisfied here. Thus

$$\begin{aligned}\Pr[\max_{i\le T}|T^{-1}S_i| > \delta] &\le T^{-2}\delta^{-2}E[\max_{i\le T}S_i^2] \\ &\le \delta^{-2}K_2T^{-2}\sum_{i=1}^{T}d_i^2 \\ &= \delta^{-2}K_2(\kappa_4K_1^4/(1-\lambda)^4)/T\end{aligned}$$

(where $K_2$ is a constant) which tends to zero for all $\delta>0$, where the second inequality obtains by Lemma 2.1 of Hall and Heyde. Thus $T^{-1}S_{[T\lambda]} \Rightarrow 0$ so $T^{-1}\Sigma_{t=1}^{[T\delta]}(\Delta y_t)^2 \Rightarrow \delta\gamma_0$.

The final term is the p×1 vector $V_{13T}$, of which consider the i-th element, $(V_{13T}(\delta))_i$. Recall that by assumption $Z_t=0$, $t\le 0$. For $s=[T\delta]$,

$$(V_{13T}(\delta))_i = T^{-3/2}\sum_{t=1}^{s}y_{t-1}\Delta y_{t-i}$$

$$= T^{-3/2}\sum_{t=1}^{s}(y_{t-i-1} + \sum_{r=1}^{i}\Delta y_{t-r})\Delta y_{t-i}$$
$$= T^{-3/2}\sum_{t=1}^{s-i}y_{t-1}\Delta y_t + \sum_{r=1}^{i}T^{-3/2}\sum_{t=1}^{s-r}\Delta y_t\Delta y_{t-i+r} .$$

Thus $T^{½}(V_{13T}(\delta))_i = \nu_{iT}^1(\delta) + \nu_{iT}^2(\delta)$, $i=1,\ldots,p$, where

$$\nu_{iT}^1(\delta) = ½((T^{-½}y_{[T\delta]-i})^2 - T^{-1}\sum_{t=1}^{[T\delta]-i}(\Delta y_t)^2)$$
$$= ½((T^{-½}y_{[T\delta]-i})^2 - (V_{11T}(([T\delta]-i)/T))_{11})$$
$$\nu_{iT}^2(\delta) = \sum_{r=1}^{i}T^{-1}\sum_{t=1}^{s-r}\Delta y_t\Delta y_{t-i+r}$$
$$= \sum_{j=1}^{i}(V_{11T}(([T\delta]-i+j)/T))_{1j} .$$

It follows from $V_{11T}(\delta) \Rightarrow V_{11}(\delta)$ (shown above) and (for fixed i) $T^{-½}y_{[T\delta]-i} \Rightarrow b\sigma W(\delta)$ that $\nu_{iT}^1(\delta) \Rightarrow ½(b^2\sigma^2W(\delta)^2 - (V_{11}(\delta))_{11})$ and $\nu_{iT}^2(\delta) \Rightarrow \sum_{j=1}^{i}(V_{11}(\delta))_{1j}$. Thus $(V_{13}(\cdot))_i \Rightarrow 0$, , $i=1,\ldots,p$, so $V_T(\cdot) \Rightarrow V(\cdot)$.

(b) Given the convergence results in (a) and the moment conditions in Assumption A, it follows that $\hat{\sigma}^2(\delta) \Rightarrow \sigma^2$. The asymptotic representation in Theorem 1 follows directly from the results in (a) and from Theorem 2 of Sims, Stock and Watson (1990).

(c) Part (c) follows directly from (a) and (b). □

Proof of Theorem 2

This follows from the proof of Theorem 2.1 of Stock (1988) and from arguing, as in the proof of Theorem 1(a), that $g(\delta;f) = \int_0^\delta f(r)^2dr$ is a continuous mapping from $D[0,1]$ to $D[0,1]$. □

Proof of Theorem 3

(a) The proof is similar to the proof of Theorem 1. First consider $\Gamma$. The convergence of $\Gamma_{11T}$ and $\Gamma_{13T}$ is by Assumption B(ii). The results involving exclusively deterministic terms ($\Gamma_{ijT}$, i,j = 2,4,5) obtain by direct calculation. The remaining limits, which involve $y_t$, obtain by noting that under $H_0$ ($\alpha$=1, $\mu_1$=$\mu_2$=0), $\Delta y_t - \bar{\mu}_0 = C(L)\omega_0' x_{t-1} + C(L)\epsilon_t$, where $C(L) = (1-L\beta(L))^{-1}$. Thus by assumption B and the fact that C(L) is 1-summable,

$$T^{-1/2}\sum_{t=1}^{[T\lambda]}(\Delta y_t - \bar{\mu}_0) \Rightarrow b\pi H(\lambda) + b\sigma W(\lambda) = J(\lambda).$$

The results for the remaining terms follow from this limit and Assumption B.

Next consider $\Psi$. The terms $\Psi_{2t}$ and $\Psi_{4T}$ follow directly. For example, in case A, $\Psi_{4T}(\delta) = T^{-3/2}\sum_{t=1}^{T}(t-k)1(t>k)\epsilon_t = T^{-3/2}\sum_{t=k}^{T}(t-k)\epsilon_t$. The result $\Psi_{4T}(\delta) \Rightarrow \Psi_4(\delta)$ (for $k/T \to \delta$) obtains from this final expression by applying Assumption B and the FCLT. Because $Z_{t-1}^1\epsilon_t$ is a MDS with $2+\gamma$ moments, $T^{-1}\sum_{t=1}^{T}Z_{t-1}^1\epsilon_t \Rightarrow \sigma\Sigma(\delta)^{1/2}B(1)$. The independence of B(1) and (H,W) follows from Chan and Wei (1988), Theorem 2.2. The convergence $\Psi_{3T} = T^{-1}\sum_{t=1}^{T}(y_{t-1} - \bar{\mu}_0(t-1))\epsilon_t \Rightarrow \int_0^1 J(\lambda)dW(\lambda)$ follows from Chan and Wei (1988), Theorem 2.4.

(b) The argument for Case B is analogous to that for Case A.

These calculations formally show convergence of the processes $\Gamma_T(\cdot)$ and $\Psi_T(\cdot)$. The step from these results to results for $T_T(\hat{\theta}(\cdot)-\theta)$ and, e.g., $\inf_{\delta_0\leq\delta\leq 1} t_{DF}(\delta)$ requires showing that these are continuous mappings from $D[0,1] \to D[0,1]$ and $D[0,1] \to \Re^1$, respectively. This argument is made in Zivot and Andrews (1990). □

## Appendix B: Data sources

Data for the United States are GNP from Citibase, for 1947:I to 1989:II. The data for the other six countries come from two sources, the OECD Main Economic Indicators database maintained by Data Resources, Inc. (DRI), and Moore and Moore (1985). In most cases, two series have been spliced together to construct a longer time series of data. Where this has involved an adjustment because the real series are indexed to different base years, they have been adjusted using the earliest available ratio of the two series.

The Canadian data are GNP, with 1948:I to 1960:IV from Moore and Moore and 1961:I to 1989:II from DRI. The French data are GDP, 1963:I to 1989:II, and are from DRI. The French data contain a large negative spike (a strike) in 1968:II; we eliminated this spike by linearly interpolating the value for this quarter. The data for Germany are GNP, with 1950:I to 1959:IV from Moore and Moore and 1960:I to 1989:II from DRI. The data for Italy from DRI were nominal rates, so we have used GDP from Moore and Moore for 1952:I to 1982:IV. The GNP data for Japan is from Moore and Moore for 1952:I to 1964:IV and from DRI for 1965:I to 1989:II. The data for the UK are GDP at Factor Cost and are from DRI for 1960:I to 1989:II.

All data were seasonally adjusted at the source.

## Footnotes

1. Since writing the first draft of this paper, we have learned of six independent contemporaneous treatments of various aspects of the recursive coefficients and trend break problems: Andrews (1989), Bates (1990), Chu (1989), Hansen (1990), Perron (1990b), and Zivot and Andrews (1989). Discussion of these papers has been incorporated into the relevant sections of this revision.

## References


Anderson, T.W. and D.A. Darling (1954), "A Test of Goodness of Fit," *Journal of the American Statistical Assocation* 49, 765-769.

Andrews, Donald W.K. (1989), "Tests for Parameter Instability and Structural Change with Unknown Change Point," manuscript, Cowles Foundation, Yale University.

Banerjee, Anindya, Juan Dolado, and John W. Galbraith (1990), "Recursive Tests for Unit Roots and Structural Breaks in Long Annual GNP Series," manuscript, University of Florida.

Bates, Charles E. (1990), "Structural Stability and Trends in Economic Time Series," manuscript, Johns Hopkins University.

Bhargava, A. (1986), "On the Theory of Testing for Unit Roots in Observed Time Series," *Review of Economic Studies* 53, no. 174, 369-384.

Blanchard, Olivier J., and Mark W. Watson (1986), "Are Business Cycles All Alike?," in R. Gordon (ed.), *The American Business Cycle: Continuity and Change*. Chicago: University of Chicago Press.

Brown, R.L., J. Durbin, and J.M. Evans (1975), "Techniques for Testing the Constancy of Regression Relationships over Time with Comments," *Journal of the Royal Statistical Society*, Series B, 37, 149-192.

Campbell, John Y., and N. Gregory Mankiw (1989), "International Evidence on the Persistence of Economic Fluctuations," *Journal of Monetary Economics* 23, 319-333.

Chan, N.H. and C.Z. Wei (1988), "Limiting Distributions of Least Squares Estimates of Unstable Autoregressive Processes," *The Annals of Statistics* 16, no. 1, 367-401.

Christiano, Lawrence J. (1988), "Searching for a Break in GNP," Federal Reserve Bank of Minneapolis Working Paper No. 416.

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

Clark, Peter K. (1989), "Trend Reversion in Real Output and Unemployment," *Journal of Econometrics* 40, 15-32.

Cochrane, John H. (1988), "How Big is the Random Walk in GNP?" *Journal of Political Economy* 96, 893-920.

Cogley, Timothy (1990), "International Evidence on the Size of the Random Walk in Output," *Journal of Political Economy* 98, 501-518.

Dickey, D.A., and W.A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association* 74, no. 366, 427-431.

Dufour, Jean-Marie (1982), "Recursive Stability Analysis of Linear Regression Relationships: An Exploratory Methodology," *Journal of Econometrics* 19, 31-76.

Evans, G.W. (1989), "Output and Unemployment Dynamics in the United States: 1950-1985," *Journal of Applied Econometrics* 4, 213-237.

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

Hall, P., and C.C. Heyde (1980), *Martingale Limit Theory and its Application.* San Diego: Academic Press, Inc.

Hamilton, J. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica* 57, 357-384.

Harvey, A.C. (1985), "Trends and Cycles in Macroeconomic Time Series," *Journal of Business and Economic Statistics*, 3, 216-227.

Hansen, Bruce E. (1990), "Testing for Structural Change of Unknown Form in Models with Non-Stationary Regressors," manuscript, Department of Economics, University of Rochester.

Hendry, D.F. (1987), *PC-GIVE: An Interactive Menu-Driven Econometric Modelling Program.* User's Manual, Version 5.0, Institute of Economics and Statistics, Oxford.

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

James, B., K.L. James, and D. Siegmund (1987), "Tests for a Change-Point," *Biometrika* 74, 71-83.

Kormendi, Roger C. and Philip Meguire (1988), "A Multi-Country Characterization of the Non-stationarity of Aggregate Output," manuscript, School of Business Administration, University of Michigan.

Kramer, W., W. Ploberger, and R. Alt (1988), "Testing for Structural Change in Dynamic Models," *Econometrica* 565, 1355-1370.

McCabe, B.P.M. and M.J. Harrison (1980), "Testing the Constancy of Regression Relationships over Time Using Least Squares Residuals," *Applied Statistics 29*, 142-148.

Moore, Geoffrey H., and Melita H. Moore (1985), *International Economic Indicators. A Sourcebook.* Westport, CN: Greenwood Press, 1985.

Nelson, C.R., and C.I. Plosser (1982), "Trends and Random Walks in Macro-economic Time Series: Some Evidence and Implications," *Journal of Monetary Economics* 10, 139-162.

Perron, Pierre (1989), "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Econometrica* 57, 1361-1401.

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

Perron, Pierre (1990b), "Further Evidence on Breaking Trend Functions in Macroeconomic Variables," manuscript, Department of Economics, Princeton University.

Phillips, P.C.B. and S. Ouliaris (1990), "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica* 58, 165-194.

Ploberger, W., W. Kramer, and K. Kontrus (1989), "A New Test for Structural Stability in the Linear Regression Model," *Journal of Econometrics 40*, 307-318.

Quandt, Richard E. (1960), "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes," *Journal of the American Statistical Association* 55, 324-330.

Sargan, J.D., and A. Bhargava (1983), "Testing for Residuals from Least Squares Regression for Being Generated by the Gaussian Random Walk," *Econometrica* 51, 153-174.

Sen, P.K. (1980), "Asymptotic Theory of Some Tests for a Possible Change in the Regression Slope Occurring at an Unknown Time Point," *Zeitschrift fur Wahrescheinlich. und ver. Gebiete* 52, 203-218.

Sen, P.K. (1982), "Invariance Principles for Recursive Residuals," *Annals of Statistics 10*, 307-312.

Sims, Christopher A., James H. Stock, and Mark W. Watson (1990), "Inference in Linear Time Series Models with Some Unit Roots," *Econometrica* 58, 113-144.

Stock, James H. (1987), "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors." *Econometrica* 55, 1035-1056.

Stock, James H. (1988), "A Class of Tests for Integration and Cointegration," manuscript, Kennedy School of Government, Harvard University.

Stock, J.H. and M.W. Watson (1988), "Testing for Common Trends," *Journal of the American Statistical Association* 83, 1097-1107.

Watson, M.W. (1986), "Univariate Detrending Methods with Stochastic Trends", *Journal of Monetary Economics*, 18, 49-75.

Zivot, E. and Donald W.K. Andrews (1989), "Further Evidence on the Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," manuscript, Cowles Foundation, Yale University.

**Table 1**
**Recursive Unit Root Tests: Critical Values**

10% (5%) critical values

| T | $t_{DF}$ | $t_{DF}^{max}$ | $t_{DF}^{min}$ | $t_{DF}^{diff}$ | $\hat{R}^{min}$ |
|---|---|---|---|---|---|
| 100 | -3.15<br>(-3.45) | -1.93<br>(-2.21) | -3.88<br>(-4.13) | 2.95<br>(3.37) | .0195<br>(.0165) |
| 250 | -3.13<br>(-3.43) | -1.88<br>(-2.14) | -3.80<br>(-4.07) | 2.98<br>(3.36) | .0199<br>(.0170) |
| 500 | -3.13<br>(-3.42) | -1.88<br>(-2.14) | -3.82<br>(-4.10) | 3.01<br>(3.45) | .0198<br>(.0173) |

Notes: The Dickey-Fuller statistic critical values ($t_{DF}$) were taken from Fuller (1976), Table 8.5.2, and apply to the case of (1,t) being included as regressors. The remaining entries were computed using data generated by the null model $\Delta y_t = \epsilon_t$, $\epsilon_t$ NIID(0,1), using the regression equation (2.1) with no lags of $\Delta y_t$ in the Dickey-Fuller regressions and setting $2\pi\hat{s}_{\Delta y}(0) = 1$ in the construction of $\hat{R}(\delta)$. The recursive statistics were computed for $0.25 \le \delta \le 1$. The Monte Carlo simulations involved 2000 replications. The statistics are defined in the text.

**Table 2**
**Recursive Unit Root Tests: Size**

True Model: $\Delta y_t = \beta \Delta y_{t-1} + \epsilon_t$, $\epsilon_t$ NIID(0,1)

Recursive regression: $\Delta y_t = \mu_0 + \mu_1 t + \alpha y_{t-1} + \beta \Delta y_{t-1}$ + error

Percent rejections at 10% critical values, T=100

| $\beta$ | $t_{DF}$ | $t_{DF}^{max}$ | $t_{DF}^{min}$ | $t_{DF}^{diff}$ | $\hat{R}^{min}$ |
|---|---|---|---|---|---|
| 0.4 | 9.0% | 9.6% | 9.2% | 14.6% | 24.6% |
| 0.6 | 8.6 | 8.8 | 11.0 | 19.8 | 34.4 |

Notes: The T=100 critical values from Table 1 were used to evaluate the percent rejections. The recursive t-statistic $t_{DF}(\delta)$ was computed for $0.25 \le \delta \le 1$. $\hat{R}^{min}$ was computed using the autoregressive spectral estimator $\hat{s}_{\Delta y}(0) = (2\pi(1-\hat{\beta}(1))^2)^{-1}\hat{\sigma}^2$ from the regression $\Delta y_t = \mu_0 + \alpha y_{t-1} + \beta(L)\Delta y_{t-1} + \epsilon_t$, where 4 lags of $\Delta y_{t-1}$ were included. Based on 500 replications.

**Table 3**
**Recursive Unit Root Tests: Power**

True Model: $y_t = \alpha_t y_{t-1} + \epsilon_t$, $\epsilon_t$ NIID(0,1)

where
$$\alpha_t = \begin{cases} \alpha_1, & t \le \tfrac{1}{2}T \\ \alpha_2, & t > \tfrac{1}{2}T \end{cases}$$

Recursive regression: $\Delta y_t = \mu_0 + \mu_1 t + \alpha y_{t-1} + \beta \Delta y_{t-1} + \text{error}$

Percent rejections at 10% critical values, T=100

| $\alpha_1$ | $\alpha_2$ | $t_{DF}$ | $t_{DF}^{max}$ | $t_{DF}^{min}$ | $t_{DF}^{diff}$ | $\hat{R}^{min}$ |
|---|---|---|---|---|---|---|
| 0.9 | 1.0 | 15.2% | 11.5% | 15.4% | 13.2% | 18.2% |
| 0.8 | 1.0 | 21.2 | 18.0 | 29.2 | 14.2 | 22.6 |
| 1.0 | 0.9 | 4.4 | 4.2 | 8.6 | 36.2 | 37.8 |
| 1.0 | 0.8 | 6.8 | 4.4 | 10.8 | 42.8 | 49.2 |

Notes: See the notes to Table 2.

Table 4

Sequential Unit Root Tests, Trend Shift (Case A): Critical Values

| | 10% (5%) critical values | | |
|---|---|---|---|
| T | $F_T^{max}$ | $t_{DF}(\hat{\delta})$ | $t_{DF}^{min*}$ |
| 100 | 14.30 (16.74) | -4.20 (-4.51) | -4.20 (-4.51) |
| 250 | 12.96 (15.69) | -4.10 (-4.41) | -4.11 (-4.42) |
| 500 | 13.20 (15.29) | -4.09 (-4.38) | -4.11 (-4.38) |

Notes: The entries were computed using data generated by the null model $\Delta y_t = \epsilon_t$, $\epsilon_t$ NIID(0,1), using the regression equation (3.1) and specification (3.2) with no lags of $\Delta y_t$. The series of sequential F- and t-statistics were computed for $0.15 \le \delta \le 0.85$. The Monte Carlo simulations involved 2000 replications. The statistics are described in the text.

Table 5

Sequential Unit Root Tests, Trend Shift (Case A): Size

True Model: $\Delta y_t = \beta \Delta y_{t-1} + \epsilon_t$, $\epsilon_t$ NIID(0,1)

Recursive regression: $y_t = \mu_0 + \mu_1 \tau_{1t}(k) + \mu_2 t + \alpha y_{t-1} + \beta \Delta y_{t-1} + \text{error}$,

$$\tau_{1t}(k) = (t-k)1(t>k)$$

Percent rejections at 10% critical values, T=100

| $\beta$ | $F_T^{max}$ | $t_{DF}(\hat{\delta})$ | $t_{DF}^{min*}$ |
|---|---|---|---|
| 0.4 | 10.6% | 12.2% | 12.2% |
| 0.6 | 12.4 | 12.6 | 12.6 |

Notes: The T=100 critical values from Table 4 were used to evaluate the percent rejections. The sequential statistics $F_T(\delta)$ and $t_{DF}(\delta)$ were computed for $0.15 \le \delta \le 0.85$. Based on 500 replications.

**Table 6**
**Sequential Unit Root Tests, Trend Shift (Case A): Power**

True Model: $y_t = \mu_1 \tau_{1t}([T\delta^*]) + \alpha y_{t-1} + \epsilon_t$, $\epsilon_t$ NIID(0,1),

$$\tau_{1t}([T\delta^*]) = (t-[T\delta^*])1(t>[T\delta^*])$$

Sequential regression: $y_t = \mu_0 + \mu_1 \tau_{1t}(k) + \mu_2 t + \alpha y_{t-1} + \beta \Delta y_{t-1}$ + error,

$$\tau_{1t}(k) = (t-k)1(t>k)$$

Percent rejections at 10% critical values, T=100

| $(\alpha,\delta^*,\mu_1)$ | $t_{DF}$ | $F_T^{max}$ | $t_{DF}(\hat{\delta})$ | $t_{DF}^{min*}$ | Percent $\hat{k}$'s within $k \pm .05T$ |
|---|---|---|---|---|---|
| (.9,.25,.2) | 24.2% | 68.8% | 80.8% | 82.2% | 68.8% |
| (.9,.50,.2) | 0.0 | 90.2 | 70.6 | 71.4 | 76.6 |
| (.9,.75,.2) | 0.0 | 94.0 | 30.0 | 32.8 | 83.4 |
| (.9,.25,.4) | 84.2 | 99.6 | 100.0 | 100.0 | 95.4 |
| (.9,.50,.4) | 0.0 | 100.0 | 99.6 | 99.6 | 99.4 |
| (.9,.75,.4) | 0.0 | 99.8 | 58.8 | 66.2 | 98.4 |
| (.8,.25,.2) | 8.0 | 65.4 | 77.2 | 78.2 | 81.4 |
| (.8,.50,.2) | 0.0 | 91.6 | 76.4 | 76.4 | 92.0 |
| (.8,.75,.2) | 0.0 | 97.6 | 57.0 | 57.8 | 90.2 |

Notes: see the notes to Table 5.

Table 7
Sequential Unit Root Tests, Mean Shift (Case B): Critical Values

10% (5%) critical values

| T | $F(\hat{k})$ | $t_{DF}(\hat{\delta})$ | $t_{DF}^{min*}$ | $\max_{\delta}\lvert t_{\tau_1}(\delta)\rvert$, restricted |
|---|---|---|---|---|
| 100 | 15.91<br>(18.40) | -4.51<br>(-4.82) | -4.52<br>(-4.83) | 2.62<br>(2.94) |
| 250 | 16.42<br>(18.61) | -4.49<br>(-4.75) | -4.51<br>(-4.75) | 2.66<br>(2.92) |
| 500 | 16.70<br>(19.03) | -4.53<br>(-4.79) | -4.55<br>(-4.81) | 2.64<br>(2.86) |

Notes: The entries were computed using data generated by the null model $\Delta y_t=\epsilon_t$, $\epsilon_t$ NIID(0,1), using the regression equation (3.1) and specification (3.3) with no lags of $\Delta y_t$. The series of sequential F- and t-statistics were computed for $0.15\leq\delta\leq0.85$. The Monte Carlo simulations involved 2000 replications. The statistics are described in the text. The t-statistic for $\tau_1(\delta)$ in the final column is computed under the restriction that $\mu_2=0$ and $\alpha=1$ in Model II, Case B.

Table 8
Recursive Statistics (p-values) Across Countries

| Country | Sample | $t_{DF}$ | $t_{DF}^{max}$ | $t_{DF}^{min}$ | $Min_{\delta}\hat{R}(\delta)$ |
|---|---|---|---|---|---|
| Canada | 48:1-89:2 | -1.96 | -1.17 | -3.70 | 0.023 |
| | | (.62) | (.41) | (.14) | (.20) |
| France | 63:1-89:2 | -1.74 | 1.39 | -2.44 | 0.047 |
| | | (.73) | (.99) | (.80) | (.74) |
| Germany | 50:1-89:2 | -1.96 | 0.20 | -1.96 | 0.110 |
| | | (.62) | (.94) | (.97) | (.99) |
| Italy | 52:1-82:4 | 0.26 | 0.26 | -3.87 | 0.018 |
| | | (.99) | (.95) | (.10) | (.07) |
| Japan | 52:1-89:2 | -0.09 | 0.72 | -2.44 | 0.072 |
| | | (.99) | (.99) | (.80) | (.93) |
| UK | 60:1-89:2 | -1.88 | -0.48 | -3.09 | 0.025 |
| | | (.66) | (.75) | (.42) | (.26) |
| US | 47:1-89:2 | -2.60 | -0.00 | -2.60 | 0.037 |
| | | (.27) | (.91) | (.71) | (.56) |

Notes: The data and statistics are described in the text. P-values testing the constant-drift unit root hypothesis, based on critical values computed for T=100, are given in parentheses. The sample period refers to the full sample of data used to compute the statistics, including initial values for lags in the autoregressions. The statistic $t_{DF}$ was computed using the full sample; the rest of the statistics are recursive. $\hat{R}^{min}$ was computed using the autoregressive spectral estimator $\hat{s}_{\Delta y}(0) = (2\pi(1-\hat{\beta}(1))^2)^{-1}\hat{\sigma}^2$, from the regression $\Delta y_t = \mu_0 + \alpha y_{t-1} + \beta(L)\Delta y_{t-1} + \epsilon_t$, where 4 lags of $\Delta y_{t-1}$ were included. The recursive statistics were computed starting at observation [.25T], where T is the full sample length given in the second column.

Table 9
Sequential Statistics (p-values) Across Countries

| | Case A: Trend Shift | | | Case B: Mean Shift | | | Case B, Restricted | |
|---|---|---|---|---|---|---|---|---|
| Country | $\hat{k}$ | $t_{DF}(\hat{\delta})$ | $t_{DF}^{min*}$ | $\hat{k}$ | $t_{DF}(\hat{\delta})$ | $t_{DF}^{min*}$ | $\hat{k}$ | $t_{\tau_1}(\hat{\delta})$ |
| Canada | 76:2 | -4.14 (.12) | -4.14 (.12) | 81:3 | -5.14 (.02) | -5.14 (.02) | 76:3 | -2.11 (.29) |
| France | 72:4 | -3.89 (.19) | -3.89 (.20) | 68:1 | -3.55 (.50) | -3.55 (.54) | 74:2 | -4.45 (.00) |
| Germany | 55:4 | -2.79 (.75) | -2.79 (.77) | 80:2 | -2.84 (.83) | -2.84 (.90) | 60:4 | -3.28 (.02) |
| Italy | 70:4 | -3.67 (.28) | -3.67 (.29) | 74:3 | -1.60 (.98) | -1.60 (.99) | 74:2 | -3.48 (.01) |
| Japan | 69:4 | -4.78 (.03) | -4.81 (.02) | 73:2 | -0.69 (.99) | -2.23 (.98) | 73:2 | -4.85 (.00) |
| UK | 64:2 | -2.49 (.88) | -2.49 (.91) | 79:3 | -3.98 (.27) | -3.98 (.29) | 82:4 | 1.55 (.64) |
| US | 68:2 | -3.27 (.50) | -3.27 (.50) | 63:1 | -3.76 (.39) | -3.76 (.41) | 53:2 | -2.04 (.32) |

Notes: The data and statistics are described in the text. P-values testing the constant-drift unit root hypothesis, based on critical values computed for T=100, are given in parentheses. The sequential statistics were computed for $[.15T] \le k \le [.85T]$, where T is the full sample length given in the second column of Table 8.

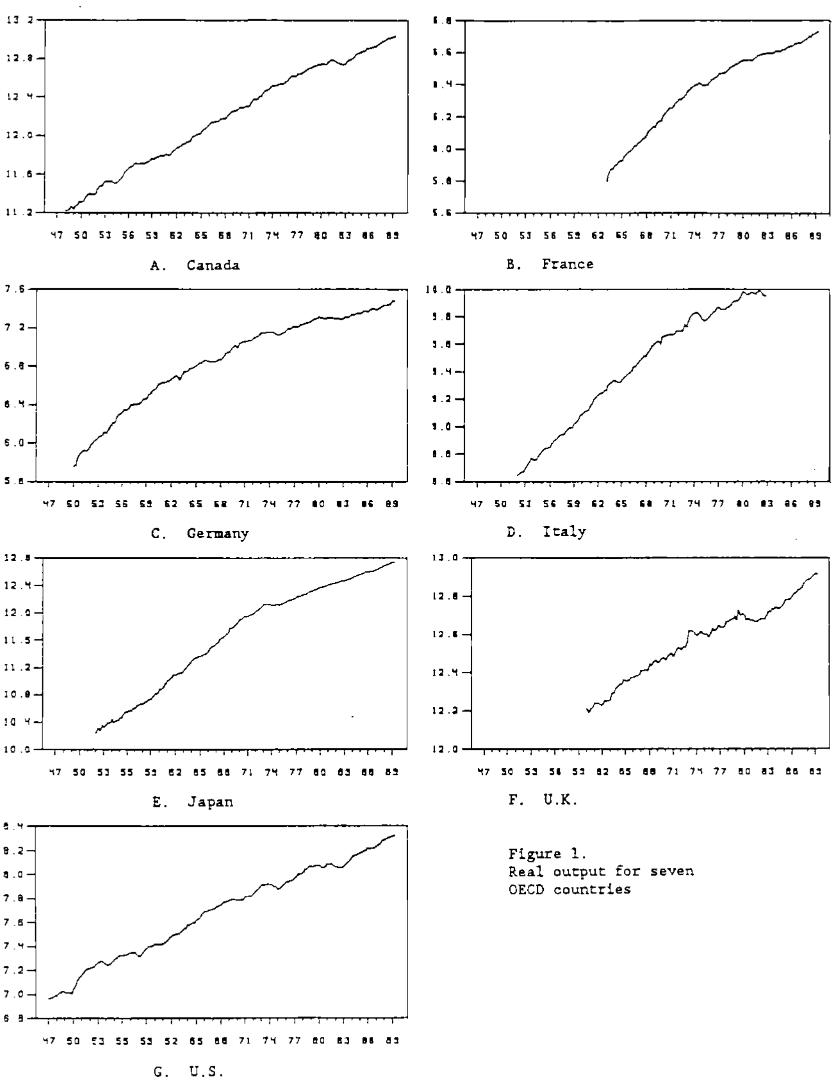



Figure 1.
Real output for seven
OECD countries

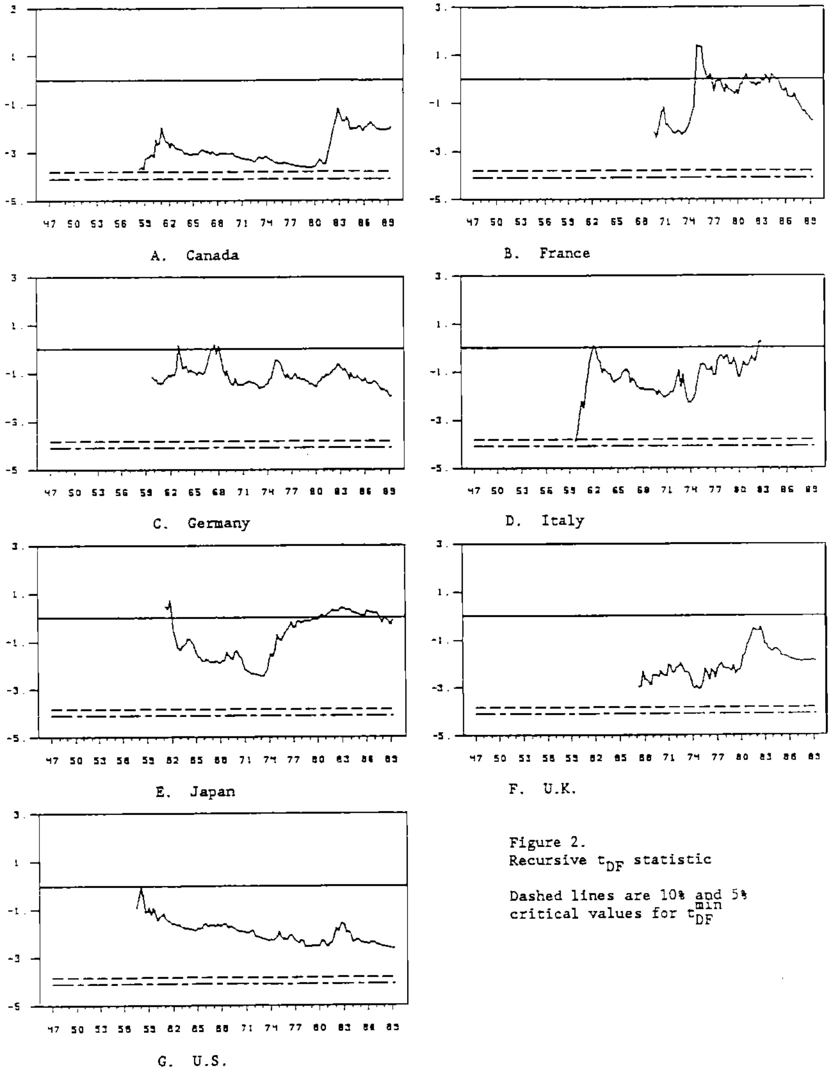
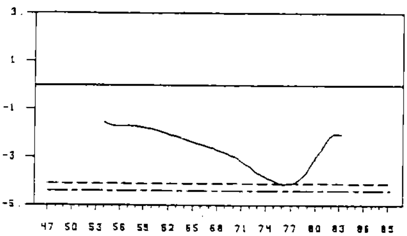
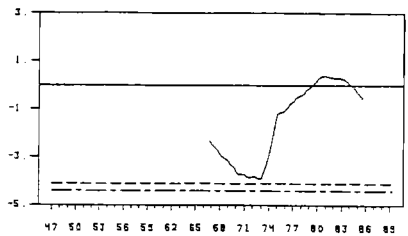



Figure 2.
Recursive $t_{DF}$ statistic

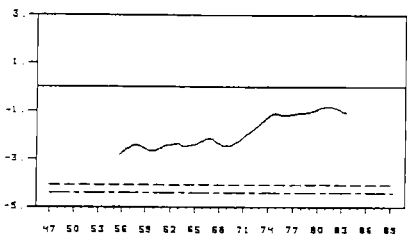
Dashed lines are 10% and 5% critical values for $t_{DF}^{min}$



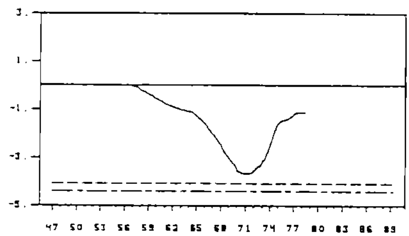

A. Canada



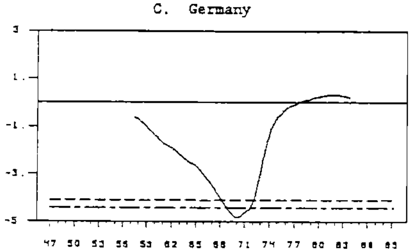

B. France



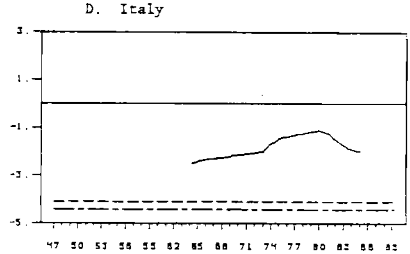

C. Germany



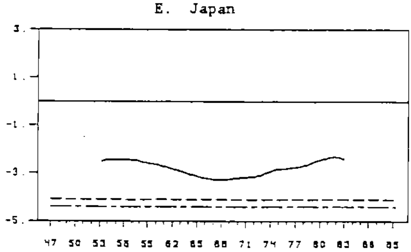

D. Italy




E. Japan



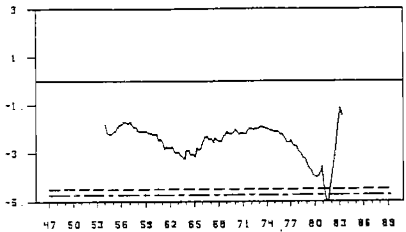

F. U.K.



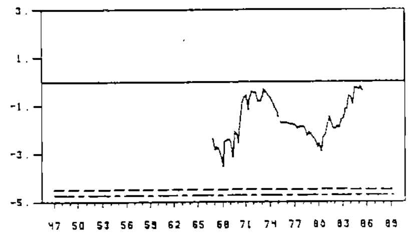

G. U.S.

Figure 3.
Sequential $t_{DF}$ statistic,
case A (trend shift)

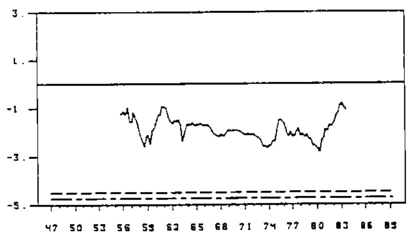
Dashed lines are 10% and 5%
critical values for $t_{DF}^{min*}$



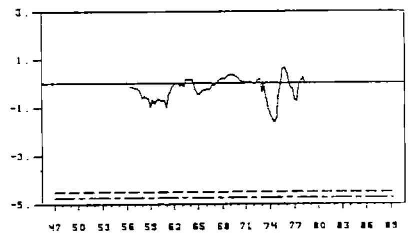

A. Canada



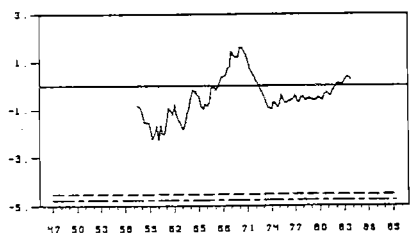

B. France



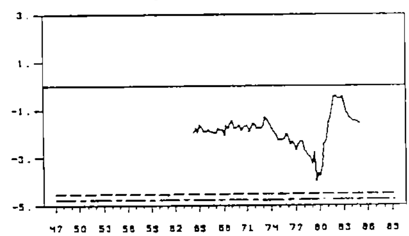

C. Germany



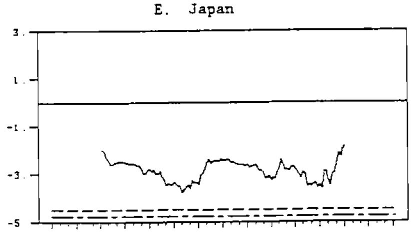

D. Italy




E. Japan



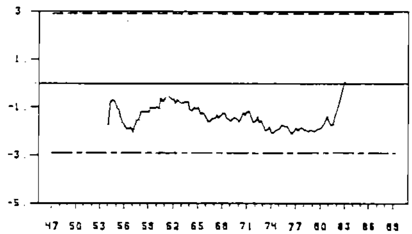

F. U.K.



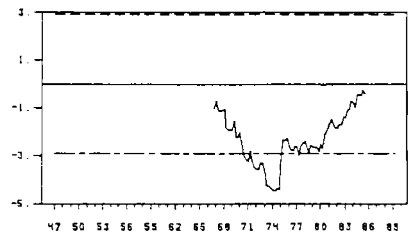

G. U.S.

Figure 4.
Sequential $t_{DF}$ statistic, case B (mean shift)

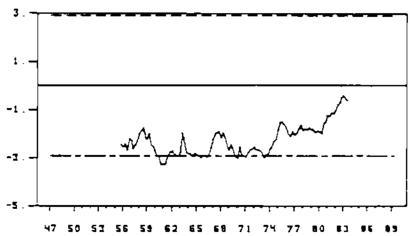
Dashed lines are 10% and 5% critical values for $t_{DF}^{min*}$



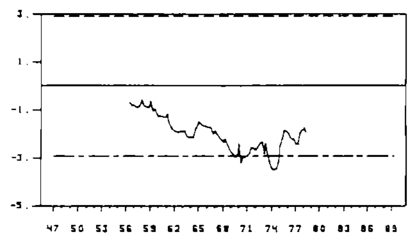
A. Canada



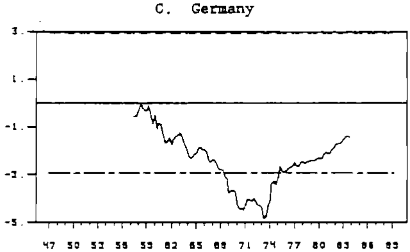
B. France



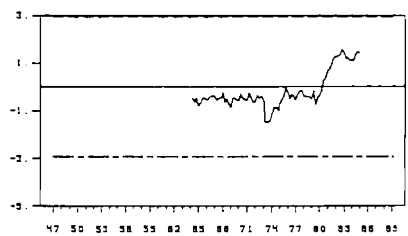
C. Germany



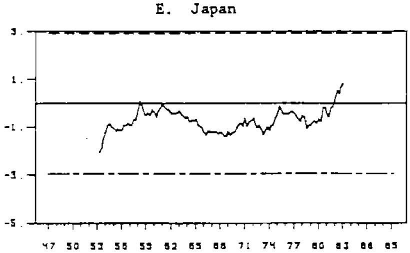
D. Italy



E. Japan



F. U.K.



G. U.S.

Figure 5.
Sequential t-statistic on $\tau_{1t}(k)$, case B (mean shift), with $\alpha=1$ and zero linear time trend imposed.

Dashed lines are 5% critical values for $\max_{[T\delta_0]\le k\le T-[T\delta_0]} |t_{\tau_1}(k)|$