TESTING FOR THE PRESENCE OF A RANDOM WALK IN SERIES WITH STRUCTURAL BREAKS

by

Fabio Busetti
London School of Economics and Political Science

and

Andrew Harvey
Faculty of Economics and Politics, University of Cambridge

Contents:

Abstract
1. Introduction
2. Testing against the presence of a random walk component
3. LBI tests when breaks are present
4. A simplified test
5. Unknown breakpoint
6. Serial correlation
7. Examples
8. Seasonality
9. Conclusions
Acknowledgements
Appendix
References
Tables 1 – 7
Figures 1 – 3

Discussion Paper
No. EM/98/365
December 1998

The Suntory Centre
Suntory and Toyota International Centres for Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 0171-405 7686

Abstract

The paper considers tests for the presence of a random walk component in a stationary or trend stationary time series and extends them to series which contain structural breaks. The locally best invariant (LBI) test is derived and the asymptotic distribution obtained. Then a modified test statistic is proposed. The advantage of this statistic is that its asymptotic distribution is not dependent on the location of the breakpoint and its form is that of the generalised Cramér-von Mises distribution, with degrees of freedom depending on the number of breakpoints. The performance of this modified test is shown, via some simulation experiments, to be comparable to that of the LBI test. An unconditional test, based on the assumption that there is a single break at an unknown point is also examined. The use of the tests is illustrated with data on the flow of the Nile and US Gross National Product.

Keywords: Brownian bridge; Cramér-von Mises distribution; intervention analysis; locally best invariant test; structural time series model; unobserved components.

JEL Nos.: C12, C22.

© by Fabio Busetti and Andrew Harvey. All rights reserved. Short sections of test, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

1 Introduction

A test for the presence of a random walk component in series which would otherwise be white noise about a constant level was developed by Nyblom and Mäkeläinen (1983) and MacNeill (1978). In a Gaussian model this test is locally best invariant (LBI) and under the null hypothesis it has a known asymptotic distribution, namely the Cramér-von Mises distribution. The tests may also be applied when the model contains a deterministic time trend, but with a different asymptotic distribution under the null hypothesis.

If the model is extended so that the underlying series is generated by any indeterministic stationary stochastic process, a nonparametric correction can be made so that a statistic with the same asymptotic distribution is obtained. This is known as the KPSS test after Kwiatkowski, Phillips, Schmidt and Shin (1992). Parametric procedures have been proposed by Leybourne and McCabe (1994) and Harvey and Streibel (1997).

The above tests have high power against alternative hypotheses in which the level or trend contains a small number of breaks but is otherwise deterministic. Indeed Gardner (1969) and Nyblom (1989) provide a formal justification for their use against such alternatives. Thus if structural breaks are known to be present it is vital to take account of them if a test against a random walk component is to be carried out. This article develops LBI tests when there are breaks in the level and/or the slope at known points in time and the asymptotic distributions of the test statistics are derived under the null hypothesis. When there are no restrictions on the form of the breaks, for example the requirement that the deterministic trend be piecewise continuous, and they are equally spaced, it is shown that the asymptotic distributions belong to the Cramér-von Mises family. Of course breaks are not equally spaced in general but a slight modification gives test statistics with the same asymptotic distributions as in the equally spaced case. Thus the critical values depend only on the number of breaks and on whether or not a time trend is included. This is important, since although it may be feasible to tabulate critical values for a single break at different points in the sample, constructing tables when there are two or more breaks is impractical.

The tests here are for situations in which the model is (trend) stationary under the null and nonstationary under the alternative. Such tests are to be contrasted with those in the unit root literature where the null of nonstationarity is tested against an alternative of stationarity. The unit root tests are normally carried out within an autoregressive framework, with the augmented Dickey-Fuller (ADF) test being, in some sense, analogous to KPSS. Perron (1989) modified the Dickey-Fuller test for models containing exogenous structural breaks and derived the relevant asymptotic distribu-

tions. Although our tests are not directly comparable with those of Perron, because of the reversal of the null and alternative hypotheses, it is worth noting that there is no modified version of his tests for which the asymptotic distribution depends only on the number of break points.

All the tests mentioned so far assume that the locations of the break points are known. This may well be the case if there are known interventions or policy changes. However, if possible breaks are detected from a statistical analysis or a visual inspection of the data, the tests are no longer valid. Proceeding in this way biases the tests in favour of the stationarity hypothesis. If, on the other hand, it is known *a priori,* that there is a break, but at an unknown location, Zivot and Andrews (1992) showed that Perron's testing procedure could be legitimately applied by endogenously selecting the breakpoint so as to give the least favourable result for the null hypothesis of a unit root. A similar idea can be applied for testing the null of stationarity.

The plan of the paper is as follows. After reviewing the LBI tests for a random walk component in section 2, we derive the tests when structural breaks are present in section 3. Section 4 looks at the modified tests and reports some simulation experiments which show that their size and power are comparable to those of the corresponding LBI tests. Section 5 looks at tests when the there is a single break at an unknown point and section 6 discusses how serial correlation may be handled. Section 7 illustrates the tests with two data sets, the flow of the Nile and US GNP. Section 8 discusses how the tests can be carried out in the presence of seasonality and then investigates the effects of breaks on the Canova-Hansen test against nonstationary stochastic seasonality. Conclusions and extensions are in section 9.

2 Testing against the presence of a random walk component

Consider a model containing a time trend, a set of strictly exogenous regressors, $\mathbf{x}_t$, and a random walk, that is

$$\begin{aligned} y_t &= \mu_t + \beta t + \mathbf{x}_t'\boldsymbol{\delta} + \varepsilon_t, \qquad \varepsilon_t \sim \text{NID}\left(0, \sigma^2\right) \qquad t = 1, ..., T, \qquad (1) \\ \mu_t &= \mu_{t-1} + \eta_t, \qquad \eta_t \sim \text{NID}\left(0, \sigma_\eta^2\right) \end{aligned}$$

where μ_0 is fixed and ε_t and η_t are mutually independent. The notation $\text{NID}(0, \sigma^2)$ denotes normally and independently distributed with mean zero and variance σ^2. The locally best invariant (LBI) test for $H_0 : \sigma_\eta^2 = 0$

against $H_A : \sigma_\eta^2 > 0$ can be written as

$$\xi = \frac{\sum_{t=1}^{T} (\sum_{s=1}^{t} e_s)^2}{T^2 \widehat{\sigma}^2} > c \tag{2}$$

where c is a critical value and $\widehat{\sigma}^2 = T^{-1} \sum_{t=1}^{T} e_t^2$ with the e_t's being residuals from regressing y_t on a constant, time trend and explanatory variables. The test is invariant to $\beta, \boldsymbol{\delta}$ and μ_0 and it is best in the sense that the power function has maximum slope at the origin among all invariant tests of the same hypothesis. The test can also be interpreted as a one-sided Lagrange multiplier (LM) test.

The distribution of ξ depends on whether or not a time trend is included and on the form of the explanatory variables in the vector $\mathbf{x}_t$. If the observations consist simply of a random walk plus noise then $e_t = y_t - \overline{y}, \quad t = 1, ..., T$, where $\overline{y}$ is the sample mean. The asymptotic distribution of the statistic is found by first observing that the partial sum of deviations from the mean converges weakly to a standard Brownian bridge, that is

$$\sigma^{-1} T^{-\frac{1}{2}} \sum_{s=1}^{[T\cdot]} e_s \Rightarrow B(\cdot), \tag{3}$$

where $B(r) = W(r) - rW(1)$, $r \in [0, 1]$, with $W(\cdot)$ being a standard Wiener process or Brownian motion. Hence

$$\xi \Rightarrow \int_0^1 B(r)^2 dr \tag{4}$$

since $\widehat{\sigma}^2 = T^{-1} \sum_{t=1}^{T} (y_t - \overline{y})^2 \xrightarrow{p} \sigma^2$. This is the Cramér-von Mises distribution; see Nyblom and Mäkeläinen (1983). We will denote it as CvM. Note that it is sufficient for the observations to be independent and identically distributed to yield this asymptotic distribution; see Nabeya and Tanaka (1988).

When a time trend is included in the model we have the analogous result that the partial sum of residuals from a first order polynomial regression weakly converges to a second level Brownian bridge denoted $B_2(\cdot)$ where, as in McNeill (1978),

$$B_2(r) = W(r) - rW(1) + 6r(1-r) \left\{ \tfrac{1}{2} W(1) - \int_0^1 W(s) ds \right\}. \tag{5}$$

Then

$$\xi \Rightarrow \int_0^1 B_2(r)^2 dr \ . \tag{6}$$

We will refer to this asymptotic distribution as a second level Cramér-von Mises distribution, and denote it as CvM_2. In the case of any ambiguity we will refer to the distribution in (4) as CvM_1.

Percentage points for the Cramér-von Mises distribution have been tabulated by and Anderson and Darling (1952), MacNeill (1978), Nyblom and Mäkeläinen (1983), Nyblom (1986) and Kwiatkowski *et al.* (1992).

Kwiatkowski *et al.* (1992) extended the above tests to allow ε_t to follow a stationary indeterministic process. In fact ε_t can be even more general, satisfying the assumptions in Phillips and Perron (1988) or Phillips and Solo (1992). The test statistic, which has the same asymptotic distribution as the corresponding LBI statistic for white noise disturbances, is

$$\xi(\ell) = \frac{\sum_{t=1}^{T}(\sum_{s=1}^{t} e_s)^2}{T^2 s^2(\ell)}, \tag{7}$$

where

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

is a consistent estimator of the long run variance

$$\sigma_L^2 = \lim_{T\to\infty}\frac{1}{T}\mathrm{Var}\left(\sum_{t=1}^{T}\varepsilon_t\right).$$

The weighting function can be chosen in a variety of ways, as long as $l \to \infty$ as $T \to \infty$. A simple option, which guarantees the non-negativity of $s^2(\ell)$, is to let $w(j,\ell) = 1 - j/(\ell+1)$. Other possibilities are examined in Andrews (1991).

3 LBI tests when breaks are present

The asymptotic distribution of the ξ statistic is still Cramér-von Mises if the explanatory variables in (1) satisfy certain conditions, for example weak dependence conditions if they are stochastic. However, the dummy variables needed to account for structural breaks do change the distribution.

Assume there is a structural break in the trend at time $\tau + 1$, and let $\lambda = \tau/T$ denote the fraction of the sample before the break occurs. We consider the following models:

$$\begin{aligned}
[1] \quad y_t &= \mu_t + \delta w_t + \varepsilon_t && (9)\\
[2] \quad y_t &= \mu_t + \beta t + \delta_\mu w_t + \delta_\beta (w_t t) + \varepsilon_t && (10)\\
[2a] \quad y_t &= \mu_t + \beta t + \delta_\mu w_t + \varepsilon_t && (11)\\
[2b] \quad y_t &= \mu_t + \beta t + \delta_\beta z_t + \varepsilon_t, && (12)
\end{aligned}$$

where μ_t is a random walk as in (1), ε_t is white noise and

$$w_t = \begin{cases} 0 & \text{for } t \leq \tau \\ 1 & \text{for } t > \tau \end{cases} \quad \textit{and} \quad z_t = \begin{cases} 0 & \text{for } t \leq \tau \\ t-\tau & \text{for } t > \tau \end{cases}$$

There is no slope in model [1] and so the only break is in the level. The other models all contain a time trend. In model [2] there is a structural change in both the level and the slope. Models [2a] and [2b] contain respectively a break in the level only and in the slope only. The second of these corresponds to a piecewise linear trend.

Under Gaussianity the LBI (and one-sided LM) test statistic for H_0 : $\sigma_\eta^2 = 0$ against $H_A : \sigma_\eta^2 > 0$ is of the same form as (2), that is

$$\xi_i(\lambda) = \frac{\sum_{t=1}^{T} \left(\sum_{s=1}^{t} e_s \right)^2}{T^2 \widehat{\sigma}^2}, \quad i = 1, 2, 2a, 2b, \tag{13}$$

where the e_t's are the residuals from regressing the observations on the appropriate set of regressors. The subscript i in $\xi_i(\lambda)$ indicates that the residuals depend on the model, 1,2, 2a or 2b, while λ denotes that the statistic has been constructed for a specific value of the breakpoint location parameter and that its asymptotic distribution depends on it.

As in the previous section, the limiting distribution can be derived by looking at the asymptotic properties of the process followed by the partial sum of residuals $S_i(\cdot, \lambda) = \sigma^{-1} T^{-\frac{1}{2}} \sum_{s=1}^{[T\cdot]} e_s$, $r \in [0,1]$, $i = 1, 2, 2a, 2b$. This will converge to a limiting process -defined on an underlying Wiener process- that will depend on λ and collapse to a (second level) Brownian bridge when $\lambda = 0$ or $\lambda = 1$. The asymptotic distribution of $\xi_i(\lambda)$ is then the integral on the unit interval of the square of this process.

The following proposition states the asymptotic distributions of the $\xi_i(\lambda)$ statistics under the null hypothesis $H_0 : \sigma_\eta^2 = 0$. The proof is given in the appendix[1].

Proposition 1. *Let $\{y_t\}$ be generated under the null hypothesis of model i, $i = 1, 2, 2a, 2b$, with $\varepsilon_t \sim IID(0, \sigma^2)$. Then*

$$\xi_i(\lambda) \Rightarrow \int_0^1 [B_i(r, \lambda)]^2 \, dr,$$

where

$$B_1(r, \lambda) = \begin{cases} W(r) - \frac{r}{\lambda} W(\lambda) & \textit{for } r \leq \lambda \\ (W(r) - W(\lambda)) - \frac{r-\lambda}{1-\lambda} (W(1) - W(\lambda)) & \textit{for } r > \lambda \end{cases}$$

[1]Hao (1996) gives a related result for a structural break in the level only in the context of co-integration. Integrated regressors are included in a model which would otherwise be of the form 1. His results do not specialise to the ones presented here.

$$B_2(r,\lambda) = \begin{cases} W(r) - \frac{r}{\lambda}W(\lambda) \\ \qquad -\frac{6}{\lambda^3}r(r-\lambda)\left[\int_0^\lambda r dW(r) - \frac{\lambda}{2}W(\lambda)\right] & for\ r \le \lambda \\ \\ (W(r) - W(\lambda)) - \frac{r-\lambda}{1-\lambda}(W(1) - W(\lambda)) \\ \qquad -\frac{6}{(1-\lambda)^3}(r-1)(r-\lambda)\left[\int_\lambda^1 r dW(r) - \frac{1+\lambda}{2}(W(1) - W(\lambda))\right] & for\ r > \lambda \end{cases}$$

$$B_{2a}(r,\lambda) = \begin{cases} W(r) - \frac{r}{\lambda}W(\lambda) \\ \qquad -\frac{6}{1-3\lambda+3\lambda^2}r(r-\lambda) \\ \qquad \cdot\left[\int_0^1 r dW(r) - \frac{\lambda}{2}W(\lambda) - \frac{1+\lambda}{2}(W(1) - W(\lambda))\right] & for\ r \le \lambda \\ \\ (W(r) - W(\lambda)) - \frac{r-\lambda}{1-\lambda}(W(1) - W(\lambda)) \\ \qquad -\frac{6}{1-3\lambda+3\lambda^2}r(r-\lambda) \\ \qquad \cdot\left[\int_0^1 r dW(r) - \frac{\lambda}{2}W(\lambda) - \frac{1+\lambda}{2}(W(1) - W(\lambda))\right] & for\ r > \lambda \end{cases}$$

$$B_{2b}(r,\lambda) = \begin{cases} W(r) - rW(1) - \frac{3}{\lambda^3(1-\lambda)^3} \\ \quad \cdot\left\{\left(a\frac{r^2}{2} - a\lambda r + \frac{r}{2}\left(a\lambda^2 - b(1-\lambda)^2\right)\right)J_1 \right. \\ \quad \left. + \left(b\frac{r^2}{2} - b\lambda r + \frac{r}{2}\left(b\lambda^2 - c(1-\lambda)^2\right)\right)J_2\right\} & for\ r \le \lambda \\ \\ W(r) - rW(1) - \frac{3}{\lambda^3(1-\lambda)^3} \\ \quad \cdot\left\{\left(-a\frac{\lambda^2}{2} + b\frac{r^2-\lambda^2}{2} - b\lambda(r-\lambda) + \frac{r}{2}\left(a\lambda^2 - b(1-\lambda)^2\right)\right)J_1 \right. \\ \quad \left. + \left(-b\frac{\lambda^2}{2} + c\frac{r^2-\lambda^2}{2} - c\lambda(r-\lambda) + \frac{r}{2}\left(b\lambda^2 - c(1-\lambda)^2\right)\right)J_2\right\} & for\ r > \lambda \end{cases}$$

with

$$\begin{aligned} a &= (1-\lambda)^3(1+3\lambda), \\ b &= -3\lambda^2(1-\lambda)^2, \\ c &= \lambda^3(4-3\lambda), \\ J_1 &= \int_0^\lambda r dW(r) - \lambda W(\lambda) + \frac{\lambda^2}{2}W(1), \\ J_2 &= \int_\lambda^1 r dW(r) - \lambda\left(W(1) - W(\lambda)\right) - \frac{(1-\lambda)^2}{2}W(1). \end{aligned}$$

The upper tail percentage points of the above asymptotic distributions are reported in table 1 for different values of λ. As in Kwiatkowski et al (1992), they are obtained by simulating empirical approximations to Brownian motions with samples size of 1000 and 10000 replications. We use the random numbers generator of the matrix programming language OX; see Doornik (1996). The figures for $\lambda \to 0$ or $\lambda \to 1$ correspond to the critical values for the Cramér-von Mises distributions of the previous section. As expected, the percentage points -as functions of λ- are symmetric around $\lambda = 1/2$, which is also the minimum for models $1, 2$ and 2b.

For models [1] and [2] the asymptotic distributions can be characterized in terms of two independent Cramér-von Mises distribution. To see that this is the case, first notice that we can rewrite the statistic as

$$\xi_i(\lambda) = \lambda^2 \frac{\sum_{t=1}^{\tau}(\sum_{s=1}^{t} e_s)^2}{\tau^2\widehat{\sigma}^2} + (1-\lambda)^2 \frac{\sum_{t=\tau+1}^{T}(\sum_{s=\tau+1}^{t} e_s)^2}{(T-\tau)^2\widehat{\sigma}^2}, \quad i = 1, 2, \tag{14}$$

since from the OLS orthogonality conditions $\sum_{s=1}^{T} e_s = \sum_{s=\tau+1}^{T} e_s = 0$, and this implies $\sum_{s=1}^{\tau} e_s = 0$. Then it is easy to see that for model [1]

$$e_s = \begin{cases} y_s - \overline{y}_1 & \text{for } s \leq \tau \\ y_s - \overline{y}_2 & \text{for } s > \tau \end{cases}$$

where $\overline{y}_1 = \tau^{-1}\sum_{t=1}^{\tau} y_t$ and $\overline{y}_2 = (T-\tau)^{-1}\sum_{s=\tau+1}^{T} y_t$ are the averages in the first and second subsamples respectively. A similar result holds for model [2] where the two sets of residuals are obtained from regressing on a constant and a time trend. Thus the residuals are independent across subsamples and the following proposition holds.

Proposition 2. *Let $\{y_t\}$ be generated under the null hypothesis of model i, $i = 1, 2$, with $\varepsilon_t \sim IID(0, \sigma^2)$. Then*

$$\xi_i(\lambda) \Rightarrow \lambda^2 \int_0^1 [B_i(r)]^2\, dr + (1-\lambda)^2 \int_0^1 [B_i'(r)]^2\, dr, \quad i = 1, 2, \tag{15}$$

where $B_1(\cdot)$ and $B_1'(\cdot)$ are independent Brownian bridges and $B_2(\cdot)$ and $B_2'(\cdot)$ are independent second level Brownian bridges. Hence the statistics are weighted averages of two independent Cramér-von Mises distributions.

This is a very simple way to characterize the asymptotic distribution and it is trivially generalizable to the case of more than one break. Note that if the breaks are equispaced the distribution of the statistic (when multiplied by four) converges to the sum of two random variables with independent Cramér-von Mises distributions. Of course assuming equispaced breaks is not appropriate in general. However, the same additivity property can be obtained after a slight modification of the test statistic, as suggested in the following section. By doing this we can eliminate the dependence on the parameter λ in the asymptotic distribution.

4 A simplified test

The Cramér-von Mises distribution can be represented by a series expansion of independent $\chi^2(1)$ variables, that is

$$CvM = \int_0^1 B(r)^2 dr = \sum_{j=1}^{\infty} (\pi j)^{-2} \chi_j^2(1). \tag{16}$$

The proof follows from the argument in Gihman-Skorohod (1974, p. 229-230); see also Nyblom (1989). This allows us to characterise the generalised Cramér-von Mises distribution with k degrees of freedom as

$$CvM(k) = \sum_{j=1}^{\infty} (\pi j)^{-2} \chi_j^2(k), \tag{17}$$

with the interesting corrolary that, because of the additive property of chi-square distributions, the sum of k independent random variables with $CvM(1)$ distributions is $CvM(k)$. The same additivity property holds for a second level Cramér-von Mises distribution, $CvM_2(k)$. The series expansion is

$$\int_0^1 B_2(r)^2 dr = \sum_{j=1}^{\infty} \varphi_j^{-2} \chi_j^2(1),$$

where $B_2(\cdot)$ is a second level Brownian bridge as in (9), and φ_j is defined by $\varphi_{2j-1} = 2\pi j$ and φ_{2j} being the root of $\tan(\varphi/2) = \varphi/2$ on $(2\pi j, 2\pi(j+1))$, $j = 1, 2, ...$; see Nyblom (1986).

Bearing the above in mind, we propose the following test statistics for models [1] and [2]:

$$\xi_i^* = \frac{\sum_{t=1}^{\tau} (\sum_{s=1}^{t} e_s)^2}{\tau^2 \widehat{\sigma}^2} + \frac{\sum_{t=\tau+1}^{T} (\sum_{s=\tau+1}^{t} e_s)^2}{(T-\tau)^2 \widehat{\sigma}^2}, \qquad i = 1, 2. \tag{18}$$

Thus we eliminate the weights in (15). The statistics still depend on the location of the breakpoint, but their asymptotic distributions do not since

$$\xi_i^* \Rightarrow \begin{cases} \mathrm{CvM}(2) & \text{for } i = 1 \\ \mathrm{CvM}_2(2) & \text{for } i = 2. \end{cases} \tag{19}$$

Not having to consult a table giving the distribution of the test statistic for all the possible values of λ is a big advantage; compare the unit root tests in Perron (1989). Furthermore the test immediately generalises to cases where there are several structural breaks. If there are k breaks at times $\tau_1 = \lambda_1 T < ... < \tau_k = \lambda_k T$ the test statistic is

$$\xi_i^*(k) = \sum_{j=1}^{k+1} \frac{\sum_{t=\tau_{j-1}+1}^{\tau_j} \left(\sum_{s=\tau_{j-1}+1}^{t} e_s \right)^2}{(\tau_j - \tau_{j-1})^2 \widehat{\sigma}^2}, \quad i = 1, 2, \tag{20}$$

where $\tau_0 = 0$ and $\tau_{k+1} = T$. The distribution of this statistic converges to a (second-level) generalised Cramér-von Mises distribution with $k+1$ degrees of freedom. The advantage is now even greater since constructing tables for all patterns of k breakpoints would be extremely cumbersome.

The upper tail percentage points for generalised Cramér-von Mises distributions are tabulated in Nyblom (1989), Canova and Hansen (1995) and Nyblom and Harvey (1997), and reported in table 2.

How good is this simplified test? Table 3 compares the LBI test, based on (13), and the simplified test, (18), in terms of size and power by a Monte Carlo experiment. The model with a break in level and slope, model 2, was simulated 5000 times for different values of λ and $q = \sigma_\eta^2/\sigma^2$, the test statistics were computed and the number of rejections was counted for 5% asymptotic critical values obtained from table 1b for the LBI test and from table 2b for the simplified test. For $\lambda = 0.5$ the two tests are the same, except that the critical value of the LBI test is one quarter of the critical value of the modified test. For other values of λ the size and power are comparable, with the LBI being clearly superior only in the region close to the null hypothesis and for the break point near the beginning or end of the sample. The conclusions are similar for samples of 100 and 200.

The above experiment was repeated for a data generating process with two structural breaks, and no slope, with the breaks located in a variety of positions. The 5% asymptotic critical values for the LBI statistic, $\xi_1(\lambda_1, \lambda_2)$, are reported in table 4 and then the performance of the simplified test, $\xi_1(2)$, is compared with the LBI test in table 5. The conclusions are similar to those reached for the case of a single break, with the simplified test having a size close to the nominal and power comparable with the LBI test.

5 Unknown breakpoint

The tests of the last two sections depend on knowing the location of the breakpoints. In some instances one would like to test for trend stationarity under the assumption that there may be a single break in an unknown position.

For a single structural break at an unknown point, we consider a set of unconditional tests, obtained by following the argument in Zivot and Andrews (1992). The idea is to choose the breakpoint that gives the most favourable result for the null hypothesis of trend stationarity using the $\xi_i(\lambda)$ statistic, that is

$$\widetilde{\xi}_i = \inf_{\lambda \in \Lambda} \xi_i(\lambda), \quad i = 1, 2, 2a, 2b, \tag{21}$$

where Λ is a closed subset of the interval (0,1).

The distribution of $\widetilde{\xi}_i$ will depend not only on the location of the true breakpoint, denoted λ_0, but also on the magnitude of the level and/or slope shift, δ, because each $\xi_i(\lambda)$ statistic will depend on the latter when it is computed for a breakpoint different from λ_0. The following assumption on the magnitude of the shift allows us to derive the asymptotic distribution of

$\widetilde{\xi}_i$.

Shift assumption. *The magnitude of the shifts δ, δ_μ and δ_β in (9) to (12) decreases to zero with the sample size at a rate faster than $T^{-1/2}$ for δ and δ_μ and at a rate faster than $T^{-3/2}$ for δ_β.*

Whether this assumption is a reasonable one is open to question. However, in the literature on breakpoint estimation, Bai (1994, 1997) assumes that the magnitude of the shift shrinks to zero at a rate slower than $T^{-1/2}$ in order to derive the asymptotic distribution of the breakpoint estimator. In our case, the rate is faster. Note that the assumption covers the case of no break actually occuring.

Proposition 3. *Let $\{y_t\}$ be generated under the null hypothesis of model i, $i = 1, 2, 2a, 2b$, with $\varepsilon_t \sim IID(0, \sigma^2)$. Under the above shift assumption*

$$\widetilde{\xi}_i \Rightarrow \inf_{\lambda \in \Lambda} \int_0^1 [B_i(r, \lambda)]^2 \, dr, \qquad i = 1, 2, 2a, 2b,$$

where $B_i(r, \lambda)$ is defined as in proposition 1.

The proposition is proved in the appendix. First we prove that, under shift assumption, the asymptotic distribution of proposition 3 still holds when the location of the breakpoint is wrongly assumed. Then it is sufficient to apply the continuous mapping theorem as in Zivot and Andrews (1992) to get the result. Note that a co-integration test corresponding to case [1] is proposed by Hao(1996), but he does not apparently make the shift assumption.

Table 6 provides the asymptotic critical values for the statistic $\widetilde{\xi}_i$, $i = 1, 2, 2a, 2b$, obtained by simulation for a sample size of T =500 using 5000 replications. Each replication yielded one value of $\xi_i(\lambda)$ from its asymptotic distribution for all possible breakpoints (2 to 499). The minimum of these values was taken as a realization from the distribution of the inf-statistic, $\widetilde{\xi}_i$.

6 Serial correlation

The break point tests can be extended to models where ε_t is serially correlated. The nonparametric correction of Kwiatkowski et al. (1992) can be applied straightforwardly to yield statistics

$$\xi_i(\lambda, \ell) = \frac{\sum_{t=1}^{T} (\sum_{s=1}^{t} e_s)^2}{T^2 s^2(\ell)}, \qquad i = 1, 2, 2a, 2b, \tag{22}$$

with the same limiting distributions as the corresponding $\xi_i(\lambda)$ tests. The consistency of the tests follow from the argument in Kwiatkowski et al.

(1992) provided ℓ is o(T). The simplified statistics proposed in section 4 and the unknown breakpoint statistics of section 5 can be amended in the same way. Note that it might sometimes be sensible to estimate the long run variance separately in each subsample.

As was noted earlier the KPSS correction can lead to size distortion and/or low power. An alternative strategy, proposed by Harvey and Streibel (1997) is to formulate a structural time series model and then construct the test statistic using the innovations obtained by running the Kalman filter under the null hypothesis with the nuisance parameters estimated under the alternative. The same strategy was advocated by Leybourne and McCabe (1994), except that they work in a slightly less parametric framework based on modelling ε_t by an AR(p) process. The test statistic is constructed from the innovation series $y_t^* = y_t - \widehat{\phi}_1 y_{t-1} - ... - \widehat{\phi}_p y_{t-p}$, where the autoregressive parameters are ML estimates obtained from fitting an ARMA(p,1) model to first differences. The asymptotic distribution of the test statistic is unchanged and Monte Carlo simulations show that in finite samples the procedure has better size and power properties than KPSS .

Both the Harvey and Streibel (1997) and Leybourne and McCabe (1994) procedures can be used with structural breaks.

7 Examples

Annual data on the volume of the flow of the Nile (in cubic metres $\times 10^8$) is shown in Figure 1; see Koopman et al (1995). Fitting a mean and computing the test statistic (2) gives a value of $\xi = 2.527$, indicating a clear rejection of the null hypothesis that there is no random walk component; the asymptotic 5% critical value is 0.461. The KPSS test gives the same result with the statistics for $\ell = 3$ and 7 being 1.100 and 0.735 respectively. However, it is known that the first Aswan dam was constructed in 1899 and if a level intervention is included, neither the LBI nor the simplified test rejects the null hypothesis, since $\xi_1(\lambda) = 0.088$ and $\xi_1^* = 0.301$. In fact a simple random walk plus noise model with a break in the level provides a good fit to the data; see Harvey, Koopman and Penzer (1998). There is a possible outlier in 1913, but including a dummy variable in the model for this year had little impact on the test statistics. Since the stationary part of the model appears to be white noise, the KPSS correction is unnecessary, but the test statistics for $\ell = 3$ and 7 are 0.074 and 0.096 and so it appears to have little adverse effect.

The unconditional test also does not reject the null hypothesis, the test statistic taking the values 0.058, 0.045 and 0.052 for $\ell = 0, 3$, and 7 respectively. However, it is interesting to note that the break point is located at 1897 rather than 1899.

As a second illustration we apply the tests to annual data on US real GNP for the period 1909-1970. These data were used in the well-known article by Nelson and Plosser (1982). On the basis of augmented Dickey-Fuller tests, Nelson and Plosser (1982) did not reject the null hypothesis of a unit root. Subsequently, Perron (1989) observed that a structural break is likely to have occured at the start of the Great Depression, and using his testing procedure he was able to reject the unit root hypothesis in favour of a trend stationary process. Zivot and Andrews (1992) then modified Perron's test by endogenizing the breakpoint, but reached the same conclusion as Perron.

Our own view is that there are a quite a number of places where an argument can be made for the introduction of a break, or a set of breaks, into an economic time series like GNP. Thus we are not dealing with a situation, as in the case of the Nile, where there is a well defined event at a particular point in time which one would expect to give rise to a break in the series. Nevertheless, suppose, following Perron (1989) that we assume there is a known break after 1929 (that is $\tau = 1929$). Table 7a shows the results of applying tests with the KPSS correction in the following cases: no break, break in the level, and break in both level and slope. The columns of the table labelled ℓ=0 to ℓ=8 refer to the lag length in the KPSS correction, while the last two columns report the 5% critical values for $\lambda = 0.3$ and for $\lambda = 0.4$, since the break occurs in between, at $\lambda = 0.34$. The no break case was reported in Kwiatkowski et al. (1992), where it was felt that the evidence favoured the null hypothesis of trend stationarity, though it was noted that the outcome of the test depends on the lag length in the nonparametric correction. Fitting the "Great Crash" model, that is a break in the level as in model 2a, leads to non-rejection of trend stationarity, though the outcome of the test is unclear when we consider a break in both the level and the slope. This finding, then, is in line with Perron (1989). However, the unconditional test, reported in table 7b for the case of model 2a, does not lead to a rejection of trend stationarity, though the implied breakpoint is rather misplaced.

Looking at the graph of the series, in figures 2 and 3, one might equally plausibly assume that there are two breaks: one at the time of the Great Crash and one immediately after the Second World War. Figure 3 shows the series with a fitted deterministic trend and breaks in both the level and the slope in 1929 and 1945. Table 7c shows results obtained with the simplified test, $\xi_2^*(2)$ in (20). These indicate a clear rejection of the trend stationarity hypothesis. Since the breaks have been chosen by examining the data, it could be argued that an unconditional test would be more appropriate. However, this test is redundant if the null hypothesis has already been rejected on the assumption of known breakpoints.

8 Seasonality

The presence of seasonal dummies will not affect the asymptotic distributions of the test statistics described so far. If the seasonal pattern evolves according to a nonstationary process with complex unit roots, it can be modelled explicitly as suggested by Harvey and Streibel (1997) or rendered stationary by an appropriate transformation.

Canova and Hansen (1995) developed a procedure, analogous to the KPSS test, for testing against the hypothesis that a series contains a nonstationary seasonal component. If all the seasonal frequencies are included, the asymptotic distribution of the resulting test statistic under the null is $CvM(s-1)$. As shown in Harvey and Streibel (1997), the inclusion of a time trend does not affect this asymptotic distribution. Thus $CvM_2(s-1)$ is not appropriate unless seasonal slopes, as used, for example, by Proetti (1998), are included in the model; this can be shown by an extension of the argument in Harvey and Streibel (1997, appendix).

The presence of breaks in the level and/or slope will not affect the asymptotic distribution of the Canova-Hansen test. However, a break in the seasonal pattern will. The asymptotic distribution of the LBI test statistic could be derived. This would be tedious to do, but the simplified test statistic, based on the residuals obtained before and after the break, has a $CvM(2s-2)$ asymptotic distribution. With k breaks in the seasonal pattern the degrees of freedom would be $k(s-1)$.

9 Conclusions

This paper has derived the LBI test for the presence of a random walk component in a series with structural breaks at known points. We then proposed a simplified test in which the asymptotic distribution of the test statistic is not dependent on the location of the breakpoint. The asymptotic distribution is the generalised Cramér-von Mises distribution, with degrees of freedom depending on the number of break points. The performance of this simplified test was shown, via some simulation experiments, to be comparable to that of the LBI test, with the LBI test being superior only in the region close to the null hypothesis and for the break point near the beginning or end of the sample.

An unconditional test, based on the assumption that there is a single break at an unknown point, was formulated and critical values computed. However, the examples illustrate that the breakpoint located by the test may not necessarily be the correct or most appropriate one.

Serial correlation can be handled by a nonparametric correction to the test staistics or by explicitly modelling the series. Seasonal effects can also

be allowed for in this way. The tests can also be applied in the context of co-integration when a given co-integrating vector is hypothesized.

The asymptotic distribution of the test for the presence of nonstationary stochastic seasonality is not affected by structural breaks in the trend. It is affected by a break in the seasonal pattern but again the simplified test has a Cramér-von Mises distribution.

ACKNOWLEDGEMENTS

We would like to thank Jukka Nyblom and all the participants in the LSE Econometrics Workshop for helpful comments on an earlier draft. Of course we bear sole responsibility for any errors.

10 APPENDIX

10.1 Proof of Proposition 1

To prove the proposition we use the following two lemmas (whose proofs are trivial). The first lemma contains asymptotic approximations of some functions of time, while the second one deals with simple applications of the invariance principle and continuous mapping theorem. Introduce the notations $\sum\limits_1$ for $\sum_{t=1}^{\tau}$, $\sum\limits_2$ for $\sum_{t=\tau+1}^{T}$, $\overline{t}_1 = \tau^{-1}\sum\limits_1 t$, $\overline{t}_2 = (T-\tau)^{-1}\sum\limits_2 t$ and $\approx$ for asymptotic equivalence.

Lemma 4. *Let $\tau = [\lambda T]$ for $\lambda \in (0,1)$. Then*

- $\sum\limits_1 t \approx \frac{\lambda^2}{2}T^2$,
- $\overline{t}_1 \approx \frac{\lambda}{2}T$,
- $\sum\limits_2 t \approx \frac{1-\lambda^2}{2}T^2$,
- $\overline{t}_2 \approx \frac{1+\lambda}{2}T$,
- $\sum\limits_1 t^2 \approx \frac{\lambda^3}{3}T^3$,
- $\sum\limits_1 (t-\overline{t}_1)^2 \approx \frac{\lambda^3}{12}T^3$,
- $\sum\limits_2 t^2 \approx \frac{1-\lambda^3}{3}T^3$,
- $\sum\limits_2 (t-\overline{t}_2)^2 \approx \frac{(1-\lambda)^3}{12}T^3$,
- $\sum\limits_{s=1}^{t} (s-\overline{t}_1) \approx \frac{r(r-\lambda)}{2}T^2 \quad \textit{with } r = t/T$,
- $\sum\limits_{s=\tau+1}^{t} (s-\overline{t}_2) \approx \frac{(r-1)(r-\lambda)}{2}T^2 \quad \textit{with } r = t/T$.

Lemma 5. *Let $\varepsilon_t \sim iid(0,\sigma^2)$ and $\alpha, \beta \in [0,1]$. Then*

- $\sigma^{-1}T^{-1/2}\sum_{t=[\alpha T]+1}^{[\beta T]} \varepsilon_t \Rightarrow W(\beta) - W(\alpha)$,
- $\sigma^{-1}T^{-3/2}\sum_{t=[\alpha T]+1}^{[\beta T]} t\varepsilon_t \Rightarrow \int_{\alpha}^{\beta} r\,dW(r)$,
- $\sigma^{-1}T^{-3/2}\sum\limits_1 (t-\overline{t}_1)\,\varepsilon_t \Rightarrow \int_0^{\lambda} r\,dW(r) - \frac{\lambda}{2}W(\lambda)$,

• $\sigma^{-1}T^{-3/2}\sum_{2}(t-\overline{t}_2)\,\varepsilon_t \Rightarrow \int_{\lambda}^{1} r\,dW(r) - \frac{1+\lambda}{2}\left(W(1)-W(\lambda)\right),$

where $W(.)$ is a standard Wiener process (Brownian motion).

Consider model 2 first (the proof for model 1 is a special case). Under $H_0 : \sigma_\eta^2 = 0$, we can reparametrize the model as

$$y_t = \mu_1 d_{1t} + \mu_2 d_{2t} + \beta_1 (d_{1t}t) + \beta_2 (d_{2t}t) + \varepsilon_t,$$

where $d_{1t} = 1(t \leq \tau)$, $d_{2t} = 1(t > \tau)$. Denote by $(\cdot)^*$ the transformed variables after partialing out the effect of the two level dummies d_{1t}, d_{2t} (which corresponds to taking deviations from the subsample averages), i.e. transform the model into

$$y_t^* = \beta_1 (d_{1t}t)^* + \beta_2 (d_{2t}t)^* + \varepsilon_t^*.$$

Let $\mathbf{P}$ be the matrix that projects onto the space of these new regressors, with entries

$$p_{st} = \begin{cases} A_1(s-\overline{t}_1)(t-\overline{t}_1) & 1 \leq s \leq \tau,\ 1 \leq t \leq \tau, \\ A_2(s-\overline{t}_2)(t-\overline{t}_2) & \tau < s \leq T,\ \tau < t \leq T, \\ 0 & \text{otherwise}, \end{cases}$$

where $A_1 = \left(\sum_1 (t-\overline{t}_1)^2\right)^{-1}$, $A_2 = \left(\sum_2 (t-\overline{t}_2)^2\right)^{-1}$. Thus the regression residuals are

$$e_s = \begin{cases} \varepsilon_s - \overline{\varepsilon}_1 - A_1 (s-\overline{t}_1) \sum_1 (t-\overline{t}_1)(\varepsilon_t - \overline{\varepsilon}_1) & \text{for } s \leq \tau \\ \varepsilon_s - \overline{\varepsilon}_2 - A_2 (s-\overline{t}_2) \sum_2 (t-\overline{t}_2)(\varepsilon_t - \overline{\varepsilon}_2) & \text{for } s > \tau, \end{cases}$$

where the notation $\overline{x}_i$ stands for the average of the variable x in the i-th subsample, $i = 1, 2$.

Then, using the results in the lemmas above, it is easy to see that

$$\sigma^{-1}T^{-\frac{1}{2}}\sum_{s=1}^{[T\cdot]} e_s \Rightarrow B_2(\cdot, \lambda),$$

where $B_2(r, \lambda)$ is defined in proposition 1. This, the continuous mapping theorem and $\widehat{\sigma}^2 \xrightarrow{p} \sigma^2$ imply the proposition.

For model 2a we use the following parametrization under H_0:

$$y_t = \mu_1 d_{1t} + \mu_2 d_{2t} + \beta t + \varepsilon_t.$$

Taking deviations from the subsample averages, we obtain $y_t^* = \beta t^* - \varepsilon_t^*$. The projection onto the space spanned by t^* is now given by the matrix $\mathbf{P}$,

with entries

$$p_{st} = \begin{cases} A(s-\overline{t}_1)(t-\overline{t}_1), & 1 \le s \le \tau,\ 1 \le t \le \tau, \\ A\left(s-\overline{t}_1\right)\left(t-\overline{t}_2\right), & 1 \le s \le \tau,\ \tau < t \le T, \\ A\left(s-\overline{t}_2\right)\left(t-\overline{t}_1\right), & \tau < s \le T,\ 1 \le t \le \tau, \\ A(s-\overline{t}_2)(t-\overline{t}_2), & \tau < s \le T,\ \tau \le t \le T, \end{cases}$$

where $A = \left[\sum_1 (t-\overline{t}_1)^2 + \sum_2 (t-\overline{t}_2)^2\right]^{-1}$. This yields the following regression residuals

$$e_s = \begin{cases} \varepsilon_s - \overline{\varepsilon}_1 - A\left(s-\overline{t}_1\right)\left(\sum_1 (t-\overline{t}_1)(\varepsilon_t - \overline{\varepsilon}_1) + \sum_2 (t-\overline{t}_2)(\varepsilon_t - \overline{\varepsilon}_2)\right) & \text{for } s \le \tau \\ \varepsilon_s - \overline{\varepsilon}_2 - A\left(s-\overline{t}_2\right)\left(\sum_1 (t-\overline{t}_1)(\varepsilon_t - \overline{\varepsilon}_1) + \sum_2 (t-\overline{t}_2)(\varepsilon_t - \overline{\varepsilon}_2)\right) & \text{for } s > \tau. \end{cases}$$

Then, using the lemmas above, we have that the partial sum process for these residuals weakly converges to $B_{2a}(\cdot, \lambda)$ and therefore the result follows.

In the case of 2b we parametrize the model as

$$y_t = \mu + \beta_1 z_{1t} + \beta_2 z_{2t} + \varepsilon_t,$$

where $z_{1t} = (t-\tau)1(t \le \tau)$ and $z_{2t} = (t-\tau)1(t > \tau)$. Now, first we take whole sample total averages (i.e. project off the constant) and call z^*_{1t}, z^*_{2t} the resulting transformed variables. Then the projection onto the space of the new regressors is defined by the matrix $\mathbf{P} = \mathbf{Z}^* \left(\mathbf{Z}^{*\prime}\mathbf{Z}^*\right)^{-1} \mathbf{Z}^{*\prime}$, with entries

$$p_{st} \approx \frac{3}{\lambda^3(1-\lambda)^3} T^{-3} \left[\left(a z^*_{1s} + b z^*_{2s}\right) z^*_{1t} + \left(b z^*_{1s} + c z^*_{2s}\right) z^*_{2t}\right],$$

where we have already replaced $\left(\mathbf{Z}^{*\prime}\mathbf{Z}^*\right)^{-1}$ with its asymptotic counterpart

$$\frac{3}{\lambda^3(1-\lambda)^3} T^{-3} \begin{bmatrix} a & b \\ b & c \end{bmatrix},$$

with a, b, c defined in the proposition. Using $\mathbf{P}$ we can again compute the residuals and show that the partial sum process weakly converges to $B_{2b}(\cdot, \lambda)$, and the result follows from the same arguments as in the previous cases.

10.2 Proof of proposition 3

Let $\lambda_0 = \tau_0/T$ be the true breakpoint parameter. Consider each of the models (9) to (12) under the null hypothesis of no random walk and rewrite them more compactly as

$$\begin{array}{ll} y_t(\lambda_0) = \mathbf{x}_t'\boldsymbol{\beta} + \varepsilon_t & \text{for } t \le \lambda_0 T, \\ y_t(\lambda_0) = \mathbf{x}_t'\boldsymbol{\beta} + \mathbf{w}_t'\boldsymbol{\delta} + \varepsilon_t & \text{for } t > \lambda_0 T, \end{array} \tag{23}$$

where for instance for case [2] $\mathbf{x}_t = \mathbf{w}_t = (1,t)'$, $\boldsymbol{\beta} = (\mu_0, \beta)'$, $\boldsymbol{\delta} = (\delta_\mu, \delta_\beta)$. We can also rewrite (23) as

$$y_t(\lambda_0) = \mathbf{x}_t'\boldsymbol{\beta} + \mathbf{w}_t(\lambda_0)'\boldsymbol{\delta} + \varepsilon_t, \tag{24}$$

where $\mathbf{w}_t(\lambda_0) = \mathbf{w}_t \cdot 1(t > \lambda_0 T)$.

When we compute the statistic (21) we regress $y_t(\lambda_0)$ on $\mathbf{x}_t$ and $\mathbf{w}_t(\lambda)$ for all λ. So it is useful to express the true model in terms of $\mathbf{w}_t(\lambda)$ with λ not necessarily equal to λ_0. Consider the case $\lambda > \lambda_0$ for example. Then

$$\mathbf{w}_t(\lambda_0) = \mathbf{w}_t(\lambda) + \boldsymbol{\varphi}_t(\lambda_0, \lambda), \tag{25}$$

where $\boldsymbol{\varphi}_t(\lambda_0, \lambda) = \mathbf{w}_t \cdot 1(\lambda_0 T < t \leq \lambda T)$. Therefore we may rewrite (24) as

$$\begin{aligned} y_t(\lambda_0) &= \mathbf{x}_t'\boldsymbol{\beta} + \mathbf{w}_t(\lambda)'\boldsymbol{\delta} + \varepsilon_t^*(\lambda, \lambda_0) && (26) \\ &= \mathbf{z}_t(\lambda)' \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\delta} \end{pmatrix} + \varepsilon_t^*(\lambda, \lambda_0), && (27) \end{aligned}$$

where $\varepsilon_t^*(\lambda, \lambda_0) = \varepsilon_t + \boldsymbol{\varphi}_t(\lambda, \lambda_0)'\boldsymbol{\delta}$ and $\mathbf{z}_t(\lambda) = (\mathbf{x}_t', \mathbf{w}_t(\lambda)')'$. The regression (27) corresponds to the sequence (over λ) of models that we estimate to compute the statistic (21).

Let $\mathbf{M}_Z(\lambda) = \mathbf{I} - \mathbf{Z}(\lambda)\left(\mathbf{Z}(\lambda)'\mathbf{Z}(\lambda)\right)^{-1}\mathbf{Z}(\lambda)'$. The vector of OLS residuals from regression (27) is then

$$\begin{aligned} \mathbf{e}(\lambda, \lambda_0) &= \mathbf{M}_Z(\lambda)\boldsymbol{\varepsilon}^*(\lambda, \lambda_0) && (28) \\ &= \mathbf{M}_Z(\lambda)\left(\boldsymbol{\varepsilon} + \boldsymbol{\varphi}(\lambda, \lambda_0)\boldsymbol{\delta}\right) && (29) \\ &= \mathbf{e}(\lambda, \lambda) + \mathbf{M}_Z(\lambda)\boldsymbol{\varphi}(\lambda, \lambda_0)\boldsymbol{\delta}. && (30) \end{aligned}$$

From this expression we deduce that, under the shift assumption, fitting a model with a misspecified breakpoint does not alter the asymptotic distribution given by proposition 1. In fact $\mathbf{M}_Z(\lambda)\boldsymbol{\varphi}(\lambda, \lambda_0)$ are the residuals from removing the broken trend $\mathbf{z}_t(\lambda)$ from the piecewise linear variable $\varphi_t(\lambda, \lambda_0)$. Obviously these residuals are $O(1)$ when $\varphi_t(\lambda, \lambda_0) = 1(\lambda_0 T < t \leq \lambda T)$ and $O(t)$ when $\varphi_t(\lambda, \lambda_0) = t \cdot 1(\lambda_0 T < t \leq \lambda T)$. Then the partial sum of the elements of $\mathbf{M}_Z(\lambda)\boldsymbol{\varphi}(\lambda, \lambda_0)\boldsymbol{\delta}$ is $o(T^{1/2})$ under the shift assumption. Therefore

$$\begin{aligned} T^{-1/2}\sum_{t=1}^{[T\cdot]} e_t(\lambda, \lambda_0) &= T^{-1/2}\sum_{t=1}^{[T\cdot]} e_t(\lambda, \lambda) + o(1) && (31) \\ &\Rightarrow B_i(\cdot, \lambda), && (32) \end{aligned}$$

where the process $B_i(\cdot, \lambda)$, $r \in [0,1]$, has been defined in proposition 1 for the cases $i = 1, 2, 2a, 2b$.[2]

[2] Alternatively we could have obtained the same result by keeping the magnitude of the

Then, since the statistic (21) is defined as

$$\widetilde{\xi} = \inf_{\lambda \in \Lambda} \left\{ T^{-2} \widehat{\sigma}^{-2} \sum_{t=1}^{T} \left(\sum_{s=1}^{t} e_s(\lambda, \lambda_0) \right)^2 \right\}, \tag{33}$$

where the subscript i has been dropped, applying the continuous mapping theorem as in Zivot and Andrews (1992) yields the result. This is because the statistic can be expressed as a continuous functional of stochastic processes defined on the underlying innovations, ε_t, where continuity is achieved when Λ is a closed subset of the interval (0,1).

shift fixed and assuming that $|\lambda - \lambda_0| \to 0$ sufficiently fast as the sample size increases. For example for a break in the level the partial sum of the elements of $\mathbf{M}_Z(\lambda)\boldsymbol{\varphi}(\lambda, \lambda_0)$ is $O\left(|\lambda - \lambda_0| T\right)$, so the assumption $|\lambda - \lambda_0| = o\left(T^{-1/2}\right)$ would be adequate. Incidentaly, this shows that the LBI test of section 3 is still valid when we don't get the breakpoint exactly but are close to it. However, in this section we are concerned with λ spanning an arbitrary closed subset of the interval (0,1), so assumption 1 is more appropriate.

References

[1]Anderson, T.W. and D.A. Darling (1952), Asymptotic theory of certain 'Goodness of Fit' criteria based on stochastic processes, Annals of Mathematical Statistics 23, 193-212.

[2]Andrews, D.J. (1991), Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica, 59, 817-858.

[3]Bai, J. (1994), Least squares estimation of a shift in linear processes, Journal of Time Series Analysis 15, 453-472.

[4]Bai, J. (1997), Estimation of a change point in multiple regression models, Review of Economics and Statistics, 551-563.

[5]Canova, F. and B.E. Hansen (1995), Are seasonal patterns constant over time? A test for seasonal stability, Journal of Business and Economic Statistics, 2, 292-349.

[6]Doornik, J.A. (1996), OX: an object-oriented matrix language. London: International Thomson Business Press.

[7]Gardner, L.A. (1969), On detecting changes in the mean of normal variates, Annals of Mathematical Statistics 40, 116-126.

[8]Hao, K. (1996), Testing for structural change in cointegrated regression models: some comparisons and generalizations, Econometric Reviews 15, 401-429.

[9]Harvey, A.C., Koopman, S.J. and J. Penzer (1998), Messy time series, Advances in Econometrics, 13 (forthcoming).

[10]Harvey, A.C. and M. Streibel (1997), Testing for nonstationary unobserved components, mimeo.

[11]King, M.L. and G.H. Hillier (1985), Locally best invariant tests of the error covariance matrix of the linear regression model, Journal of The Royal Statistical Society, Series B, 47, 98-102.

[12]Koopman, S.J., Harvey, A.C., Doornik J.A. and N. Shephard (1995), STAMP 5.0 Structural time series analyser, modeller and predictor. London: Chapman and Hall.

[13]Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Y. Shin (1992), Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?, Journal of Econometrics 44, 159-178.

[14]Leybourne, S.J. and B.P.M. McCabe (1994), A consistent test for a unit root, Journal of Business and Economic Statistics 12, 157-166.

[15]MacNeill, I. (1978), Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times, Annals of Statistics 6, 422-433.

[16]Nabeya, S. and K. Tanaka (1988), Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative, Annals of Statistics 16, 218-235.

[17]Nelson, C.R. and C.I. Plosser (1982), Trends versus random walks in macroeconomic time series: some evidence and implications, Journal of Monetary Economics 10, 139-162.

[18]Nyblom, J. (1986), Testing for deterministic linear trend in time series, Journal of the American Statistical Association 81, 545-549.

[19]Nyblom, J. (1989), Testing for the constancy of parameters over time, Journal of the American Statistical Association 84, 223-230.

[20]Nyblom, J. and A.C. Harvey (1997), Tests of common stochastic trends, Mimeo.

[21]Nyblom, J. and T. Mäkeläinen (1983), Comparison of tests for the presence of random walk coefficients in a simple linear model, Journal of the American Statistical Association 78, 856-864.

[22]Perron, P. (1989), The Great Crash, the oil price shock, and the unit root hypothesis, Econometrica 57, 1361-1401.

[23]Phillips, P.C.B. and P. Perron (1988), Testing for a unit root in time series regression, Biometrika 75, 335-346.

[24]Phillips, P.C.B and V. Solo (1992), Asymptotics for linear processes, Annals of Statistics 20, 971-1001.

[25]Proetti, T. (1998), Seasonal heteroscedasticity and trends, Journal of Forecasting, 17, 1-17.

[26]Zivot, E. and D.W.K. Andrews (1992), Further evidence on the Great Crash, the oil-price shock, and the unit-root hypothesis, Journal of Business and Economic Statistics 10, 251-270.

TABLE 1

Upper Tail percentage points of the distribution of the statistic (14) for i=1,2,2a,2b.
(simulated with T=1000, #replications=10000)

i=1: BREAK IN THE LEVEL WITH NO SLOPE.

Lambda	*0.900*	*0.950*	*0.975*	*0.990*
0.01	0.339	0.456	0.559	0.716
0.1	0.285	0.378	0.471	0.607
0.2	0.225	0.293	0.368	0.478
0.3	0.189	0.246	0.302	0.379
0.4	0.161	0.204	0.245	0.303
0.5	0.150	0.187	0.223	0.264
0.6	0.164	0.207	0.251	0.314
0.7	0.191	0.242	0.295	0.378
0.8	0.231	0.305	0.388	0.484
0.9	0.283	0.378	0.484	0.606
0.99	0.345	0.463	0.581	0.748

i=2: BREAK IN BOTH LEVEL AND SLOPE

Lambda	*0.900*	*0.950*	*0.975*	*0.990*
0.01	0.119	0.146	0.172	0.213
0.1	0.095	0.120	0.145	0.175
0.2	0.079	0.097	0.114	0.137
0.3	0.064	0.079	0.095	0.112
0.4	0.056	0.066	0.076	0.091
0.5	0.053	0.062	0.071	0.084
0.6	0.056	0.067	0.078	0.092
0.7	0.065	0.079	0.095	0.118
0.8	0.079	0.095	0.115	0.140
0.9	0.097	0.119	0.142	0.173
0.99	0.117	0.145	0.174	0.213

i=2a: BREAK IN THE LEVEL ONLY

Lambda	*0.9*	*0.95*	*0.975*	*0.99*
0.01	0.119	0.146	0.172	0.213
0.1	0.096	0.122	0.146	0.177
0.2	0.085	0.103	0.122	0.143
0.3	0.086	0.105	0.121	0.142
0.4	0.097	0.123	0.144	0.178
0.5	0.105	0.133	0.162	0.209
0.6	0.097	0.121	0.145	0.177
0.7	0.085	0.102	0.120	0.144
0.8	0.084	0.103	0.123	0.147
0.9	0.098	0.120	0.143	0.173
0.99	0.117	0.145	0.174	0.213

i=2b: BREAK IN THE SLOPE ONLY

Lambda	*0.9*	*0.95*	*0.975*	*0.99*
0.01	0.119	0.146	0.176	0.213
0.1	0.101	0.126	0.152	0.186
0.2	0.088	0.108	0.130	0.155
0.3	0.078	0.097	0.114	0.139
0.4	0.072	0.086	0.100	0.119
0.5	0.070	0.083	0.098	0.116
0.6	0.073	0.089	0.104	0.126
0.7	0.078	0.096	0.116	0.145
0.8	0.087	0.109	0.131	0.161
0.9	0.101	0.126	0.151	0.187
0.99	0.117	0.146	0.175	0.213

TABLE 2

(a) Upper tail percentage points of the $CvM_1(k)$ distribution (no time trend)

k	*90%*	*95%*	*99%*
1	0.347	0.461	0.743
2	0.607	0.748	1.074
3	0.841	1.000	1.359
4	1.063	1.237	1.623

(b) Upper tail percentage points of the $CvM_2(k)$ distribution (time trend)

k	*90%*	*95%*	*99%*
1	0.119	0.149	0.218
2	0.211	0.247	0.329
3	0.296	0.332	0.428
4	0.377	0.423	0.521

Source: Nyblom and Harvey (1997).

TABLE 3

Size and power comparison between the LBI statistic (13) and the simplified statistic (18) for i=2.

i) T=100, #replications=5000.

	lambda / *q*	*0.1*	*0.3*	*0.5*	*0.7*	*0.9*
	0	0.048	0.054	0.048	0.053	0.047
	0.01	0.313	0.221	0.177	0.230	0.325
LBI	0.1	0.884	0.824	0.830	0.819	0.896
	1	0.999	0.999	1	1	0.999
	10	1	1	1	1	1
	0	0.050	0.051	0.048	0.047	0.047
	0.01	0.256	0.194	0.177	0.200	0.260
Simplified	0.1	0.852	0.832	0.830	0.831	0.853
	1	0.999	1	1	1	0.999
	10	1	1	1	1	1

ii) T=200, #replications=5000.

	lambda / *q*	*0.1*	*0.3*	*0.5*	*0.7*	*0.9*
	0	0.050	0.056	0.042	0.052	0.047
	0.01	0.695	0.576	0.534	0.577	0.699
LBI	0.1	0.994	0.989	0.994	0.990	0.993
	1	1	1	1	1	1
	10	1	1	1	1	1
	0	0.049	0.049	0.042	0.041	0.043
	0.01	0.634	0.550	0.534	0.560	0.626
Simplified	0.1	0.991	0.991	0.994	0.993	0.990
	1	1	1	1	1	1
	10	1	1	1	1	1

TABLE 4

Upper tail percentage points for the statistic (14) for i=1 and generalised to the case of 2 structural breaks

lambdas ***perc. points***	***1/8, 1/4***	***1/8, 3/8***	***1/8, 5/8***	***1/4, 1/2***	***1/4, 7/12***	***1/3, 2/3***
0.9	0.205	0.153	0.124	0.112	0.100	0.093
0.95	0.269	0.197	0.151	0.140	0.119	0.110
0.99	0.418	0.300	0.218	0.208	0.167	0.148

TABLE 5

Size and power comparison between the LBI statistic (14) generalised to the case of two structural breaks (and for i=1) and the simplified statistic (20)

	lambdas	***1/8, 1/4***	***1/8, 3/8***	***1/8, 5/8***	***1/4, 1/2***	***1/4, 7/12***	***1/3, 2/3***
	q						
	0	0.048	0.047	0.048	0.051	0.052	0.051
	0.01	0.775	0.723	0.759	0.708	0.735	0.756
LBI	**0.1**	0.989	0.995	0.995	0.997	0.996	0.999
	1	1	1	1	1	1	1
	10	1	1	1	1	1	1
	0	0.047	0.050	0.043	0.055	0.051	0.048
	0.01	0.748	0.753	0.755	0.747	0.743	0.754
Simplified	**0.1**	0.994	0.997	0.998	0.998	0.996	0.999
	1	1	1	1	1	1	1
	10	1	1	1	1	1	1

Note: Simulation with T=200 and #replications=5000.

TABLE 6

Upper tail percentage points for the unconditional breakpoint test

	0.9	*0.95*	*0.99*
i=1	0.071	0.087	0.134
i=2	0.033	0.041	0.054
i=2a	0.071	0.089	0.125
i=2b	0.050	0.060	0.084

Note: simulation with T=500, #replications=5,000.

TABLE 7

A) Structural break in 1929 (lambda=0.34)

	ℓ =0	*ℓ =1*	*ℓ =2*	*ℓ =3*	*ℓ =4*	*ℓ =5*	*ℓ =6*	*ℓ =7*	*ℓ =8*	***5%*** **λ=0.3**	***5%*** **λ=0.4**
No break	0.630	0.337	0.242	0.198	0.173	0.158	0.148	0.141	0.137	**0.149**	**0.149**
Break in the level	0.322	0.182	0.138	0.118	0.107	0.101	0.096	0.093	0.091	**0.105**	**0.123**
Break in level and slope	0.195	0.111	0.086	0.075	0.070	0.068	0.068	0.068	0.070	**0.079**	**0.066**
Simplified statistic	0.529	0.301	0.232	0.204	0.191	0.186	0.184	0.186	0.191	**0.247**	**0.247**

B) Unconditional test (structural break in the level)

	ℓ =0	*ℓ =1*	*ℓ =2*	*ℓ =3*	*ℓ =4*	*ℓ =5*	*ℓ =6*	*ℓ =7*	*ℓ =8*	***5%***	***1%***
Inf-statistic	0.194	0.108	0.081	0.071	0.066	0.065	0.064	0.064	0.066	**0.089**	**0.133**
Breakpoint	1926	1926	1926	1926	1925	1920	1920	1920	1920	***/***	***/***

C) 2 structural breaks: 1929 and 1945

	ℓ =0	*ℓ =1*	*ℓ =2*	*ℓ =3*	*ℓ =4*	*ℓ =5*	*ℓ =6*	*ℓ =7*	*ℓ =8*	***5%***	***1%***
Simplified statistic	0.889	0.552	0.468	0.449	0.452	0.463	0.479	0.501	0.548	**0.332**	**0.428**

Figure 1.

Volume of the flow of the Nile (cubic metres x10^8)

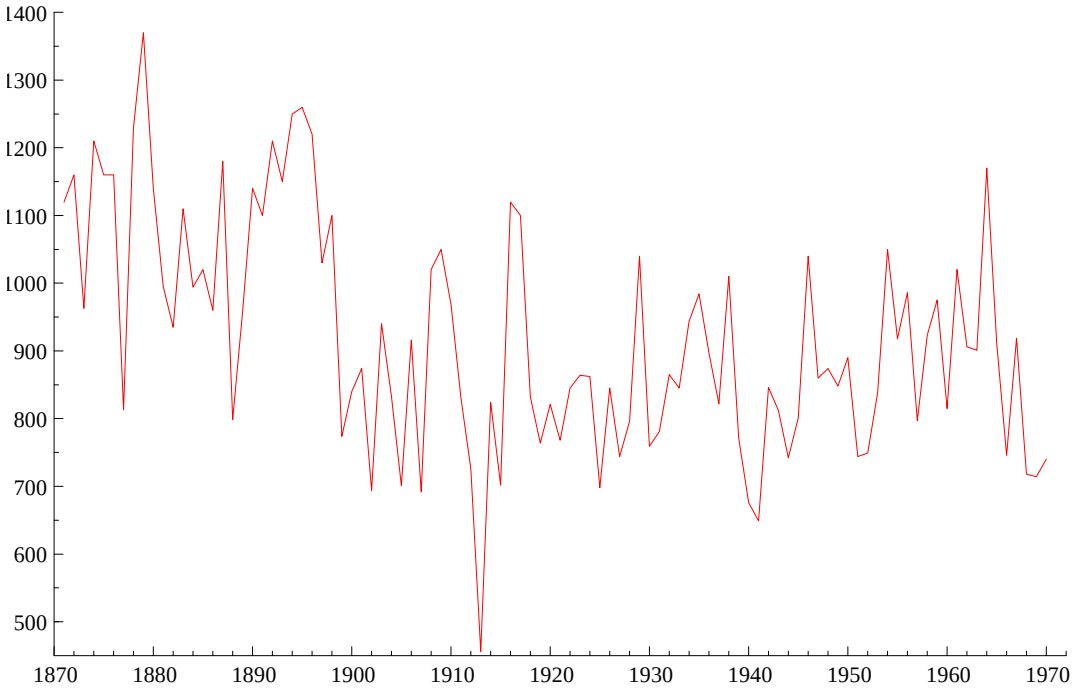

Figure 2.

US gnp and fitted trend with a break in the level.

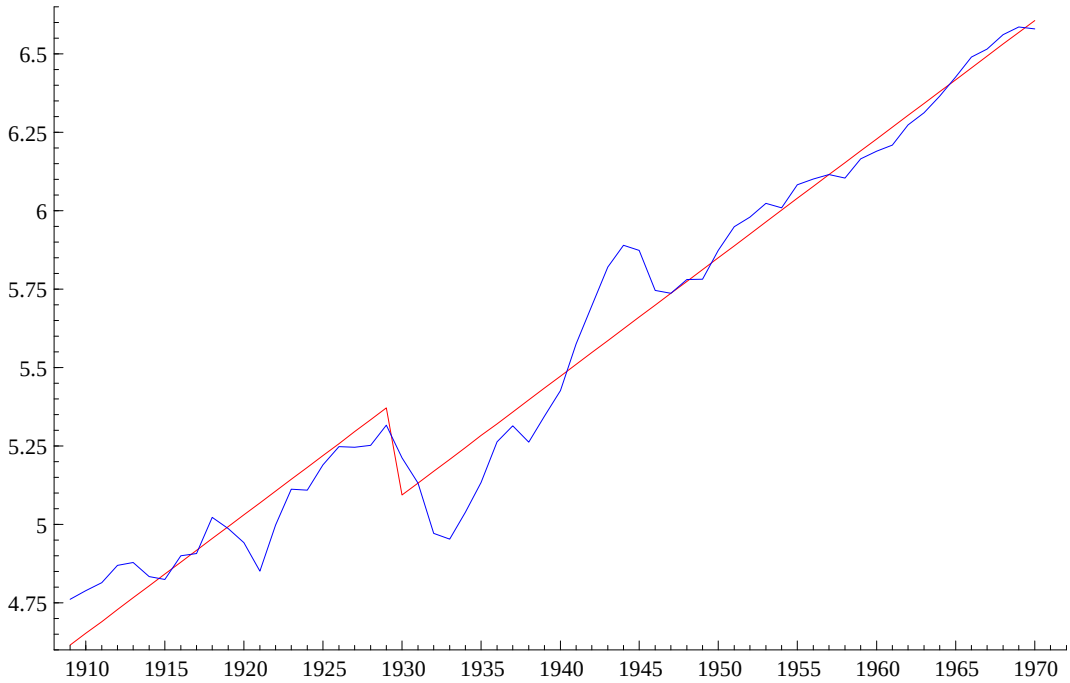
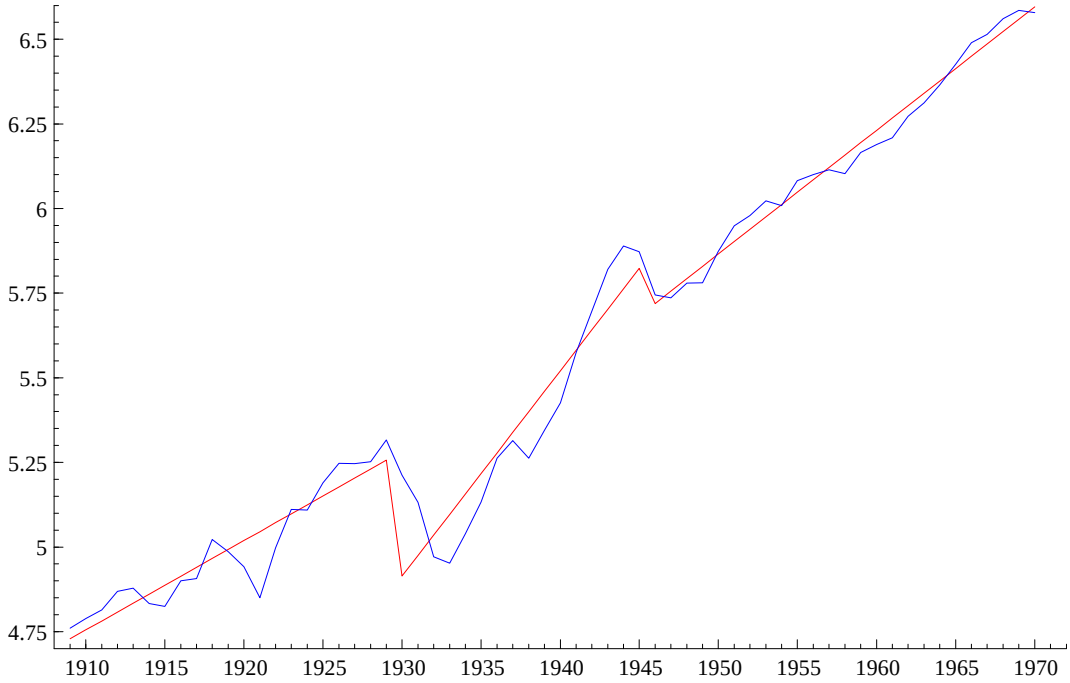
Figure 3.
US gnp and fitted trend with 2 structural breaks.