REVISTA DE MÉTODOS CUANTITATIVOS PARA LA ECONOMÍA Y LA EMPRESA (9). Páginas 3–16.
Junio de 2010. ISSN: 1886-516X. D.L: SE-2927-06.
URL: http://www.upo.es/RevMetCuant/art34.pdf

Spurious Rejections by Dickey-Fuller Tests in the Presence of an Endogenously Determined Break under the Null

Badillo Amador, Rosa
Departamento de Economía
Universidad Politécnica de Cartagena
Correo electrónico: Rosa.Badillo@upct.es

Belaire Franch, Jorge
Departamento de Fundamentos del Análisis Económico
Universidad de Valencia
Correo electrónico: Jorge.Belaire@uv.es

Reverte Maya, Carmelo
Departamento de Economía Financiera y Contabilidad
Universidad Politécnica de Cartagena
Correo electrónico: Carmelo.Reverte@upct.es

ABSTRACT

Leybourne *et al.* (1998) have proved the possibility of a 'converse Perron phenomenon' when conventional Dickey-Fuller tests are applied to determine the order of integration of a time series. That is, if the true generating process is $I(1)$ but with a break, frequent spurious rejections of the null hypothesis can occur. Although Leybourne *et al.* (1998) suggest it would be appropriate to use procedures in which the break date was treated as endogenous, they consider it as exogenous. Thus, this paper analyses whether their results change when the structural break is identified endogenously, that is, if the break point is gleaned from the data. In this sense, applying a recursive t_{DF} test to a unit root process which has a break in its level, there is no virtually evidence of the 'converse Perron phenomenon'. For the rest of the endogeneization procedures (i.e., rolling and sequential) and for the two types of breaks considered (in level or in drift), we find, in line with Leybourne *et al.* (1998), some distortion in the Dickey-Fuller t_{DF} test size, which depends on the break size, the location of the break point in the sample and the sample size.

Keywords: unit roots; structural breaks; Dickey-Fuller tests.
JEL classification: C12; C15; C22.
MSC2010: 62P20.

Artículo recibido el 16 de noviembre de 2009 y aceptado el 29 de enero de 2010.

Rechazos espurios de los test de Dickey-Fuller en presencia de una ruptura bajo la hipótesis nula endógenamente determinada

RESUMEN

Leybourne *et al.* (1998) muestran el cumplimiento del denominado "fenómeno inverso de Perron" cuando se aplican los test convencionales de Dickey-Fuller para determinar el orden de integración de una serie temporal. Este fenómeno consiste en que, si el verdadero proceso generador es $I(1)$ pero con una ruptura, pueden producirse rechazos espurios frecuentes de la hipótesis nula. Aunque Leybourne *et al.* (1998) sugieren que sería apropiado utilizar procedimientos en los que la ruptura sea tratada como endógena, ellos la consideran como exógena. Así, este trabajo analiza si sus resultados cambian cuando la ruptura estructural se determina endógenamente, es decir, a partir de los datos. En este sentido, aplicando el procedimiento t_{DF} recursivo a un proceso de raíz unitaria con una ruptura en el nivel, no encontramos prácticamente evidencia del "fenómeno inverso de Perron". Para el resto de procedimientos de endogeneización (rolling y secuencial) y para los dos tipos de rupturas considerados (en nivel o en deriva) encontramos, en línea con Leybourne *et al.* (1998), alguna distorsión en el tamaño del test t_{DF} de Dickey-Fuller, la cual depende de la magnitud de la ruptura, de su ubicación en la muestra y del tamaño de la misma.

Palabras clave: raíces unitarias; cambios estructurales; test Dickey-Fuller.
Clasificación JEL: C12; C15; C22.
MSC2010: 62P20.

I. INTRODUCTION

Much conventional asymptotic theory for least-squares estimation assumes stationarity, I(0), of the explanatory variables. However, Nelson and Plosser (1982) argue that almost all macroeconomic time series are non-stationary, and typically do have a unit root (I(1) series). The presence or absence of unit roots helps to identify some features of the underlying data generating process of a series. If the series is stationary, it tends to return to its mean value and fluctuate around it within a more-or-less constant range (i.e., it has a finite variance which does not depend on time). On the other hand, non-stationary series have a mean and/or variance depending on time and thus have no tendency to return to long-run deterministic path.

The method of estimation of the standard regression model, Ordinary Least Square (OLS) method, is based on the assumption that the means and variances of these variables being tested are constant over the time. One illustration of the difficulties that can arise when performing an OLS regression with clearly non-stationary series is the problem of *nonsense regression*, so named by Yule (1926), or *spurious regression* in the terminology of Granger and Newbold (1974). That is, given two completely unrelated but integrated series, regression of one on the other will tend to produce statistically significant relationships between the variables when the fact all that is obtained is evidence of contemporaneous correlations rather than meaningful causal relations. Instead, if variables are non-stationary, the estimation of long-run relationship between those variables should be based on the cointegration method. Since the testing of the unit roots of a series is a precondition to the existence of cointegration relationship, Dickey and Fuller (1979) devised a procedure to formally test for non-stationarity (DF test). The simplest form of the DF test amounts to estimating:

$$y_t = \rho y_{t-1} + \varepsilon_t, \qquad (1)$$

with the null being H_0: $\rho=1$ (unit root) against the alternative H_1: $\rho < 1$. The standard approach to testing such a hypothesis is to construct a *t*-test, however, under non-stationarity, the statistic computed does not follow a standard *t*-distribution but, rather, a Dickey-Fuller distribution. This fact justifies the use of Monte Carlo techniques[1], which are developed in Sections II, III and IV of the paper.

[1] These Monte Carlo techniques involve taking (1) as the underlying data generating process (DGP), imposing the null hypothesis by fixing $\rho=1$, and randomly drawing samples of the ε_t from the normal distribution; this then generates thousands of samples of y_t, all of which are consistent with the DGP (1). Then for each of the y_t a regression based on (1) is undertaken, with ρ now free to vary, in order to compute (on the basis of thousands of replications) the percentage of times the model will reject the null hypothesis of a unit root when the null is true. These are the critical values for rejecting the null of a unit root at various significance levels based on the DF distribution of ($\hat{\rho}$ -1)/ $\sigma_{\hat{\rho}}$, been $\sigma_{\hat{\rho}}$ the standard deviation of $\hat{\rho}$.

Following the work of Perron (1989), it is well known, however, that the usual DF test of the unit root null hypothesis can have low power when the true generating process is stationary around a broken linear trend. Perron (1989)'s study was criticized on the grounds that he treated the date of the break as known. Subsequent works used a variety of tests endogenizing the break point (Christiano, 1992; Zivot and Andrews, 1992; Banerjee *et al.*, 1992; Lumsdaine and Stock, 1992; Perron and Vogelsang, 1992; Perron, 1994, 1997 and Vogelsang and Perron, 1998, *inter alia*). The summary picture one gets from these studies is that endogenizing the break point reverses the conclusions arrived at by Perron (1989).

Leybourne *et al.* (1998) have also proved the possibility of the so-called 'converse Perron phenomenon', that is, if the true generating process is I(1) but with a break, frequent spurious rejections of the null hypothesis can occur. They also proved that this phenomenon can lead to a very serious problem of spurious rejections of the unit root null hypothesis, especially if the break occurs early in the series. Leybourne *et al.* (1998) also point out that the practice of using data further back in time to enlarge the series, presumably in search of additional power and more precise estimates, could easily lead to erroneous conclusions if incorporating the additional data introduces a break. These authors consider, as in Perron (1989), the date of the break as known, that is, as an exogenous event.

In this context, the main goal of this paper is to re-examine the Monte Carlo analysis of Leybourne *et al.* (1998) in order to analyse whether the 'converse Perron phenomenon' also holds when the break point is chosen endogenously. In other words, we focus our attention in analysing if endogenizing the break point reverses the conclusions arrived at by Leybourne *et al.* (1998). Therefore, this article considers the presumption that, if there is a break, its date is not known *a priori* but rather is gleaned from the data, as it would be appropriate if there was no strong exogenous reason to suspect a break at a particular time.

Following Banerjee *et al.* (1992), we carry out in this paper a set of t_{DF} tests that control endogenously for structural breaks. These are known as *recursive*, *rolling* and *sequential* tests. Not surprisingly, we obtain that the empirical critical values are well below the full-sample standard t_{DF} test. In addition, we obtain, in some cases, proportions of rejections of the unit root null hypothesis, when it is true, lower than those obtained by Leybourne *et al.* (1998) when the break date is treated as exogeneous. One of these cases is when a break in level is occurred under the null and a recursive t_{DF} test is applied. In this case, the spurious rejection of the null is so low that we can consider that there is virtually no evidence of the 'converse Perron phenomenon'. For the rest of the endogeneization procedures (i.e., rolling and sequential) and for the two types of breaks considered (in level or in drift), we find, in line with Leybourne *et* *al.* (1998), some distortion in the Dickey-Fuller t_{DF} test size, which depends on the break size, the location of the break point in the sample and the sample size.

The rest of the paper is organized as follows. Section II reviews a variety of tests, based on the standard t_{DF} statistics, which treat the break date as unknown *a priori*. Section III reports finite critical values calculated by Monte Carlo experiments for these tests. In Section IV and V we analyse the possibility of spurious rejection of the unit root null hypothesis when an I(1) time series presents a structural change in either its level or its drift and when the tests analysed in Section II are applied. In Section VI the data of Leybourne *et al.* (1998) are re-examined to empirically illustrate the simulation results. Section VII concludes.

II. THE MODELS AND STATISTICS

We begin with a briefly review about the statistical procedures used to test for a unit root allowing for the presence of a structural change in the I(1) generated process. Three classes of standard DF statistics that control endogenously for structural breaks are considered. These are known as *recursive*, *rolling* and *sequential* tests[2].

A traditional DF regression, like this:

$$y_t = \mu + \beta t + \rho y_{t-1} + \varepsilon_t \qquad t = 1, ..., T, \qquad (2)$$

is estimated in this paper. We take subsamples $t = 1,\ldots,\kappa$, where $\kappa = \kappa_0, \kappa_{0+1},\ldots,T$, and using as criteria the minimum values[3] of the t-ratio evaluating $\rho = 1$. κ_0 is the starting value of the recursive estimation and T is the size of the full sample. This test is known as the *recursive* $t_{DF}^{\min}$ *test* ($\hat{t}_{DF}^{\min}$ test).

The *rolling* $t_{DF}^{\min}$ test ($\overline{t}_{DF}^{\min}$) is based on subsamples of fixed size T_s, rolling through the sample. We choose the $t_{DF}^{\min}$ statistic between all subsamples.

Finally, the *sequential* test statistic ($t_{DF}^{*\min}$ test) is computed using the full sample and sequentially incrementing the date of the hypothetical break using a dummy variable and

---

[2] For more details, see Banerjee *et al.* (1992).

[3] We consider the minimal *t*-statistic criteria in all t_{DF} tests that control endogenously for structural breaks due to the fact that we are interested in obtaining the highest spurious rejection frequency. It is clear that the use of other criteria used in Banerjee *et al.* (1992), such as the maximum Dickey Fuller *t*-statistic or a *t*-statistic based on the difference between its maximum and minimum values, would result in lower spurious rejection frequencies.

choosing the lowest value of the statistic. We consider a shift in mean, which is referred by Perron (1989, 1990) as the 'crash' model:

$$y_t = \mu + dD_t(\tau) + \beta t + \rho y_{t-1} + \varepsilon_t, \quad t = 1,...,T \qquad (3)$$

where:

$$D_t(\tau) = \begin{cases} 1, & \text{if } t > \tau T \\ 0, & \text{otherwise} \end{cases} \quad \tau \in (0,1) \qquad (4)$$

and the break fraction is denoted as $\tau=\kappa/T$. The t-stastistic testing $d=0$ provides information about whether there has been a break or jump in the mean. The t_{DF} test evaluating $\rho = 1$ is used to test for the order of integration of the series.

III. CRITICAL VALUES FOR THE RECURSIVE, ROLLING AND SEQUENTIAL TESTS

This section reports finite critical values of recursive, rolling and sequential $t_{DF}^{\min}$ tests. All the calculations have been programmed in Ox 4.1 (http://www.doornik.com). The critical values are computed using data generated for the null model $\Delta y_t = \varepsilon_t$, ε_t *iid* $N(0,1)$ and are based on 10,000 Monte Carlo replications[4] for the following finite sample sizes[5]: T=100, 75 and 50 (see Table 1). The *recursive* statistic, $\hat{t}_{DF}^{\min}$, is computed by estimating (2), under both assumptions: β=0 or $\beta\neq$0 (see in Table 1 *break in level* or *break in drift* columns, respectively), over $t=1,\ldots,\kappa$, for $\kappa = \kappa_0,...,T$, with the following trimming parameter: τ_0=0.25. The *rolling* statistic, $\overline{t}_{DF}^{\min}$, is computed by estimating (2), also under both assumptions: β=0 or $\beta\neq$0, over $t=\kappa-[T\tau_0]+1,\ldots,\kappa$, $\kappa=[T\tau_0],\ldots,T$, being the trimming parameter: τ_0=1/3. The *sequential* statistic, $t_{DF}^{*\min}$, is computed by estimating (3) sequentially, for $\kappa = \kappa_0,...,T-\kappa_0$, under both assumptions: β=0 or $\beta\neq$0, with $D_t(\tau)$ given by (4). For the $t_{DF}^{*\min}$ statistic the trimming parameter is τ_0=0.01. As pointed out by Banerjee *et al.* (1992), the choice of τ_0 for the previous statistics implies a trade-off between needing enough observations in the shortest regression and wanting to capture possible breaks early and late in the sample. As shown in Table 1, recursive, rolling and sequential critical values are, not surprisingly, well below the full-sample standard DF critical values[6].

[4] The use of Monte Carlo method is justified in the Introduction of the paper.
[5] We consider T= 50 and T=75 because a great number of annual macroeconomic time series have small sample sizes. T=100 is also chosen to compare our results with those obtained by Leybourne *et al.* (1998).
[6] See Fuller (1976).

Table 1. Recursive, Rolling and Sequential t_{DF}^{min} Statistics: Critical Values.

	Break in level			
T	Percentile	Recursive $\hat{t}_{DF}^{min}$	Rolling $\bar{t}_{DF}^{min}$	Sequential t_{DF}^{*min}
100	0.010	-4.2865	-5.2763	-4.9464
	0.025	-3.9356	-4.8736	-4.6571
	0.050	-3.6393	-4.5247	-4.3866
	0.100	-3.3446	-4.2341	-4.0984
75	0.010	-4.3915	-5.3762	-5.0532
	0.025	-3.9982	-4.9242	-4.7051
	0.050	-3.7039	-4.5907	-4.4135
	0.100	-3.3678	-4.2551	-4.1251
50	0.010	-4.5631	-6.0037	-5.1457
	0.025	-4.1142	-5.3701	-4.7674
	0.050	-3.7624	-4.9114	-4.4688
	0.100	-3.4080	-4.4835	-4.1421

	Break in drift			
T	Percentile	Recursive $\hat{t}_{DF}^{min}$	Rolling $\bar{t}_{DF}^{min}$	Sequential t_{DF}^{*min}
100	0.010	-4.9516	-5.8392	-5.4076
	0.025	-4.6063	-5.4476	-5.1584
	0.050	-4.3453	-5.1752	-4.9072
	0.100	-4.0021	-4.8272	-4.6183
75	0.010	-5.1104	-6.1137	-5.4617
	0.025	-4.7386	-5.6565	-5.1791
	0.050	-4.4036	-5.2671	-4.9080
	0.100	-4.0559	-4.8914	-4.6110
50	0.010	-5.4902	-6.9147	-5.6003
	0.025	-4.8853	-6.2614	-5.2342
	0.050	-4.5261	-5.7501	-4.9747
	0.100	-4.1104	-5.2707	-4.6490

IV. UNIT ROOT WITH A BREAK IN LEVEL

We next analyse the possibility of spurious rejection of the unit root null hypothesis when recursive, rolling and sequential tests are applied, and when there is a break in an I(1) generating process. In line with Perron (1989) and Leybourne *et al.* (1998), we permit just a single break and we shall concentrate on *additive outlier models*, implying that the break in trend is abrupt. Particularly, we discuss in this section the simplest possible case, where monotonic trend or drift is assumed to be absent. In that case the alternative would be stationarity about a fixed mean, and the null would be I(1) with zero mean change.

The experimental design of Leybourne *et al.* (1998) was employed. Thus, we consider a time series y_t with the following data generation process (DGP):

$$y_t = \alpha s_t(\tau) + v_t, \quad v_t = v_{t-1} + \varepsilon_t, \quad t = 1,...,T \, , \qquad (6)$$

where ε_t~i.i.d.N(0,1). In Eq. (6):

$$s_t(\tau) = \begin{cases} 0, & \mathrm{t} \le \tau \mathrm{T} \\ 1, & \mathrm{t} > \tau \mathrm{T}, \end{cases} \quad \tau \in (0,1) \qquad (7)$$

All simulations are based on 5,000 replications using sample sizes of 100 observations[7]. An additional initial 100 observations were discarded to remove the influence of the initial

[7] Due to space restrictions, we report only the results for T=100. Results for T=75 and T=50 are available from the authors upon request.

condition $y_0 = 0$. In order to compare our results with those of Leybourne *et al.* (1998), the values $\alpha \in \{2.5, 5, 10\}$ were chosen for the break size. The break in level was therefore imposed after observation $\tau T=\kappa$.

For each replication, the $\hat{t}_{DF}^{\min}$ and $\overline{t}_{DF}^{\min}$ tests are estimated using regression (2), under the assumption β=0, and the $t_{DF}^{*\min}$ statistic is estimated using regression (3), under the same assumption and for $D_t(\tau)$ defined in Eq. (4).

The (false) rejections of the unit root hypothesis are noted at the 5% level of significance using the critical values calculated in Section III (see Table 1, *break in level* columns). The resulting empirical rejection frequencies are presented in Tables 2 to 4 for T=100 in *level-break* columns. It can be seen that, using $\hat{t}_{DF}^{\min}$ test, the spurious rejection of the null hypothesis is below the nominal size, and is independent of the location of the break and its magnitude, but not of the sample size, as the higher is T the lower is the spurious rejection rate of the null hypothesis. However, in the case of the $\overline{t}_{DF}^{\min}$ and $t_{DF}^{*\min}$ tests, ignoring the possibility of a break produces many rejections of the null, especially when α increases and when T decreases. For these two tests, the break location also influences on the spurious rejection rate. For example, in the case of the $\overline{t}_{DF}^{\min}$ test, the spurious rejection rate is lower when $\tau > T - \tau_0 T$, since only the subsample from the $T - \tau_0 T$ observation to the last observation is the one capturing the break[8]. Regarding the $t_{DF}^{*\min}$ test, there is a higher rejection rate when the break point is closer to the middle of the sample.

Comparing our results with those obtained in Leybourne *et al.* (1998) where the break point is considered as exogenous, we obtain a lower proportion of rejections of the unit root null hypothesis only when using the $\hat{t}_{DF}^{\min}$ test, and this lower rejection frecuency is occurred for all magnitudes of the break and for the different τ values considered in our study. In this case, the spurious rejection of the null is so low that we can consider that there is virtually no evidence of the 'converse Perron phenomenon'. This finding suggests the use of the $\hat{t}_{DF}^{\min}$ test when there is a break in the level of the series y_t and when its DGP could be given by expression (6).

[8] This phenomenon is inherent to this procedure, which is based on subsamples of fixed size rolling through the sample.

Table 2. Proportion of rejections (at the nominal 5% level) of the unit root null hypothesis when it is true, but there is a break. T=100 (*Recursive t_{DF} test*).

	Level-break			Drift-break		
τ	α=2.5	α=5.0	α=10.0	α=0.5	α=1.0	α=2.0
0.01	0.0040	0.0044	0.0034	0.0558	0.0742	0.1546
0.05	0.0034	0.0032	0.0012	0.0620	0.1388	0.5604
0.10	0.0018	0.0012	0.0004	0.0502	0.1214	0.7320
0.15	0.0024	0.0016	0.0002	0.0344	0.0824	0.7218
0.20	0.0026	0.0024	0.0014	0.0304	0.0420	0.5712
0.25	0.0034	0.0038	0.0040	0.0326	0.0322	0.3322
0.30	0.0030	0.0038	0.0038	0.0350	0.0304	0.1134
0.35	0.0034	0.0038	0.0032	0.0364	0.0306	0.0400
0.40	0.0032	0.0034	0.0028	0.0376	0.0316	0.0304
0.45	0.0032	0.0028	0.0034	0.0386	0.0346	0.0334
0.50	0.0026	0.0028	0.0038	0.0396	0.0362	0.0354
0.55	0.0030	0.0028	0.0028	0.0416	0.0386	0.0376
0.60	0.0032	0.0028	0.0028	0.0416	0.0398	0.0394
0.65	0.0028	0.0026	0.0028	0.0446	0.0418	0.0414
0.70	0.0026	0.0032	0.0028	0.0458	0.0428	0.0426
0.75	0.0028	0.0030	0.0030	0.0464	0.0450	0.0454
0.80	0.0028	0.0032	0.0036	0.0462	0.0458	0.0454
0.85	0.0032	0.0036	0.0044	0.0474	0.0468	0.0466
0.90	0.0030	0.0032	0.0042	0.0482	0.0478	0.0474
0.95	0.0028	0.0034	0.0036	0.0490	0.0484	0.0484
0.99	0.0030	0.0034	0.0036	0.0490	0.0492	0.0492

Table 3. Proportion of rejections (at the nominal 5% level) of the unit root null hypothesis when it is true, but there is a break. T=100 (*Rolling t_{DF} test*).

	Level-break			Drift-break		
τ	α=2.5	α=5.0	α=10.0	α=0.5	α=1.0	α=2.0
0.01	0.0706	0.1998	0.6492	0.0492	0.0500	0.0590
0.05	0.0718	0.1974	0.6424	0.0512	0.0596	0.1360
0.10	0.0688	0.1978	0.6462	0.0512	0.0608	0.1464
0.15	0.0686	0.2040	0.6498	0.0486	0.0574	0.1378
0.20	0.0668	0.1948	0.6476	0.0466	0.0566	0.1410
0.25	0.0662	0.1906	0.6580	0.0460	0.0530	0.1334
0.30	0.0626	0.1956	0.6512	0.0462	0.0490	0.1356
0.35	0.0608	0.1952	0.6446	0.0456	0.0540	0.1324
0.40	0.0602	0.1918	0.6370	0.0442	0.0504	0.1354
0.45	0.0644	0.1878	0.6388	0.0478	0.0488	0.1368
0.50	0.0624	0.1898	0.6372	0.0484	0.0500	0.1396
0.55	0.0656	0.1980	0.6262	0.0434	0.0528	0.1326
0.60	0.0662	0.1908	0.6306	0.0458	0.0472	0.1328
0.65	0.0618	0.1934	0.6386	0.0466	0.0514	0.1330
0.70	0.0638	0.1914	0.6344	0.0462	0.0506	0.1366
0.75	0.0412	0.0360	0.0342	0.0472	0.0432	0.0676
0.80	0.0466	0.0386	0.0372	0.0448	0.0392	0.0372
0.85	0.0472	0.0420	0.0410	0.0450	0.0424	0.0414
0.90	0.0474	0.0454	0.0442	0.0466	0.0452	0.0434
0.95	0.0496	0.0480	0.0480	0.0484	0.0484	0.0462
0.99	0.0512	0.0510	0.0516	0.0490	0.0490	0.0490

Table 4. Proportion of rejections (at the nominal 5% level) of the unit root null hypothesis when it is true, but there is a break. T=100 (*Sequential t_{DF} test*).

	Level-break			Drift-break		
τ	α=2.5	α=5.0	α=10.0	α=0.5	α=1.0	α=2.0
0.01	0.0860	0.2210	0.6678	0.0446	0.0540	0.0834
0.05	0.0540	0.0868	0.3074	0.0646	0.1418	0.5588
0.10	0.0526	0.0748	0.2988	0.0646	0.1746	0.7700
0.15	0.0550	0.0906	0.3488	0.0570	0.1512	0.7978
0.20	0.0578	0.0944	0.3716	0.0478	0.1150	0.7838
0.25	0.0596	0.1024	0.3802	0.0400	0.0760	0.6802
0.30	0.0558	0.1062	0.3999	0.0272	0.0454	0.4492
0.35	0.0558	0.1044	0.4104	0.0226	0.0210	0.1932
0.40	0.0550	0.1048	0.4124	0.0172	0.0072	0.0520
0.45	0.0556	0.0992	0.4128	0.0120	0.0032	0.0054
0.50	0.0544	0.1008	0.4104	0.0096	0.0008	0.0002
0.55	0.0508	0.0958	0.4018	0.0088	0.0000	0.0000
0.60	0.0512	0.0978	0.4012	0.0096	0.0002	0.0000
0.65	0.0498	0.0892	0.3822	0.0116	0.0004	0.0000
0.70	0.0496	0.0812	0.3594	0.0110	0.0006	0.0000
0.75	0.0510	0.0818	0.3336	0.0116	0.0004	0.0002
0.80	0.0488	0.0716	0.3044	0.0112	0.0006	0.0002
0.85	0.0484	0.0584	0.2682	0.0166	0.0016	0.0000
0.90	0.0462	0.0472	0.2220	0.0234	0.0032	0.0006
0.95	0.0448	0.0332	0.1522	0.0348	0.0162	0.0016
0.99	0.0462	0.0312	0.0092	0.0434	0.0436	0.0420

V. UNIT ROOT WITH A BREAK IN DRIFT

In this section we examine the behaviour of the previous tests assuming a different case, i.e., the trend is permitted under the alternative hypothesis and a drift is allowed under the null. Specifically, we generate data from an I(1) process where the mean experiences a single abrupt shift, corresponding under the alternative to the two segments of the trend function joined at the break point. Thus, as Leybourne *et al.* (1998), we consider a time series y_t with the following DGP[9]:

$$y_t = \alpha s_t(\tau) + y_{t-1} + \varepsilon_t, \quad t = 1,...,T, \tag{8}$$

where ε_t~i.i.d.N(0,1) and $s_t(\tau)$ is defined as in Eq. (7)

Again, as Leybourne *et al.* (1998), the sizes of the drift break considered are $\alpha \in \{0.5, 1, 2\}$. For each replication, the $\hat{t}_{DF}^{\min}$, $\overline{t}_{DF}^{\min}$ and $t_{DF}^{*\min}$ tests are estimated under the assumption $\beta \neq 0$ in equation (2) for the $\hat{t}_{DF}^{\min}$, $\overline{t}_{DF}^{\min}$ tests and the same assumption in equation (3) for the $t_{DF}^{*\min}$ test.

The (false) rejections of the unit root hypothesis are noted at the 5% level of significance using the critical values calculated in Section III (see Table 1, *break in drift* columns). The

[9] The treatment of initial conditions, sample size, number of replications and discarding observations for the break in drift experiments are the same as for the previous level break experiments.

resulting empirical rejection frequencies are presented in Tables 2 to 4 for *T*=100 in the *drift-break* columns[10]. It can be seen that, when the break size is low ($\alpha = 0.5$), the three tests ($\hat{t}_{DF}^{\min}$, $\overline{t}_{DF}^{\min}$ and $t_{DF}^{*\min}$) obtain, in general, spurious rejection rates below the significance level. However, for $\hat{t}_{DF}^{\min}$ and $t_{DF}^{*\min}$ tests, a severe phenomenon of spurious rejection of the null emerges when α increases and, contrary to the above section, when *T* increases. For both tests, the size distortion is higher for a break relatively early in the time series, evaporating this problem as τ increases, especially for the $t_{DF}^{*\min}$ test. Comparing our results with those obtained in Leybourne *et al.* (1998) where the break point is considered exogenous, we obtain a lower proportion of rejections of the unit root null hypothesis when using the three tests ($\hat{t}_{DF}^{\min}$, $\overline{t}_{DF}^{\min}$ and $t_{DF}^{*\min}$), for all magnitudes of the break and when the break occurs early in the series. For the rest of τ values, the proportion of rejections of the null hypothesis is below the nominal size. This finding suggests the use of these tests that endogeneize structural breaks when there is a break in the drift of the series y_t and when its DGP could be given by expression (8).

VI. APPLICATION TO GDP

In order to analyze the behaviour of the t_{DF} statistic under the presence of structural breaks, Leybourne *et al.* (1998) study the convergence phenomenon in the economies of a group of west European countries. In particular, they focus on the series of the natural logarithm of the ratio of real output per capita (in U.S. dollars) of Denmark and Germany over the period 1950-1994. Both in levels and first differences (see Figures 1 and 2), it can be observed the possibility of an abrupt break early in the series. With the purpose of re-examining the sensitivity of the findings of Leybourne *et al.* (1998) to the endogeneization of the break through the tests analyzed in this paper, the $\hat{t}_{DF}^{\min}$ and $\overline{t}_{DF}^{\min}$ statistics are obtained when the following augmented DF testing equation was applied:

$$y_t = \mu + \beta t + \rho y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \varepsilon_t \qquad (9)$$

In order to compute the $t_{DF}^{*\min}$ statistic, we additionally include in (9) the regressor $D_t(\tau)$, defined in equation (4). As Leybourne *et al.* (1998), the value for the lag truncation parameter *p* chosen for the above equation is zero (*p=0*), using the general-to specific approach of Ng and Perron (1995) testing at the 5% level (with a maximum possible value of five). Application of

[10] Due to space restrictions, we report only the results for *T*=100. Results for *T*=75 and *T*=50 are available from the authors upon request.

the minimum recursive, rolling and sequential t_{DF} tests ($\hat{t}_{DF}^{\min}$, $\bar{t}_{DF}^{\min}$ and $t_{DF}^{*\min}$) to the series of levels of y_t yields the statistics values -5,96, -4,58 and -5,99, respectively. Compared with the critical values shown in Table 5 corresponding to our sample size of 45 observations, there is an general indication of trend stationarity on the series[11], as in Leybourne *et al.* (1998), except for the rolling test ($\bar{t}_{DF}^{\min}$) where the unit root null hypothesis is not rejected, even at the significance level of 10%. This test behaves better for this particular example in which the series have an abrupt break in drift located early in the sample. Visual inspection of Figures 1 and 2 reveals that this can be effectively the case, supporting the simulation results of the previous sections. This suggests that the procedures in which the break date is treated as endogenous can, in some of the cases shown in this paper, obtain lower spurious rejection rates of the unit root null hypothesis than for the case where the break date is treated as exogenous.

Figure 1. Log ratio of real output p.c. of Denmark and Germany (y_t).

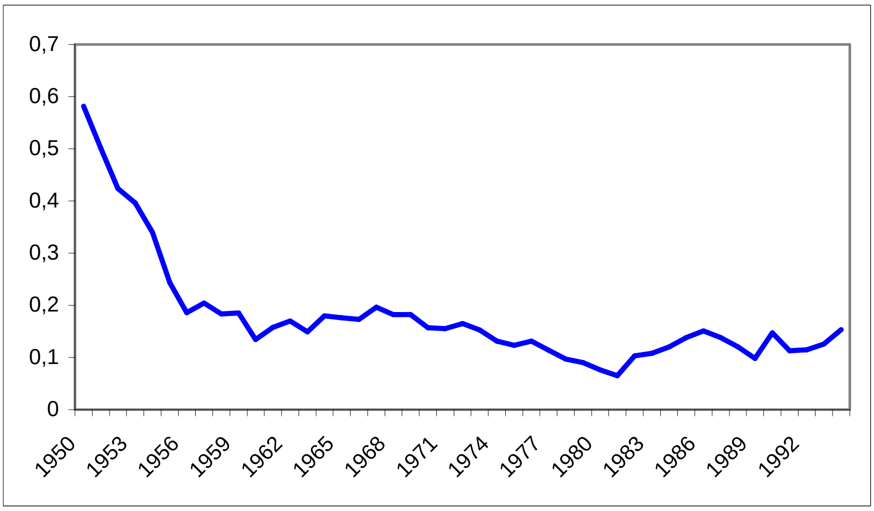
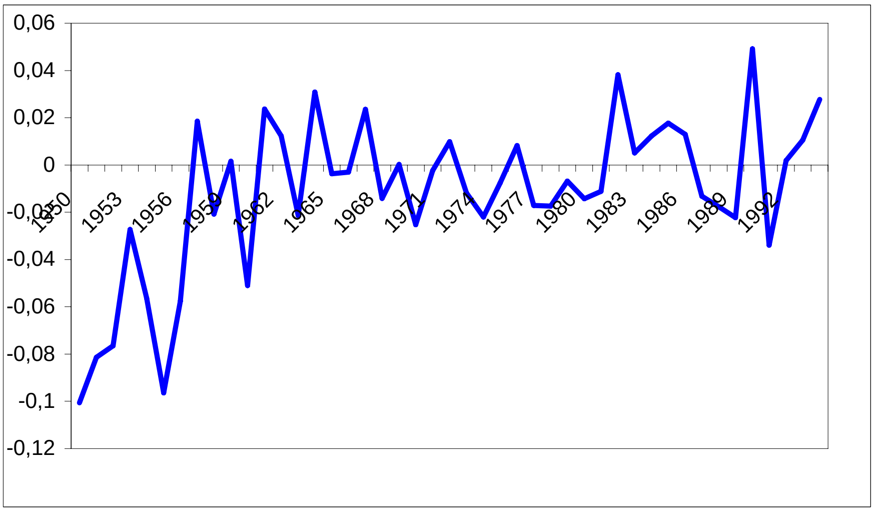


Figure 2. Differences of log ratio of real output p.c. of Denmark and Germany.



[11] Leybourne *et al.* (1998) find the same result for the case of an exogeneous break.

Table 5. Recursive, Rolling and Sequential t_{DF}^{min} Statistics: Critical Values.

	Break in drift			
T	Percentile	Recursive $\hat{t}_{DF}^{min}$	Rolling $\bar{t}_{DF}^{min}$	Sequential t_{DF}^{*min}
45	0.010	-5.7908	-6.8872	-5.6501
	0.025	-5.0958	-6.3216	-5.2675
	0.050	-4.6367	-5.8401	-4.9687
	0.100	-4.2020	-5.3210	-4.6357

VII. CONCLUSIONS

Leybourne *et al.* (1998) have proved the possibility of a 'converse Perron phenomenon' when conventional Dickey-Fuller tests are applied to determine the order of integration of a time series. That is, if the true generating process is I(1) but with a break, frequent spurious rejections of the null hypothesis can occur. However, they suggest it would be appropriate to use procedures in which the break date was treated as endogenous. Thus, the main goal of this paper is to analyse whether their results change when the structural break is identified endogenously, that is, if the break point is related to the data. Applying recursive, rolling and sequential DF type tests that control endogenously for structural breaks, we find no evidence of the 'converse Perron phenomenon' when there is a break in level under the unit root null hypothesis and the recursive procedure is used, contrary to Leybourne *et al.* (1998). However, in line with them, we find some distortion in the t_{DF} test size when using both rolling and sequential procedures for the two types of breaks (in level or in drift) and even when we apply the recursive tests when there is a drift-break in a unit root time series. In those cases in which we find some size distortion, the spurious rejection of the null depends on the break type (in level or in drift), the break size, the location of the break point in the sample (τ) and the sample size. It is noted, however, that, when there is a break in drift of a unit root process, we obtain a lower proportion of rejections of the unit root null hypothesis than Leybourne *et al.* (1998) when the break occurs early in the series and proportions of rejections of the null below the nominal size for the rest of τ values. These findings suggest, in general, the use of these tests that endogeneize structural breaks.

Acknowledgements

This work is an outcome of the research projects: 05838/PHCS/07, financed by '*Programa de Generación de Conocimiento Científico de Excelencia de la Fundación Séneca, Agencia de Ciencia y Tecnología de la Región de Murcia*', and ECO2008-06238-C02-01/ECON, funded by the Spanish Ministry of Education and Science.

REFERENCES

Banerjee, A., Lumsdaine, R.L. and Stock, J.H. (1992) Recursive and sequential tests of the unit-root and trend-break hypotheses: Theory and international evidence, *Journal of Business & Economic Statistics*, **10**, 271–87.

Christiano, L. J. (1992) Searching for a break in GNP, *Journal of Business and Economic Statistics*, **10**, 237–250.

Dickey, D. and Fuller, W. (1979) Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Satistical Association*, **74**, 427–431.

Fuller, W. A. (1976) *Introduction to statistical time series.* New York, Willey.

Granger, C.W.J. and Newbold, P. (1974) Spurious regressions in Econometrics, *Journal of Econometrics,* **2**, 111–120.

Leybourne, S., Mills, T. and Newbold, P. (1998) Spurious rejections by DF tests in the presence of a break under the null, *Journal of Econometrics*, **87**, 191–203.

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

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

Perron, P. (1994) *Trend, unit root and structural change in macroeconomic time series*. In Rao, B. B. (ed), Cointegration for the Applied Economists. MacMillan, New York, NY, 113–146.

Perron, P. (1997) Further evidence on breaking trend functions in macroeconomic variables, *Journal of Econometrics*, **80**, 355–385.

Perron, P., Vogelsang, T. J., (1992) Nonstationarity and level shifts with an application to purchasing power parity, *Journal of Business and Economic Statistics* **10**, 301–320.

Vogelsang, T. J., Perron, P. (1998) Additional tests for a unit root allowing for a break in the trend function at an unknown time, *International Economic Review* **39(4)**, 1073–1100.

Yule, G.U. (1926) Why do we sometimes get nonsense correlations between time series? A study in sampling and the nature of time series, *Journal of the Royal Statistical Society,* **89**, 1–64.

Zivot, E., Andrews, D.W.K., (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.