Structural changes, common stochastic trends, and unit roots in panel data

Jushan Bai
Department of Economics
New York University

Josep Lluís Carrion-i-Silvestre
Department of Econometrics, Statistics and Spanish Economy
University of Barcelona

Abstract

In this paper we propose a new test statistic that considers multiple structural breaks to analyse the non-stationarity of a panel data set. The methodology is based on the common factor analysis in an attempt to allow for some sort of dependence across the individuals. Thus allowing for multiple structural breaks in the "Panel Analysis of Non-stationarity in Idiosyncratic and Common components" (PANIC) methodology increases the degree of heterogeneity when assessing the stochastic properties of the panel data set.

Keywords: multiple structural breaks, common factors, panel data unit root test, principal components

JEL codes: C12, C22, C3, C5

1   Introduction

Nowadays, the increasing application of the panel data techniques to the determination of time series' stochastic properties has led to the development of a wide range of new proposals in the econometric literature. The short time period's coverage that offer most of the available macroeconomic time series may be thought as the main reason behind this exploding phenomenon. This lack of information, in terms of time observations, implies a loss in the power of unit root, stationarity and cointegration tests. The combination of the information in the time and cross-section dimensions to compose a panel data set of individuals, *i.e.* countries or regions, onto which perform the analysis of the stochastic properties has revealed as a promising way to increase the power of these tests.

Thus, a gain in power is expected when performing a statistical inference –unit root, stationarity or cointegration test– using a panel data set made up of individuals that share, at first, some similarities. Breitung and Meyer (1994), Im, Pesaran and Shin (1997), Maddala and Wu (1999) and Levin, Lin and Chu (2002), on the unit root tests, and Pedroni (1995) and Phillips and Moon (1999), on the cointegration analysis, are some of the most relevant papers. Comprehensive surveys of the field can be found in Banerjee (1999), Baltagi (2001) and Baltagi and Kao (2001).

Although the deterministic component should not be of interest when analysing the order of integration of the time series, its misspecification can drive to misleading conclusions. Thus, a stationary time series that evolves around a breaking-trend model might be characterized as a non-stationary process if the order of integration analysis fails to consider the structural breaks –see Perron (1989) for the univariate time series framework and Carrion-i-Silvestre, del Barrio and López-Bazo (2001) for the panel data framework. Our proposal focus on the presence of multiple structural breaks affecting the panel data set, so that taking into account the presence of these structural breaks overcomes the interferences that can cause the misspecification error in the stochastic properties of the panel. In this paper we analyse the presence of multiple structural breaks when testing for the unit root hypothesis in a panel data framework. Some of the recent proposals in the panel data based unit root and stationarity tests have addressed this question by developing suitable tests –see Im, Lee and Tieslau (2002) for the LM test and Carrion-i-Silvestre *et al.* (2001) for the DF with one structural break, and Carrion-i-Silvestre *et al.* (2002) for the KPSS tests with multiple structural breaks. However, our approach overcomes the criticism that has raised the assumption of cross-section independence in which most of the panel data based tests rely, and models the cross-section dependence in terms of the common factors as in Bai and Ng (2001, 2004). Briefly speaking, the idea is to establish a distinction between comovements and idiosyncratic shocks that may be affecting the individual time series. Filtering out the comovements will reduce the noise in the system, so that, the analysis will focus on those shocks that are specific for each individual. Moreover, note that the cross-section independence is more likely to be fulfilled when using these idiosyncratic shocks than when using the raw data.

The rest of the paper is organized as follows. Section 2 describes the model and the two deterministic specifications that are considered along the paper. These models arise because of the different effects that the structural breaks may cause on the deterministic part of the model. Section 4 presents different pooled tests, while in Section ?? we analyse the finite sample performance. Finally, Section 6 concludes. All proofs are presented in the Appendix.

2 Panel unit root test with multiple structural breaks

Let us define the panel data model given by:

$$X_{i,t} = D_{i,t} + F_t \pi_i + e_{i,t}; \tag{1}$$

$$(I - L) F_t' = C(L) u_t; \tag{2}$$

$$(1 - \rho_i L) e_{i,t} = H_i(L) \varepsilon_{i,t}, \tag{3}$$

$t = 1, \ldots, T$, $i = 1, \ldots, N$, where $C(L) = \sum_{j=0}^{\infty} C_j L^j$ and $H_i(L) = \sum_{j=0}^{\infty} H_{i,j} L^j$. $D_{i,t}$ denotes the deterministic part of the model, F_t is a $(l \times 1)$-vector that accounts for the common factors that are present in the panel and $e_{i,t}$ is the idiosyncratic disturbance term. Our analysis is based on the same set of assumptions in Bai and Ng (2004). Let $M < \infty$ be a generic positive number, not depending on T and N:

Assumption A: (i) for non-random π_i, $\|\pi_i\| \leq M$; for random π_i, $E\|\pi_i\|^4 \leq M$, (ii) $\frac{1}{N}\sum_{i=1}^{N}\pi_i\pi_i' \xrightarrow{p} \Sigma_\Pi$, a $(l \times l)$ positive matrix.

Assumption B: (i) $u_t \sim iid(0, \Sigma_u)$, $E\|u_t\|^4 \leq M$, and (ii) $Var(\Delta F_t') = \sum_{j=0}^{\infty} C_j \Sigma_u C_j' > 0$, (iii) $\sum_{j=0}^{\infty} j\|C_j\| < M$; and (iv) $C(1)$ has rank l_1, $0 \leq l_1 \leq l$.

Assumption C: (i) for each i, $\varepsilon_{i,t} \sim iid(0, \Sigma_\varepsilon)$, $E|\varepsilon_{i,t}|^8 \leq M$, $\sum_{j=0}^{\infty} j|H_{i,j}| < M$, $\omega_i^2 = H_i(1)^2 \sigma_i^2 > 0$; (ii) $E(\varepsilon_{i,t}\varepsilon_{j,t}) = \tau_{i,j}$ with $\sum_{i=1}^{N}|\tau_{i,j}| \leq M$ for all j; (iii) $E\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[\varepsilon_{i,s}\varepsilon_{i,t} - E(\varepsilon_{i,s}\varepsilon_{i,t})]\right|^4 \leq M$, for every (t, s).

Assumption D: The errors $\varepsilon_{i,t}$, u_t, and the loadings π_i are three mutually independent groups.

Assumption E: $E\|F_0\| \leq M$, and for every $i = 1, \ldots, N$, $E|e_{i,0}| \leq M$.

Assumption A ensures that the factor loadings are identifiable. Assumption B establishes the conditions on the short and long-run variance of ΔF_t –i.e. positive definite short-run variance and long-run variance that can be of reduced rank in order to accomodate linear combinations of $I(1)$ factors to be stationay. Assumption C(i) allows for some weak serial correlation in $(1 - \rho_i L) e_{i,t}$, whereas C(ii) and C(iii) allow for weak cross-section correlation. Finally, Assumption E defines the initial condition on $e_{i,t}$.

This model expresses the stochastic process $X_{i,t}$ as the sum of up to three different components, so that we can focus on each of these components to characterize $X_{i,t}$ in terms of its stochastic properties. Note that the non-stationarity of $X_{i,t}$ can be due to the non-stationarity of either F_t or $e_{i,t}$, so that we have two potential sources of non-stationarity with different economic interpretations. Thus, the matrix F_t collects the common effects that are present across the cross-section dimension and, therefore, the non-stationarity of F_t will mean that all individuals in the panel are *common* non-stationary. These effects affect the individuals with different magnitude (π_i). However, even if $X_{i,t}$ is driven by a common non-stationary component (F_t), the idiosyncratic effect may be

$e_{i,t} \sim I(0)$. This will mean that the stochastic shocks that only affect each individual are stationary. Hence, the non-stationarity analysis can be performed through the application of unit root tests on F_t and $e_{i,t}$.

Regarding the deterministic component, the specification that is adopted in the model is quite general to allow for the presence of multiple structural breaks. Specifically, we formulate:

$$D_{i,t} = \mu_i + \beta_i t + \sum_{k=1}^{m_i} \theta_{i,k} DU_{i,k,t} + \sum_{k=1}^{m_i} \gamma_{i,k} DT_{i,k,t}^*, \qquad (4)$$

that is, we allow for m_i structural breaks affecting the mean of the time series. The dummy variables are defined as $DU_{i,k,t} = 1$ and $DT_{i,k,t}^* = \left(t - T_{b,k}^i\right)$ for $t > T_{b,k}^i$ and 0 elsewhere, where $T_{b,k}^i$ denotes the k-th date of the break for the i-th individual, $k = 1, \ldots, m_i$, $m_i \geq 1$. In fact, equation (4) nests two different specifications depending on the effect of the structural breaks on the deterministic components. On the one hand, we can introduce the constraint $\beta_i = \gamma_{i,k} = 0$, $\forall i, k$, in (4) to analyse the stochastic properties of panel data sets formed by non-trended variables –for instance, the PPP hypothesis ought to be tested using this specification. Hereafter, the constrained model is denoted as Model 1. Formally speaking, Model 1 implies the following deterministic specification:

$$D_{i,t} = \mu_i + \sum_{k=1}^{m_i} \theta_{i,k} DU_{i,k,t},$$

which includes individual effects and individual shifting effects. On the other hand, we will denote the unconstrained model given by (4) as Model 2, a specification that is suitable for trended variables that may be affected by structural breaks that shift both the individual and the specific time trend –for instance, the analysis of the unit root hypothesis in GDP should be based on this specification.

Notice that both models assume that the structural breaks are idiosyncratic for the individuals, since (i) they can be positioned at different dates for each individual, (ii) they may have different magnitude and (iii) each individual may have different number of structural breaks. Therefore, our specification takes into account a high degree of individual's heterogeneity. Once the model have been defined in a general way, now we are going to address the unit root null hypothesis testing through the consideration of two situations: first, we assume that there are no common factors, $\pi_i = 0$ $\forall i$ in (1) and, second, we allow for the presence of such common factors, $\pi_i \neq 0$ in (1), $i = 1, \ldots, N$. For ease of exposition, at first we take the date of the breaks as known. Once the limit distributions are derived, we introduce the discussion about the procedures that can be applied in order to estimate them.

2.1 Individuals are assumed to be independent across i

From a theoretical point of view, it is of interest to consider the simplified situation in which $\pi_i = 0\ \forall i$ in (1) and $\{e_{i,t}\}$ is a stochastic process independent across $i = 1, \dots, N$. In order to test the null hypothesis that $X_{i,t} \sim I(1)$, $\forall i,\ i = 1, \dots, N$, we suggest to compute the square of the modified Sargan-Bhargava (MSB) test statistic defined in Stock (1999):

$$MSB_i\left(\lambda_i\right) = \frac{T^{-2} \sum_{t=1}^{T} \tilde{X}_{i,t-1}^2}{\tilde{\sigma}_i^2}, \qquad (5)$$

where $\tilde{X}_{i,t} = X_{i,t} - \tilde{D}_{i,t}$ and $\tilde{\sigma}_i^2$ is the long-run variance of $\Delta \tilde{X}_{i,t}$. We have made explicit the dependency of the test on the structural breaks through the consideration of λ_i in the notation, where $\lambda_i = \left(\lambda_{i,1}, \dots, \lambda_{i,m_i}\right)'$, $\lambda_{i,k} = T_{b,k}^i/T$, $k = 1, \dots, m_i$, is the so-called vector of break fraction parameters. The limit distribution of (5) for the two different models considered in the paper is given in the following Theorem.

Theorem 1 *Let $X_{i,t}$, $i = 1, \dots, N$, $t = 1, \dots, T$, be the stochastic process generated by (1) with $\pi_i = 0\ \forall i$ and $\rho_i = 1$ in (3). As T, $T_{b,k}^i \to \infty$ in a way that $\lambda_{i,k} = T_{b,k}^i/T$ remains constant, $\forall i, k,\ i = 1, \dots, N,\ k = 1, \dots, m_i$, then the test in (5) converges to:*

(1) Model 1: $\quad MSB_i\left(\lambda_i\right) \Rightarrow \sum_{k=1}^{m_i+1} \left(\lambda_{i,k} - \lambda_{i,k-1}\right)^2 \int_0^1 V_{i,k}^{\mu}\left(b\right)^2 db$
(2) Model 2: $\quad MSB_i\left(\lambda_i\right) \Rightarrow \sum_{k=1}^{m_i+1} \left(\lambda_{i,k} - \lambda_{i,k-1}\right)^2 \int_0^1 V_{i,k}^{\tau}\left(b\right)^2 db$

where $\Rightarrow$ denotes weak convergence of the associated measure of probability, $V_{i,k}^{\mu}\left(b\right) = W_{i,k}\left(b\right) - \int_0^1 W_{i,k}\left(s\right) ds$, and $V_{i,k}^{\tau}\left(b\right) = W_{i,k}\left(b\right) - \left(4 - 6b\right) \int_0^1 W_{i,k}\left(s\right) ds - \left(-6 + 12b\right) \int_0^1 s W_{i,k}\left(s\right) ds$, with $W_{i,k}\left(b\right)$ the standard Brownian motion, and $\lambda_{i,0} = 0$ and $\lambda_{i,m_i+1} = 1$.

Theorem 1 shows that the limit distribution of the $MSB_i\left(\lambda_i\right)$ test is function of Brownian motions and two nuisance parameters –*i.e.* the break fraction parameters $\left(\lambda_i\right)$ and the number of structural breaks $\left(m_i\right)$. Moreover and as shown in the Appendix, when there is only one structural break, $m_i = 1$, the limit distribution of the test is symmetric around $\lambda_i = 0.5$. Finally, note that for $m_i = 0$ the limit distributions in Theorem 1 coincide with the ones given in Stock (1999). Besides, although the situation in which $N = 1$ can be understood as a special case, this is of great interest provided that it generalises the proposal in Perron (1997) and Lumsdaine and Papell (1997) through the consideration of multiple structural breaks in the non-stationarity analysis. Thus, our can be applied to test the null hypothesis of unit root on a single time series allowing for the presence of multiple structural breaks both under the null and alternative hypotheses.

As mentioned above, Theorem 1 indicates that the limit distribution of the MSB test depends both on the number of structural breaks $\left(m_i\right)$ and their location $\left(\lambda_i\right)$. This gives rise to two possible situations. First, practitioners should be

willing to assume that the number and dates of the structural breaks are known. For instance, the German reunification and the Euro currency's birth are two events for which the exogenous nature of the structural breaks can be assumed. However, this situation is rarely found in practice, so that the computation of the MSB will require the application of a consistent estimation procedure to determine the number of structural breaks and the respective vector of break fraction parameters. This defines the second situation of interest.

Let us now focus on the first situation in which both the number and the position of the structural breaks are known. The MSB test can be computed and compared to the critical values drawn from the limit distributions in Theorem 1. However, we believe that the availability of the assossiated p-value could be more informative when performing the statistical inference. Provided that the MSB test has a non-standard limit distribution, the p-values have to be approximated by simulations. MacKinnon (1994), Adda and Gonzalo (1996), Hansen (1997), and Bai and Ng (2003) computed asymptotic p-values for test statistic with non-standard distribution. Here we follow MacKinnon (1994) and estimate a set of response surfaces to approximate the p-values of the MSB test. However, we generalise the previous proposals and estimate response surfaces for the p-values that take into account the sample size. The estimation is made assuming a probit model for the p-value (p_i) as a function of powers of the quantile (q_i), the sample size and the break fraction parameters, $\log\left(\frac{p_i}{1-p_i}\right) = g\left(q_i, T, \lambda_i\right)$. We have essayed different functional forms using the Newey-West robust covariance estimator to analyse the individual significance of the parameters. In concrete, for the situation in which $m_i = 0$ the response surface is given by:

$$\log\left(\frac{p_i}{1-p_i}\right) = \sum_{j=0}^{1}\left(\zeta_{0j} + \zeta_{1j}q_i + \zeta_{2j}q_i^{-1/2} + \zeta_{3j}q_i^{-1/3} + \zeta_{4j}q_i^{-1/4}\right)\left(\frac{1}{T}\right)^j + u_i,$$

(6)

where, for each sample size (T), 1,000 quantiles, $i = 1, \ldots, 1000$, has been computed from the empirical distribution to estimate the model. We have conducted a Monte Carlo experiment to obtain the empirical distribution of the MSB test for $T = \{30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 125, 150, 175, 200, 225, 250, 300, 350, 400, 450, 500, 2000\}$ using 50,000 replications. The p-values response surfaces are collected in Panel A of Table 1. Similar response surfaces are presented in Panel A of Table 2 for $m_i = 1$. Note that (6) does not produce a direct estimate of p_i. The estimate of p_i is obtained from

$$\hat{p}_i = \frac{\exp\left\{\hat{g}\left(q_i, T, \lambda_i\right)\right\}}{1 + \exp\left\{\hat{g}\left(q_i, T, \lambda_i\right)\right\}}.$$

Let us now focus on the procedures that are based on the endogenous determination of the breaking points. The proposal described in Bai and Perron (1998) is very convenient for the specification in Model 2, provided that both the number and dates of the breaks can be consistently estimated under the null hypothesis taking the first difference of y_t. Therefore, the problem reduces to the

Table 1: Response surfaces for the p-values estimation for $m = 0$

	Panel A		Panel B	
	Model 1	Model 2	Model 1	Model 2
$\hat{\zeta}_{00}$	19.6465	51.4380	10.7071	20.6574
$\hat{\zeta}_{10}$	3.8634	12.9479	1.0586	3.7891
$\hat{\zeta}_{20}$	-13.0719	-24.3411	-8.9767	-13.5487
$\hat{\zeta}_{30}$	76.9588	168.6843	46.7634	80.4365
$\hat{\zeta}_{40}$	-81.4440	-189.8098	-47.6969	-85.3760
$\hat{\zeta}_{01}$	178.3351	968.0575	-41.0104	
$\hat{\zeta}_{11}$	29.6877	123.1754	9.5077	41.7794
$\hat{\zeta}_{21}$	-191.4098	-586.9201	-31.2927	-130.8045
$\hat{\zeta}_{31}$	1098.9950	3963.1880	63.8206	587.8564
$\hat{\zeta}_{41}$	-1093.0750	-4307.8610		-484.0637

Panel A corresponds to Theorem 1 and Panel B corresponds to Theorem 2. The functional form is given by $\log\left(\frac{p_i}{1-p_i}\right) = \sum_{j=0}^{1} \left(\zeta_{0j} + \zeta_{1j}q_i + \zeta_{2j}q_i^{-1/2} + \zeta_{3j}q_i^{-1/3} + \zeta_{4j}q_i^{-1/4}\right)\left(\frac{1}{T}\right)^j + u_i$. The $\bar{R}^2$ of all these estimations were 0.99. The included parameters were significant at the 5% level - we used the Newey-West robust estimator to compute the s.e.

identifycation of level shifts on Δy_t, a stationary variable, on which the dynamic optimization algorithm in Bai and Perron (1998) can be applied. Notwithstanding, for the Model 1 we have to follow a different approach given that taking the first difference of y_t will imply dating impulse outliers –additive outliers (AO)– and this situation is not covered in Bai and Perron (1998). The standard way to deal with AO outliers requires the estimation of a fully parametrised ARMA model on which the outlier detection analysis is performed using a t statistic in an iterative fashion –see Tsay (1986) and Chen and Liu (1993), among others. This iterative approach was followed in Franses and Haldrup (1994) to allow for AO outliers in the ADF test. However, two main drawbacks can be highlighted. First, it requires to control the dynamic structure –i.e. estimation of a fully parametrised ARMA model– and, second, the t statistic that is used to detect the presence of outliers relies on the distributional assumptions about the error term.

Instead, we could estimate the shift dates using the proposals in Perron and Vogelsang (1992) and Vogelsang (1998). Briefly speaking, Perron and Vogelsang (1992) date the breaking points in the additive specification through the minimisation of the significance test of the dummy parameters. On the other hand, Vogelsang (1998) uses the $\sup PS_T$ test which does not rely on the dynamic of the system and, hence, serial-correlation parameters does not have to be esti-

Table 2: Response surfaces for the p-values estimation for $m = 1$

| | Panel A | | Panel B |
	Model 1	Model 2	Model 2
$\hat{\zeta}_{00}$	35.490	111.3526	29.997
$\hat{\zeta}_{10}$	13.715	2.593527	7.791
$\hat{\zeta}_{20}$	-19.677	-38.27770	-17.269
$\hat{\zeta}_{30}$	129.744	299.7168	110.703
$\hat{\zeta}_{40}$	-142.760	-356.3048	-120.846
$\hat{\zeta}_{01}$	1029.652	6.959164	592.798
$\hat{\zeta}_{11}$	3.955	-16.29266	-12.452
$\hat{\zeta}_{21}$	-571.744	9.835470	-411.448
$\hat{\zeta}_{31}$	3928.595	235.5653	2625.848
$\hat{\zeta}_{41}$	-4326.511	-591.7060	-2787.674
$\hat{\varphi}_{10}$	10.181	374.3920	9.407
$\hat{\varphi}_{20}$	-33.434	18898.49	-33.580
$\hat{\varphi}_{30}$	46.504	-44424.27	48.346
$\hat{\varphi}_{40}$	-23.252	28018.83	-24.173
$\hat{\varphi}_{11}$	73.704	-5029.990	39.579
$\hat{\varphi}_{21}$	-213.648	12255.89	-132.099
$\hat{\varphi}_{31}$	279.888	-8638.537	185.039
$\hat{\varphi}_{41}$	-139.944	28817.84	-92.519
$\hat{\vartheta}_{1}$	-96.699	-69590.05	-76.999
$\hat{\vartheta}_{2}$	555.555		444.876
$\hat{\vartheta}_{3}$	-917.712		-735.753
$\hat{\vartheta}_{4}$	458.856		367.876

Panel A corresponds to Theorem 1 and Panel B corresponds to Theorem 2. The functional form is given by $\log\left(\frac{p_i}{1-p_i}\right) = \sum_{j=0}^{1}\left(\zeta_{0j} + \zeta_{1j}q_i + \zeta_{2j}q_i^{-1/2} + \zeta_{3j}q_i^{-1/3} + \zeta_{4j}q_i^{-1/4}\right)\left(\frac{1}{T}\right)^j + \sum_{j=0}^{1}\left(\varphi_{1j}\lambda_i + \varphi_{2j}\lambda_i^2 + \varphi_{3j}\lambda_i^3\right)\left(\frac{1}{T}\right)^j + \sum_{j=1}^{3}\vartheta_j q_i \lambda_i^j + u_i$. The $\bar{R}^2$ of all these estimations were 0.99. The included parameters were significant at the 5% level - we used the Newey-West robust estimator to compute the s.e.

8

mated. However, these proposals do not provide a good approximation. On the one hand, Perron and Vogelsang (1992) show that the date of the break is not identified under the null alternative of unit root. On the other hand, the test in Vogelsang (1998) is not consistent when $y_t \sim I(1)$ since it has the same limiting distribution under the null and the alternative hypothesis. Therefore, this test should not be used to estimate the location of the level shift. To overcome these limitations we propose the use of the procedure defined in Carrion-i-Silvestre (2003), which consists on the identification of AO's in the first differenced time series without having to specify a fully parametrised model as required in the existing proposals.

Finally, for further purposes it would be useful to derive the mean and variance of the limit distribution of MSB for Models 1 and 2. Specifically, these two moments are used to define one of the pooled tests in Section 4. They are presented in the following Proposition.

Proposition 1 *Let $MSB_i(\lambda_i) = \tilde{\sigma}_i^{-2} T^{-2} \sum_{t=1}^{T} \tilde{X}_{i,t-1}^2$ be the test statistic with limit distribution given in Theorem 1. Moreover, let $\xi_i = E(MSB_i(\lambda_i))$ and $\varsigma_i^2 = V(MSB_i(\lambda_i))$ be the mean and variance of $MSB_i(\lambda_i)$ respectively. Then, as T, $T_{b,k}^i \to \infty$ in a way that $\lambda_{i,k} = T_{b,k}^i/T$ remains constant, $\forall i, k$, $i = 1, \ldots, N$, $k = 1, \ldots, m_i$,*

$$
\begin{aligned}
\text{(1) Model 1:} \quad & \xi_i = \tfrac{1}{6} \sum_{k=1}^{m_i+1} (\lambda_{i,k} - \lambda_{i,k-1})^2 \ and \\
& \varsigma_i^2 = \tfrac{1}{45} \sum_{k=1}^{m_i+1} (\lambda_{i,k} - \lambda_{i,k-1})^4, \\
\text{(2) Model 2:} \quad & \xi_i = \tfrac{1}{15} \sum_{k=1}^{m_i+1} (\lambda_{i,k} - \lambda_{i,k-1})^2 \ and \\
& \varsigma_i^2 = \tfrac{11}{6300} \sum_{k=1}^{m_i+1} (\lambda_{i,k} - \lambda_{i,k-1})^4,
\end{aligned}
$$

where $\lambda_{i,0} = 0$ and $\lambda_{i,m_i+1} = 1$.

Note that these moments are function of the break fraction parameters. Besides, when there are no structural breaks they coincide with the mean and the variance of the limit distribution in Stock (1999). These results agree with the limit distributions in Theorem 1.

2.2 Allowing for common factors

Let us now weaken the framework that has been considered in the previous section taking into account the presence of common factors in the panel data. Obviously, the main difficulty comes from the fact that the factors and the idiosyncratic components are unobserved so that, the first step of the analysis lies in getting a consistent estimate of both components. Following Bai and Ng (2001, 2004), in order to estimate these unobserved common factors we apply the principal components technique to the differenced-detrended model which, expressed in matrix notation, is given by:

$$
\begin{aligned}
M_i \, \Delta X_i &= M_i \, \Delta F \pi_i + M_i \, \Delta e_i \\
x_i &= f \pi_i + z_i,
\end{aligned} \tag{7}
$$

where $\Delta X_i = (\Delta X_{i,2}, \Delta X_{i,3}, \ldots, \Delta X_{i,T})'$ and $\Delta e_i = (\Delta e_{i,2}, \Delta e_{i,3}, \ldots, \Delta e_{i,T})'$ are two $((T-1) \times 1)$-vectors for the i-th individual, $\Delta F = [\Delta F_1 \ \Delta F_2 \ \ldots \ \Delta F_l]$ is a $((T-1) \times l)$-matrix being $\Delta F_j = (\Delta F_{j,2}, \Delta F_{j,3}, \ldots, \Delta F_{j,T})'$, $j = 1, \ldots, l$, a $((T-1) \times 1)$-vector, and $\pi_i = (\pi_{i,1}, \ldots, \pi_{i,l})'$ is the $(l \times 1)$-vector of loading parameters for the i-th individual, $i = 1, \ldots, N$. On the other hand, we define $M_i = I_{T-1} - a_i \left(a_i' a_i\right)^{-1} a_i'$, with $a_{i,t} = \left[D\left(T_{b,1}^i\right)_t, \ldots, D\left(T_{b,m_i}^i\right)_t\right]$ for Model 1, being $D\left(T_{b,k}^i\right)_t = 1$ for $t = T_{b,k}^i + 1$ and 0 elsewhere, $k = 1, \ldots, m_i$, and $a_{i,t} = \left[D\left(T_{b,1}^i\right)_t, \ldots, D\left(T_{b,m_i}^i\right)_t, DU_{i,1,t}, \ldots, DU_{i,m_i,t}\right]$ for Model 2. M_i is the usual idempotent matrix of projection into the space spanned by $a_{i,t}$. The estimated factors $\hat{f}_{1,t}, \ldots, \hat{f}_{l,t}$ are the l eigenvectors that corresponds to the l largest eigenvalues of the $(T-1 \times T-1)$ matrix xx', being $x = [x_1, \ldots, x_N]$. The matrix of estimated weights, $\hat{\Pi} = (\hat{\pi}_1, \ldots, \hat{\pi}_N)'$, is given by $\hat{\Pi} = x' \hat{f}_t$.

As a result, we can obtain an estimate of z_i from $\hat{z}_i = x_i - \hat{f} \hat{\pi}_i$, that, after computing its cumulated sum, produces a consistent estimation of the idiosyncratic disturbance term, $\tilde{e}_{i,t} = \sum_{j=1}^{t} \hat{z}_{i,j} = \sum_{j=1}^{t} \left(M_i \Delta \hat{e}_i\right)_j$. Now, the null hypothesis of unit root in the idiosyncratic stochastic element, *i.e.* $e_{i,t} \sim I(1)$, can be tested through the computation of the MSB test using $\tilde{e}_{i,t}$:

$$MSB_i\left(\lambda_i\right) = \frac{T^{-2} \sum_{t=1}^{T} \tilde{e}_{i,t-1}^2}{\tilde{\sigma}_i^2}, \tag{8}$$

where $\tilde{\sigma}_i^2$ is an estimation of the long-run variance of $\{\Delta \tilde{e}_{i,t}\}$. The following Theorem gives the asymptotic distribution of (8).

Theorem 2 *Let $\{X_{i,t}\}_{i=1,t=1}^{N,T}$ the stochastic process generated by (1) with $\pi_i \neq 0 \ \forall i$. If $\rho_i = 1$ in (3), and T, $T_{b,k}^i \to \infty$ in a way that $\lambda_{i,k} = T_{b,k}^i / T$ remains constant, $\forall i, k$, $i = 1, \ldots, N$, $k = 1, \ldots, m_i$, then the test in (8) converges to:*

(1) Model 1: $\ MSB_i\left(\lambda_i\right) \Rightarrow \int_0^1 W_i^2\left(r\right) dr$
(2) Model 2: $\ MSB_i\left(\lambda_i\right) \Rightarrow \sum_{k=1}^{m_i+1} \left(\lambda_{i,k} - \lambda_{i,k-1}\right)^2 \int_0^1 V_{i,k}^2\left(b\right) db,$

where $W_i\left(r\right)$ is the standard Brownian motion, $V_{i,k}\left(b\right) = W_{i,k}\left(b\right) - bW_{i,k}\left(1\right)$ is a Brownian bridge, and $\lambda_{i,0} = 0$ and $\lambda_{i,m_i+1} = 1$.

Theorem 2 shows that the limiting distribution of the MSB test for Model 1 does not depend on the presence of the structural breaks, since the effect of the impulse dummy is asymptotically negligible. This result is also found in Im *et al.* (2002) for the LM panel data based unit root test. However, this is not true for the model that allow for structural breaks affecting the time trend. Thus, the asymptotic distribution of the test for Model 2 depends on the set of nuisance parameters defined by the break fraction parameters. Moreover, the asymptotic distribution of the MSB test for $m_i = 1$ is symmetric around $\lambda_i = 0.5$ for Model 2, a feature that has also been highlighted in the previous section. The response surfaces for the p-values estimation are collected in Panel

B of Tables 1 and 2 for $m_i = 0$ and $m_i = 1$ respectively. The mean and variance of the limit distribution of MSB for Models 1 and 2 are presented in the following Proposition.

Proposition 2 *Let* $MSB_i(\lambda_i) = \tilde{\sigma}_i^{-2} T^{-2} \sum_{t=1}^{T} \tilde{e}_{i,t-1}^2$ *be the test statistic with limit distribution given in Theorem 2. Moreover, let* $\xi_i = E(MSB_i(\lambda_i))$ *and* $\varsigma_i^2 = V(MSB_i(\lambda_i))$ *be the mean and variance of* $MSB_i(\lambda_i)$ *respectively. Then, as* T, $T_{b,k}^i \to \infty$ *in a way that* $\lambda_{i,k} = T_{b,k}^i / T$ *remains constant,* $\forall i, k$, $i = 1, \ldots, N$, $k = 1, \ldots, m_i$,

> *(1) Model 1:* $\xi_i = \frac{1}{2}$ *and* $\varsigma_i^2 = \frac{1}{3}$;
> *(2) Model 2:* $\xi_i = \frac{1}{6} \sum_{k=1}^{m_i+1} (\lambda_{i,k} - \lambda_{i,k-1})^2$ *and*
> $\varsigma_i^2 = \frac{1}{45} \sum_{k=1}^{m_i+1} (\lambda_{i,k} - \lambda_{i,k-1})^4$,

where $\lambda_{i,0} = 0$ *and* $\lambda_{i,m_i+1} = 1$.

See Levin and Lin (1992) for the proof of statement 1 and the Appendix for the proof of the statement 2 of Proposition 2.

3 A simplified test statistic

In this Section we propose a simplified test that exploits the fact that the limiting distributions in Theorems 1 and 2 are weigthed sums of independent functionals of Brownian motions. We follow Busetti and Harvey (2001) and compute the MSB test as a weighted sum of partial sum processes so that we get rid of the break fraction parameters in the limit distributions. This simplification reduces the amount of computation effort that has to be made to provide practitioners with suitable sets of p-values for large m_i. However, this approach is primarily addressed for panels with large T provided that the approximation is for the limit distribution. First of all, let us focus on the situation where there are not common factors, that is, $\pi_i = 0 \ \forall i$, $i = 1, \ldots, N$. The weighted MSB test, $MSB_i^*(\lambda_i)$, is given by:

$$ MSB_i^*(\lambda_i) = \frac{\sum_{k=1}^{m_i+1} \left(\left(T_{b,k}^i - T_{b,k-1}^i \right)^{-2} \sum_{t=T_{b,k-1}^i}^{T_{b,k}^i} \tilde{X}_{i,t-1}^2 \right)}{\tilde{\sigma}_i^2}, \qquad (9) $$

$i = 1, \ldots, N$, with $T_{b,0}^i = 0$ and $T_{b,m_i+1}^i = T$. Now the computation of the test distinguishes among $m_i + 1$ subperiods which are rescaled by the square of the corresponding number of observations. The limit distribution of the $MSB_i^*(\lambda_i)$ test for the models without common factors is presented in the following Corollary.

Corollary 1 *Let* $X_{i,t}$, $i = 1, \ldots, N$, $t = 1, \ldots, T$, *be the stochastic process generated by (1) with* $\pi_i = 0 \ \forall i$ *and* $\rho_i = 1 \ \forall i$ *in (3). As* T, $T_{b,k}^i \to \infty$ *in a way*

that $\lambda_{i,k} = T_{b,k}^i / T$ remains constant, $\forall i, k$, $i = 1, \ldots, N$, $k = 1, \ldots, m_i$, then the test in (9) converges to:

$$(1) \text{ Model 1: } MSB_i^* (\lambda_i) \Rightarrow \sum_{k=1}^{m_i+1} \int_0^1 V_{i,k}^\mu (b)^2 db$$
$$(2) \text{ Model 2: } MSB_i^* (\lambda_i) \Rightarrow \sum_{k=1}^{m_i+1} \int_0^1 V_{i,k}^\tau (b)^2 db$$

where $\Rightarrow$ denotes weak convergence of the associated measure of probability, $V_{i,k}^\mu (b) = W_{i,k} (b) - \int_0^1 W_{i,k} (s) \, ds$, and $V_{i,k}^\tau (b) = W_{i,k} (b) - (4 - 6b) \int_0^1 W_{i,k} (s) \, ds - (-6 + 12b) \int_0^1 s W_{i,k} (s) \, ds$, with $W_{i,k} (b)$ the standard Brownian motion.

The proof follows from Theorem 1 and, hence, is omitted. Similar developments can be made for the specification in Model 2 with common factors. For this model, the $MSB_i^* (\lambda_i)$ test should be computed as:

$$MSB_i^* (\lambda_i) = \frac{\sum_{k=1}^{m_i+1} \left(\left(T_{b,k}^i - T_{b,k-1}^i \right)^{-2} \sum_{t=T_{b,k-1}^i}^{T_{b,k}^i} \tilde{e}_{i,t-1}^2 \right)}{\tilde{\sigma}_i^2} \tag{10}$$

with the limiting distribution given in the following Corollary.

Corollary 2 Let $\{X_{i,t}\}_{i=1,t=1}^{N,T}$ the stochastic process generated by (1) with $\pi_i \neq 0 \; \forall i$. If $\rho_i = 1 \; \forall i$ in (3), and T, $T_{b,k}^i \to \infty$ in a way that $\lambda_{i,k} = T_{b,k}^i / T$ remains constant, $\forall i, k$, $i = 1, \ldots, N$, $k = 1, \ldots, m_i$, then the test in (10) converges to:

$$MSB_i^* (\lambda_i) \Rightarrow \sum_{k=1}^{m_i+1} \int_0^1 V_{i,k}^2 (b) \, db,$$

where $V_{i,k} (b) = W_{i,k} (b) - b W_{i,k} (1)$ is a Brownian bridge.

The proof follows from Theorem 2 and, hence, is omitted. Note that the definition of the weighted MSB test makes free the limit distribution of the break fraction parameters, although it still depends on the number of structural breaks –in fact, they belong to the family of Cramér-von Mises distributions with $(m_i + 1)$-degrees of freedom. The asymptotic p-values of the limit distributions in Corollaries 1 and 2 can be computed from the response surfaces in Table 3 –Panel A for the limit distributions in Corollary 1 and Panel B for the one in Corollary 2. They are computed using the methodology described above using up to $m_i = 15$ structural breaks with $T = 2,000$ to approach the steps and 50,000 replications.

It can be shown that the response surfaces in Table 3 provides a good approximation of the critical values for the Cramér-von Misses distribution computed in Canova and Hansen (1995) and Nyblom and Harvey (2000). For instance, for the Cramér-von Misses distribution with two degrees of freedom defined by demeaned Brownian motions, these authors set the 95% quantile as 0.749 - see the second row of Table 1 in Canova and Hansen (1995). Using this quantile ($\hat{q}_i = 0.749$) with $m_i = 1$ in the response surface for the Model 2 –Panel B of Table 3 –we obtain $\hat{p}_i = 0.94965$.

Table 3: Response surfaces for the p-values estimation for the simplified test statistics

| | Panel A | | Panel B |
	Model 1	Model 2	Model 2
$\hat{\zeta}_{00}$			31.087
$\hat{\zeta}_{01}$	7.288	34.762	3.789
$\hat{\zeta}_{02}$	-1.826	-15.523	-27.683
$\hat{\zeta}_{03}$	2.174	73.602	154.984
$\hat{\zeta}_{04}$		-62.667	-156.524
$\hat{\zeta}_{10}$	9.934	-22.413	-9.316
$\hat{\zeta}_{11}$		2.489	0.301
$\hat{\zeta}_{12}$	-18.193	-6.255	
$\hat{\zeta}_{13}$	88.838		-16.629
$\hat{\zeta}_{14}$	-81.857	22.627	24.046
$\hat{\zeta}_{20}$	2.597	4.862	2.872
$\hat{\zeta}_{21}$		-0.135	-0.037
$\hat{\zeta}_{22}$	-6.817	-8.695	-6.145
$\hat{\zeta}_{23}$	28.813	40.734	26.669
$\hat{\zeta}_{24}$	-24.517	-36.618	-23.212
$\hat{\zeta}_{30}$	-0.041	-0.098	0.001
$\hat{\zeta}_{31}$	0.001	0.007	
$\hat{\zeta}_{32}$	-0.114	-0.184	-0.219
$\hat{\zeta}_{33}$	0.151	0.272	0.559
$\hat{\zeta}_{34}$			-0.349

Panel A corresponds to Corollary 1 and Panel B corresponds to Corollary 2. The functional form is given by $\log\left(\frac{p_i}{1-p_i}\right) = \sum_{j=0}^{3}\left(\zeta_{0j} + \zeta_{1j}q_i + \zeta_{2j}q_i^{-1/2} + \zeta_{3j}q_i^{-1/3} + \zeta_{4j}q_i^{-1/4}\right)m_i^j + u_i$. The $\bar{R}^2$ of all these estimations were 0.99. The included parameters were significant at the 5% level - we used the Newey-West robust estimator to compute the s.e.

The performance of the simplified test in finite samples might not show good properties. The statements in Corollaries 1 and 2 are valid as $T \to \infty$, which prevent the use of the P-value functions that have been estimated above in finite samples. The value of T for which the asymptotic results are of appliance is something to be addressed in the Monte Carlo analysis, but we should mention in advance that the simplified test shows an empirical size distortion even for $T = 300$. Thus, we would like to make available a test statistic that can be applied in finite samples, allowing for multiple structural breaks, and for which the computation of suitable p-values (or critical values) would not represent a high cost. The point here is the computation of these finite sample p-values.

Note that the limiting distributions in Corollaries 1 and 2 do not depend on the break fraction parameters, but just on the number of breaks. This is because as T, $T_{b,k}^i \to \infty$ in a way that $\lambda_{i,k} = T_{b,k}^i/T$ remains constant, $\forall i, k$, $i = 1, \ldots, N$, $k = 1, \ldots, m_i$, then the limiting distributions can be expressed as the sum of m_i independent functionals of Brownian motions. When applying this strategy to the finite sample framework we find that it is impossible to get rid of the number of observations that are involved in each regime. Thus, we should compute the finite moments using finite values for T. One possible solution consist on the use of an approximate finite sample distribution. Thus, we can define by $T_{approx}^i = T/(m_i + 1)$ the finite sample size for the i-th individual and approximate the finite sample distribution using T_{approx}^i. This simplification is specially appealing provided that this finite sample distribution will converge to the limiting distribution as in Corollaries 1 and 2 $T \to \infty$. Table ?? presents the estimates for the P-value functions that can be used to obtain the corresponding finite sample p-values for up to $m_i = 15$ structural breaks.

4   Pooling the individual tests

The results contained in Propositions 1 and 2 define the first way of pooling the individual information, which gives rise to the following test statistic:

$$ Z = \sqrt{N} \frac{\overline{MSB(\lambda)} - \bar{\xi}}{\bar{\varsigma}} \sim N(0,1), $$

where $\overline{MSB(\lambda)} = N^{-1} \sum_{i=1}^N MSB_i(\lambda_i)$, with $\bar{\xi} = N^{-1} \sum_{i=1}^N \xi_i$ and $\bar{\varsigma}^2 = N^{-1} \sum_{i=1}^N \varsigma_i^2$ computed using the statements in Propositions 1 and 2. The standard normal distribution is obtained from the application of the Lindberg-Lévy Central Limit Theorem (CLT). As mentioned in Bai and Ng (2001), this way of pooling can drive to unsatisfactory results, specifically when the asymptotic distribution of the individual tests is skewed, as this is the case. Instead, they suggest to follow the proposal in Maddala and Wu (1999) and Choi (2001) that pool the p-values associated to the individual tests - henceforth, we denote these p-values as p_i, $i = 1, \ldots, N$. Under the assumption of cross-section independence, $-2 \ln p_i \sim \chi_2^2$, a results that was used in Maddala and Wu (1999) to

define the Fisher-type test statistic:

$$P = -2 \sum_{i=1}^{N} \ln p_i \sim \chi_{2N}^2.$$

Notice that this statement does not require $N \to \infty$ to be satisfied, so this test statistic is of appliance for panels with small cross-section dimension. Besides, Choi (2001) proposes the following test when $N \to \infty$:

$$P_m = \frac{-2 \sum_{i=1}^{N} \ln p_i - 2N}{\sqrt{4N}} \sim N(0, 1),$$

where the standard normal limit distribution is obtained from the application of the Lindberg-Lévy CLT. As a result, the P_m test is suitable for those panels with large N. This specification was chosen in Bai and Ng (2004) to test the null hypothesis of non-stationary panel using the DF test. While the main advantage of the p-values pooling strategy comes from the fact that the definition of the test can be adapted to the cross-section dimension, its main drawback relies on the availability of the p-values. They are provided by the response surfaces estimated in Section 2.

5 Finite sample performance

We analyse the performance of the panel data unit root test in two different situations. First we consider the case in which there are no common factors, that is, we study the properties of the test assuming that the individuals are cross-section independent. After that, we will focus on those panels where the cross-section dependence is driven by the presence of up to three common factors. In all these simulations we assume that the date of the breaks are known. Three values for the number of individuals $N = \{20, 40, 100\}$ have been considered, with a sample size equal to $T = 100$. The number of replications is $r = 5,000$.

The DGP is given by equations (1) to (3) with $\mu_i \sim U[0,1]$, $\beta_i \sim U[0.2, 0.5]$, $\theta_{i,k} \sim U[-10, -3]$ and $\gamma_{i,k} \sim U[0.3, 0.9]$, where $U[\cdot]$ denotes the Uniform distribution. We have allowed one structural break randomly positioned according to $\lambda_i \sim U[0.15, 0.85]$. Under the null hypothesis $e_{i,t} \sim I(1)$ have been generated as a random walk without drift defined by the cumulated sum of iid $N(0,1)$ processes. The common factors are defined following the AR(1) model:

$$F_j = \alpha F_{j-1} + \sigma_F u_t,$$

where $\alpha = \{0.5, 0.8, 0.9, 0.95\}$ and $\sigma_F^2 = \{0.5, 1, 10\}$, $j = 1, \dots, l$, with the factor loadings given by $\pi_j \sim N(1, 1)$. The simulations have specified $l = 1$ and $l = 3$ common factors. The number of common factors are fixed using the panel BIC information criterion in Bai and Ng (2002) with $l_{\max} = 6$ as the maximum number of factors.

Table 4 reports the sample size of the three different statistics when there are no common factors. The test based on the standarisation present a size

Table 4: Empirical size. Known breaks and no common factors

Model 1

N	Z	P_m	P	Z	P_m	P
				Simplified test		
20	0.540	0.051	0.042	0.475	0.058	0.048
40	0.860	0.051	0.045	0.808	0.059	0.053
100	0.997	0.044	0.040	0.993	0.058	0.053

Model 2

N	Z	P_m	P	Z	P_m	P
				Simplified test		
20	0.132	0.066	0.058	0.150	0.061	0.050
40	0.271	0.073	0.062	0.265	0.059	0.051
100	0.568	0.075	0.066	0.548	0.053	0.051

distortion that increses with the number of individuals. This is in accordance with Bai and Ng (2004), where it is mentioned that pooling in this way can lead to unsatisfactory results specially when the asymptotic distribution of the individual tests is skewed, as this is the case. On the contrary, the tests based on the combination of the individual p-values show have an empirical size close to the nominal one. Note that this is also true for the simplified test, which indicates the usefulness of our proposal in applied research.

The picture changes when we analyse the panel data set that allows for common factors. For Model 1 all three test statistics show good performance in terms of empirical size. The exception is the P test, which in some situations presents empirical size distortions that lead to under reject the null hypothesis –see Table 5. For Model 2 the P_m test is the one with the most stable empirical size, provided that the Z and P tests under reject the null hypothesis. This is also true for all the version of the simplified tests –see Table 6.

6 Conclusions

In this paper, we have proposed new procedures for testing non-stationarity of panel data in the presence of multiple structural breaks and dynamic common factors. In the absence of common factors, the limiting distributions are shown to be weighted sum of independent and identically distributed Brownian motions (demeaned or detrended). These results are of special interest for the single time series analysis –i.e. panels with $N = 1$ individual– provided that they extend the proposals in Perron (1997) and Lumsdaine and Papell (1997), among others, and allow to test the unit root hypothesis with multiple structural changes. When dynamic factors are present, the PANIC approach of Bai and Ng (2004) is used to estimate the model. The limiting distributions of the test statistics are invariant to mean breaks. For breaks in the linear trend, the limiting distributinos are

Table 5: Empirical size for model 1 and $N = 40$. Known breaks with common factors

σ_F^2	α	Z	P_m	P	Z	P_m	P
			$r=1$			$r=3$	
0.5	0	0.04	0.04	0.04	0.04	0.03	0.03
0.5	0.5	0.05	0.04	0.03	0.04	0.03	0.03
0.5	0.8	0.05	0.05	0.04	0.04	0.04	0.04
0.5	0.9	0.05	0.05	0.04	0.05	0.04	0.04
0.5	0.95	0.04	0.05	0.04	0.05	0.05	0.04
1	0	0.04	0.05	0.04	0.04	0.04	0.03
1	0.5	0.04	0.04	0.04	0.03	0.04	0.03
1	0.8	0.05	0.04	0.04	0.04	0.05	0.04
1	0.9	0.05	0.05	0.04	0.05	0.04	0.03
1	0.95	0.04	0.04	0.04	0.04	0.04	0.03
10	0	0.06	0.05	0.04	0.03	0.03	0.03
10	0.5	0.05	0.04	0.04	0.04	0.04	0.03
10	0.8	0.06	0.05	0.04	0.04	0.04	0.04
10	0.9	0.05	0.05	0.04	0.04	0.04	0.04
10	0.95	0.05	0.05	0.04	0.05	0.05	0.04

shown to be weighted sum of iid Brownian bridges. We further introduced a simplified test statistic, and showed that the limiting distribution is invariant to both mean and trend breaks. Pooled test statistic is also studied. Response surfaces for p-values of all test statistics are computed.

Table 6: Empirical size for model 2 and $N = 40$. Known breaks with common factors

| | | | | | Simplified tests | | |
| | | | | | $r = 1$ | | |
σ_F^2	α	Z	P_m	P	Z	P_m	P
0.5	0	0.02	0.05	0.04	0.01	0.03	0.03
0.5	0.5	0.02	0.06	0.06	0.02	0.04	0.03
0.5	0.8	0.03	0.07	0.06	0.02	0.04	0.03
0.5	0.9	0.03	0.06	0.05	0.02	0.03	0.03
0.5	0.95	0.02	0.06	0.06	0.01	0.04	0.03
1	0	0.03	0.08	0.07	0.02	0.05	0.05
1	0.5	0.02	0.06	0.05	0.02	0.03	0.03
1	0.8	0.04	0.05	0.04	0.02	0.03	0.03
1	0.9	0.02	0.06	0.05	0.02	0.03	0.03
1	0.95	0.02	0.06	0.05	0.02	0.03	0.03
10	0	0.03	0.06	0.06	0.02	0.04	0.03
10	0.5	0.03	0.06	0.06	0.02	0.04	0.03
10	0.8	0.02	0.06	0.05	0.02	0.03	0.03
10	0.9	0.03	0.06	0.05	0.02	0.03	0.03
10	0.95	0.03	0.07	0.06	0.01	0.04	0.03

| | | | | | Simplified tests | | |
| | | | | | $r = 3$ | | |
σ_F^2	α	Z	P_m	P	Z	P_m	P
0.5	0	0.02	0.05	0.04	0.02	0.02	0.02
0.5	0.5	0.02	0.04	0.04	0.01	0.02	0.02
0.5	0.8	0.03	0.05	0.05	0.02	0.03	0.02
0.5	0.9	0.03	0.05	0.05	0.01	0.03	0.03
0.5	0.95	0.03	0.06	0.05	0.02	0.04	0.03
1	0	0.01	0.04	0.03	0.01	0.02	0.02
1	0.5	0.02	0.04	0.04	0.01	0.02	0.02
1	0.8	0.02	0.05	0.04	0.02	0.03	0.03
1	0.9	0.02	0.06	0.06	0.01	0.03	0.03
1	0.95	0.02	0.05	0.05	0.01	0.04	0.03
10	0	0.02	0.05	0.04	0.01	0.02	0.02
10	0.5	0.02	0.05	0.04	0.01	0.03	0.02
10	0.8	0.03	0.06	0.05	0.02	0.02	0.02
10	0.9	0.02	0.05	0.05	0.01	0.03	0.03
10	0.95	0.03	0.08	0.07	0.02	0.05	0.05

7 Appendix: Proof of Theorem 1

7.1 Proof of statement (1)

The model that is considered in this statement is the one for non-trended variables where $\beta_i = \gamma_{i,k} = 0$, $\forall i, k$ in (4). In addition, the constraint in $\pi_i = 0$ $\forall i$ is imposed in order to avoid the presence of common factors that drive the behaviour of the individual time series.

From this specification, the estimated OLS residuals of the model are obtained from $\tilde{e}_i = M_i e_i$, with $M_i = I_T - a_i \left(a_i' a_i\right)^{-1} a_i'$. Note that for m_i structural changes the deterministic part of the model given in (4) can be expressed in terms of orthogonal regressors defining a block diagonal matrix. The elements in the diagonal are given by vectors $\iota_k = (1, \ldots, 1)'$ of dimension $\left(\left(T_{b,k}^i - T_{b,k-1}^i\right) \times 1\right)$, $k = 1, \ldots, m_i+1$, with $T_{b,0}^i = 0$ and $T_{b,m_i+1}^i = T$. Thus, the cross-product matrix of regressors $a_i' a_i$ is given by

$$
a_i' a_i = \begin{bmatrix}
T_{b,1}^i & & & & 0 \\
& T_{b,2}^i - T_{b,1}^i & & & \\
& & \ddots & & \\
& & & T_{b,m_i}^i - T_{b,m_i-1}^i & \\
0 & & & & T - T_{b,m_i}^i
\end{bmatrix},
$$

Using the fact that $T_{b,k}^i = \lambda_{i,k} T$ and defining the $(m_i \times m_i)$-diagonal rescaling matrix $P_i = diag\left(T^{-1/2}, \ldots, T^{-1/2}\right)$, $P_i' a_i' a_i P_i$ can be expressed as $P_i' a_i' a_i P_i = diag\left(\lambda_{i,1}, \left(\lambda_{i,2} - \lambda_{i,1}\right), \ldots, \left(\lambda_{i,m_i} - \lambda_{i,m_i-1}\right), \left(1 - \lambda_{i,m_i}\right)\right)$.

On the other hand, under the null hypothesis that $e_i \sim I(1)$, $T^{-1} P_i' a_i' e_i \Rightarrow \left(\sigma_i \int_0^{\lambda_{i,1}} W_i(s)\, ds, \sigma_i \int_{\lambda_{i,1}}^{\lambda_{i,2}} W_i(s)\, ds, \ldots, \sigma_i \int_{\lambda_{i,m_i-1}}^{\lambda_{i,m_i}} W_i(s)\, ds, \sigma_i \int_{\lambda_{i,m_i}}^{1} W_i(s)\, ds\right)'$. This means that for $t \leq T_{b,1}^i$

$$
T^{-1/2}\tilde{e}_{i,t} \Rightarrow \sigma_i W_i(r) - \sigma_i \frac{1}{\lambda_{i,1}} \int_0^{\lambda_{i,1}} W_i(s)\, ds;\ 0 < r < \lambda_{i,1},
$$

for $T_{b,1}^i < t \leq T_{b,2}^i$

$$
T^{-1/2}\tilde{e}_{i,t} \Rightarrow \sigma_i W_i(r) - \sigma_i \frac{1}{(\lambda_{i,2} - \lambda_{i,1})} \int_{\lambda_{i,1}}^{\lambda_{i,2}} W_i(s)\, ds;\ \lambda_{i,1} < r < \lambda_{i,2},
$$

and so on, so that, it can be established for $T_{b,k-1}^i < t \leq T_{b,k}^i$

$$
T^{-1/2}\tilde{e}_{i,t} \Rightarrow \sigma_i W_i(r) - \sigma_i \frac{1}{(\lambda_{i,k} - \lambda_{i,k-1})} \int_{\lambda_{i,k-1}}^{\lambda_{i,k}} W_i(s)\, ds;\ \lambda_{i,k-1} < r < \lambda_{i,k},
$$

$k = 1, \ldots, m_i + 1$, with $\lambda_{i,0} = 0$ and $\lambda_{i,m_i+1} = 1$.

The goal is to show that limit distribution of the test can be expressed as a sum of a set of independent integrals of detrended Brownian motions. To do so we rescale the Brownian motions so that we ensure that the index of the Brownian motion in each subsample belongs to $[0,1]$. Thus, notice that in the first subsample $0 < r < \lambda_{i,1}$ can be rescaled as $0/\lambda_{i,1} < r/\lambda_{i,1} < \lambda_{i,1}/\lambda_{i,1}$, so that $b_1 = (r/\lambda_{i,1}) \in [0,1]$. In general, for $\lambda_{i,k-1} < r < \lambda_{i,k}$ it can be defined $(\lambda_{i,k-1} - \lambda_{i,k-1})/(\lambda_{i,k} - \lambda_{i,k-1}) < (r - \lambda_{i,k-1})/(\lambda_{i,k} - \lambda_{i,k-1}) < (\lambda_{i,k} - \lambda_{i,k-1})/(\lambda_{i,k} - \lambda_{i,k-1})$, so that $b = (r - \lambda_{i,k-1})/(\lambda_{i,k} - \lambda_{i,k-1}) \in [0,1]$ –Lee (1996), Lee and Strazicich (2001) and Bartley, Lee and Strazicich (2002) use similar developments when deriving the limit distribution of the KPSS test with one structural break. Therefore, for $0 < b < 1$ we have

$$T^{-1/2}\tilde{e}_{i,t} \Rightarrow \sigma_i\sqrt{(\lambda_{i,k} - \lambda_{i,k-1})}\left(W_{i,k}(b) - \int_0^1 W_{i,k}(s)\,ds\right)$$

$$= \sigma_i\sqrt{(\lambda_{i,k} - \lambda_{i,k-1})}V_{i,k}^{\mu}(b),$$

where $V_{i,k}^{\mu}(b)$ denotes the demeaned Brownian motion. Thus, the limit distribution of the $MSB_i(\lambda_i)$ test is given by

$$MSB_i(\lambda_i) \Rightarrow \lambda_{i,1}^2 \int_0^1 V_{i,1}^{\mu}(b)^2\,db + \cdots + (\lambda_{i,k} - \lambda_{i,k-1})^2 \int_0^1 V_{i,k}^{\mu}(b)^2\,db$$

$$+ \cdots + (1 - \lambda_{i,m_i})^2 \int_0^1 V_{i,m_i+1}^{\mu}(b)^2\,db,$$

with $V_{i,k}^{\mu}(b) = W_{i,k}(b) - \int_0^1 W_{i,k}(s)\,ds$, $k = 1,\dots,m_i + 1$, independent demeaned Brownian motions and provided that $\tilde{\sigma}_i^2 \to \sigma_i^2$ - see Stock (1999). Notice that for $m_i = 1$ the limit distribution of the test is given by

$$MSB_i(\lambda_i) \Rightarrow \lambda_i^2 \int_0^1 V_{i,1}^{\mu}(b)^2\,db + (1 - \lambda_i)^2 \int_0^1 V_{i,2}^{\mu}(b)^2\,db \qquad (11)$$

which it is shown to be symmetric around $\lambda_i = 0.5$, provided that we can interchange λ_i and $(1 - \lambda_i)$ in (11) and obtain the same asymptotic distribution.

Finally, note that the limit distribution of $MSB_i(\lambda_i)$ is the weighted sum of $(m_i + 1)$ independent Cramér-von Mises distributions -see Harvey (2001). The expectations of these Cramér-von Mises distributions are $E\left[\int_0^1 V_{i,k}^2(b)\,db\right] = 1/6$ where the variance are $V\left[\int_0^1 V_{i,k}^2(b)\,db\right] = 1/45$, $\forall k = 1,\dots,m_i + 1$ –see Levin and Lin (1992). Therefore, $E[MSB_i(\lambda_i)] = (1/6)\sum_{k=1}^{m_i+1}(\lambda_{i,k} - \lambda_{i,k-1})^2$ and $V[MSB_i(\lambda_i)] = (1/45)\sum_{k=1}^{m_i+1}(\lambda_{i,k} - \lambda_{i,k-1})^4$.

7.2 Proof of statement (2)

This statement presents the limit distribution of the test for trended variables. As before, we assume $\pi_i = 0\ \forall i$. Following the steps on the previous proof, notice

that for m_i structural changes the deterministic part of the model given in (4) can be expressed in terms of orthogonal regressors defining a block diagonal matrix. Now, the elements in the diagonal are given by vectors $\iota_k = (1, \dots, 1)'$ and $t_k = \left(1, 2, \dots, \left(T^i_{b,k} - T^i_{b,k-1}\right)\right)'$ both of dimension $\left(\left(T^i_{b,k} - T^i_{b,k-1}\right) \times 1\right)$, $k = 1, \dots, m_i+1$, with $T^i_{b,0} = 0$ and $T^i_{b,m_i+1} = T$. Thus, the cross-product matrix of regressors $a'_i a_i$ is block diagonal matrix, with the k-th block given by the (2×2)-matrix $a'_{i,k} a_{i,k}$ with elements $a'_{i,k} a_{i,k}[1,1] = \left(T^i_{b,k} - T^i_{b,k-1}\right)$, $a'_{i,k} a_{i,k}[1,2] = a'_{i,k} a_{i,k}[2,1] = 1/2 \left(T^i_{b,k} - T^i_{b,k-1}\right)\left(\left(T^i_{b,k} - T^i_{b,k-1}\right) + 1\right)$ and $a'_{i,k} a_{i,k}[2,2] = 1/6 \left(T^i_{b,k} - T^i_{b,k-1}\right)\left(\left(T^i_{b,k} - T^i_{b,k-1}\right) + 1\right)\left(2\left(T^i_{b,k} - T^i_{b,k-1}\right) + 1\right)$. If we define the rescaling diagonal matrix $P_k = diag\left(T^{-1/2}, T^{-3/2}\right)$, then

$$\left(P'_k a'_{i,k} a_{i,k} P_k\right)^{-1} \rightarrow \frac{1}{\Delta} \left[\begin{array}{cc} 1/3 \left(\lambda_{i,k} - \lambda_{i,k-1}\right)^3 & -1/2 \left(\lambda_{i,k} - \lambda_{i,k-1}\right)^2 \\ -1/2 \left(\lambda_{i,k} - \lambda_{i,k-1}\right)^2 & \left(\lambda_{i,k} - \lambda_{i,k-1}\right) \end{array} \right],$$

where $\Delta = (1/12) \left(\lambda_{i,k} - \lambda_{i,k-1}\right)^4$. Hence, it can be shown that $\left(P' a'_i a_i P\right)^{-1} = diag\left(\left(P' a'_{i,1} a_{i,1} P\right)^{-1}, \dots, \left(P' a'_{i,m_i} a_{i,m_i} P\right)^{-1}\right)$.

On the other hand, $T^{-1} P' a'_i e_i$ is a $(2m_i \times 1)$-vector defined by stacking the m_i (2×1)-vectors given by

$$\left[\begin{array}{c} T^{-3/2} \sum_{t=T^i_{b,k-1}+1}^{T^i_{b,k}} e_{i,t} \\ T^{-5/2} \sum_{t=T^i_{b,k-1}+1}^{T^i_{b,k}} t e_{i,t} \end{array} \right] \Rightarrow \left[\begin{array}{c} \sigma_i \int_{\lambda_{i,k-1}}^{\lambda_{i,k}} W(s)\, ds \\ \sigma_i \int_{\lambda_{i,k-1}}^{\lambda_{i,k}} s W(s)\, ds \end{array} \right],$$

$k = 1, \dots, m_i + 1$. Therefore, for $T^i_{b,k-1} < t \leq T^i_{b,k}$ it can be established that

$$T^{-1/2} \tilde{e}_{i,t} \Rightarrow \sigma_i W_i(r) - \frac{\sigma_i}{\Delta} \left[\left(\left(1/3 \left(\lambda_{i,k} - \lambda_{i,k-1}\right)^3\right) \right.\right.$$
$$\left. -1/2 \left(\lambda_{i,k} - \lambda_{i,k-1}\right)^2 r\right) \int_{\lambda_{i,k-1}}^{\lambda_{i,k}} W(s)\, ds$$
$$\left. - \left(1/2 \left(\lambda_{i,k} - \lambda_{i,k-1}\right)^2 - \left(\lambda_{i,k} - \lambda_{i,k-1}\right) r\right) \int_{\lambda_{i,k-1}}^{\lambda_{i,k}} s W(s)\, ds \right],$$

$\lambda_{i,k-1} < r < \lambda_{i,k}$, $k = 1, \dots, m_i + 1$, with $\lambda_{i,0} = 0$ and $\lambda_{i,m_i+1} = 1$. Rescaling the index of the Brownian motions in a way that $b_k = (r - \lambda_{i,k-1}) / (\lambda_{i,k} - \lambda_{i,k-1})$ so that $0 < b_k < 1$, is straightforward to see that

$$T^{-1/2} \tilde{e}_{i,t} \Rightarrow \sigma_i \sqrt{(\lambda_{i,k} - \lambda_{i,k-1})} \left[W_i(b_k) - (4 - 6b_k) \int_0^1 W_i(s_k)\, ds_k \right.$$
$$\left. - (-6 + 12b_k) \int_0^1 s_k W_i(s_k)\, ds_k \right]$$
$$= \sigma_i \sqrt{(\lambda_{i,k} - \lambda_{i,k-1})} V^\tau_{i,k}(b),$$

21

where $V_{i,k}^\tau (b)$ denotes the detrended Brownian motion. Notice that $V_{i,k}^\tau (b)$ is equivalent to the detrended Brownian motion involved in the limit distribution of the MSB test in Stock (1999) Theorem 1.

Thus, the limit distribution of the $MSB_i (\lambda_i)$ test

$$MSB_i (\lambda_i) \Rightarrow \lambda_{i,1}^2 \int_0^1 V_{i,1}^\tau (b)^2 \, db + \cdots + (\lambda_{i,k} - \lambda_{i,k-1})^2 \int_0^1 V_{i,k}^\tau (b)^2 \, db$$
$$+ \cdots + (1 - \lambda_{i,m_i})^2 \int_0^1 V_{i,m_i+1}^\tau (b)^2 \, db,$$

with $V_{i,k}^\tau (b) = W_{i,k} (b) - (4 - 6b) \int_0^1 W_{i,k} (s) \, ds - (-6 + 12b) \int_0^1 s W_{i,k} (s) \, ds$, $k = 1, \ldots, m_i + 1$, independent detrended Brownian motions and provided that $\tilde{\sigma}_i^2 \to \sigma_i^2$ –see Stock (1999). Notice that, as before, for $m_i = 1$ the limit distribution of the test is symmetric around $\lambda_i = 0.5$.

The mean and the variance of the limit distribution is given by can be computed from the moment generating function in Tanaka (1996), which for our test is given by

$$m (\varphi) = \left[- \frac{3 \left(\sqrt{2} \varphi \sin \sqrt{2\varphi} + 2\sqrt{\varphi} \cos \sqrt{2\varphi} - 2\sqrt{\varphi} \right)}{\left(\sqrt{\varphi} \right)^5} \right]^{-1/2} .$$

The first derivative of the moment generating function evaluated at $\varphi = 0$ will provide us the first moment of the limit distribution:

$$\xi_i = \lim_{\varphi \to 0} \frac{dm (\varphi)}{d\varphi} = \frac{1}{15},$$

whereas the variance is obtained from

$$\varsigma_i^2 = \lim_{\varphi \to 0} \frac{d^2 m (\varphi)}{d\varphi^2} - \left(\frac{1}{15} \right)^2 = \frac{11}{6300}.$$

Therefore, $E [MSB_i (\lambda_i)] = (1/15) \sum_{k=1}^{m_i+1} (\lambda_{i,k} - \lambda_{i,k-1})^2$ and $V [MSB_i (\lambda_i)] = (11/6300) \sum_{k=1}^{m_i+1} (\lambda_{i,k} - \lambda_{i,k-1})^4$.

8 Appendix: Proof of Theorem 2

8.1 Proof of statement (1)

Statement (1) in Theorem 1 is concerned with Model 1, that is, the model for non-trended variables where $\beta_i = \gamma_{i,k} = 0$, $\forall i, k$ in (4). The estimation of the (differenced and detrended) model produces the following result:

$$x_{i,t} = \hat{f}_t \hat{\pi}_i + \hat{z}_{i,t}. \tag{12}$$

Subtracting (12) from (7) we obtain

$$\hat{z}_{i,t} = z_{i,t} + f_t \pi_i - \hat{f}_t \hat{\pi}_i.$$

Following Bai and Ng (2003), we can express the model as

$$
\begin{aligned}
\hat{z}_{i,t} &= z_{i,t} + f_t H H^{-1} \pi_i - \hat{f}_t H^{-1} \pi_i + \hat{f}_t H^{-1} \pi_i - \hat{f}_t \hat{\pi}_i \\
&= z_{i,t} + \left(f_t H - \hat{f}_t \right) H^{-1} \pi_i - \hat{f}_t \left(\hat{\pi}_i - H^{-1} \pi_i \right) \\
&= z_{i,t} + v_t H^{-1} \pi_i - \hat{f}_t d_i,
\end{aligned}
\tag{13}
$$

where $v_t = \left(f_t H - \hat{f}_t \right)$ and $d_i = \left(\hat{\pi}_i - H^{-1} \pi_i \right)$. Let us define the partial sum process using the estimated residuals as $\tilde{e}_{i,t} = \sum_{s=2}^{t} \hat{z}_{i,s} = \sum_{s=2}^{t} \left(M_i \ \Delta \hat{e}_i \right)_s$. By Lemmas 3 and C1 in Bai and Ng (2003), $T^{-1/2} \left\| \sum_{s=2}^{t} v_s H^{-1} \pi_i \right\| = o_p(1)$ and $T^{-1/2} \left\| \sum_{s=2}^{t} \hat{f}_s d_i \right\| = o_p(1)$, so that

$$T^{-1/2} \tilde{e}_{i,t} = T^{-1/2} \sum_{s=2}^{t} \left(M_i \ \Delta e_i \right)_s + o_p(1).$$

The partial sum process can be expressed in terms of the population residuals as:

$$T^{-1/2} \tilde{e}_{i,t} = T^{-1/2} \sum_{s=2}^{t} \left(M_i \ \Delta e_i \right)_s + o_p(1), \tag{14}$$

with

$$\sum_{s=2}^{t} \left(M_i \ \Delta e_i \right)_s = \sum_{s=2}^{t} \left(\Delta e_{i,s} - \left(P_i \Delta e_i \right)_s \right),$$

where $P_i = a_i \left(a_i' a_i \right)^{-1} a_i'$. Note that $\sum_{s=1}^{t} \Delta e_{i,s} = e_{i,t} - e_{i,1}$ and $P_i \Delta e_i$ is a vector of zeros except for the $\left(T_{b,k}^i + 1 \right)$-th positions, $k = 1, \ldots, m_i$. The cumulated process is equal to:

$$\tilde{e}_{i,t} = \left(e_{i,t} - e_{i,1} \right) - \left(e_{i,T_{b,1}^i+1} - e_{i,T_{b,1}^i} \right) DU_{i,1,t} - \ldots - \left(e_{i,T_{b,m_i}^i+1} - e_{i,T_{b,m_i}^i} \right) DU_{i,m_i,t}.$$

If we assume that $e_{i,t} \sim I(1)$:

$$
\begin{aligned}
T^{-1/2} \tilde{e}_{i,t} &= T^{-1/2} \left(e_{i,t} - e_{i,1} \right) - T^{-1/2} \left(e_{i,T_{b,1}^i+1} - e_{i,T_{b,1}^i} \right) DU_{i,1,t} - \ldots \\
&\quad - T^{-1/2} \left(e_{i,T_{b,m_i}^i+1} - e_{i,T_{b,m_i}^i} \right) DU_{i,m_i,t} + o_p(1) \\
&\Rightarrow \sigma_i W_i(r) - \sigma_i W_i(0) - \sigma_i dW_i(\lambda_1) du_1 - \ldots - \sigma_i dW_i(\lambda_{m_i}) du_{m_i} \\
&= \sigma_i W_i(r) - \sigma_i dW_i(\lambda_1) du_1 - \ldots - \sigma_i dW_i(\lambda_{m_i}) du_{m_i}, \tag{15}
\end{aligned}
$$

where $W_i(r)$ denotes the standard Brownian motion and $du_k = 1$ for $r > \lambda_k$ and 0 elsewhere, with $\lambda_k = T_{b,k}^i/T$, $k = 1, \dots, m_i$. The limit expression of $T^{-1/2}\tilde{e}_{i,t}$ given by (15) involves two different kind of elements: (i) the Brownian motion, $W_i(r)$, and (ii) the difference of Brownian motions, $dW_i(\lambda_k)$, $k = 1, \dots, m_i$. Following Perron (1997), the effect of these differences can be understood as negligible compared to $W_i(r)$, so that, we can consider that $T^{-1/2}\tilde{e}_{i,t} \Rightarrow \sigma_i W_i(r)$. Therefore, the test statistic converges to:

$$MSB_i(\lambda_i) \Rightarrow \int_0^1 W_i^2(r)\,dr,$$

provided that $\tilde{\sigma}_i^2 \to \sigma_i^2$. Notice that after considering the negligible effect of the $dW_i(\lambda_k)$ terms, $k = 1, \dots, m_i$, the asymptotic distribution of the test does not depend on the break fraction parameters λ_k, that is, the test is invariant to the presence of structural breaks affecting the mean of the time series.

8.2   Proof of statement (2)

Let us now focus on the specification given by Model 2, that is, the model for trended regressors where $\beta_i \neq \gamma_{i,k} \neq 0$, $\forall i, k$ in (4). As in the previous proof, the computation of the partial sum process can be done from (13). However, we have to assess that $T^{-1/2}\left\|\sum_{s=2}^t \hat{f}_s d_i\right\| = o_p(1)$. Note that $T^{-1/2}\left\|\sum_{s=2}^t \hat{f}_s d_i\right\| \leq T^{-1/2}\left\|\sum_{s=2}^t \hat{f}_s\right\|\|d_i\|$. From Bai and Ng (2003), $\|d_i\| = o_p(1)$, and

$$
\begin{aligned}
T^{-1/2}\sum_{s=2}^t \hat{f}_s &= T^{-1/2}\sum_{s=2}^t \left(\hat{f}_s - Hf_s + Hf_s\right) \\
&= T^{-1/2}\sum_{s=2}^t v_s + HT^{-1/2}\sum_{s=2}^t f_s \\
&= o_p(1) + HT^{-1/2}\sum_{s=2}^t f_s.
\end{aligned}
$$

To determine the order in probability of $T^{-1/2}\sum_{s=2}^t f_s$ we rewrite the matrix of deterministic elements a_i in a $(T \times (2m_i + 1))$ *quasi* block diagonal matrix:

$$a_i = \left[DU_i^1 \; D\left(T_{b,1}^i\right) \; DU_i^2 \; D\left(T_{b,2}^i\right) \; DU_i^3 \; \dots \; DU_i^{m_i} D\left(T_{b,m_i}^i\right) \; DU_i^{m_i+1}\right],$$

where $DU_i^k = 1$ for $T_{b,k-1}^i < t \leq T_{b,k}^i$ and 0 elsewhere, $k = 1, \dots, m_i + 1$, with $T_{b,0}^i = 1$ and $T_{b,m_i+1}^i = T$. Now the elements that defines the a_i matrix can be grouped in two different sets, the first one compounded by the DU_i^k regressors and the second one composed by the $D\left(T_{b,k}^i\right)$ regressors, $k = 1, \dots, m_i + 1$. Note the orthogonality property that characterize the elements of each set.

Moreover, this transformation makes the P_i matrix to be block diagonal. Without loose of generality and in order to simplify cumbersome algebraic manipulations, we derive the order in probability of $T^{-1/2} \sum_{s=2}^{t} f_s$ assuming that $m_i = 1$, although the results are valid when $m_i \geq 1$. Thus, when $m_i = 1$ the P_i matrix is given by:

$$
\begin{aligned}
P_i &= \frac{1}{T\lambda} DU^1 DU^{1\prime} + \frac{(-1+\lambda)T}{1-T+T\lambda} D\left(T_b^i\right) D\left(T_b^i\right)' + \frac{1}{1-T+T\lambda} D\left(T_b^i\right) DU^{2\prime} \\
&\quad + \frac{1}{1-T+T\lambda} DU^2 D\left(T_b^i\right)' - \frac{1}{1-T+T\lambda} DU^2 DU^{2\prime}.
\end{aligned}
\tag{16}
$$

The effect of P_i on ΔF can be analysed by parts. When multiplying the first element of (16) by ΔF produces:

$$
\frac{1}{T\lambda} DU^1 DU^{1\prime} \Delta F = \left\{ \begin{array}{cc} \frac{1}{T\lambda}\left(F_{T_b^i} - F_1\right) & t \leq T_b^i \\ 0 & t > T_b^i \end{array} \right. .
$$

The computation of the partial sum process involves:

$$
T^{-1/2} \sum_{s=2}^{t} \frac{1}{T\lambda} \left(DU^1 DU^{1\prime}\right)_s \Delta F_s = \left\{ \begin{array}{cc} T^{-3/2}\lambda^{-1}\left(F_{T_b^i} - F_1\right)t & t \leq T_b^i \\ T^{-1/2}\left(F_{T_b^i} - F_1\right) & t > T_b^i \end{array} \right. ,
$$

which is $O_p(1)$. The same result is found for the product involving the fifth element of (16). The second element of (16) gives:

$$
\frac{(-1+\lambda)T}{1-T+T\lambda} D\left(T_b^i\right) D\left(T_b^i\right)' \Delta F = \left\{ \begin{array}{cc} \frac{(-1+\lambda)T}{1-T+T\lambda} \Delta F_{T_b^i+1} & t = T_b^i + 1 \\ 0 & t \neq T_b^i + 1 \end{array} \right. ,
$$

so that the partial sum process is

$$
T^{-1/2} \sum_{s=2}^{t} \frac{(-1+\lambda)T}{1-T+T\lambda} \left(D\left(T_b^i\right) D\left(T_b^i\right)'\right)_s \Delta F_s = \left\{ \begin{array}{cc} 0 & t \leq T_b^i \\ \frac{(-1+\lambda)T^{1/2}}{1-T+T\lambda} \Delta F_{T_b^i+1} & t > T_b^i \end{array} \right. ,
$$

with $O_p(1)$ as order in probability. For the third element we have

$$
\frac{1}{1-T+T\lambda} D\left(T_b^i\right) DU^{2\prime} \Delta F = \left\{ \begin{array}{cc} F_T - F_{T_b^i+1} & t = T_b^i + 1 \\ 0 & t \neq T_b^i + 1 \end{array} \right. ,
$$

so that

$$
T^{-1/2} \sum_{s=2}^{t} \frac{1}{1-T+T\lambda} \left(D\left(T_b^i\right) DU^{2\prime}\right)_s \Delta F_s = \left\{ \begin{array}{cc} 0 & t \leq T_b^i \\ \frac{\left(F_T - F_{T_b^i+1}\right)}{(1-T+T\lambda)T^{1/2}} & t > T_b^i \end{array} \right. ,
$$

25

which is also $O_p(1)$. The fourth element is

$$\frac{1}{1-T+T\lambda}DU^2D\left(T_b^i\right)'\Delta F = \begin{cases} 0 & t \le T_b^i \\ \frac{1}{1-T+T\lambda}\Delta F_{T_b^i+1} & t > T_b^i \end{cases}.$$

Thus,

$$T^{-1/2}\sum_{s=2}^{t}\frac{1}{1-T+T\lambda}\left(DU^2D\left(T_b^i\right)'\right)_s\Delta F_s = \begin{cases} 0 & t \le T_b^i \\ \frac{\Delta F_{T_b^i+1}\left(t-T_b^i\right)}{(1-T+T\lambda)T^{1/2}} & t > T_b^i \end{cases},$$

which is $O_p(1)$. Finally, the fifth element

$$-\frac{1}{1-T+T\lambda}DU^2DU^{2\prime}\Delta F = \begin{cases} 0 & t \le T_b^i \\ -\frac{1}{1-T+T\lambda}\left(F_T - F_{T_b^i+1}\right) & t > T_b^i \end{cases},$$

with cumulated sum

$$T^{-1/2}\sum_{s=2}^{t}-\frac{1}{1-T+T\lambda}\left(DU^2DU^{2\prime}\right)_s\Delta F_s = \begin{cases} 0 & t \le T_b^i \\ -\frac{\left(F_T-F_{T_b^i+1}\right)\left(t-T_b^i\right)}{(1-T+T\lambda)T^{1/2}} & t > T_b^i \end{cases},$$

which is also $O_p(1)$. Therefore, all the partial sum processes involving $P_i\Delta F$ are $O_p(1)$, a result that can be straightforwardly extended to those situations that allow for multiple breaks. Consequently,

$$T^{-1/2}\left\|\sum_{s=2}^{t}\hat{f}_s d_i\right\| \le T^{-1/2}\left\|\sum_{s=2}^{t}\hat{f}_s\right\|\|d_i\|$$
$$\le O_p(1)o_p(1),$$

which means that $T^{-1/2}\left\|\sum_{s=2}^{t}\hat{f}_s d_i\right\| = o_p(1)$.

As in the previous proof, the partial sum process of the estimated residuals is given by (14). Now, the cumulative process are given by the previous expressions but replacing F by e_i. The first element of the partial sum process, which involves the first set of step dummy variables, converges to $T^{-1/2}\sum_{s=2}^{t}\frac{1}{T\lambda}\left(DU^1DU^{1\prime}\right)_s\Delta e_{i,s} \Rightarrow (r/\lambda)W_i(\lambda_i)$ for $t \le T_b^i$ and $T^{-1/2}\sum_{s=2}^{t}\frac{1}{T\lambda}\left(DU^1DU^{1\prime}\right)_s\Delta e_{i,s} \Rightarrow W_i(\lambda_i)$ for $t > T_b^i$. The second element produces $\frac{(-1+\lambda)T^{1/2}}{1-T+T\lambda}\Delta e_{i,T_b^i+1} \Rightarrow dW_i(\lambda_i)$, an element that vanish asymptotically. The third element is $o_p(1)$, whereas the fourth element, which involves the second set of step dummy variables, $\frac{1}{1-T+T\lambda}\Delta e_{i,T_b^i+1}\left(t-T_b^i\right) \Rightarrow -\frac{r-\lambda_i}{1-\lambda_i}dW_i(\lambda_i)$, another element that vanish asymptotically. Finally, the fifth element is $T^{-1/2}\sum_{s=2}^{t}-\frac{1}{1-T+T\lambda}\left(DU^2DU^{2\prime}\right)_s\Delta e_{i,s} \Rightarrow \frac{r-\lambda_i}{1-\lambda_i}\left(W_i(1)-W_i(\lambda_i)\right)$ for $t > T_b^i$ and 0 elsewhere.

Therefore, $T^{-1/2}\tilde{e}_{i,t} \Rightarrow \sigma_i W_i(r) - \sigma_i(r/\lambda_i) W_i(\lambda_i)$ for $r \leq \lambda_i$ and $T^{-1/2}\tilde{e}_{i,t} \Rightarrow \sigma_i W_i(r) - \sigma_i W_i(\lambda_i) - \sigma_i \frac{r-\lambda_i}{1-\lambda_i}(W_i(1) - W_i(\lambda_i))$ for $r > \lambda_i$, and the $MSB_i(\lambda_i)$ test statistic converges to:

$$\frac{T^{-2}\sum_{t=1}^{T} \tilde{e}_{i,t-1}^2}{\tilde{\sigma}_i^2} = \frac{T^{-2}\sum_{t=1}^{T_b^i} \tilde{e}_{i,t-1}^2}{\tilde{\sigma}_i^2} + \frac{T^{-2}\sum_{t=T_b^i+1}^{T} \tilde{e}_{i,t-1}^2}{\tilde{\sigma}_i^2}$$

$$\Rightarrow \int_0^{\lambda_i} [W_i(r) - (r/\lambda_i) W_i(\lambda_i)]^2 dr$$

$$+ \int_{\lambda_i}^1 \left[W_i(r) - W_i(\lambda_i) - \frac{r-\lambda_i}{1-\lambda_i}(W_i(1) - W_i(\lambda_i))\right]^2 dr,$$

provided that $\tilde{\sigma}_i^2 \to \sigma_i^2$. However, the limit distribution of $MSB_i(\lambda_i)$ can be expressed as the sum of two independent integrals. Let us define $b = r/\lambda_{i,1}$ so that $0 < b < 1$. Using the properties of the Brownian motions the limit distribution can be written in terms of b as $W_i(r) - (r/\lambda_i) W_i(\lambda_i) = \sqrt{\lambda_i}W_{i,1}(b) - b\sqrt{\lambda_i}W_{i,1}(1) = \sqrt{\lambda_i}(W_{i,1}(b) - bW_{i,1}(1))$, so that

$$\int_0^{\lambda_i} [W_i(r) - (r/\lambda_i) W_i(\lambda_i)]^2 dr = \lambda_i^2 \int_0^1 [W_{i,1}(b) - bW_{i,1}(1)]^2 db$$

$$= \lambda_i^2 \int_0^1 V_{i,1}^2(b) db,$$

where $V_{i,1}(b)$ denotes the demeaned Brownian motion. For the second integral, let us now define $b = (r - \lambda_i)/(1 - \lambda_i)$, so that $0 < b < 1$. Now, the limit distribution can be reexpressed in terms of b as $W_i(r) - W_i(\lambda_i) - \frac{r-\lambda_i}{1-\lambda_i}(W_i(1) - W_i(\lambda_i)) = W_{i,2}(b) - bW_{i,2}(1)$, which implies that

$$\int_{\lambda_i}^1 \left[W_i(r) - W_i(\lambda_i) - \frac{r-\lambda_i}{1-\lambda_i}(W_i(1) - W_i(\lambda_i))\right]^2 dr = (1 - \lambda_i)^2 \int_0^1 V_{i,2}^2(b) db,$$

where $V_{i,2}(b)$ denotes the demeaned Brownian motion. Therefore, the asymptotic distribution of the test when $m_i = 1$ is given by

$$MSB_i(\lambda_i) \Rightarrow \lambda_i^2 \int_0^1 V_{i,1}^2(b) db + (1 - \lambda_i)^2 \int_0^1 V_{i,2}^2(b) db, \qquad (17)$$

where $V_{i,1}(b)$ and $V_{i,2}(b)$ are two independent Brownian bridges. Note also the symmetry of the asymptotic distribution around $\lambda_i = 0.5$. As shown above, we can interchange λ_i and $(1 - \lambda_i)$ in (17) and obtain the same asymptotic distribution.

In general, for $k = 1, \ldots, m_{i+1}$ we have $T_{i,k-1}^b < t \leq T_{i,k}^b$ and the partial sum processes converges to $\tilde{\sigma}_i^{-1}T^{-1/2}\tilde{e}_{i,t} \Rightarrow W_i(r) - (r - \lambda_{i,k-1})/(\lambda_{i,k} - \lambda_{i,k-1})(W_i(\lambda_{i,k}) - W_i(\lambda_{i,k-1}))$, with $\lambda_{i,0} = 0$ and $\lambda_{i,m_i+1} = 1$. Let us now define $b = (r - \lambda_{i,k-1})/(\lambda_{i,k} - \lambda_{i,k-1})$ so that $0 < b < 1$. As before, the

limit distribution of the partial sum processes is given by $\tilde{\sigma}_i^{-1} T^{-1/2} \tilde{e}_{i,t} \Rightarrow$ $\sqrt{\lambda_{i,k} - \lambda_{i,k-1}} (W_{i,k}(b) - bW_{i,k}(1))$, and the test statistic $MSB_i(\lambda_i) = \tilde{\sigma}_i^{-2} T^{-2} \sum_{t=1}^{T} \tilde{e}_{i,t-1}^2$
$= \tilde{\sigma}_i^{-2} T^{-2} \left[\sum_{t=1}^{T_{b,1}^i} \tilde{e}_{i,t-1}^2 + \cdots \cdots + \sum_{t=T_{b,k-1}^i+1}^{T_{b,k}^i} \tilde{e}_{i,t-1}^2 + \cdots + \sum_{t=T_{b,m_i}^i+1}^{T} \tilde{e}_{i,t-1}^2 \right]$
with limit distribution given by:

$$MSB_i \quad \Rightarrow \quad \lambda_{i,1}^2 \int_0^1 V_{i,1}^2(b)\, db + \cdots + (\lambda_{i,k} - \lambda_{i,k-1})^2 \int_0^1 V_{i,k}^2(b)\, db$$

$$+ \cdots + (1 - \lambda_{i,m_i})^2 \int_0^1 V_{i,m_i+1}^2(b)\, db,$$

where $V_{i,k}(\cdot)$, $k = 1, \ldots, m_{i+1}$, denotes the demeaned Brownian motion and provided that $\tilde{\sigma}_i^2 \to \sigma_i^2$ –see below the proof of the consistency of the non-parametric long-run variance estimation. The limit distribution of $MSB_i(\lambda_i)$ is the weighted sum of $(m_i + 1)$ independent Cramér-von Mises distributions. The expectations of these Cramér-von Mises distributions are $E\left[\int_0^1 V_{i,k}^2(b)\, db\right] =$ $1/6$ where the variance are $V\left[\int_0^1 V_{i,k}^2(b)\, db\right] = 1/45$, $\forall k = 1, \ldots, m_i + 1$ - see Levin and Lin (1992). Therefore, $E[MSB_i(\lambda_i)] = (1/6) \sum_{k=1}^{m_i+1} (\lambda_{i,k} - \lambda_{i,k-1})^2$ and $V[MSB_i(\lambda_i)] = (1/45) \sum_{k=1}^{m_i+1} (\lambda_{i,k} - \lambda_{i,k-1})^4$.

8.3 Proof of the consistency of the long-run variance estimation

Let us define the AR(1) regression on the estimated idiosyncratic residuals:

$$\tilde{e}_{i,t} = b_i \tilde{e}_{i,t-1} + \eta_{i,t}, \tag{18}$$

which under the null hypothesis of unit root implies that $\left(\tilde{b}_i - 1\right) = O_p(1/T)$. From (18) we can express the error term as:

$$\tilde{\eta}_{i,t} = \Delta \tilde{e}_{i,t} + \left(1 - \tilde{b}_i\right) \tilde{e}_{i,t-1},$$

where from (13) it follows that

$$\tilde{\eta}_{i,t} = z_{i,t} + v_t H^{-1} \pi_i - \hat{f}_t d_i + \left(1 - \tilde{b}_i\right) \tilde{e}_{i,t-1}$$
$$= z_{i,t} + w_{i,t},$$

with $z_{i,t} = (M_i \Delta e_i)_t$ and $w_{i,t} = v_t H^{-1} \pi_i - \hat{f}_t d_i + \left(1 - \tilde{b}_i\right) \tilde{e}_{i,t-1}$.

For arbitrary time series a_t and b_t define:

$$\widehat{NW}_{ab} = \frac{1}{T} \sum_{t=1}^{T} a_t b_t + \sum_{j=1}^{J} K(j) \left[\frac{1}{T} \sum_{t=1}^{T-j} (a_t b_{t+j} + a_{t+j} b_t) \right],$$

with $K(j) = 1 - j/(J+1)$. Then $\widehat{NW}_{z_i z_i}$ is the Newey-West estimator of the long-run variance of $z_i = M_i \Delta e_i$. In order to proof the consistency of this estimator we need to show that

$$\widehat{NW}_{\tilde{\eta}_i \tilde{\eta}_i} - \widehat{NW}_{z_i z_i} = o_p(1).$$

From $\tilde{\eta}_{i,t} = z_{i,t} + w_{i,t}$ we have that

$$\widehat{NW}_{\tilde{\eta}_i \tilde{\eta}_i} = \widehat{NW}_{z_i z_i} + 2\widehat{NW}_{z_i w_i} + \widehat{NW}_{w_i w_i}.$$

We next show that if $J \to \infty$ and $J/\delta_{NT} \to 0$, $\delta_{NT} = \min[N, T]$, then

$$\widehat{NW}_{z_i w_i} = o_p(1)$$

and

$$\widehat{NW}_{w_i w_i} = o_p(1).$$

First, notice that

$$\left| \widehat{NW}_{z_i w_i} \right| \leq \left(\frac{1}{T} \sum_{t=1}^{T} z_{i,t}^2 \right)^{1/2} \left(\frac{1}{T} \sum_{t=1}^{T} w_{i,t}^2 \right)^{1/2}$$
$$+ \sum_{j=1}^{J} K(j) \left[\left(\frac{1}{T} \sum_{t=1}^{T-j} z_{i,t}^2 \right)^{1/2} \left(\frac{1}{T} \sum_{t=1}^{T-j} w_{i,t+j}^2 \right)^{1/2} + \left(\frac{1}{T} \sum_{t=1}^{T-j} z_{i,t+j}^2 \right)^{1/2} \left(\frac{1}{T} \sum_{t=1}^{T-j} w_{i,t}^2 \right)^{1/2} \right].$$

Notice that $\frac{1}{T} \sum_{t=1}^{T} z_{i,t}^2 = O_p(1)$. On the other hand,

$$|w_{i,t}|^2 \leq 4 \|v_t\|^2 \|H^{-1}\pi_i\|^2 - 4 \|\hat{f}_t\|^2 \|d_i\|^2 + 4 \left(1 - \tilde{b}_i\right)^2 \tilde{e}_{i,t-1}^2,$$

so that

$$\frac{1}{T} \sum_{t=1}^{T} |w_{i,t}|^2 \leq 4 \frac{1}{T} \sum_{t=1}^{T} \|v_t\|^2 \|H^{-1}\pi_i\|^2 - 4 \frac{1}{T} \sum_{t=1}^{T} \|\hat{f}_t\|^2 \|d_i\|^2 + 4 \left(1 - \tilde{b}_i\right)^2 \frac{1}{T} \sum_{t=1}^{T} \hat{e}_{i,t-1}^{*2}.$$

From Lemmas 1(a) and 1(c) in Bai and Ng (2003), $\frac{1}{T} \sum_{t=1}^{T} \|v_t\|^2 = O_p\left(\delta_{NT}^{-2}\right)$ and $\|d_i\|^2 = O_p\left(\delta_{NT}^{-2}\right)$ respectively, and $\left(1 - \tilde{b}_i\right)^2 \frac{1}{T} \sum_{t=1}^{T} \tilde{e}_{i,t-1}^2 = T \left(1 - \tilde{b}_i\right)^2 \frac{1}{T^2} \sum_{t=1}^{T} \tilde{e}_{i,t-1}^2 = O_p(1/T)$. Therefore, $\frac{1}{T} \sum_{t=1}^{T} |w_{i,t}|^2 \leq O_p\left(\delta_{NT}^{-2}\right)$. These intermediate results enable us to establish that

$$\left| \widehat{NW}_{z_i w_i} \right| \leq (J+1) O_p\left(\delta_{NT}^{-1}\right) \to 0.$$

Moreover, since $\frac{1}{T} \sum_{t=1}^{T} |w_{i,t}|^2 \leq O_p\left(\delta_{NT}^{-2}\right)$ then

$$\left| \widehat{NW}_{w_i w_i} \right| \leq (J+1) O_p\left(\delta_{NT}^{-2}\right) \to 0.$$

Thus, we have shown that the long-run variance can be consistently estimated through the application of the non-parametric Newey-West estimation procedure, that is, we propose to use $\tilde{\sigma}_i^2 = \widehat{NW}_{\tilde{\eta}_i \tilde{\eta}_i}$.

References

[1] Adda, J. and J. Gonzalo (1996): "P-values for Non-standard Distributions with an Application to the DF Test". *Economics Letters*, 50, 155-160.

[2] Bai, J. and S. Ng (2001): "A New Look at Panel Testing of Stationarity and the PPP Hypothesis". *Mimeo.* http://server.econ.jhu.edu/people/ng//papers/spanic.pdf. Department of Economics. Johns Hopkins University.

[3] Bai, J. and Ng, S. (2002): "Determining the Number of Factors in Approximate Factor Models". *Econometrica*, 70, 191-221.

[4] Bai, J. and S. Ng (2004): "A PANIC Attack on Unit Root and Cointegration". *Mimeo.* http://server.econ.jhu.edu/people/ng//papers/panic.pdf. Department of Economics. Johns Hopkins University. Forthcoming in *Econometrica*.

[5] Baltagi, B. H. (2001): *Econometric Analysis of Panel Data.* John Wiley & Sons. Second edition.

[6] Baltagi, B. H. and C. Kao (2001): "Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey". *Advances in Econometrics*, 15, 7-51.

[7] Banerjee, A. (1999): Panel Data Unit Roots and Cointegration: An Overview. *Oxford Bulletin of Economics and Statistics*, Special Issue, Vol. 61, 607-629.

[8] Bartley, W. A., J. Lee and M. C. Strazicich (2001): "Testing the Null of Cointegration in the Presence of a Structural Break". *Economics Letters*, 73, 3, 315-323.

[9] Breitung, J. and W. Meyer (1994): "Testing for Unit Roots in Panel Data: Are Wages on Different Bargaining Levels Cointegrated?". *Applied Economics*, 26, 353-361.

[10] Busetti, F. and A. Harvey (2001): "Testing for the Presence of a Random Walk in Series with Structural Breaks". *Journal of Time Series Analysis*, 22, 2, 127-150.

[11] Carrion-i-Silvestre, J. Ll., T. del Barrio-Castro and E. López-Bazo (2001): "Level Shifts in a Panel Data Based Unit Root Test. An Application to the Rate of Unemployment". Proceedings of the 2001 North American Econometric Society.

[12] Carrion-i-Silvestre, J. Ll., T. del Barrio-Castro and E. López-Bazo (2002): "Breaking the panels. An Application to real per capita GDP". Mimeo. http://www.eco.ub.es/~carrion/breaking.pdf. Department of Econometrics, Statistics and Spanish Economy. University of Barcelona.

[13] Carrion-i-Silvestre, J. Ll. (2003): "A simple test of additive outlier detection for stationary Gaussian time series". Mimeo, Department of Econometrics, Statistics and Spanish Economy. University of Barcelona.

[14] Chen, C. and L. Liu (1993): "Joint Estimation of Model Parameters and Outlier Effects in the Time Series". *Journal of the American Statistical Association*, 88, 284-297.

[15] Choi, I. (2001): "Unit root tests for panel data". *Journal of Econometrics*, 20, 249-272.

[16] Franses, P. H. and N. Haldrup (1994): "The Effects of Additive Outliers on Tests for Unit Roots and Cointegration". *Journal of Business & Economic Statistics*, 12, 4, 471-478.

[17] Hansen, B. E. (1997): "Approximate Asymptotic P Values for Structural-Change Tests". *Journal of Business & Economic Statistics*, 15, 1, 60-67.

[18] Harvey, A. (2001): "A Unified Approach to Testing for Stationarity and Unit Roots". *Mimeo*. Faculty of Economics and Politics. University of Cambridge.

[19] Im, K. S., J. Lee and M. Tieslau (2002): "Panel LM Unit Root Test with Level Shifts". *Mimeo*. Department of Economics. University of Florida.

[20] Im, K. S., M. H. Pesaran and Y. Shin (1997): "Testing for Unit Roots in Heterogeneous Panels". *Mimeo*. Department of Applied Economics, University of Cambridge.

[21] Lee, J. (1996): "Minimum Statistics Testing for Stationarity in the Presence of a Structural Break". *Mimeo*. Department of Economics and Business Administration. Vanderbilt University.

[22] Lee, J. and M. C. Stratizich (2001): "Testing the Null of Stationarity in the Presence of a Structural Break". *Applied Economics Letters*, 8, 377-382.

[23] Levin, A. and C. F. Lin (1992): "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties". Working paper 92-23. Department of Economics. University of California at San Diego.

[24] Levin, A., C. F. Lin and C. S. J. Chu (2002): "Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties". *Journal of Econometrics,* 108, 1-24.

[25] Lumsdaine, R. L. and D. Papell (1997): "Multiple Trend Breaks and the Unit-Root Hypothesis". *Review of Economics and Statistics*, 79, 2, 212-218.

[26] MacKinnon, J. G. (1994): "Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration Tests". *Journal of Business and Economic Statistics*, 12, 167-76.

[27] Maddala, G. S. and S. Wu (1999): "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test". *Oxford Bulletin of Economics and Statistics*, Special Issue, 61, 631-652.

[28] Pedroni, P. (1995): "Panel Cointegration. Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP hypothesis". *Indiana University Working Papers in Economics*, num. 95-013.

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

[30] Perron, P. (1997): "Further Evidence on Breaking Trend Functions in Macroeconomic Variables". *Journal of Econometrics*, 80, 355-385.

[31] Perron, P. and T. Vogelsang (1992): "Nonstationarity and Level Shifts with an Application to Purchasing Power Parity". *Journal of Business and Economic Statistics*, 10, 3, 301-20.

[32] Phillips, P. C. B. and H. R. Moon (1999): "Linear Regression Limit Theory for Nonstationary Panel Data". *Econometrica*, Vol. 67 (5), 1057-1111.

[33] Stock, J. H. (1999): "A Class of Tests for Integration and Cointegration", in Engle, R. F. and H. White (ed.), *Cointegration, Causality and Forecasting. A Festchrift in Honour of Clive W. F. Granger*. Oxford University Press.

[34] Tanaka, K. (1996): *Time Series Analysis. Nonstationary and Noninvertible Distribution Theory*. John Wiley & Sons.

[35] Tsay, R. S. (1986): "Time Series Model Specification in the Presence of Outliers". *Journal of the American Statistical Association*, 81, 132-141.

[36] Vogelsang, T. (1998): "Testing for a Shift in Mean without Having to Estimate Serial-Correlation Parameters". *Journal of Business and Economic Statistics*, 16, 1, 73-80.