Generalized Fixed-T Panel Unit Root Tests Allowing for Structural Breaks

Karavias, Yiannis and Tzavalis, Elias

July 2012

Generalized fixed-T Panel Unit Root Tests Allowing for Structural Breaks

Yiannis Karavias and Elias Tzavalis*

Department of Economics

Athens University of Economics & Business

Athens 104 34, Greece

July 2012

Abstract

In this paper we suggest panel data unit root tests which allow for a structural breaks in the individual effects or linear trends of panel data models. This is done under the assumption that the disturbance terms of the panel are heterogenous and serially correlated. The limiting distributions of the suggested test statistics are derived under the assumption that the time-dimension of the panel (T) is fixed, while the cross-section (N) grows large. Thus, they are appropriate for short panels, where T is small. The tests consider the cases of a known and unknown date break. For the latter case, the paper gives the analytic form of the distribution of the test statistics. Monte Carlo evidence suggest that our tests have size which is very close to its nominal level and satisfactory power in small-T panels. This is true even for cases where the degree of serial correlation is large and negative, where single time series unit root tests are found to be critically oversized.

JEL classification: C22, C23

Keywords: Panel data models; unit roots; structural breaks;

*Corresponding author
Yiannis Karavias: email: ioannis.karavias@nottingham.ac.uk
Elias Tzavalis: email: etzavalis@aueb.gr, tel: 00302108203332

1 Introduction

A vast amount of work has been recently focused on drawing inference about unit roots based on dynamic panel data models (see, Hlouskova and Wagner (2006), for a more recent survey). Since many empirical panel data studies rely on short panels, of particular interest is testing for a unit root in dynamic panel data model when the time dimension of the panel, denoted as T, is fixed (finite) and its cross-section, denoted as N, grows large (see, e.g. Blundell and Bond (1998), Harris and Tzavalis (1999, 2004), Arellano and Honore (2002) and Binder et al (2005)). These tests have better small-T sample performance, compared to large-T panel unit root tests (see, e.g., Levin et al (2002)), given that they assume finite T In this paper, we extend the fixed-T panel data unit roots test statistics of Harris and Tzavalis (1999, 2004) to allow for a common structural break in the deterministic components of panel data models, namely their individual effects or linear trends of a known and unknown date. This is done in a generalized dynamic panel data framework which allows for heterogenous and serially correlated disturbance terms, for all units of the panel. This assumption makes the tests applicable under quite general panel data generating processes, observed in reality. The maximum order of serial correlation allowed is a function of T.

The extension of fixed-T panel unit root tests to allow for structural breaks is very useful given evidence supporting the view that the presence of unit roots in economic time series can be falsely attributed to the existence of structural breaks in their deterministic components (see, e.g., Perron (2006), for a survey). On this front, the panel data approach offers an interesting and unique perspective that it is not shared by single time series tests. The cross-sectional dimension of the panel can provide useful information, which can help to distinguish the type of shifts (breaks) in the deterministic components of the panel from the effects of stochastic permanent shocks. As pointed out by Bai (2010), this framework can more accurately trace out structural break points of the panel data.[1] There are a few studies in the literature which suggest fixed-T panel data unit root tests allowing for a common structural break in the deterministic components of the panel data model (see, more recently, Karavias and Tzavalis (2012)). These studies however suggest unit root tests using the simple AR(1), dynamic panel data model as an auxiliary regression model, which may not be operational in practice due to the assumption of no serial correlation in its disturbance terms. The main goal of these studies is to pass ideas how to test for unit roots in the presence of structural breaks,

[1] Detecting procedures of structural breaks for stationary panel data models have been also suggested in the literature by De Wachter and Tzavalis (2005, 2012).

considering mainly the case of a known date break point. In addition to the above, there are also studies in the literature which suggest panel unit root tests allowing for a common structural break, but they assume that T is large and grows faster than N (see, e.g., Carrion-i-Silvestre et al (2005), Bai and Carrion-i-Silvestre (2009), and Kim (2011)). These tests are appropriate for large-T panel data sets. Application of these tests to small-T panel data sets will lead to serious size distortions and critical power reductions of them. As shown in Karavias and Tzavalis (2012), the existence of a break in the data generating process requires panel data sets with a quite large time-dimension, T (e.g. $T > 150$), so as large-T panel unit root tests to have satisfactory size and power performance in short panels.

The paper suggests panel data unit root test statistics allowing for a structural break in both cases of a known and an unknown date break. The second category of test statistics relies on a sequential application of the first, such as that suggested by Zivot and Andrews (1992), Andrews (1993), Perron and Vogelsang (1998), *inter alia*), for single time series. The limiting distribution of these test statistics is obtained as the minimum value of a finite number of correlated variates; $T-2$ for the dynamic panel data model with individual effects and $T-3$ for the extension of this model allowing also for individual linear trends. This distribution is derived analytically, based on recent results of Arellano-Valle and Genton (2008) who have derived the analytic form of the probability density function of the maximum of absolutely continuous dependent random variables. The analytic form of this distribution enables us to derive critical values of our suggested test statistics without having to rely on Monte Carlo analysis. This substantially facilitates application of the tests in practice and their generalization to the case of serially correlated disturbance terms.

The paper is organized as follows. In Section 2, we derive the limiting distributions of the test statistics under the assumption that the disturbance terms of the panel data models considered are white noise processes. This analysis will helps us to better interpret the limiting distribution of the sequential version of the test statistics, in the case of an unknown date break. In Section 3, we generalize the test statistics to allow for serial correlation in the disturbance terms. In Section 4 we extend the tests to allow also for individual linear trends. In this section, we also show how to carry out the tests when there is a break in the individual effects of panel data models under the null hypothesis of unit roots. Section 5 conducts a Monte Carlo simulation study to examine the small sample performance of the tests. Section 6 concludes the paper. All the mathematical derivations are provided in the Appendix of the paper.

2 Test statistics and their limiting distribution

In this section, we present panel unit root test statistics under the assumption that the disturbance terms of the AR(1) panel data model considered are independently, identically normally distributed ($NIID$). This is done, first, for the known date break case and, then, for the unknown. Extensions of the tests to the more general case of serially correlated and heterogenous disturbance terms are made in the next section.

2.1 Known date break

Consider the following AR(1) nonlinear dynamic panel data model:

$$y_{it} = a_{it}^{(\lambda)}(1-\varphi) + \varphi y_{it-1} + u_{it}, \quad i = 1, 2, .., N, \tag{1}$$

where $\varphi \in (-1, 1]$, $a_{it}^{(\lambda)} = a_i^{(1)}$ if $t \leq T_0$ and $a_i^{(2)}$ if $t > T_0$, where T_0 denotes the time-point of the sample, referred to as break-point, where a common break in the individual effects of panel data model (1) α_i occurs, for all cross-section units of the panel i. $a_i^{(1)}$and $a_i^{(2)}$ denote the individual effects of model (1) before and after the break point T_0, respectively. Throughout the paper, we will denote the fraction of the sample that this break occurs as λ, i.e. $\lambda = \frac{T_0}{T} \in I = \left\{\frac{2}{T}, \frac{3}{T}, ..., \frac{T-1}{T}\right\}$.

Under the null hypothesis of a unit root (i.e. $\varphi = 1$), model (1) reduces to the pure random walk model $y_{it} = y_{it-1} + u_{it}$, for all i, while, under the alternative of stationarity (i.e. $\varphi < 1$), it considers a common structural break in individual effects a_i. The above specification of the null and the alternative hypotheses is very common in single time series unit root inference procedures allowing for structural breaks (see, e.g., Zivot and Andrews (1992), Andrews (1993), Perron and Vogelsang (1998). The main focus of these procedures is to diagnose whether evidence of unit roots can be spuriously attributed to the ignorance of structural breaks in nuisance parameters of the data generating processes like individual effects a_i. The common break assumption across all units of the panel i can be attributed to a monetary regime shift, which is common across all economic units, or to a structural economic shock which is independent of the disturbance terms u_{it}, like a credit crunch or an exchange rate realignment. As aptly noted by Bai (2010), even if each series of the panel data model has its own break point, the common break assumption across i is useful in practice not only for its computational simplicity, but also because it allows for estimating the mean of possibly random break points.

The AR(1) panel data model (1) can be employed to carry out unit root tests allowing for a structural break in individual effects $a_{it}^{(\lambda)}$ based on the within groups least squares (LS) estimator of autoregressive coefficient of φ, denoted as $\hat{\varphi}^{(\lambda)}$. This estimator is also known as least square dummy variable (LSDV) estimator (see, e.g., Baltagi (1995), *inter alia*). Under null hypothesis $\varphi = 1$, it implies:

$$\hat{\varphi}^{(\lambda)} - 1 = \left[\sum_{i=1}^{N} y_{i,-1}' Q^{(\lambda)} y_{i,-1}\right]^{-1} \left[\sum_{i=1}^{N} y_{i,-1}' Q^{(\lambda)} u_i\right], \tag{2}$$

where $y_i = (y_{i1}, ..., y_{iT})'$ is a $(TX1)$-dimension vector collecting the time series observations of dependent variable y_{it} of each cross-section unit of the panel i, $y_{i,-1} = (y_{i0}, ..., y_{iT-1})'$ is vector y_i lagged one period back, $u_i = (u_{i1}, ..., u_{iT})$ is a $(TX1)$-dimension vector of disturbance terms u_{it} and $Q^{(\lambda)}$ is the (TXT) "within" transformation matrix of the individual series of the panel data model, y_{it}. Let us define $X^{(\lambda)} \equiv \left(e^{(1)}, e^{(2)}\right)$ to be a matrix of deterministic components used by the LSDV estimator to demean the levels of series y_{it}, for all i, where $e^{(1)}$ and $e^{(2)}$ are $(TX1)$-column vectors whose elements are defined as follows: $e_t^{(1)} = 1$ if $t \leq T_0$ and 0 otherwise, and $e_t^{(2)} = 1$ if $t > T_0$ and 0 otherwise. Then, matrix $Q^{(\lambda)}$ will be defined as $Q^{(\lambda)} = I_T - X^{(\lambda)}(X^{(\lambda)\prime}X^{(\lambda)})^{-1}X^{(\lambda)\prime}$, where I_T is an identity matrix of dimension (TXT).

Panel data unit root testing procedures based on above LSDV estimator $\hat{\varphi}^{(\lambda)}$ have the very interesting property that, under the null hypothesis of $\varphi = 1$, are invariant (similar) to the initial conditions of the panel y_{i0} and, after appropriate specification of matrix $X^{(\lambda)}$, to the individual effects of the panel data model, as will be seen in Section 4. The latter happens if matrix $X^{(\lambda)}$ also contains broken linear trends. Similarity of the tests with respect to initial conditions y_{i0} does not require any mean or covariance stationarity conditions on the panel data processes y_{it}, as assumed by generalized method of moments (GMM), or conditional and unconditional maximum likelihood (ML) based panel data unit root inference procedures (see, e.g., Hsiao et al (2002) and Madsen (2008)). These conditions may be proved restrictive in practice.[2] However, $\hat{\varphi}^{(\lambda)}$ is an inconsistent (asymptotic biased) estimator of φ, due to the within transformation of the data which wipes off individual effects $a_{it}^{(\lambda)}$ and/or initial conditions y_{i0} under null hypothesis $\varphi = 1$. Thus, our suggested panel unit root test statistics will rely on a correction of estimator $\hat{\varphi}^{(\lambda)}$ for its inconsistency (asymptotic bias) (see, e.g., Harris and Tzavalis (1999, 2004)). To derive the limiting distribution of these tests, we make the following assumption about the sequence of disturbance terms $\{u_{it}\}$.

[2] Furthermore, the performance of the GMM estimator over the LS may detetiorates due to the inacurate estimation of the weigthing (variance-covariance) matrix. See De Wachter et al (2007) and Han and Phillips (2010).

Assumption 1: (a1) $\{u_i\}$ constitutes a sequence of independent identically distributed (IID) $(TX1)$-dimension vectors with means $E(u_i) = 0$ and variance-autocovariance matrices $\Gamma_i \equiv E(u_i u_i') = \sigma_u^2 I_T < +\infty$ and nonzero, for all i. (a2) $E(u_{it} y_{io}) = E\left(u_{it} a_{it}^{(1)}\right) = E\left(u_{it} a_{it}^{(2)}\right) = 0$ and $\forall\, i \in \{1, 2, ..., N\}$, $t \in \{1, 2, ..., T\}$. (a3) $E\left(u_{it}^4\right) < +\infty$, $E(y_{i0}^4) < +\infty$, $E\left((a_{it}^{(1)})^4\right) < +\infty$, $E\left((a_{it}^{(2)})^4\right) < +\infty$ and $E\left(y_{i0}^2 \left(a_{it}^{(1)}\right)^2\right) < +\infty$, $E\left(y_{i0}^2 \left(a_{it}^{(2)}\right)^2\right) < +\infty$.

Condition (a1) of Assumption 1 enables us to derive under null hypothesis $\varphi = 1$ the limiting distribution of a panel data unit root test statistic based on estimator $\hat{\varphi}^{(\lambda)}$ by applying standard asymptotic theory for IID processes, while (a2) and (a3) are simple regularity conditions under which the suggested test statistic can be proved that is consistent under alternative hypothesis $\varphi < 1$. The following theorem provides the limiting distribution of such a test statistic, based on estimator $\hat{\varphi}^{(\lambda)}$ corrected for its bias. For analytic convenience, this is done under the assumption that disturbance terms u_{it} are also normally distributed, i.e. $u_{it} \sim NIID(0, \sigma_u^2)$, for all i and t.

Theorem 1 *Let $u_{it} \sim NIID(0, \sigma_u^2)$, then, under null hypothesis $\varphi = 1$ and known λ, we have*

$$Z^{(\lambda)} \equiv \widehat{V}^{(\lambda)-1/2} \hat{\delta}^{(\lambda)} \sqrt{N} \left(\hat{\varphi}^{(\lambda)} - \frac{\hat{b}^{(\lambda)}}{\hat{\delta}^{(\lambda)}} - 1\right) \xrightarrow{d} N(0, 1) \tag{3}$$

as $N \to \infty$, where

$$\frac{\hat{b}^{(\lambda)}}{\hat{\delta}^{(\lambda)}} \equiv \frac{\hat{\sigma}_u^2 tr(\Lambda' Q^{(\lambda)})}{\frac{1}{N} \sum_{i=1}^N y_{i,-1}' Q^{(\lambda)} y_{i,-1}} \tag{4}$$

is a consistent estimate of the asymptotic bias of $\hat{\varphi}^{(\lambda)}$ which, under the null hypothesis, is given as $\frac{b^{(\lambda)}}{\delta^{(\lambda)}} = \frac{\sigma_u^2 tr(\Lambda' Q^{(\lambda)})}{\sigma_u^2 tr(\Lambda' Q^{(\lambda)} \Lambda)}$, $\hat{\sigma}_u^2$ is a consistent estimator of variance σ_u^2 under the null hypothesis, given as $\hat{\sigma}_u^2 = \frac{\sum_{i=1}^N \Delta y_i' \Psi^{(\lambda)} \Delta y_i}{N tr(\Psi^{(\lambda)})}$, where Δ is the difference operator and $\Psi^{(\lambda)}$ is a (TXT)-dimension matrix having in its main diagonal the corresponding elements of matrix $\Lambda' Q^{(\lambda)}$, and zeros elsewhere, and $V^{(\lambda)}$ is a variance function given as

$$V^{(\lambda)} = \sigma_u^4 F^{(\lambda)\prime} (K_{T^2} + I_{T^2}) F^{(\lambda)}, \tag{5}$$

where $F^{(\lambda)} = vec(Q^{(\lambda)} \Lambda - \Psi^{(\lambda)\prime})$, K_{T^2} is a $(T^2 X T^2)$-dimension commutation matrix and I_{T^2} is a $(T^2 X T^2)$-dimension identity matrix.

The test statistic $Z(\lambda)$, given by Theorem 1, can be easily implemented to test null hypothesis $\varphi = 1$ based on the tables of the standard normal distribution. Theorem 1 shows that the asymptotic bias of estimator $\hat{\varphi}^{(\lambda)}$ stems from the "within" transformation matrix $Q^{(\lambda)}$, which induces correlation between vectors $y_{i,-1}$ and u_i (see, e.g. Nickel (1981)). Since disturbance terms u_{it} are IID, the correlation between $y_{i,-1}$ and u_i comes only from the main diagonal elements of the variance-autocovariance matrices of u_{it}, defined by Assumption 1 as $\Gamma_i \equiv E(u_i u_i') = \sigma_u^2 I_T$, for all i. The above bias can be estimated by the nonparametric estimator $\frac{\hat{b}^{(\lambda)}}{\hat{\delta}^{(\lambda)}}$ and, thus, it can be subtracted from $\hat{\varphi}^{(\lambda)} - 1$ to obtain a test statistic which is normally distributed and is asymptotically net of nuisance parameter effects. To test null hypothesis $\varphi = 1$, this test statistic is based on the off-diagonal elements of the sample moments of variance-autocovariance matrices Γ_i which are equal to zero, i.e. $E(u_{it}u_{is}) = 0$ for $s \neq t$. This can be better seen by writing test statistic $Z^{(\lambda)}$ as

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} u_i'(\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)}) u_i = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} tr\left[(\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)}) u_i u_i'\right], \tag{6}$$

(see Appendix) where $(\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)})$ is matrix with zeros in its main diagonal due to the subtraction of matrix $\Psi^{(\lambda)}$ from $\Lambda' Q^{(\lambda)}$, which implies that $tr\left[(\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)}) E(u_i u_i')\right] = 0$, for all i.[3] Matrix $\Psi^{(\lambda)}$ allows us to capture the correlation effects between vectors $y_{i,-1}$ and u_i, which are induced by the "within" transformation of the data through matrix $Q^{(\lambda)}$ and generate the bias of LSDV estimator $\hat{\varphi}^{(\lambda)}$. Subtracting $\Psi^{(\lambda)}$ from $\Lambda' Q^{(\lambda)}$ enables us to adjust $\hat{\varphi}^{(\lambda)}$ for this bias. The adjusted LS estimator relies on sample moments of variance-autocovariance Γ_i with zero elements, i.e. $E(u_{it}u_{is}) = 0$, for $s \neq t$. These moments are weighted by the elements of matrix $\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)}$. They can be consistently estimated under null hypothesis $\varphi = 1$. Writing analytically matrix $\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)}$ can be easily seen that the elements of this matrix put more weights to sample moments of $E(u_{it}u_{is})$, for $s \neq t$, with s and t defined immediately before break point, T_0.

The next theorem establishes the consistency of test statistic $Z^{(\lambda)}$.

Theorem 2 *Under conditions (a1)-(a3) of Assumption 1, it can be proved that*

$$\lim_{N \to +\infty} P(Z^{(\lambda)} < z_a \mid H_a) = 1, \qquad \lambda \in I, \tag{7}$$

where z_a is the critical value of standard normal distribution at significance level a.

[3] Note that matrix $\Psi^{(\lambda)}$ is used to estimate σ_u^2, based on estimator $\hat{\sigma}_u^2 = \frac{\sum_{i=1}^{N} \Delta y_i' \Psi^{(\lambda)} \Delta y_i}{N tr(\Psi^{(\lambda)})}$ where Δ is the difference operator.

2.2 Unknown break point

In this section, we relax the assumption that break point T_0 is known. We propose a panel data unit root test statistic which, under the alternative hypothesis of stationarity, assumes that T_0 is unknown. As in single time series literature (see, e.g., Zivot and Andrews (1992) and Perron and Vogelsang (1998)), we will view the selection of the break point as the outcome of minimizing the standardized test statistic $Z^{(\lambda)}$, given by Theorem 1, over all possible break fractions (or break points T_0) of the sample, λ, after trimming out the initial and final parts of the time series observations of the panel data. The minimum value of test statistics $Z^{(\lambda)}$, for all $\lambda \in I$, defined as $z \equiv \min_{\lambda \in I} Z^{(\lambda)}$, will give the least favorable result of null hypothesis $\varphi = 1$. Let $\hat{\lambda}_{\min}$ denote the break point at which the minimum value of $Z^{(\lambda)}$, over all $\lambda \in I$, is obtained. Then, the null hypothesis of $\varphi = 1$ will be rejected if

$$Z^{(\hat{\lambda}_{\min})} < c_{\min}, \tag{8}$$

where $c_{\min}$ denotes the size a left-tail critical value of the limiting distribution of $\min_{\lambda \in I} Z^{(\lambda)}$. The following theorem enables us to tabulate the critical values of this distribution at any significance (size) level a.

Theorem 3 *Let condition (a1) of Assumption 1 hold and u_{it} is normally distributed. Then, under null hypothesis $\varphi = 1$ and unknown λ, we have*

$$z \equiv \min_{\lambda \in I} Z^{(\lambda)} \xrightarrow{d} \zeta \equiv \min_{\lambda \in I} N(0, \Sigma) \tag{9}$$

as $N \to \infty$, where $\Sigma \equiv [\sigma_{\lambda s}]$ is the variance-covariance matrix of the test statistics $Z^{(\lambda)}$, with elements $\sigma_{\lambda s}$ given by the following formula:

$$\sigma_{\lambda s} = \frac{F^{(\lambda)\prime}(K_{T^2} + I_{T^2})F^{(s)}}{\sqrt{F^{(\lambda)\prime}(K_{T^2} + I_{T^2})F^{(\lambda)}}\sqrt{F^{(s)\prime}(K_{T^2} + I_{T^2})F^{(s)}}}, \tag{10}$$

where λ and s denote two different fractions of the sample that the break can occur.

The result of Theorem 3 implies that critical values of the limiting distribution of the standardized test statistic $\min_{\lambda \in I} Z^{(\lambda)}$, denoted $c_{\min}$, can be obtained from the distribution of the minimum value of a fixed number of $T-2$ correlated normal variables $Z^{(\lambda)}$ with covariance matrix Σ. Since $\min\{Z^{(\frac{2}{T})}, Z^{(\frac{3}{T})}..., Z^{(\frac{T-1}{T})}\} =$

$\max\{-Z^{(\frac{2}{T})}, -Z^{(\frac{3}{T})}..., -Z^{(\frac{T-1}{T})}\}$, we can use the distribution of the maximum of normal variables $-Z^{(\lambda)}$ to calculate critical value $c_{\min}$ for a significance level a, i.e.

$$P(\zeta < c_{\min}) = P(-\zeta > -c_{\min}) = a. \tag{11}$$

The integral function $P(\zeta > -c_{\min}) = a$ can be calculated numerically based on the probability density function (pdf) of $-\zeta$. This density function has been recently derived by Arellano-Valle and Genton (2008), for the more general case of the maximum of absolutely continuous dependent random variables of elliptically contoured distributions. For the case of normal random variables, it is given as

$$f_{\zeta}(x) = \sum_{\lambda} \phi(x; \mu_{\lambda}, \Sigma_{\lambda,\lambda}) \Phi(xe_{T-3}; \mu_{-\lambda,\lambda}, \Sigma_{-\lambda-\lambda,\lambda}), \quad x \in R, \tag{12}$$

where e_{T-3} is a $(T-3)$-column vector of unities, $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the normal distribution with arguments given as follows:

$$\mu_{-\lambda,\lambda}(x) = \mu_{-\lambda} + (x - \mu_{\lambda})\Sigma_{-\lambda,\lambda}(\Sigma_{\lambda,\lambda})^{-1} \quad \text{and} \quad \Sigma_{-\lambda-\lambda,\lambda} = \Sigma_{-\lambda-\lambda} - \Sigma_{-\lambda,\lambda}\Sigma'_{-\lambda,\lambda}(\Sigma_{\lambda,\lambda})^{-1},$$

where $\mu = (\mu_{-\lambda} \vdots \mu_{\lambda})'$ and $\Sigma = \begin{bmatrix} \Sigma_{-\lambda,-\lambda} & \Sigma_{-\lambda,\lambda} \\ \Sigma_{\lambda,-\lambda} & \Sigma_{\lambda,\lambda} \end{bmatrix}$ are respectively the vector of means and the variance-autocovariance matrix of the $(T-2)$-column vector Z which consists of random variables $Z^{(\boldsymbol{\lambda})}$, for $\lambda \in I$, partitioned as $Z = (Z^{(-\boldsymbol{\lambda})} \vdots Z^{(\boldsymbol{\lambda})})'$, where $Z^{(-\boldsymbol{\lambda})}$ is a (T-3)-column vector consisting of the remaining elements of Z, which exclude $Z^{(\boldsymbol{\lambda})}$.

The above pdf of random variable $-\zeta$, defined as $f_{\zeta}(x)$, is a mixture of the normal marginal densities $\phi(x; \mu_{\lambda}, \Sigma_{\lambda,\lambda})$ corresponding to all possible break fractions of the sample λ. These densities are weighed with the cdf values of the (T-3)-column vector xe_{T-3}, given as $\Phi(xe_{T-3}; \mu_{-\lambda,\lambda}(x), \Sigma_{-\lambda-\lambda,\lambda})$. Intuitively, the pdf formula given by (12) sums up the probabilities that one random variable $-Z^{(\lambda)}$ takes its maximum value x (implying that $Z^{(\lambda)}$ takes its minimum value), while the remaining variables, collected in vector $-Z^{(-\lambda)}$, take values smaller than x.

The consistency of the test given by Theorem 3 follows immediately from Theorem 2, which proves the consistency of $Z^{(\lambda)}$ for a known date break. This can be seen by noting that if, under the alternative hypothesis of $\varphi < 1$, test statistic $Z^{(\boldsymbol{\lambda})}$ converges to minus infinity, for $\lambda \in I$, then so does their minimum.

3 The generalization of the test statistics for serially correlated and heterogenous disturbance terms

The test statistics presented in the previous section can be generalized to allow for serially correlated and heterogenous disturbance terms u_{it}, for all i. Due to the fixed-T dimension of panel data model (1) and the allowance for a common structural break in the individual effects $\alpha_{it}^{(\lambda)}$, the maximum order of serial correlation, denoted as $p_{\max}$, which will be considered by the generalized test statistics is a function of the time-dimension of the panel T. This will be assumed to be the same for both sample intervals before and after break point T_0. Later on, we will give a table of values for $p_{\max}$ which do not depend on the location of the break, T_0. These are very useful for the application of our tests, in practice.

To derive the limiting distribution of test statistics based on estimator $\hat{\varphi}^{(\lambda)}$ under the above more general assumptions, we will make the following assumption about the sequence of the disturbance terms $\{u_i\}$.

Assumption 2: (b1): $\{u_i\}$ constitutes a sequence of independent random vectors of dimension $(TX1)$ with means $E(u_i) = 0$ and variance-autocovariance matrices $E(u_i u_i') = \Gamma_i \equiv [\gamma_{i,ts}]$, for all i, where $\gamma_{i,ts} = E(u_{it}u_{is}) = 0$ for $s = t + p_{\max} + 1, ..., T$ and $t < s$. (b2): The average population covariance matrix $\Gamma_N \equiv \frac{1}{N}\sum_{i=1}^{N}\Gamma_i$ is bounded away from zero in large samples: $\bar{\gamma}_{N,tt} > \eta'$ for some $\eta' > 0$ and for all $N > N_0$, for some N_0, and for at least one $t \in \{1, ..., T\}$. (b3): The $4+\eta$-th population moments of Δy_i, $i = 1, ..., N$, are uniformly bounded. That is, for every real $(TX1)$ vector l such that $l'l = 1$, we have $E(|l'\Delta y_i|^{4+\eta}) < B < \infty$, for some B. (b4): $\frac{1}{N}\sum\limits_{i=1}^{N} l'Var(vec(\Delta y_i \Delta y_i'))l > \eta'$ for some $\eta' > 0$, and for all $N > N_1$, for some N_1 and for every real $(\frac{1}{2}T(T+1)X1)$ vector l with $l'l = 1$. (b5): $E(u_{it}y_{io}) = E\left(u_{it}a_i^{(1)}\right) = E\left(u_{it}a_i^{(2)}\right) = 0$ and $\forall$ $i \in \{1, 2, ..., N\}$, $t \in \{1, 2, ..., T\}$.

Assumption 2 enables us to derive the limiting distribution of a normalized test statistic based on $\hat{\varphi}^{(\lambda)} - 1$ by employing standard asymptotic theory under more general conditions than those of Assumption 1 (see White (2000)), which considers the simple case that $u_{it} \sim NIID(0, \sigma_u^2)$, for all i. More specifically, condition (b1) allows the variance-autocovariance matrices of disturbance terms u_{it}, $\Gamma_i = E(u_i u_i')$, to be heterogenous across the cross-sectional units of the panel i with a maximum order (degree) of serial correlation p_{max} less than T. The pattern of serial correlation considered by matrices Γ_i can capture that implied by moving average (MA) processes of u_{it}, often assumed for many economic series (see, e.g. Schwert (1989)). It can be also though of as approximating that implied by AR models of u_{it} whose autocorrelation dies out after $p_{\max}$.[4] This pattern will enable us to correct LSDV $\hat{\varphi}^{(\lambda)}$ for its inconsistency due to serial correlation in u_{it}. This can be done based on moments $E(u_{it-p_{\max}-1}u_{it})$ which are zero, across t, since disturbance terms $u_{it-p_{\max}-1}$ and u_{it} are assumed to be uncorrelated (see, e.g. Kruininger and Tzavalis (2002), De Wachter, Harris and Tzavalis (2007)).

Condition (b2) qualifies application of a central limit theorem (CLT) to derive the limiting distribution of a test statistic $\hat{\varphi}^{(\lambda)} - 1$ adjusted for the inconsistency of estimator $\hat{\varphi}^{(\lambda)}$, as $N \to \infty$, under the more general assumptions than condition (b1). More specifically, Condition (b2) along with condition (b4) guarantees that, the variance and the suggested test statistic will be different than zero. Finally, conditions (b5) and (b3) constitute weak conditions under which the consistency of the tests can be proved. These two conditions correspond to conditions (a2) and (a3) of Assumption 1.

Under the conditions of Assumption 2, the next theorem derives the limiting distribution of a normalized test statistic based on estimator $\hat{\varphi}^{(\lambda)}$ corrected for its inconsistency under $\varphi = 1$ and for a known date break point.

Theorem 4 *Let conditions (b1) - (b5) of Assumption 2 hold. Then, under null hypothesis $\varphi = 1$ and λ known, we have*

$$Z_1^{(\lambda)} \equiv \hat{V}_1^{(\lambda)-1/2}\hat{\delta}_1^{(\lambda)}\sqrt{N}\left(\hat{\varphi}^{(\lambda)} - \frac{\hat{b}_1^{(\lambda)}}{\hat{\delta}_1^{(\lambda)}} - 1\right) \xrightarrow{d} N(0,1) \tag{13}$$

as $N \to \infty$, where

$$\frac{\hat{b}_1^{(\lambda)}}{\hat{\delta}_1^{(\lambda)}} = \frac{tr(\Psi_1^{(\lambda)}\hat{\Gamma}_N)}{\frac{1}{N}\sum_{i=1}^{N} y'_{i,-1}Q^{(\lambda)}y_{i,-1}} \tag{14}$$

is a consistent estimate of the asymptotic bias of $\hat{\varphi}^{(\lambda)}$ which, under the null hypothesis, is given as

$$\frac{b_1^{(\lambda)}}{\delta_1^{(\lambda)}} = \frac{tr(\Lambda' Q^{(\lambda)}\Gamma_N)}{tr(\Lambda' Q^{(\lambda)}\Lambda\Gamma_N)}, \tag{15}$$

where matrix $\Psi_1^{(\lambda)}$ is a (TXT)-dimension matrix having in its main diagonal, and its p-lower and $p-$upper diagonals of the main diagonal the corresponding elements of matrix $\Lambda' Q^{(\lambda)}$, and zero otherwise, $\hat{\Gamma}_N =$

[4] In single time series literature, $p_{\max}$ is assumed to increase with T with an order of $o(T^{1/2})$, see Chang and Park (2002).

$\frac{1}{N}\sum_{i=1}^{N}(\Delta y_i \Delta y_i')$ is a consistent estimator of population variance-autocovariance matrix Γ_N and $V_1^{(\lambda)}$ is a variance function given as

$$V_1^{(\lambda)} = F_1^{(\lambda)\prime}\Theta F_1^{(\lambda)\prime}, \tag{16}$$

where $F_1^{(\lambda)} = vec(Q^{(\lambda)}\Lambda - \Psi_1^{(\lambda)\prime})$ and $\Theta = \frac{1}{N}\sum_{i=1}^{N} Var(vec(u_i u_i'))$ is the variance-covariance matrix of $vec(u_i u_i')$.

To implement the test statistic given by Theorem 4, $Z_1^{(\lambda)}$, we need consistent estimates of the variance-covariance matrix of vector $vec(u_i u_i')$, defined as Θ. This can be done under null hypothesis $\varphi = 1$ based on the following estimator:

$$\hat{\Theta} = \frac{1}{N}\sum_{i=1}^{N}\left(vec(\Delta y_i \Delta y_i')vec(\Delta y_i \Delta y_i')'\right). \tag{17}$$

As $\Psi^{(\lambda)}$ for $Z^{(\lambda)}$, matrix $\Psi_1^{(\lambda)}$ plays a crucial role in constructing test statistic $Z_1^{(\lambda)}$. It adjusts LS estimator $\hat{\varphi}^{(\lambda)}$ for its asymptotic bias. This bias now comes from two sources: the "within" transformation of the data through matrix $Q^{(\lambda)}$, which has been examined before, and the serial correlation of disturbance terms u_{it}.[5] Subtracting $\Psi_1^{(\lambda)}$ from $\Lambda' Q^{(\lambda)}$ enables to adjust $\hat{\varphi}^{(\lambda)}$ for the above two sources of bias. The adjusted LS estimator $\hat{\varphi}^{(\lambda)}$ enables us to test the null hypothesis of $\varphi = 1$ based on sample moments of the elements of variance-autocovariance matrices Γ_i, for all i, which are mot serially correlated, i.e. $E(u_{it}u_{is}) = 0$, for $s = t + p_{\max} + 1, ..., T$ and $t < s$. These moments are weighted by elements of matrix $\Lambda' Q^{(\lambda)} - \Psi_1^{(\lambda)}$. These assign higher weights to the moments which are immediately before the break point T_0 than those which are away from it. They can be consistently estimated under the null hypothesis through the variance-covariance estimator $\hat{\Theta}$. The weights that matrix $\Lambda' Q^{(\lambda)} - \Psi_1^{(\lambda)}$ assigns to the above elements of variance-autocovariance matrices Γ_i obviously depend on the break point and the maximum order of serial correlation $p_{\max}$ considered by test statistic $Z_1^{(\lambda)}$. Based on the specification of this matrix, Table 1 and following relationship

$$p_{\max} = \left[\frac{T}{2} - 2\right]^*, \tag{18}$$

where $[.]^*$ denotes the greatest integer function, give values of $p_{\max}$ which enable us to implement test statistic $Z_1^{(\lambda)}$ independently of the location of the break T_0, or sample fraction λ. These values are chosen so as the

[5] Note that, under conditions of Assumption 1, test statistic $Z_1^{(\lambda)}$ becomes identical to $Z^{(\lambda)}$.

elements of matrix $\Lambda' Q^{(\lambda)} - \Psi_1^{(\lambda)}$ do not assign weights to zero elements of Γ_i, which result in a zero value of variance function $V_1^{(\lambda)}$. They are useful in choosing the maximum order of serial correlation $p_{\max}$ considered by test statistic $Z_1^{(\lambda)}$, in practice, especially when the break is of an unknown date.

Table 1: Maximum order of serial correlation

T	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$p_{\max}$	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8

Note that, in the case that disturbance tests u_{it} are normally distributed, variance function $V_1^{(\lambda)}$ can be written in a more analytic form as

$$V_1^{(\lambda)} = F_1^{(\lambda)\prime}(K_{T^2} + I_{T^2})(\Gamma_N \otimes \Gamma_N)F_1^{(\lambda)}, \tag{19}$$

where $\otimes$ denotes the Kronecker product.[6] This form of $V_1^{(\lambda)}$ can be easily calculated by replacing Γ_N with its consistent estimate $\hat{\Gamma}_N = \frac{1}{N}\sum_{i=1}^{N}(\Delta y_i \Delta y_i')$.

Test statistic $Z_1^{(\lambda)}$ can be easily extended to the case of an unknown break point date, which requires a sequential application of it. Define this test statistic as $z_1 \equiv \min\limits_{\lambda \in I} Z_1^{(\lambda)}$. Following analogous steps to those for sequential test statistic z, it can be proved that the limiting distribution of z_1 is given as

$$z_1 \equiv \min_{\lambda \in I} Z_1^{(\lambda)} \xrightarrow{d} \zeta_1 \equiv \min_{\lambda \in I} N(0, \Sigma_1), \tag{20}$$

$N \to \infty$, where $\Sigma_1 \equiv [\sigma_{1,\lambda s}]$ is the variance-covariance matrix of the test statistics $Z_1^{(\lambda)}$ whose elements, defined as $\sigma_{1,\lambda s}$, are given by the following formula:

$$\sigma_{1,\lambda s} = \frac{F_1^{(\lambda)\prime}\Theta F_1^{(s)}}{\sqrt{F_1^{(\lambda)\prime}\Theta F_1^{(\lambda)}}\sqrt{F_1^{(s)\prime}\Theta F_1^{(s)}}}. \tag{21}$$

Critical values of the distribution of random variable ζ_1, denoted as $f_{\zeta_1}(x_1)$ where $x_1 \in R$, can be calculated by replacing the values of $\sigma_{\lambda s}$ in pdf formula (12) with those of $\sigma_{1,\lambda s}$, given by (21). This also requires to

[6] This can be easily seen using standard results of the variance of a quadratic form for normally distributed variates (see e.g. Schott(1996)), which imply

$$Var[vec(u_i u_i')] = Var(u_i \otimes u_i) = (I_{T^2} + K_{T^2})(\Gamma_N \otimes \Gamma_N).$$

obtain consistent estimates of variance-covariance matrix Θ, in the first step.

4 Extension of the tests to the case of deterministic trends

In this Section, we will extend the tests presented in the previous section to allow for individual linear trends in the panel data generating processes, referred to as incidental trends. We will consider two cases of AR(1) panel data models with linear trends. In the first case, we will assume that these trends are present only under the alternative hypothesis of stationarity (see, e.g., Karavias and Tzavalis (2012), and Zivot and Andews (1992) for single time series), while in the second that they are present under the null hypothesis of $\varphi = 1$ either (see, e.g., Kim (2011). The first of the above cases is more appropriate in distinguishing between nonstationary panel data series which exhibit persistent random deviations from linear trends, implied by the presence of individual effects under $\varphi = 1$, and stationary panel data series allowing for broken individual linear trends under $\varphi < 1$. The second case is more suitable when considering more explosive panel data series under $\varphi = 1$, which can exhibit both deterministic and random persistent shifts from their linear trends.

4.1 Broken trends under the alternative hypothesis of stationarity

Consider the following extension of the nonlinear AR(1) model (1):

$$y_{it} = \alpha_{it}^{(\lambda)}(1-\varphi) + \varphi\beta_i + \beta_{it}^{(\lambda)}(1-\varphi)t + \varphi y_{it-1} + u_{it}, \quad i = 1, ..., N \tag{22}$$

where $\alpha_{it}^{(\lambda)}$ are defined by equation (1) and $\beta_{it}^{(\lambda)} = \beta_i^{(1)}$ if $t \leq T_0$ and $\beta_i^{(2)}$ if $t > T_0$. Under null hypothesis $\varphi = 1$, β_i constitute individual effects of the panel data model, which capture linear trends in the level of series y_{it}, for all i. Under alternative hypothesis $\varphi < 1$, β_i are defined as $\beta_i = \beta_i^{(1)}$ if $t \leq T_0$ and $\beta_i^{(2)}$ if $t > T_0$. That is, they constitute the slope coefficients of individual linear trends t, for all i.

Let us define matrix $X_*^{(\lambda)} = (e^{(1)}, e^{(2)}, \tau^{(1)}, \tau^{(2)})$, where $\tau^{(1)}$ and $\tau^{(2)}$ are $(TX1)$-column vectors whose elements are given as $\tau_t^{(1)} = t$ if $t \leq T_0$, and zero otherwise, and $\tau_t^{(2)} = t$ if $t > T_0$, and zero otherwise. Then, the "within" transformation matrix now will be written as $Q_*^{(\lambda)} = I_T - X_*^{(\lambda)}(X_*^{(\lambda)\prime}X_*^{(\lambda)})^{-1}X_*^{(\lambda)\prime}$ and the LSDV estimator, denoted as $\hat{\varphi}_*^{(\lambda)}$, can be written under null hypothesis $\varphi = 1$ as follows:

$$\hat{\varphi}_{*}^{(\lambda)} - 1 = \left[\sum_{i=1}^{N} y'_{i,-1} Q_{*}^{(\lambda)} y_{i,-1}\right]^{-1} \left[\sum_{i=1}^{N} y'_{i,-1} Q_{*}^{(\lambda)} u_i\right]. \tag{23}$$

Following analogous steps to those for the derivation of test statistics $Z^{(\lambda)}$ and $Z_1^{(\lambda)}$, inference about unit roots can be conducted based on estimator $\hat{\varphi}_{*}^{(\lambda)}$ adjusted for its inconsistency. Under conditions of Assumption 2, this inconsistency is given as

$$\frac{b_2^{(\lambda)}}{\delta_2^{(\lambda)}} = \frac{tr(\Lambda' Q_{*}^{(\lambda)} \Gamma_N)}{tr(\Lambda' Q_{*}^{(\lambda)} \Lambda \Gamma_N)} \tag{24}$$

(see Appendix, proof of Theorem 5). However, in contrast to the case of model (1), the average population variance-autocovariance matrix Γ_N can not be consistently estimated based on estimator $\hat{\Gamma}_N = \frac{1}{N}\sum_{i=1}^{N}(\Delta y_i \Delta y'_i)$, due to the presence of individual effects β_i under null hypothesis $\varphi = 1$. It can be easily seen that, under $\varphi = 1$, $\Delta y_i = u_i + \beta_i e$, where e is a $(TX1)$-vector of unities, and thus

$$\frac{1}{N}\sum_{i=1}^{N} E(\Delta y_i \Delta y'_i) = \Gamma_N + \beta_N^2 J_T, \tag{25}$$

where J_T is a $T \times T$ matrix of ones and $\beta_N^2 = \frac{1}{N}\sum_{i=1}^{N} E((\beta_i)^2)$. The last relationship clearly shows that in order to provide consistent estimates of matrix Γ_N based on estimator $\hat{\Gamma}_N = \frac{1}{N}\sum_{i=1}^{N}(\Delta y_i \Delta y'_i)$, we need to substitute out the average of squared individual effects β_N^2 entering this estimator. This can be done with the help of a (TXT)-dimension selection matrix M, defined as follows: M has elements $m_{ts} = 0$ if $\gamma_{ts} \neq 0$ and $m_{ts} = 1$ if $\gamma_{ts} = 0$. That is, the elements of M correspond to those of matrix $\Gamma_N + \beta_N^2 J_T$ (or $\frac{1}{N}\sum_{i=1}^{N} E(\Delta y_i \Delta y'_i)$ which contain only β_N^2. Based on matrix M, which implies since $tr(M\Gamma_N) = 0$, we can derive a consistent estimator of β_N^2 under null hypothesis $\varphi = 1$, given as

$$\frac{1}{tr(MJ_T)N}\sum_{i=1}^{N} \Delta y'_i M \Delta y_i \xrightarrow{p} \beta_N^2, \tag{26}$$

where "$\xrightarrow{p}$" signifies convergence in probability. Given this estimator, we can derive a consistent estimator of the inconsistency of the LSDV estimator $\hat{\varphi}_{*}^{(\lambda)}$ for model (22), defined as $\frac{b_2^{(\lambda)}}{\delta_2^{(\lambda)}}$, as

$$\frac{\hat{b}_2^{(\lambda)}}{\hat{\delta}_2^{(\lambda)}} = \frac{tr(\Psi_2^{(\lambda)} \hat{\Gamma}_N)}{\frac{1}{N}\sum_{i=1}^{N} y'_{i,-1} Q_{*}^{(\lambda)} y_{i,-1}}, \tag{27}$$

where $\Psi_2^{(\lambda)} = \Psi_{1*}^{(\lambda)} + \frac{tr(\Lambda' Q_*^{(\lambda)} M)}{trace(MJ_T)} M$ (see Appendix, proof of Theorem 5), $\Psi_{1*}^{(\lambda)}$ is a (TXT)-dimension matrix having in its main diagonal, and its p-lower and $p-$upper diagonals of the main diagonal the corresponding elements of matrix $\Lambda' Q_*^{(\lambda)}$, and zero otherwise. It can be easily seen that $tr(\Psi_2^{(\lambda)} \hat{\Gamma}_N)$ constitutes a consistent estimator of $\hat{b}_2^{(\lambda)}$, since $tr(\Psi_2^{(\lambda)}(\Gamma_N + \beta_N^{(1)} J_T)) = tr(\Psi_2^{(\lambda)} \Gamma_N)$.

Having derived a consistent estimator of the asymptotic bias of LS estimator $\hat{\varphi}_*^{(\lambda)}$ under null hypothesis $\varphi = 1$ net of the individual effects β_i, next we derive the limiting distribution of a normalized test statistic based on this estimator adjusted for its inconsistency. This is done after trimming out two time series observations from the end of the sample, i.e. $\lambda = \frac{T_0}{T} \in I^* = \left\{\frac{2}{T}, \frac{2}{T}, ..., \frac{T-2}{T}\right\}$, due to the presence of individual effects and linear trends under alternative hypothesis of $\varphi = 1$. To derive this limiting distribution and to prove the consistency of the suggest test statistic, we rely the following assumption.

Assumption 3: Let all conditions of Assumption 2 hold and we also have: $E(u_{it}\beta_i) = 0$, $\forall\ i \in \{1, 2, ..., N\}$, $t \in \{1, 2, ..., T\}$, $E(a_{it}^{(\lambda)} \beta_{it}^{(\lambda)}) = 0$, $\forall\ i \in \{1, 2, ..., N\}$.

Theorem 5 *Let the sequence $\{y_{i,t}\}$ be generated according to model (22) and conditions (b1)-(b4) of Assumption 2 hold. Then, under the null hypothesis $\varphi = 1$ and λ known, we have*

$$Z_2^{(\lambda)} \equiv \hat{V}_2^{(\lambda)-0.5} \hat{\delta}_2^{(\lambda)} \sqrt{N} \left(\hat{\varphi}_*^{(\lambda)} - 1 - \frac{\hat{b}_2^{(\lambda)}}{\hat{\delta}_2^{(\lambda)}} \right) \xrightarrow{d} N(0, 1), \tag{28}$$

as $N \to \infty$, where $V_2^{(\lambda)} = F_2^{(\lambda)\prime} \Theta F_2^{(\lambda)}$, Θ is defined in Theorem 4, and $F_2^{(\lambda)} = vec(Q_*^{(\lambda)} \Lambda - \Psi_2^{(\lambda)\prime})$.

Apart from the initial conditions of the panel y_{i0}, the test statistic given by Theorem 5, defined $Z_2^{(\lambda)}$, is similar under null hypothesis $\varphi = 1$ to individual effects of the panel β_i, due to the allowance of broken trends in the "within" transformation matrix, $Q_*^{(\lambda)}$. To test the null hypothesis of unit roots, test statistic $Z_2^{(\lambda)}$ relies on the same moments to those assumed by statistic $Z_1^{(\lambda)}$, namely $E(u_{it}u_{is}) = 0$, for $s = t + p_{\max} + 1, ..., T$ and $t < s$. These moments now are weighted by elements of matrix $\Lambda' Q_*^{(\lambda)} - \Psi_2^{(\lambda)}$, where matrix $\Psi_2^{(\lambda)}$ is appropriately adjusted to wipe off the effects of nuisance parameters β_i on the limiting distribution of the test statistic. The maximum order of serial correlation of variance-autocovariance matrices Γ_i assumed by test statistic $Z_2^{(\lambda)}$ is the same to that assumed by test statistic $Z_1^{(\lambda)}$.

Finally, note that test statistic $Z_2^{(\lambda)}$ can be extended to the case of an unknown date break point, following an analogous procedure to that assumed for sequential tests statistics z and z_1, defined by equations (9) and

(20), respectively. This version of the test statistic is defined as $z_2 \equiv \min_{\lambda \in I^*} Z_2^{(\lambda)}$. Its limiting distribution is given as

$$z_2 \equiv \min_{\lambda \in I^*} Z_2^{(\lambda)} \xrightarrow{d} \zeta_2 \equiv \min_{\lambda \in I^*} N(0, \Sigma_2), \tag{29}$$

as $N \to \infty$, where $\Sigma_2 \equiv [\sigma_{2,\lambda s}]$ is the variance-covariance matrix of test statistics $Z_2^{(\lambda)}$ whose elements $\sigma_{2,\lambda s}$ are given by the following formula: $\sigma_{2,\lambda s} = \frac{F_2^{(\lambda)\prime} \Theta F_2^{(s)}}{\sqrt{F_2^{(\lambda)\prime} \Theta F_2^{(\lambda)}} \sqrt{F_2^{(s)\prime} \Theta F_2^{(s)}}}$. Critical values of the distribution of ζ_2 can be derived based on pdf $f_\zeta(x)$, given by (12), following analogous steps to those for test statistic z_1.

4.2 Broken trends under the null hypothesis of unit roots

To allow for a common break in the individual effects of the panel data model under the null hypothesis of $\varphi = 1$, consider the following extension of $AR(1)$ model (1):

$$y_{it} = \alpha_{it}^{(\lambda)}(1-\varphi) + \varphi \beta_{it}^{(\lambda)} + \beta_{it}^{(\lambda)}(1-\varphi)t + \varphi y_{it-1} + u_{it}, \quad i = 1, ..., N \tag{30}$$

Using vector notation, this model implies that, under hypothesis $\varphi = 1$, the first-difference of vector y_i is given as $\Delta y_i = \beta_i^{(1)} e^{(1)} + \beta_i^{(2)} e^{(2)} + u_i$. As for model (22), this means that estimator $\hat{\Gamma}_N = \frac{1}{N}\sum_{i=1}^{N}(\Delta y_i \Delta y_i^\prime)$ will not lead to consistent estimates of the average population variance-autocovariance matrix Γ_N, due to the presence of individual effects $\beta_i^{(1)}$ and $\beta_i^{(2)}$. These imply

$$\frac{1}{N}\sum_{i=1}^{N} E(\Delta y_i \Delta y_i^\prime) = \beta_N^{(1)} e^{(1)} e^{(1)\prime} + \beta_N^{(2)} e^{(2)} e^{(2)\prime} + \Gamma_N, \tag{31}$$

where $J_1 = e^{(1)} e^{(1)\prime}$ and $J_2 = e^{(2)} e^{(2)\prime}$. The allowance of a break in incidental parameters β_i under null hypothesis $\varphi = 1$ requires estimation of squared individual effects $\beta_N^{(1)}$ and $\beta_N^{(2)}$ in order to obtain consistent estimates of matrix Γ_N, net of these effects. To this end, we will adopt an analogous procedure to that following in the previous subsection, based on selection matrix M. We will define two (TXT)-dimension block diagonal selection matrices $M^{(1)}$ and $M^{(2)}$, which select square individual effects $\beta_N^{(1)}$ and $\beta_N^{(2)}$, respectively. The elements of matrix $M^{(1)}$ are defined as $m_{ts}^{(1)} = 0$ if $\gamma_{ts} \neq 0$, and $m_{ts}^{(1)} = 1$ if $\gamma_{ts} = 0$. That is, matrix $M^{(1)}$ selects the elements of matrix $\beta_N^{(1)} e^{(1)} e^{(1)\prime} + \beta_N^{(2)} e^{(2)} e^{(2)\prime} + \Gamma_N$ consisting only of effects $\beta_N^{(1)}$, for $t, s \leq T_0$. For t or $s > T_0$, all elements of $M^{(1)}$ are set to $m_{ts}^{(1)} = 0$. On the other hand, the elements of matrix $M^{(2)}$ are defined as $m_{ts}^{(2)} = 0$ if $\gamma_{ts} \neq 0$, and $m_{2ts} = 1$ if $\gamma_{ts} = 0$. Thus, $M^{(2)}$ selects the elements of matrix

$\beta_N^{(1)} e^{(1)} e^{(1)\prime} + \beta_N^{(2)} e^{(2)} e^{(2)\prime} + \Gamma_N$ consisting only of effects $\beta_N^{(2)}$, for $t, s > T_0$. For t or $s \leq T_0$, all the elements of $M^{(2)}$ are set to $m_{ts}^{(2)} = 0$.

Based on the above definitions of matrices $M^{(1)}$ and $M^{(2)}$, we can obtain the following two consistent estimators of $\beta_N^{(1)}$ and $\beta_N^{(2)}$:

$$\frac{1}{tr(M^{(1)}J_1)N} \sum_{i=1}^{N} \Delta y_i' M^{(1)} \Delta y_i \xrightarrow{p} \beta_N^{(1)} \quad \text{and} \quad \frac{1}{tr(M^{(2)}J_2)N} \sum_{i=1}^{N} \Delta y_i' M^{(2)} \Delta y_i \xrightarrow{p} \beta_N^{(2)}, \tag{32}$$

respectively, since $tr(M^{(j)}\Gamma_N) = 0$ for $j = 1, 2$ and $tr(M^{(j)}J_r) = 0$ for $j, r = 1, 2$ and $j \neq r$. These estimators can be employed to obtain consistent estimates of matrix Γ_N, which are net of square individual effects $\beta_N^{(1)}$ and $\beta_N^{(2)}$. Then, a consistent estimator of the bias of the LSDV estimator $\hat{\varphi}_*^{(\lambda)}$ for model (30), defined as $\frac{b_3^{(\lambda)}}{\delta_3^{(\lambda)}}$, can be obtained as

$$\frac{\hat{b}_3^{(\lambda)}}{\hat{\delta}_3^{(\lambda)}} = \frac{tr(\Psi_3^{(\lambda)} \hat{\Gamma}_N)}{\frac{1}{N} \sum_{i=1}^{N} y_{i,-1}' Q_*^{(\lambda)} y_{i,-1}}, \tag{33}$$

where $\Psi_3^{(\lambda)} = \Psi_{1*}^{(\lambda)} + \frac{tr(\Lambda' Q_*^{(\lambda)} M^{(1)})}{trace(M^{(1)}J_1)} M^{(1)} + \frac{tr(\Lambda' Q_*^{(\lambda)} M^{(2)})}{trace(M^{(2)}J_2)} M^{(2)}$. Adjusting $\hat{\varphi}_*^{(\lambda)}$ by the above estimator of its bias will lead to a panel unit root test statistic whose limiting distribution is be net of squared individual effects $\beta_N^{(1)}$ and $\beta_N^{(2)}$, under null hypothesis $\varphi = 1$. In the next theorem, we derive the limiting distribution of this test statistic under the assumption of a known date break. If break point T_0 is unknown, then this test statistic will rely on a consistent estimate of T_0, in a first step. This can be done based on the first differences of the individual panel data series y_{it} under null hypothesis $\varphi = 1$, i.e. $\Delta y_i = \beta_i^{(1)} e^{(1)} + \beta_i^{(2)} e^{(2)} + u_i$. As shown by Bai (2010), this estimator provides consistent estimates of T_0, which converges at $o(\sqrt{N})$ rate.

Theorem 6 *Let the sequence $\{y_{i,t}\}$ be generated according to model (30) and conditions (b1)-(b4) of Assumption 2 hold. Then, under the null hypothesis $\varphi = 1$ and λ known, we have*

$$Z_3^{(\lambda)} \equiv \hat{V}_3^{(\lambda)-0.5} \hat{\delta}_3^{(\lambda)} \sqrt{N} \left(\hat{\varphi}_*^{(\lambda)} - 1 - \frac{\hat{b}_3^{(\lambda)}}{\hat{\delta}_3^{(\lambda)}} \right) \xrightarrow{d} N(0, 1), \tag{34}$$

as $N \to \infty$, where

$$V_3^{(\lambda)} = F_3^{(\lambda)\prime} \Theta F_3^{(\lambda)} \tag{35}$$

and $F_3^{(\lambda)} = vec(Q_*^{(\lambda)} \Lambda - \Psi_3^{(\lambda)\prime})$. The proof of the theorem is given in the appendix.

As $Z_2^{(\lambda)}$, the test statistic given by Theorem 6, $Z_3^{(\lambda)}$, is similar under null hypothesis $\varphi = 1$ to individual effects $\beta_i^{(1)}$ and $\beta_i^{(2)}$, due to the inclusion of broken trends in the "within" transformation matrix $Q_*^{(\lambda)}$. Due to the presence of a break under $\varphi = 1$, the maximum order of serial correlation of the disturbance terms u_{it}, $p_{\max}$, allowed by statistic $Z_3^{(\lambda)}$ is not given by Table 1. This is given as[7]

$$p_{\max} = \begin{cases} \frac{T}{2} - 3, & \text{if } T \text{ is even and } T_0 = \frac{T}{2} \\ \min\{T_0 - 2, T - T_0 - 2\} & \text{in all other cases of } T \text{ or } T_0 \end{cases} \tag{36}$$

Based on conditions of Assumption 3, it can be proved that test statistic $Z_3^{(\lambda)}$ is consistent, following analogous steps to those for the proof of the consistency of test statistic $Z_2^{(\lambda)}$. The test is also consistent, if the break point is unknown and is estimated, in the first step, based on the procedure mentioned above.

5 Simulation Results

In this section, we conduct a Monte Carlo study to investigate the small sample performance of the test statistics suggested in the previous sections. For reasons of space, in our study, we consider only the case that the break date is unknown.[8] We consider experiments of different sample sizes of N and T, i.e. $N = \{50, 100, 200\}$ and $T = \{6, 10, 15\}$, while the fractions of sample that the break occurs are given by the following set: $\lambda = \{0.25, 0.5, 0.75\}$. These value of λ are chosen to facilitate implementation of the test statistics. For all experiments, we conduct 10000 iterations. In each iteration, we assume that the data generating processes are given by models (1) and (22), respectively, where disturbance terms u_{it} follow a MA(1) process, i.e. $u_{it} = \varepsilon_{it} + \theta\varepsilon_{it-1}$, with $\varepsilon_{it}\tilde{}NIID(0,1)$, for all i and t, and $\theta = \{-0.5, 0.0, 0.5\}$. The values of the nuisance parameters of the simulated models, namely the individual effects or the slope coefficients of individual linear trends are assumed that they are driven from the following distributions: $\alpha_i^{(1)} \sim U(-0.5, 0)$, $\alpha_i^{(2)} \sim U(0, 0.5)$, $\beta_i \sim U(0, 0.05)$, $\beta_i^{(1)} \sim U(0, 0.025)$, $\beta_i^{(2)} \sim U(0.025, 0.05)$, where $U(\cdot)$ stands for the uniform distribution, and and $y_{i0} = 0$, for all i .

[7] Again, $p_{\max}$ is chosen so as variance function $V_3^{(\lambda)}$ is different than zero. If T is even, then $p_{\max}$=min$\{T_0 - 2, T - T_0 - 2\}$, with the exception the case that $T_0 = \frac{T}{2}$ where $p_{\max}$=$\frac{T}{2} - 3$. Consider the following examples. First, $T = 10$ and $T_0 = 3$, then we have that $p_{\max} = \min\{T_0 - 2, T - T_0 - 2\} = \min\{1, 5\} = 1$. If $T_0 = \frac{T}{2} = 5$, then $p_{\max}$ becomes $p_{\max} = \frac{T}{2} - 3 = 2$. Note that, instead of the above, if we used the results of (18) to determine $p_{\max}$, implying $p_{\max} = \min\{T_0 - 2, T - T_0 - 2\} = \min\{3, 3\} = 3$, then $Z_3^{(\lambda)}$ could not be applied since $V_3^{(\lambda)} = 0$. If $T = 15$, then $p_{\max}$ becomes $p_{\max}$=min$\{T_0 - 2, T - T_0 - 2\}$. For $T_0 = 7$, this becomes $p_{\max} = \min\{5, 6\} = 5$.

[8] The results of the test statistics allowing for a known date break point are analogous. These are available upon request.

The small magnitude of individual effects $\alpha_i^{(j)}$ or slope coefficients $\beta_i^{(j)}$ assumed above correspond to evidence found in the empirical literature about them, see e.g. Hall and Mairesse (2005). The small magnitude of these effects makes our tests hard to distinguish null hypothesis of $\varphi = 1$ from its alternative of stationarity. For all simulation experiments, we assume that the order of serial correlation p is set to $p = 1$. This means that, for $\theta = 0$, we assume an order of serial correlation which is higher than the correct order. This experiment will show if the performance of our tests critically reduces when a higher order of serial correlation is assumed, which may happen in practice.

The results of our Monte Carlo analysis for test statistics z_1 and z_2, corresponding to models (1) and (22), are summarized in Tables 2(a)-(c) and 3(a)-(c), respectively. Table 4.presents results for test statistic $Z_3^{(\lambda)}$, which is based on model (30) considering a break in individual effects β_i under the null hypothesis. To implement $Z_3^{(\lambda)}$, the break point T_0 is treated as known and it is estimated, in a first step. Note that this table reports results for $\lambda = 0.5$ and $T = \{10, 15\}$, since for these cases of T and λ we can assume maximum order serial correlation $p_{\max} = 1$, according to equation (36). The above all tables present values of the size and power of statistics z_1, z_2 and $Z_3^{(\lambda)}$, for $\theta \in \{0.5, -0.5, 0.0\}$. The size of the test statistics is calculated for $\varphi = 1.00$, while the power for $\varphi \in \{0.95, 0.90\}$. Note that, in all experiments, the power is calculated at the nominal 5% significance level of the distribution of the tests.

The results of Tables 2(a)-(c), 3(a)-(c) and 4 indicate that the test statistics examined, namely z_1, z_2 and $Z_3^{(\lambda)}$, have size which is close to the nominal level 5% considered. This is true for all combinations of N and T considered. The size performance of all three test statistics is close to its nominal level. This is true even if the MA parameter θ takes a large negative value, i.e. $\theta = -0.5$. Note that, for this case of θ, single time series unit root tests are critically oversized (see, e.g., Schwert (1989)). The size of the test statistics improves as N increases relative to T. This can be attributed to the fact that, as N increases relative to T, variance-covariance matrix Θ is more precisely estimated by estimator $\hat{\Theta}$. The above results hold independently on the break fraction of the sample λ. As a final note that the size of al the test statistics does not deteriorate, if a higher order of serial correlation $p = 1$ is assumed than the true order, for $\theta = 0$. This result qualifies application of them in cases where a higher than the correct order p of serial correlation of the disturbance terms u_{it} is assumed in practice, i.e. $p = \{2, 3\}$, which may be considered as very high for short panels.[9]

[9] This has been confirmed by our Monte Carlo simulation analysis. These results are not reported for reasons of space. They

Regarding the power of the test statistics, the results of the tables indicate that, as was expected, the test statistic that has the highest power among all of them is that which corresponds to model (1), i.e. z_1, which allows for individual effects under the alternative hypothesis of stationarity. For models (22) and (30), where linear trends are considered either under alternative or null hypotheses, respectively, the power of the test statistics (i.e., z_2 and $Z_3^{(\lambda)}$) substantially reduces. This is feature of both single time series and panel data unit root tests allowing for linear trends. As is well known in the literature, the second category of tests have better power performance than the first (see, e.g., Harris and Tzavalis (2004), or Hluskova and Wagner (2006)). However, between test statistics z_2 and $Z_3^{(\lambda)}$, it is found that the first has clearly better power than the second. This is true for all cases of T and N considered. The less power of statistic $Z_3^{(\lambda)}$ than z_2 may be attributed to the fact that this test statistic relies on estimation of the break point under $\varphi = 1$. Despite the fact that the break point is estimated very accurately under the null hypothesis, $Z_3^{(\lambda)}$ depends on the nuisance parameters of the sample distribution of the estimator of the break point T_0, which may lead to a reduction of its power. Finally, note that, in contrast to the size, the power of all three test statistics examined increases faster with T rather than N. Consistently with the theory, the power of the test statistics increases also as the value of φ moves away from unity.

6 Conclusion

This paper suggests panel unit root test statistics which allow for a common structural break in the individual effects or linear trends of dynamic panel data models. Common breaks in panel data can arise in cases of a credit crunch, an oil price shock or a change in tax policy among others. The suggested test statistics assume that the time-dimension of the panel T is fixed (or finite), while the cross-section N grows large. Thus, they are appropriate for short panel applications, where T is smaller than N. Since they are based on the least squares dummy variable (LSDV) estimator of the autoregressive coefficient of the dynamic panel data model with individual effects and/or linear trends, the suggested test statistics are invariant to the initial conditions of the panel or the individual effects under the null hypothesis of unit roots. This property of the tests does not restrict their application to panel data where conditions of mean or covariance stationarity of the initial conditions or individual effects are required. To allow for serial correlation, the

are available upon request.

tests rely on the LSDV estimator which is also corrected for its inconsistency due to a high order serial correlation in the disturbance terms. This is done based on moments of the disturbance terms which are not serially correlated.

The paper derives the limiting distributions of the test statistics. When the break is unknown, it shows that the limiting distribution of the tests is calculated as the minimum of a fixed number of correlated normals. This distribution is given, analytically, as a mixture of normals. Knowledge of the analytic form of the limiting distribution of the tests considerably facilitates calculation of critical values for the implementation of the tests in practice. To examine the small sample size and power performance of the test statistics, the paper conducts a Monte Carlo study. This is done for the case that the break is of an unknown date. The results of this exercise indicate that when there is no break under the null, the tests have the correct nominal size and power which is bigger than their size. The size and power performance of the test statistics does not depend on the fraction of the sample that the break occurs. As was expected, the power of the tests is higher for the dynamic panel data models which consider individual effects rather than for the model which also allows for individual linear trends. For all cases, the power is found to increase as N and T increases.

7 Appendix

In this appendix, we provide proofs of the theorems presented in the main text of the paper.

Proof of Theorem 1: To derive the limiting distribution of the test statistic of the theorem, we will proceed into stages. We first show that the LSDV estimator $\hat{\varphi}^{(\lambda)}$ is inconsistent, as $N \to \infty$. Then, will construct a normalized statistic based on $\hat{\varphi}^{(\lambda)}$ corrected for its inconsistency (asymptotic bias) and derive its limiting distribution under the null hypothesis of $\varphi = 1$, as $N \to \infty$.

Decompose the vector $y_{i,-1}$ for model (1) under hypothesis $\varphi = 1$ as

$$y_{i,-1} = ey_{i0} + \Lambda u_i, \tag{37}$$

where the matrix Λ is is a (TXT) matrix defined as $\Lambda_{r,c} = 1$, if $r > c$ and 0 otherwise.

Premultiplying (37) with matrix $Q^{(\lambda)}$ yields

$$Q^{(\lambda)} y_{i,-1} = Q^{(\lambda)} \Lambda u_i, \tag{38}$$

since $Q^{(\lambda)} e = (0, 0, ..., 0)'$. Substituting (38) into (2) yields

$$\hat{\varphi}^{(\lambda)} - 1 = \frac{\frac{1}{N}\sum_{i=1}^{N} y'_{i,-1} Q^{(\lambda)} u_i}{\frac{1}{N}\sum_{i=1}^{N} y'_{i,-1} Q^{(\lambda)} y_{i,-1}} = \frac{\frac{1}{N}\sum_{i=1}^{N} u'_i \Lambda' Q^{(\lambda)} u_i}{\frac{1}{N}\sum_{i=1}^{N} u'_i \Lambda' Q^{(\lambda)} \Lambda u_i}. \tag{39}$$

By Kitchin's Weak Law of Large Numbers (KWLLN), we have

$$\frac{1}{N}\sum_{i=1}^{N} u'_i \Lambda' Q^{(\lambda)} u_i \xrightarrow{p} b^{(\lambda)} = \sigma_u^2 tr(\Lambda' Q^{(\lambda)}) \text{ and } \frac{1}{N}\sum_{i=1}^{N} u'_i \Lambda' Q^{(\lambda)} \Lambda u_i \xrightarrow{p} \delta^{(\lambda)} = \sigma_u^2 tr(\Lambda' Q^{(\lambda)} \Lambda), \tag{40}$$

where "$\xrightarrow{p}$" signifies convergence in probability. Using the last results, the yet non standardized statistic $Z^{(\lambda)}$ can be written by (39) as

$$\begin{aligned} \sqrt{N}\hat{\delta}^{(\lambda)} \left(\hat{\varphi}^{(\lambda)} - 1 - \frac{\hat{b}^{(\lambda)}}{\hat{\delta}^{(\lambda)}} \right) &= \sqrt{N}\hat{\delta}^{(\lambda)} \left(\frac{\frac{1}{N}\sum_{i=1}^{N} y'_{i,-1} Q^{(\lambda)} u_i}{\hat{\delta}^{(\lambda)}} - \frac{\hat{\sigma}_u^2 tr(\Lambda' Q^{(\lambda)})}{\hat{\delta}^{(\lambda)}} \right) \\ &= \sqrt{N} \left(\frac{1}{N}\sum_{i=1}^{N} y'_{i,-1} Q^{(\lambda)} u_i - tr(\Lambda' Q^{(\lambda)}) \frac{\sum_{i=1}^{N} \Delta y'_i \Psi^{(\lambda)} \Delta y_i}{N tr(\Psi^{(\lambda)})} \right). \end{aligned} \tag{41}$$

Since, under the null hypothesis $\varphi = 1$, we have $u_i = \Delta y_i$, (l41) can be written as follows:

$$\begin{aligned} & \sqrt{N}\hat{\delta}^{(\lambda)} \left(\hat{\varphi}^{(\lambda)} - 1 - \frac{\hat{b}^{(\lambda)}}{\hat{\delta}^{(\lambda)}} \right) \\ = & \sqrt{N} \left(\frac{1}{N}\sum_{i=1}^{N} u'_i \Lambda' Q^{(\lambda)} u_i - \frac{tr(\Lambda' Q^{(\lambda)})}{tr(\Psi^{(\lambda)})} \frac{1}{N}\sum_{i=1}^{N} u'_i \Psi^{(\lambda)} u_i \right) \\ = & \frac{1}{\sqrt{N}}\sum_{i=1}^{N} u'_i (\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)}) u_i = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} tr\left[(\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)}) u_i u'_i \right] \\ = & \frac{1}{\sqrt{N}}\sum_{i=1}^{N} W_i^{(\lambda)}, \end{aligned} \tag{42}$$

where $W_i^{(\lambda)}$ constitute random variables with mean

$$\begin{aligned} E(W_i^{(\lambda)}) &= E[u_i'(\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)})u_i] = tr[(\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)})E(u_i u_i')] \\ &= \sigma_u^2 tr(\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)}) = 0, \text{ for all } i, \end{aligned}$$

since $tr(\Lambda' Q^{(\lambda)}) = tr(\Psi^{(\lambda)})$ (or $tr(\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)}) = 0$), and variance

$$\begin{aligned} Var(W_i^{(\lambda)}) &= Var(u_i'(\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)})u_i) = Var[F^{(\lambda)\prime} vec(u_i u_i')] = \\ &= F^{(\lambda)} Var[vec(u_i u_i')] F^{(\lambda)\prime}, \text{ for all } i. \end{aligned}$$

The results of Theorem 1 follows by applying Lindeberg-Levy central limit theorem (CLT) to the sequence of IID random variables $W_i^{(\lambda)}$. Following standard linear algebra results (see e.g. Schott(1997), variance $Var[vec(u_i u_i')]$ can be analytically written as $Var[vec(u_i u_i')] = Var(u_i \otimes u_i) = \sigma_u^4 (I_{T^2} + K_{T^2})$, where $\otimes$ denotes the Kroenecker product.

Proof of Theorem 2: Assume that the break point T_0 is known. Define vector $w = (1, \varphi, \varphi^2, ..., \varphi^{T-1})'$ and matrix

$$\Omega = \begin{pmatrix} 0 & . & . & . & . & . & 0 \\ 1 & 0 & & & & & . \\ \varphi & 1 & . & & & & . \\ \varphi^2 & \varphi & . & . & & & . \\ . & & . & . & . & & . \\ . & & & . & 1 & 0 & . \\ \varphi^{T-2} & \varphi^{T-3} & . & . & \varphi & 1 & 0 \end{pmatrix}$$

Under null hypothesis $\varphi = 1$, we have $\Omega = \Lambda$. Based on the above definitions of w and Ω, vector $y_{i,-1}$ can be written as

$$y_{i,-1} = w y_{i0} + \Omega X^{(\lambda)} \gamma_i^{(\lambda)} + \Omega u_i, \tag{43}$$

where $d_i^{(\lambda)} = (a_i^{(1)}(1-\varphi), a_i^{(2)}(1-\varphi))'$. Using last expression of $y_{i,-1}$, test statistic $Z^{(\lambda)}$ can be written under alternative hypothesis $\varphi < 1$ as follows:

$$\begin{aligned}
Z^{(\lambda)} &= \sqrt{N}\hat{V}^{(\lambda)-1/2}\hat{\delta}^{(\lambda)}\left(\hat{\varphi}^{(\lambda)} - 1 - \frac{\hat{b}^{(\lambda)}}{\hat{\delta}^{(\lambda)}}\right) \\
&= \sqrt{N}\hat{V}^{(\lambda)-1/2}\hat{\delta}^{(\lambda)}\left(\varphi + \frac{\frac{1}{N}\sum_{i=1}^{N} y'_{i,-1}Q^{(\lambda)}u_i}{\frac{1}{N}\sum_{i=1}^{N} y'_{i,-1}Q^{(\lambda)}y_{i,-1}} - 1 - \frac{\hat{\sigma}_u^2 tr(\Lambda' Q^{(\lambda)})}{\frac{1}{N}\sum_{i=1}^{N} y'_{i,-1}Q^{(\lambda)}y_{i,-1}}\right) \\
&= \sqrt{N}\hat{V}^{(\lambda)-1/2}\hat{\delta}^{(\lambda)}(\varphi - 1) + \sqrt{N}\hat{V}^{(\lambda)-1/2}\left(\frac{1}{N}\sum_{i=1}^{N} y'_{i,-1}Q^{(\lambda)}u_i - \hat{\sigma}_u^2 tr(\Lambda' Q^{(\lambda)})\right) \qquad (44) \\
&= \underbrace{\left\{\sqrt{N}\hat{V}^{(\lambda)-1/2}\hat{\delta}^{(\lambda)}(\varphi - 1)\right\}}_{(I)} + \underbrace{\left\{\hat{V}^{(\lambda)-1/2}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(y'_{i,-1}Q^{(\lambda)}u_i - \Delta y'_i \Psi^{(\lambda)}\Delta y_i)\right\}}_{(II)}.
\end{aligned}$$

Next, we will show that summand (I) diverges to $-\infty$ and summand (II) is bounded in probability. These two results imply that, as $N \to \infty$, test statistic $Z^{(\lambda)}$ converges to $-\infty$, which proves its consistency. To prove the above results, we will use the following identities:

$$u_i = y_i - \varphi y_{i-1} - X^{(\lambda)}d_i^{(\lambda)} \qquad (45)$$

and

$$\Delta y_i = u_i + (\varphi - 1)y_{i-1} + X^{(\lambda)}d_i^{(\lambda)}, \qquad (46)$$

which hold under alternative hypothesis $\varphi < 1$.

To prove that summand (I), defined by (44), diverges to $-\infty$, it is sufficient to show that $p\lim \hat{\delta}^{(\lambda)}$ is $O_p(1)$ and positive, and $p\lim \hat{\sigma}_u^2 = O_p(1)$ and nonzero. The last result implies that variance function $\hat{V}^{(\lambda)} = \hat{\sigma}_u^4 F^{(\lambda)\prime}(K_{T^2} + I_{T^2})F^{(\lambda)}$ is bounded in probability. Using equations (43), (45) and (46), it can be seen that $\hat{\delta}^{(\lambda)}$ is $O_p(1)$ as follows:

$$\begin{aligned}
\hat{\delta}^{(\lambda)} &= \frac{1}{N}\sum_{i=1}^{N} y'_{i,-1}Q^{(\lambda)}y_{i,-1} = \frac{1}{N}\sum_{i=1}^{N}(wy_{i0} + \Omega X^{(\lambda)}d_i^{(\lambda)} + \Omega u_i)'Q^{(\lambda)}(wy_{i0} + \Omega X^{(\lambda)}d_i^{(\lambda)} + \Omega u_i) \quad (47) \\
&= \frac{1}{N}\sum_{i=1}^{N}(y_{i0}^2 w'Q^{(\lambda)}w + y_{i0}w'Q^{(\lambda)}\Omega X^{(\lambda)}d_i^{(\lambda)} + y_{i0}wQ^{(\lambda)}\Omega u_i + ... + u'_i\Omega' Q^{(\lambda)}\Omega u_i \\
&\xrightarrow{p} E(y_{i0}^2)w'Q^{(\lambda)}w + tr(X^{(\lambda)\prime}\Omega' Q^{(\lambda)}\Omega X^{(\lambda)}\Sigma_d) + \sigma_u^2 tr(\Omega' Q^{(\lambda)}\Omega) = O_p(1),
\end{aligned}$$

where $\Sigma_d = E(d_i^{(\lambda)}d_i^{(\lambda)\prime})$. The last result holds by condition a3 of Assumption 1. All quantities involved in the above limit are positive, since they are either variances or quadratic forms. Based on condition a3 of Assumption 1, we can also show that the following result also holds:

$$\begin{aligned}
\hat{\sigma}_u^2 &= \frac{1}{tr(\Psi^{(\lambda)})}\frac{1}{N}\sum_{i=1}^{N}\Delta y_i'\Psi^{(\lambda)}\Delta y_i \qquad (48)\\
&= \frac{1}{tr(\Psi^{(\lambda)})}\frac{1}{N}\sum_{i=1}^{N}(u_i+(\varphi-1)y_{i-1}+X^{(\lambda)}d_i^{(\lambda)})'\Psi^{(\lambda)}(u_i+(\varphi-1)y_{i-1}+X^{(\lambda)}d_i^{(\lambda)})\\
&= O_p(1).
\end{aligned}$$

This limit is a nonzero quantity, since $\sigma_u^2 > 0$. The remaining terms entered into this limit are zero or positive quantities.

To prove that summand (II) is bounded in probability note that, by Assumption 1, we have

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(y_{i,-1}'Q^{(\lambda)}u_i-\Delta y_i'\Psi^{(\lambda)}\Delta y_i)=O_p(1). \qquad (49)$$

See also proof of Theorem 1.

Proof of Theorem 3: The proof of this theorem follows as an extension of Theorem 1, by applying the continuous mapping theorem to the joint limiting distribution of standardized test statistic $Z^{(\lambda)}$, for all $\lambda \in I$. The elements of the covariance matrix between random variables $Z^{(\lambda)}$ and $Z^{(\mu)}$, for all $\lambda \neq \mu$, can be derived by writing

$$\begin{aligned}
Z^{(\lambda)}Z^{(\mu)} &= \sqrt{N}\left(\frac{\hat{\delta}^{(\lambda)}}{\sqrt{V^{(\lambda)}}}\right)\left(\hat{\varphi}-1-\frac{\hat{b}^{(\lambda)}}{\hat{\delta}^{(\lambda)}}\right)\sqrt{N}\left(\frac{\hat{\delta}^{(\mu)}}{\sqrt{V^{(\mu)}}}\right)\left(\hat{\varphi}-1-\frac{\hat{b}^{(\mu)}}{\hat{\delta}^{(\mu)}}\right)\\
&= \frac{\hat{\delta}^{(\lambda)}\hat{\delta}^{(\mu)}}{\sqrt{V^{(\lambda)}}\sqrt{V^{(\mu)}}}N\frac{(\frac{1}{N}\sum_{i=1}^{N}W_i^{(\lambda)})(\frac{1}{N}\sum_{i=1}^{N}W_i^{(\mu)})}{\hat{\delta}^{(\lambda)}\hat{\delta}^{(\mu)}}\\
&= \frac{1}{\sqrt{V^{(\lambda)}}\sqrt{V^{(\mu)}}}\frac{1}{N}\sum_{i=1}^{N}W_i^{(\lambda)}\sum_{i=1}^{N}W_i^{(\mu)}. \qquad (50)
\end{aligned}$$

By the definition of $W_i^{(\lambda)}$ (see (42)) and assumption of cross-section independence between $W_i^{(\lambda)}$ and $W_j^{(m)}$, for $i \neq j$, we have $E(W_i^{(\lambda)}W_j^{(\mu)})=0$, for $i \neq j$. Based on this result, we can show that

$$\underset{N\to\infty}{p\lim}\frac{1}{N}\sum_{i=1}^{N}W_i^{(\lambda)}\sum_{i=1}^{N}W_i^{(\mu)} = \underset{N\to\infty}{p\lim}\frac{1}{N}\sum_{i=1}^{N}W_i^{(\lambda)}W_i^{(\mu)}=E(W_i^{(\lambda)}W_i^{(\mu)}). \qquad (51)$$

$E(W_i^{(\lambda)}W_i^{(\mu)})$ can be analytically derived as

$$\begin{aligned}
E(W_i^{(\lambda)}W_i^{(\mu)}) &= E[u_i'(\Lambda' Q^{(\lambda)} - \Psi^{(\lambda)})u_i u_i'(\Lambda' Q^{(\mu)} - \Psi^{(\mu)})u_i] \\
&= E[F^{(\lambda)\prime} vec(u_i u_i') vec(u_i u_i')' F^{(\mu)}] \\
&= F^{(\lambda)\prime} E[vec(u_i u_i') vec(u_i u_i')'] F^{(\mu)} \\
&= F^{(\lambda)\prime} E[vec(u_i u_i') vec(u_i u_i')'] F^{(\mu)}, \qquad (52)
\end{aligned}$$

or

$$E(W_i^{(\lambda)}W_i^{(\mu)}) = \sigma_u^4 F^{(\lambda)\prime}[(I_{T^2} + K_{T^2}) + vec(I_T)vec(I_T)']F^{(\mu)}, \qquad (53)$$

using the following result:

$$\begin{aligned}
E[vec(u_i u_i') vec(u_i u_i')'] &= Var(u_i \otimes u_i) + E(vec(u_i u_i'))E(vec(u_i u_i'))' \qquad (54) \\
&= \sigma_u^4[(I_{T^2} + K_{T^2}) + vec(I_T)vec(I_T)'].
\end{aligned}$$

Based on (53), it can be shown that the probability limit of (50) is given as

$$\begin{aligned}
&E(Z^{(\lambda)}Z^{(\mu)}) \qquad (55) \\
= \; & \frac{F^{(\lambda)\prime}\sigma_u^4[(I_{T^2} + K_{T^2}) + vec(I_T)vec(I_T)']F^{(\mu)}}{\sqrt{F^{(\lambda)\prime}\sigma_u^4[(I_{T^2} + K_{T^2} + vec(I_T)vec(I_T)']F^{(\lambda)}}\sqrt{F^{(\mu)\prime}\sigma_u^4[(I_{T^2} + K_{T^2}) + vec(I_T)vec(I_T)']F^{(\mu)}}} \\
= \; & \frac{F^{(\lambda)\prime}(I_{T^2} + K_{T^2})F^{(\mu)}}{\sqrt{F^{(\lambda)\prime}(I_{T^2} + K_{T^2})F^{(\lambda)}}\sqrt{F^{(\mu)\prime}(I_{T^2} + K_{T^2})F^{(\mu)}}},
\end{aligned}$$

where the result of the last row follows directly from $F^{(\lambda)\prime}vec(I_T)vec(I_T)' = 0$.

Proof of Theorem 4: The theorem can be proved following analogous steps to those for the proof of Theorem 1.

Proof of Theorem 5: The theorem can be proved following analogous steps to those for the proof of

Theorem 1 and using the following results:

$$\frac{1}{N}\sum_{i=1}^{N} u_i'\Lambda' Q^{(\lambda)} u_i \xrightarrow{p} tr(\Lambda' Q^{(\lambda)}\Gamma_N) \text{ and } \frac{1}{N}\sum_{i=1}^{N} u_i'\Lambda' Q^{(\lambda)}\Lambda u_i \to^p tr(\Lambda' Q^{(\lambda)}\Lambda\Gamma_N). \tag{56}$$

Based on the definition of matrix $\Psi_1^{(\lambda)}$ and conditions $(b1)$ and $(b2)$ of Assumption 2, it can be easily seen that

$$E(W_i^{(\lambda)}) = tr((\Lambda' Q^{(\lambda)} - \Psi_1^{(\lambda)})\Gamma_N) = 0, \text{ for all } i. \tag{57}$$

Proof of Theorem 6: It can be proved following analogous steps to those followed for the proof of Theorem 1. Under null hypothesis $\varphi = 1$, vector $y_{i,-1}$ can be decomposed as

$$y_{i,-1} = y_{i0}e + \Lambda e\beta_i + \Lambda u_i. \tag{58}$$

Multiplying both sides of the last relationship by $Q_*^{(\lambda)}$ yields

$$Q_*^{(\lambda)} y_{i,-1} = Q_*^{(\lambda)}\Lambda u_i, \tag{59}$$

since $Q_*^{(\lambda)}e = 0$ and $Q_*^{(\lambda)}\Lambda e = 0$. Also, note that, under $\varphi = 1$, the following relationships hold:

$$\Delta y_i = u_i + e\beta_i \tag{60}$$

and

$$Q_*^{(\lambda)}\Delta y_i = Q_*^{(\lambda)} u_i \quad \text{and} \quad Q_*^{(\lambda)}\Lambda\Delta y_i = Q_*^{(\lambda)}\Lambda u_i. \tag{61}$$

Using (61), the numerator and denominator of $\hat{\varphi}_*^{(\lambda)} - 1$ become

$$y_{i,-1}' Q_*^{(\lambda)} u_i = u_i'\Lambda' Q_*^{(\lambda)} u_i = \Delta y_i'\Lambda' Q_*^{(\lambda)}\Delta y_i \text{ and} \tag{62}$$

$$y_{i,-1}' Q_*^{(\lambda)} y_{i,-1} = u_i'\Lambda' Q_*^{(\lambda)}\Lambda u_i = \Delta y_i'\Lambda' Q_*^{(\lambda)}\Lambda\Delta y_i, \tag{63}$$

respectively. By Kitchin's LLN, it can be shown that the inconsistency of estimator $\hat{\varphi}_*^{(\lambda)}$ is given as

$$\hat{\varphi}_*^{(\lambda)} - 1 = \frac{\sum_{i=1}^N y'_{i,-1} Q_*^{(\lambda)} u_i}{\sum_{i=1}^N y'_{i,-1} Q_*^{(\lambda)} y_{i,-1}} \xrightarrow{p} \frac{b_2^{(\lambda)}}{\delta_2^{(\lambda)}} = \frac{tr(\Lambda' Q_*^{(\lambda)} \mathbf{\Gamma}_N)}{tr(\Lambda' Q_*^{(\lambda)} \Lambda \mathbf{\Gamma}_N)}. \tag{64}$$

The last result holds because, as $N \to +\infty$, we have

$$\frac{1}{N}\sum_{i=1}^N y'_{i,-1} Q_*^{(\lambda)} u_i - tr(\Lambda' Q_*^{(\lambda)}(\Gamma_N + \beta_N^2 J_T)) \xrightarrow{p} 0, \text{ where } \beta_N^2 = \frac{1}{N}\sum_{i=1}^n E((\beta_i)^2), \tag{65}$$

$$\text{or } \frac{1}{N}\sum_{i=1}^N y'_{i,-1} Q_*^{(\lambda)} u_i - tr(\Lambda' Q_*^{(\lambda)} \Gamma_N) \xrightarrow{p} 0,$$

since $tr(\Lambda' Q_*^{(\lambda)} J_T) = 0$, and

$$\frac{1}{N}\sum_{i=1}^N y'_{i,-1} Q_*^{(\lambda)} y_{i,-1} - tr(\Lambda' Q_*^{(\lambda)} \Lambda(\Gamma_N + \beta_N^2 J_T)) \xrightarrow{p} 0, \tag{66}$$

since $tr(\Lambda' Q_*^{(\lambda)} \Lambda J_T) = 0$.

The remaining of the proof follows the same steps with those of the proof of Theorem 1. That is, subtract the consistent estimator of $\frac{b_2^{(\lambda)}}{\delta_2^{(\lambda)}}$, given by (33), from $\hat{\varphi}_*^{(\lambda)} - 1$ and, then, apply standard asymptotic theory.

References

[1] Andrews, Donald W K, 1993. Tests for Parameter Instability and Structural Change with Unknown Change Point. Econometrica, Econometric Society, vol. 61(4), pages 821-56, July.

[2] Arelano E. and B. Honoré (2002). Panel data models: Some recent developments, Handbook of Econometrics, edited by Heckman J. and E. Leamer editors, Vol 5, North Holland.

[3] Arellano-Valle, R. B., and Genton, M. G. (2008). On the exact distribution of the maximum of absolutely continuous dependent random variables. Statistics and Probability Letters, 78, 27-35.

[4] Bai J.(2010). Common breaks in means and variances for panel data. Journal of Econometrics 157, 78-92.

[5] Bai J. and J.L. Carrion-i-Silvestre (2009). Structural changes, common stochastic trends and unit roots in panel data. Review of Economic Studies, 76, 471-501.

[6] Baltagi B.H. (1995). Econometric analysis of panel data, Wiley, Chichester.

[7] Blundell R. and Bond S. 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, Elsevier, vol. 87(1), pages 115-143, August.

[8] Binder M., Hsiao C. and Pesaran H. (2005). Estimation and inference in short panel vector autoregression with unit roots and cointegration. Econometric Theory, 21, 795-837.

[9] Carrion-i-Silvestre J.L., Del Barrio-Castro, and E. Lopez-Bazo (2005). Breaking the panels. An application to real per capita GDP. Econometrics Journal, 8, 159-175.

[10] De Wachter S., Harris R.D.F., Tzavalis E. (2007). Panel unit root tests: the role of time dimension and serial correlation. Journal of Statistical Inference and Planning, 137, 230-244.

[11] De Wachter S., and Tzavalis E. (2005). Monte Carlo comparison of model and moments selection and classical inference approach to break detection in panel data models", Economic Letters (2005), 88, 91-96.

[12] De Wachter S., and Tzavalis E. (2012).Detection of structural breaks in linear panel data models, Computational Statistics & Data Analysis, 56, 3020-3034.

[13] Hall, B. H. and Mairesse J. (2005). Testing for unit roots in panel data: an exploration using real and simulated data. In D. Andrews and J. Stock (Eds.), Identification and Inference in Econometric Models: Essays in Honor of Thomas J. Rothenberg.

[14] Han, C and Phillips P.C.B (2012). GMM estimation for dynamic panels with fixed effects and strong instruments at unity. Econometric Theory, 26, 2010, 119–151.

[15] Harris R.D.F. and Tzavalis E. (1999). Inference for unit roots in dynamic panels where the time dimension is fixed. Journal of Econometrics, 91, 201-226.

[16] Harris R.D.F and Tzavalis E. (2004). Inference for unit roots for dynamic panels in the presence of deterministic trends: Do stock prices and dividends follow a random walk?, Econometric Reviews 23, 149-166.

[17] Hlouskova J. and Wagner M. (2006). The performance of panel unit root and stationary tests: Results from a large scale simulation study. Econometric Reviews, 25, 85-117.

[18] Hsiao C., Pesaran H. and Tahmiscioglu K., (2002). Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. Journal of Econometrics, Elsevier, vol. 109(1), pages 107-150, July.

[19] Karavias Y. and Tzavalis E. (2012). Testing for unit roots in short panels allowing for structural breaks. Computational Statistics and Data Analysis, http://dx.doi.org/10.1016/j.csda.2012.10.014.

[20] Kim, D. (2011). Estimating a common deterministic time break in large panels with cross-sectional dependence. Journal of Econometrics, 164 (2), 310-330.

[21] Kruininger H. and Tzavalis E. (2002). Testing for unit roots in short dynamic panels with serially correlated and heteroscedastic disturbance terms. Working Papers 459, Queen Mary, University of London, School of Economics and Finance.

[22] Levin A., Lin C.F. and Chu C-S. J, (2002). Unit root tests in panel data: asymptotic and finite-sample properties. Journal of Econometrics, 108, 1-24.

[23] Madsen E. (2008). GMM Estimators and Unit Root Tests in the AR(1) Panel Data Model. unpublished.

[24] Moon, H. R., Perron B. and Phillips P.C.B. (2007). Incidental trends and the power of panel unit root tests. Journal of Econometrics, Elsevier, vol. 141(2), pages 416-459, December.

[25] Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica 49, 1417-1426.

[26] Perron P. (2006). Dealing with structural breaks. Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory, Mills. T and K. Patterson (eds), Palgrave MacMillan, 278-352.

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

[28] Schott J.R. (1996). Matrix Analysis for Statistics, Wiley-Interscience.

[29] Stock J.H. (1994). Unit roots, structural breaks and trends. In Handbook of Econometrics, Vol. IV, Engle R.F. and D.L. McFadden (eds), Elsevier, Amsterdam.

[30] Schwert, G.W., (1989). Tests for unit roots:a Monte Carlo investigation. Journal of Business and Economic Statistics 7, 147–160.

[31] Vogelsang, T. and 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.

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

[33] White H. (2000). Asymptotic theory for econometricians, Academic Press.

Table 2(a): Size and power of $z_1 \equiv \min Z_1^{(\lambda)}$, for $\theta = 0.50$

N	50	50	50	100	100	100	200	200	200
T	6	10	15	6	10	15	6	10	15
				$\lambda = 0.25$					
$\varphi = 1.00$	0.059	0.062	0.065	0.053	0.056	0.063	0.056	0.053	0.053
$\varphi = 0.95$	0.211	0.236	0.222	0.332	0.360	0.295	0.514	0.572	0.461
$\varphi = 0.90$	0.445	0.449	0.328	0.714	0.699	0.504	0.945	0.934	0.759
				$\lambda = 0.50$					
$\varphi = 1.00$	0.060	0.064	0.065	0.053	0.055	0.063	0.052	0.050	0.058
$\varphi = 0.95$	0.215	0.241	0.223	0.321	0.359	0.297	0.512	0.587	0.462
$\varphi = 0.90$	0.452	0.440	0.330	0.712	0.698	0.505	0.947	0.935	0.766
				$\lambda = 0.75$					
$\varphi = 1.00$	0.060	0.060	0.065	0.052	0.054	0.065	0.051	0.050	0.054
$\varphi = 0.95$	0.214	0.245	0.213	0.324	0.365	0.293	0.528	0.585	0.465
$\varphi = 0.90$	0.463	0.452	0.342	0.711	0.703	0.500	0.942	0.934	0.760

Table 2(b): Size and power of $z_1 \equiv \min Z_1^{(\lambda)}$, for $\theta = -0.50$

N	50	50	50	100	100	100	200	200	200
T	6	10	15	6	10	15	6	10	15
					$\lambda = 0.25$				
$\varphi = 1.00$	0.059	0.064	0.076	0.054	0.060	0.066	0.057	0.053	0.065
$\varphi = 0.95$	0.076	0.075	0.078	0.079	0.074	0.068	0.090	0.079	0.071
$\varphi = 0.90$	0.083	0.076	0.078	0.092	0.075	0.075	0.109	0.083	0.078
					$\lambda = 0.50$				
$\varphi = 1.00$	0.057	0.064	0.072	0.056	0.061	0.066	0.051	0.053	0.069
$\varphi = 0.95$	0.082	0.070	0.073	0.074	0.072	0.068	0.087	0.079	0.071
$\varphi = 0.90$	0.083	0.073	0.079	0.093	0.079	0.072	0.116	0.082	0.073
					$\lambda = 0.75$				
$\varphi = 1.00$	0.059	0.064	0.073	0.056	0.061	0.065	0.051	0.057	0.069
$\varphi = 0.95$	0.076	0.069	0.077	0.074	0.070	0.073	0.088	0.078	0.071
$\varphi = 0.90$	0.083	0.074	0.076	0.093	0.078	0.074	0.116	0.086	0.078

Table 2(c): Size and power of $z_1 \equiv \min Z_1^{(\lambda)}$, for $\theta = 0.00$

N	50	50	50	100	100	100	200	200	200
T	6	10	15	6	10	15	6	10	15
					$\lambda = 0.25$				
$\varphi = 1.00$	0.057	0.061	0.076	0.051	0.058	0.063	0.060	0.052	0.064
$\varphi = 0.95$	0.140	0.169	0.167	0.184	0.225	0.217	0.282	0.355	0.307
$\varphi = 0.90$	0.234	0.258	0.228	0.365	0.393	0.322	0.575	0.629	0.488
					$\lambda = 0.50$				
$\varphi = 1.00$	0.056	0.060	0.076	0.054	0.058	0.063	0.053	0.052	0.064
$\varphi = 0.95$	0.139	0.159	0.165	0.182	0.224	0.216	0.291	0.357	0.311
$\varphi = 0.90$	0.235	0.246	0.232	0.365	0.395	0.326	0.592	0.632	0.490
					$\lambda = 0.75$				
$\varphi = 1.00$	0.057	0.060	0.070	0.053	0.059	0.069	0.049	0.052	0.064
$\varphi = 0.95$	0.144	0.161	0.165	0.192	0.229	0.211	0.280	0.357	0.312
$\varphi = 0.90$	0.234	0.248	0.230	0.362	0.401	0.326	0.589	0.635	0.495

Table 3(a): Size and power of $z_2 \equiv \min Z_2^{(\lambda)}$, for $\theta = 0.50$

N	50	50	50	100	100	100	200	200	200
T	6	10	15	6	10	15	6	10	15
				$\lambda = 0.25$					
$\varphi = 1.00$	0.058	0.084	0.093	0.051	0.071	0.072	0.048	0.060	0.068
$\varphi = 0.95$	0.056	0.085	0.117	0.055	0.083	0.106	0.056	0.081	0.118
$\varphi = 0.90$	0.057	0.111	0.207	0.056	0.129	0.259	0.058	0.153	0.358
				$\lambda = 0.50$					
$\varphi = 1.00$	0.060	0.078	0.093	0.054	0.072	0.072	0.051	0.060	0.068
$\varphi = 0.95$	0.055	0.090	0.117	0.054	0.084	0.106	0.054	0.081	0.120
$\varphi = 0.90$	0.062	0.120	0.205	0.058	0.125	0.254	0.057	0.154	0.350
				$\lambda = 0.75$					
$\varphi = 1.00$	0.033	0.077	0.093	0.052	0.069	0.071	0.052	0.060	0.067
$\varphi = 0.95$	0.058	0.091	0.117	0.054	0.083	0.105	0.053	0.081	0.116
$\varphi = 0.90$	0.059	0.118	0.205	0.057	0.127	0.256	0.052	0.152	0.349

Table 3(b): Size and power of $z_2 \equiv \min Z_2^{(\lambda)}$, for $\theta = -0.50$

N	50	50	50	100	100	100	200	200	200
T	6	10	15	6	10	15	6	10	15
				$\lambda = 0.25$					
$\varphi = 1.00$	0.055	0.061	0.073	0.051	0.057	0.066	0.050	0.055	0.064
$\varphi = 0.95$	0.057	0.075	0.102	0.052	0.080	0.109	0.052	0.090	0.135
$\varphi = 0.90$	0.053	0.092	0.131	0.050	0.113	0.162	0.051	0.148	0.220
				$\lambda = 0.50$					
$\varphi = 1.00$	0.053	0.061	0.076	0.051	0.058	0.066	0.050	0.055	0.064
$\varphi = 0.95$	0.050	0.073	0.104	0.052	0.080	0.107	0.052	0.086	0.133
$\varphi = 0.90$	0.055	0.093	0.131	0.051	0.105	0.153	0.051	0.131	0.195
				$\lambda = 0.75$					
$\varphi = 1.00$	0.053	0.065	0.076	0.051	0.063	0.066	0.050	0.058	0.064
$\varphi = 0.95$	0.052	0.074	0.103	0.053	0.085	0.108	0.051	0.086	0.132
$\varphi = 0.90$	0.055	0.088	0.132	0.050	0.101	0.154	0.052	0.136	0.202

Table 3(c): Size and power of $z_2 \equiv \min Z_2^{(\lambda)}$, for $\theta = 0$

N	50	50	50	100	100	100	200	200	200
T	6	10	15	6	10	15	6	10	15
				$\lambda = 0.25$					
$\varphi = 1.00$	0.056	0.076	0.089	0.051	0.066	0.075	0.051	0.064	0.070
$\varphi = 0.95$	0.050	0.079	0.117	0.051	0.079	0.123	0.049	0.083	0.128
$\varphi = 0.90$	0.054	0.103	0.184	0.053	0.119	0.231	0.055	0.141	0.322
				$\lambda = 0.50$					
$\varphi = 1.00$	0.055	0.068	0.089	0.052	0.064	0.075	0.051	0.063	0.068
$\varphi = 0.95$	0.051	0.082	0.118	0.049	0.078	0.122	0.050	0.086	0.125
$\varphi = 0.90$	0.055	0.107	0.181	0.058	0.111	0.226	0.050	0.145	0.309
				$\lambda = 0.75$					
$\varphi = 1.00$	0.053	0.076	0.091	0.055	0.061	0.075	0.051	0.064	0.069
$\varphi = 0.95$	0.051	0.078	0.114	0.052	0.078	0.107	0.052	0.082	0.123
$\varphi = 0.90$	0.055	0.107	0.189	0.050	0.110	0.229	0.059	0.137	0.306

Table 4: Size and power of $Z_3^{(\lambda)}$, when $\lambda = 0.50$ and T_0 estimated.

N	50	50	100	100	200	200
T	10	15	10	15	10	15
			$\theta = 0.50$			
$\varphi = 1.00$	0.054	0.062	0.053	0.052	0.058	0.061
$\varphi = 0.95$	0.055	0.050	0.052	0.060	0.049	0.061
$\varphi = 0.90$	0.062	0.076	0.063	0.090	0.062	0.100
			$\theta = -0.50$			
$\varphi = 1.00$	0.044	0.051	0.046	0.048	0.048	0.055
$\varphi = 0.95$	0.051	0.056	0.050	0.061	0.055	0.061
$\varphi = 0.90$	0.045	0.063	0.055	0.077	0.061	0.080
			$\theta = 0.00$			
$\varphi = 1.00$	0.050	0.055	0.052	0.050	0.054	0.062
$\varphi = 0.95$	0.049	0.056	0.054	0.072	0.055	0.068
$\varphi = 0.90$	0.065	0.079	0.063	0.098	0.071	0.112