ARTICLE IN PRESS

Computational Statistics and Data Analysis ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Testing for unit roots in short panels allowing for a structural break

Yiannis Karavias, Elias Tzavalis *

Department of Economics, Athens University of Economics & Business, Greece

ARTICLE INFO

Article history:
Received 16 April 2011
Received in revised form 18 October 2012
Accepted 19 October 2012
Available online xxxx

Keywords:
Panel data models
Unit roots
Structural breaks
Sequential tests
Bootstrap
Trade openness

ABSTRACT

Panel data unit root tests which allow for a common structural break in the individual effects or linear trends of the AR(1) panel data model are suggested. These allow the date of the break to be unknown. The tests assume that the time-dimension of the panel (T) is fixed (finite) while the cross-section (N) is large. Under the null hypothesis of unit roots, they are similar to the initial conditions of the model and its individual effects. Extensions of the tests to the AR(2) model are provided. These highlight the difficulties in extending the tests to higher order serial correlation of the error terms. Monte Carlo experiments indicate that the small sample performance of the tests is very satisfactory. Application of the tests to the trade openness variable of the non-oil countries indicates that evidence of persistence of this variable can be attributed to trade liberalization policies adopted by many developing countries since the early nineties.

1. Introduction

The autoregressive panel data model of lag order one (denoted as AR(1)), which assumes that the time dimension of the data (denoted as T) is fixed (finite) while its cross-sectional (denoted as N) is large, has been extensively used in the literature to study the dynamic behavior of many economic time series across different units, i.e. countries or industries (see Arellano, 2003, Arellano and Honoré, 2002 and Baltagi and Kao, 2000, *inter alia*). Of particular interest is the use of this model to examine if economic series contain a unit root in their autoregressive component (see Hlouskova and Wagner, 2006, for a recent survey). Recent economic applications of panel data unit root tests include investigation of the following: the economic growth convergence hypothesis (see de la Fuente, 1997, for a survey), the random walk hypothesis of stock prices and dividends (see, e.g., Harris and Tzavalis, 2004 and Lo and MacKinlay, 1995), the long-run validity of purchasing power parity (see Culver and Papell, 1999, *inter alia*) and, finally, the permanent effects of liberalization policies on trade (see, e.g., Wacziarg and Welch, 2004).

This paper extends Harris' and Tzavalis (1999) panel unit root tests assuming fixed-T to allow for a potential structural break in the deterministic components of the AR(1) panel data model, namely its individual effects and/or linear trends, at a known or unknown date. This is a very useful extension given recent evidence suggesting that the presence of a unit root in the autoregressive component of many economic series can be attributed to the existence of structural breaks in their deterministic components, which are ignored by standard unit root testing procedures (see Perron, 1989, 1990, for single time series analysis). The panel data approach offers an interesting and unique perspective to investigate if evidence of unit roots can be falsely attributed to the existence of structural breaks, which is not shared by single series methods. The cross-sectional units of panel data can provide important sample information which can help to distinguish permanent stochastic shifts of economic time series from changes in their deterministic components (see, e.g., Bai, 2010).

In contrast to the vast literature for single time series, there are few studies that consider panel data unit root tests allowing for structural breaks (see Bai and Carrion-i-Silvestre, 2009, Carrion-i-Silvestre et al., 2005 and Chan and Pauwels,

* Correspondence to: Department of Economics, Athens University of Economics & Business, Patission 76, Athens 104 34, Greece. Tel.: +30 2108203332.
E-mail addresses: jkaravia@aueb.gr (Y. Karavias), e.tzavalis@aueb.gr (E. Tzavalis).

0167-9473/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2012.10.014

2011). However, these tests assume that the time-dimension of the panel model T is large and, more importantly, that it grows larger than N. Thus, they are more appropriate for panel data sets where T is bigger than N, referred to as large panels. As shown in Harris and Tzavalis (1999), application of large-T panel unit root tests to short panels, which assume fixed-T, leads to serious size distortions and power reductions, since their sample distribution is not well approximated in panels with small T. In the literature which assumes fixed-T panel data unit root tests, there are also few studies which suggest unit root tests allowing for structural breaks (see Carrion-i-Silvestre et al., 2002, Tzavalis, 2002 and, more recently, Hadri et al., forthcoming). These studies are mainly interested in pursuing ideas on how to test for a unit root in the AR(1) panel data model allowing for a common break in its individual effects. They mainly consider the case of a known date break and they assume that the error terms of the AR(1) panel data model are normally distributed.

The main goal of this paper is to extend the above fixed-T panel unit root tests, considering a common break in the deterministic components of the AR(1) panel data model, to allow for an unknown break point. This is done under quite general distributional assumptions of the error terms. The proposed test statistics are similar (invariant) under the null hypothesis to the initial conditions and/or the individual effects of the panel data model. This property of the tests is very useful for the following two reasons. First, it does not require any assumption about the initial conditions of the panel and, second, it does not involve estimation of its individual effects. As recently shown by Kim and Perron (2009), unit root testing procedures relying on estimation of these effects in a first step perform poorly in small samples. Our suggested tests can be applied to the case of a two-way error component panel data model, which allows for cross-correlation across the error terms. This can be done by taking deviations of the individual series of the panel from their cross-section mean at each point in time t (see O'Connell, 1998), in the first step.

To apply the tests in the case of an unknown break point, the paper relies on the sequential testing procedure recommended in single time series analysis by Zivot and Andrews (1992) (see also Andrews, 1993 and Perron, 1997). This procedure calculates the minimum value of one-sided standardized test statistics which assume a known date break. These statistics are sequentially computed for each possible break point of the sample that the break can occur. The limiting distribution of these sequential test statistics is that of the minimum value of a fixed number of correlated standard normals. The paper derives analytically the correlation matrix of these variables and tabulates critical values of the distribution of their minimum based on Monte Carlo simulations. Finally, the paper shows how to extend the tests to the case of a panel data autoregressive model of lag order two, which may describe some economic series (see, e.g., Cati et al., 1999). This extension highlights some of the difficulties of generalizing the tests to allow for serially correlated error terms when T is fixed.

The paper is organized as follows. Section 2 derives the limiting distributions of the test statistics suggested by the paper for the cases that the break point is considered as known and unknown. This section also proves the consistency of the tests under the alternative hypothesis of stationarity. Section 3 extends the tests to the case of AR(2) panel data model. Section 4 carries out a Monte Carlo study which evaluates the small sample performance of the tests. Section 5 implements the tests to investigate if trade liberalization policies introduced at the end of eighties and/or the early nineties in the non-oil countries have permanently fostered international trade. Section 6 concludes the paper.

2. The test statistics and their limiting distribution

2.1. The date of the break point is known

Consider the following non-linear AR(1) panel data models, denoted as $m = \{M1, M2\}$, allowing for a common structural break in their deterministic components at time point T_0:

$$M1 : y_i = \varphi y_{i-1} + (1-\varphi)(a_i^{(\lambda)} e^{(\lambda)} + a_i^{(1-\lambda)} e^{(1-\lambda)}) + u_i, \quad i = 1, 2, \ldots, N \quad \text{and} \tag{1}$$

$$M2 : y_i = \varphi y_{i-1} + \varphi \beta_i e + (1-\varphi)(a_i^{(\lambda)} e^{(\lambda)} + a_i^{(1-\lambda)} e^{(1-\lambda)}) + (1-\varphi)(\beta_i^{(\lambda)} \tau^{(\lambda)} + \beta_i^{(1-\lambda)} \tau^{(1-\lambda)}) + u_i, \tag{2}$$

where $y_i = (y_{i1}, \ldots, y_{iT})'$ is a vector which collects the time series observations of panel data series y_{it}, for $t = 1, 2, \ldots, T$, across the cross-sectional units of the panel $i = 1, 2, \ldots, N$, $y_{i-1} = (y_{i0}, \ldots, y_{iT-1})'$ is vector y_i lagged one period back, $u_i = (u_{i1}, \ldots, u_{iT})'$ is the vector of error terms u_{it}, for all t, $\beta_i e$ is defined as $\beta_i e = \beta_i^{(\lambda)} e^{(\lambda)} + \beta_i^{(1-\lambda)} e^{(1-\lambda)}$, where e is a $(TX1)$-dimension vector of unities, and $e^{(\lambda)}$ and $e^{(1-\lambda)}$ are $(TX1)$-dimension vectors defined, respectively, as follows: $e_t^{(\lambda)} = 1$ if $t \le T_0$ and 0 otherwise, and $e_t^{(1-\lambda)} = 1$ if $t > T_0$ and 0 otherwise. These vectors are appropriately designed to capture a possible common break in the individual effects of models (1) and (2), a_i, before and after the break occurs, denoted respectively as $a_i^{(\lambda)}$ and $a_i^{(1-\lambda)}$, where λ denotes the fraction of the sample that this break occurs. λ is defined as $\lambda \in I = \left\{\left[\frac{1}{T}\right], \left[\frac{2}{T}\right], \ldots, \left[\frac{T-1}{T}\right]\right\}$ for model $M1$ and $\lambda \in I = \left\{\left[\frac{2}{T}\right], \left[\frac{3}{T}\right], \ldots, \left[\frac{T-2}{T}\right]\right\}$ for model $M2$, where $[\cdot]$ denotes the integer part. In addition to a common break in individual effects a_i, model (2) also allows for a common break in the slope coefficients of the individual linear trends of the panel, β_i, denoted as $\beta_i^{(\lambda)}$ and $\beta_i^{(1-\lambda)}$, for all i. Vectors $\tau^{(\lambda)}$ and $\tau^{(1-\lambda)}$ collect the time points of these trends. More specifically, the elements of these vectors are defined as follows: $\tau_t^{(\lambda)} = t$ if $t \le T_0$ and 0 otherwise, while $\tau_t^{(1-\lambda)} = t$ if $t > T_0$ and 0 otherwise.

The AR(1) panel data models $M1$ and $M2$, given by Eqs. (1) and (2), can be employed to obtain panel data unit root test statistics which are similar (invariant) under null hypothesis $H_0 : \varphi = 1$ to the initial conditions of the panel y_{i0} and/or

Please cite this article in press as: Karavias, Y., Tzavalis, E., Testing for unit roots in short panels allowing for a structural break. Computational Statistics and Data Analysis (2012), doi:10.1016/j.csda.2012.10.014