Unit root tests allowing for breaks in panels with fixed T

Ioannis Karavias and Elias Tzavalis

July 2010

Unit root against structural break

- In this paper we wish to test the null hypothesis of a unit root against the alternative of stationarity with a single structural break in the individual effects of the series.
- We extend Harris and Tzavalis (1999) and Kruiniger and Tzavalis (2002) unit root tests to allow for a one time break under the alternative.
- Standard tests of the unit root hypothesis are biased towards nonrejection by the existence of shifts in the deterministic components (Perron 1989; Carrion-i-Silvestre, del Barrio Castro and Lopez-Bazo, 2001).
- Similar tests in the literature (Levin and Lin 2002), consider the case of large T panels which has been shown to lead to serious size distortions and power decreases in small T Panels (Hadri 2000, Harris and Tzavalis 1999).

Applications in Economics

- Unit root tests are used to test hypothesis such as

1. Economic growth convergence hypothesis (de la Fuente 1997)
2. Purchasing Power Parity hypothesis (Culver Papell 1999)
3. Hypothesis that stock prices and dividents follow a random walk (Lo MacKinlay 1995)
4. Effects of trade liberization policies (Wacziarg Welch 2004)

The Model

- The model is

$$y_i = \phi y_{i,-1} + X_i^{(\lambda)} \gamma_i^{(\lambda)} + u_i, \quad i = 1, 2, ..., N$$

- where $y_i = (y_{i1}, ..., y_{iT})'$ $y_{i,-1} = (y_{i0}, ..., y_{iT-1})'$ and $u_i = (u_{i1}, ..., u_{iT})$ $(TX1)$-dimension vectors.
- Index λ denotes the date where the break T_0 occurs and $\lambda \in I = \{2, 3, ..., T - 1\}$.
- $X_i^{(\lambda)} \equiv \left(e_t^{(\lambda)}, e_t^{(1-\lambda)} \right)$ where $e_t^{(\lambda)} = 1$ if $t \leq T_0$ and 0 otherwise, and $e_t^{(1-\lambda)} = 1$ if $t > T_0$ and 0 otherwise.
- $\gamma_i^{(\lambda)} = (a_i^{(\lambda)}(1 - \varphi), a_i^{(1-\lambda)}(1 - \varphi))'$

The estimator

- The estimator is

$$\hat{\phi} = \left[\sum_{i=1}^{N} y'_{i,-1} Q^{(\lambda)} y_{i,-1} \right]^{-1} \left[\sum_{i=1}^{N} y'_{i,-1} Q^{(\lambda)} y_i \right]$$

- where $Q^{(\lambda)} = \left[I - X_i^{(\lambda)} \left(X_i^{(\lambda)\prime} X_i^{(\lambda)} \right)^{-1} X_i^{(\lambda)\prime} \right]$ is the (TXT) "within transformation" matrix.

Assumption 1:

- b1) $\{u_i\}$ constitutes a sequence of independent random $T-$vectors with means $E(u_i) = 0$ and $(T \times T)$ autocovariance matrices $E(u_i u_i') = \Gamma_i \equiv [\gamma_{i,rs}]$ of unknown form apart from $\gamma_{i,1T} = \gamma_{i,T1} = 0$.
- b2) The smallest eigenvalue of the average population covariance matrix $\bar{\Gamma}_N \equiv \frac{1}{N} \sum_{i=1}^{N} \Gamma_i$ is bounded away from zero for sufficiently large N.
- b3)$E(u_{it} y_{io}) = E(u_{it} a_i^m) = 0$ for m$=\lambda$, $(1-\lambda)$and $\forall$ $i \in \{1, 2, ..., N\}$, $t \in \{1, 2, ..., T\}$.
- b4) $E(u_{it}^4) < +\infty$, $E(y_{i0}^4) < +\infty$, $E((a_i^m)^4) < +\infty$, $E(y_{i0}^2 \gamma_i^m \gamma_i^{m\prime}) < +\infty$.

Known break with serial correlation

Theorem

Let the sequence $\{y_{i,t}\}$ be generated according to the previous model and let assumption 2 hold. Then under the null hypothesis $\phi = 1$, as $N \to \infty$

$$Z_1 \equiv \widehat{V}^{-0.5}\hat{\delta}\sqrt{N}\left(\hat{\phi} - 1 - \frac{\hat{b}}{\hat{\delta}}\right) \xrightarrow{d} N(0,1)$$

although V is not known as before, but is consistenly estimated.

Known break with serial correlation

- where

$$\hat{b} = vec(Q^{(\lambda)}\Lambda)'S\left(\frac{1}{N}\sum_{i=1}^{N} vec(\Delta y_i \Delta y_i')\right)$$

-

$$V = vec(Q^{(\lambda)}\Lambda)'(I_{T^2} - S)\left(\frac{1}{N}\sum_{i=1}^{N} \mathcal{V}(vec(\Delta y_i \Delta y_i'))\right)(I_{T^2} - S)vec(Q^{(\lambda)}\Lambda)$$

- where the middle term is consistently estimated by

$$\hat{\Omega} = \frac{1}{N}\sum_{i=1}^{N}\left(vec(\Delta y_i \Delta y_i')vec(\Delta y_i \Delta y_i')'\right) - \left(\frac{1}{N}\sum_{i=1}^{N} vec(\Delta y_i \Delta y_i')\right)\left(\frac{1}{N}\sum_{i=1}^{N} vec(\Delta y_i \Delta y_i')\right)'$$

Case of an unknown break point

- In this case the break point is unknown.The goal is to estimate the break point that gives the most weight to the alternative hypothesis therefore we choose λ so that it minimizes $Z(\lambda, T)$ across all λ (Andrews and Zivot 1990). Then the null hypothesis would be rejected when

$$\min_{\lambda \in I} Z_1(\lambda, T) < c_{\min},$$

where $c_{\min}$ denotes the size a left-tail critical value of the limiting distribution of $\min_{\lambda \in I} Z(\lambda, T)$.

Theorem

Let Assumption 1 hold. Then, as $N \rightarrow \infty$, we have

$$\min_{\lambda \in I} Z_1(\lambda, T) \xrightarrow{d} \min_{\lambda \in I} N(0, \Sigma)$$

where $\Sigma \equiv [\sigma_{\lambda s}]$ is a consistently estimated.

Distribution of the minimum of correlated random variables

- The distribution of the minimum is unknown but we noticing that $\min\{X_1, ..., X_k\} = -\max\{-X_1, ..., -X_k\}$ and for a significance level α, we equivalently need the distribution of the maximum

$$\begin{aligned} P(\min_{\lambda \in I} Z_1(\lambda, T) &< c_{\min}) = a \\ P(\max_{\lambda \in I} - Z_1(\lambda, T) &> -c_{\min}) = a \end{aligned}$$

- The probability density function of the maximum of correlated random variables that belong to the family of elliptically contoured distributions is known (Arellano-Vale and Genton 2007) and is given by $(x \in R)$

$$f_{X_{(n)}}(x) = \sum_{i=1}^{n} f_1(x; \mu_i, \Sigma_{ii}, h^{(1)}) F_{n-1}(x1_{n-1}; \mu_{-i,i}(x), \Sigma_{-i-i,i}, h^{(n-1)}_{z_i^2})$$

Consistency of the test

Theorem

Under assumption 1 the test is consistent:

$$\lim_{n \to +\infty} P\left(\min_{\lambda \in I} Z_1(\lambda, T) < c_{\min} \mid H_a\right) = 1$$

Reduction: No serial correlation

- a1) $\{u_{it}\}$ is a sequence of independently and identically distributed (IID) random variables with $E(u_{it}) = 0$, $Var(u_{it}) = \sigma_u^2 < \infty$, $E(u_{it}^4) = k + 3\sigma_u^4$, $\forall\ i \in \{1, 2, ..., N\}$ and $\forall\ t \in \{1, 2, ..., T\}$, where $k < \infty$.
- a2) $E(u_{it} y_{io}) = E(u_{it} a_i^m) = 0$ for m=$\lambda, (1 - \lambda)$and $\forall$ $i \in \{1, 2, ..., N\}$, $t \in \{1, 2, ..., T\}$.
- a3) $E(u_{it}^4) < +\infty$, $E(y_{i0}^4) < +\infty$, $E((a_i^m)^4) < +\infty$, $E(y_{i0}^2 \gamma_i^m \gamma_i^{m\prime}) < +\infty$.

Case of a known break point

Theorem

Let the sequence $\{y_{it}\}$ be generated according to the previous model and the break-point T_0 be known. Then, under the null hypothesis $\phi = 1$ and Assumption 1, we have:

$$Z_2(\lambda, T) \equiv C(k, \sigma_u^2, \lambda, T)^{-1/2}\sqrt{N}(\hat{\phi} - 1 - B(\lambda, T)) \xrightarrow{L} N(0,1)$$

as $N \to \infty$

- Theorem 1 can be extended to the case that the disturbance terms u_{it} are heterogenous across i and thus have $ID(0, \sigma_{u_i}^2)$

Unknown break without serial correlation

Theorem

Let Assumption 2 hold. Then, as $N \to \infty$, we have

$$\min_{\lambda \in I} Z_2 \xrightarrow{d} \min_{\lambda \in I} N(0, \Sigma)$$

where $\Sigma \equiv [\sigma_{\lambda s}]$ is consistently estimated.

Theorem

Under assumption 2 the test is consistent:

$$\lim_{n \to +\infty} P(\min_{\lambda \in I} Z_2 < c_{\min} \mid H_a) = 1$$

Monte Carlo Simulations

- The data generating process is $y_{it} = \phi y_{it-1} + (1-\varphi)a_i^{(\lambda)} + \varepsilon_{it} + \theta\varepsilon_{it-1}$ before the break and $y_{it} = \phi y_{it-1} + (1-\varphi)a_i^{(1-\lambda)} + \varepsilon_{it} + \theta\varepsilon_{it-1}$ after. All random variables $a_i^{(1-\lambda)}, a_i^{(\lambda)}, \varepsilon_{it}, \varepsilon_{it-1}, y_{i0}$ have a standard normal distribution and each result is taken after 10000 repetitions.
- Size and power of nominal level 5% for the test of theorem 5, $\theta = 0.5$

N	25	25	50	50	100	100	100
T	10	15	15	25	10	15	25
$\phi = 1.00$	0.06	0.08	0.06	0.07	0.05	0.06	0.06
				Power			
$\phi = 0.99$	0.09	0.11	0.09	0.10	0.09	0.10	0.09
$\phi = 0.95$	0.17	0.16	0.18	0.13	0.30	0.24	0.15
$\phi = 0.90$	0.24	0.23	0.28	0.19	0.55	0.40	0.24

Monte Carlo Simulations

- Size and power of nominal level 5% for the test when, $\theta = -0.5$

N	25	25	50	50	100	100	100
T	10	15	15	25	10	15	25
$\phi = 1.00$	0.07	0.07	0.06	0.07	0.05	0.05	0.05
				Power			
$\phi = 0.99$	0.08	0.07	0.06	0.08	0.06	0.05	0.05
$\phi = 0.95$	0.09	0.08	0.07	0.08	0.08	0.07	0.07
$\phi = 0.90$	0.09	0.09	0.07	0.08	0.08	0.09	0.09

Monte Carlo Simulations

Size and power of nominal level 5% for the test , $\theta = 0$ but we falsely assume ma(1) errors

N	25	25	50	50	100	100	100
T	10	15	15	25	10	15	25
$\phi = 1.00$	0.07	0.09	0.07	0.07	0.06	0.07	0.05
				Power			
$\phi = 0.99$	0.08	0.10	0.09	0.09	0.08	0.11	0.09
$\phi = 0.95$	0.13	0.14	0.17	0.13	0.20	0.20	0.16
$\phi = 0.90$	0.17	0.17	0.21	0.17	0.32	0.29	0.22

Monte Carlo Simulations

Size and power of nominal level 5% for the test under normality, homoscedasticity and no serial correlation

N	25	25	50	50	100	100	100
T	10	15	15	25	10	15	25
$\phi = 1.00$	0.07	0.08	0.08	0.07	0.05	0.07	0.07
				Power			
$\phi = 0.99$	0.10	0.14	0.15	0.22	0.12	0.18	0.33
$\phi = 0.95$	0.26	0.42	0.61	0.87	0.56	0.83	0.99
$\phi = 0.90$	0.48	0.74	0.93	0.99	0.92	0.99	1

Empirical Application

- We now apply the first statistic to examine whether there is high persistence in the consumption of the Eurogroup countries or the introduction of euro led to a structural break in the series.
- Final consumption expenditure of households is broken down by consumption purposes in twelve categories.
- Annual data are collected for a time span of eleven years, from 1996 to 2006.
- Fifteen countries of the eurogroup (Greece is not included due to missing data).
- All variables were divided by the respective country's gdp to eliminate the trend.
- To control for cross section correlation of additive form the individual series of our panel data set were taken in deviations from their cross-section mean at each point in time (O' Connel 1998).

Empirical Application

- The results of the table clearly indicate that the null hypothesis of a unit root in the level of final consumpion variable is rejected in favour of its stationary alternative.
- Break estimated in 2003.
-

Year	1998	1999	2000	2001	2002	2003	2004	2005
Statistic	3.2152	0.8988	2.1654	0.9685	0.1090	-0.2375	0.7072	1.0030

- This can be due to uncertainty and consumer reservation in the first years of the new currency.