Cointegration Testing in Dependent Panels with Breaks

Francesca Di Iorio
University of Naples Federico II

Stefano Fachin
University of Rome "La Sapienza"

Abstract

In this paper we propose panel cointegration tests allowing for breaks and cross-section dependence based on the Continuos-Path Block bootstrap. Simulation evidence shows that the proposed panel tests have satisfactory size and power properties, hence improving considerably on asymptotic tests applied to individual series. As an empirical illustration we examined investment and saving for a panel of 14 European countries over the 1960-2002 period, finding, contrary the results of most individual tests, that the hypothesis of a long-run relationship is compatible with the data.

Keywords: Panel cointegration, continuos-path block bootstrap, breaks, Feldstein-Horioka Puzzle.

JEL codes: C23, C15

October 2006

1 Introduction[1]

The literature on non-stationary panels is rapidly growing, with the most recent contributions aiming at tackling the issues of dependent panels and structural breaks (for a recent review see Breitung and Pesaran, 2006). Restricting the attention to cointegration testing, Gutierrez (2005), retaining the cross-section independence assumption, proposes to combine the p-values of Gregory and Hansen (1996) cointegration tests wit breaks computed for the individual units, while Westerlund (2006a,b) derives panel cointegration tests allowing for breaks in the deterministic kernel (level and trend) of the cointegration regression. The attempts to develop panel cointegration tests allowing for both breaks and cross-section dependence are not many. Acknowledging the importance of the dependence issue, Westerlund (2006b) proposes a bootstrap procedure complementing the asymptotic test mentioned above. However, some caution is suggested by the fact that it is based upon simple resampling of the FMOLS or DOLS cointegrating residuals, weakly dependent if cointegration holds and non-stationary if it does not[2]. Taking a completely different route, both Banerjee and Carrion-i-Silvestre (2006) and Westerlund and Edgerton (2006b) tackle the problem modelling the dependence across units with a common factor approach. Unfortunately, in both cases some restrictive assumptions on either the breaks or the form of cross-section dependence are needed. More precisely, Banerjee and Carrion-I-Silvestre's asymptotics requires the break to fall in the same period in all units, while Westerlund and Edgerton's excludes any sort of short- and long-run dependence of the explanatory variables across units. Further, in both cases large sample sizes are required. Westerlund and Edgerton's LM test has acceptable power for $T = 200$ and $N = 20$, but it is disappointing (power generally lower than 50%) for $T = 100$. In Banerjee and Carrion-I-Silvestre's framework the usual single-equation definition of cointegration (stationary residuals in the cointegrating equation) is accepted if the null of non-stationarity is rejected both for the estimated common factor and the idiosyncratic residuals, an event which in their simulations has a frequency generally much lower than 50% for rather large sample sizes such as $T = 100$ and $N = 40$.

Hence, we must conclude that a panel cointegration test fully robust to possibly heterogenous breaks and cross-section dependence and with an ac-

[1]Financial support from the Department of Statistics of the University of Naples Federico II, University of Rome "La Sapienza" and MIUR is gratefully acknowledged. We are grateful to Anindya Banerjee and Josep Carrion-i-Silvestre for kindly providing the investment and savings data set. Correspondence to: s.fachin@caspur.it, fdiiorio@unina.it.

[2]Incidentally, very much the same remark applies to the panel cointegration bootstrap test proposed in Westerlund and Edgerton (2006a), where an AR model is fitted to the cointegration residuals in order to carry out the sieve bootstrap. If cointegration does not hold the procedure hinges crucially upon precise estimation of the unit root, a notoriously difficult task (on this issue, see Fachin, 2004).

ceptable small sample performance is not available yet. Building on Fachin (2006), our proposal is to construct a panel generalisation of the Gregory and Hansen (1996), henceforth GH, test exploiting Paparoditis and Politis' (2001) Continuous-Path Block Bootstrap (CBB). As we will see, the proposed procedure can account for any form of dependence and, to some extent, heterogenous breaks, delivering satisfactory small sample size and power properties. Since on the basis of GH's simulation evidence their test in small samples ($T = 50$) appears to suffer from both slight positive size bias and low power (less than 50%), the panel extension may prove of considerable help in applied work.

We shall now in section 2 introduce the set-up and outline the testing procedure and in section 3 present the design and results of a Monte Carlo experiment. An empirical illustration on the relationship between the investment/GDP and savings/GDP ratios, the so-called Feldstein-Horioka puzzle, already examined by Banerjee and Carrion-i-Silvestre (2004) and Gutierrez (2005), is presented in section 4. Some conclusions and suggestions for future research are finally discussed in section 5.

2 Bootstrap Panel Cointegration Testing with Breaks: Set-up

Let us consider for simplicity a standard bivariate panel cointegration set-up, with the right- and left-handside variables, denoted as usual by X and Y, observed over T time periods and N units. In each unit X and Y are linked by a linear, not necessarily cointegrating, relationship with a slope break in period t_i^b. This set-up, essentially a generalisation of the Engle and Granger (1987) classical bivariate Data Generation Process (DGP) to the case of dependent panels, has been used by Fachin (2006), and is very similar to that considered by Kao (1999).

$$x_{it} = u_{it}^x \tag{1a}$$

$$y_{it} = \begin{cases} \mu_{0i} + \beta_0 x_{it} + u_{it}^y, \ t \leq t_i^b \\ \mu_{0i} + \beta_1 x_{it} + u_{it}^y, \ t > t_i^b \end{cases} \tag{1b}$$

where $i = 1, \dots, N$, $t = 1, \dots, T$, and both errors u^j, $j = x, y$, are assumed to be the linear combinatio of a common component, $f^j, j = x, y$, and an idiosyncratic one, ϵ^j, $j = x, y$,:

$$\begin{cases} u_{it}^x = \gamma_i^x f_t^x + \epsilon_{it}^x \\ u_{it}^y = \gamma_i^y f_t^y + \epsilon_{it}^y \end{cases} \tag{2}$$

$$\begin{cases} \epsilon_{it}^x = e_{it}^x + \theta \\ \epsilon_{it}^y = \phi_i \epsilon_{it-1}^y + e_{it}^y \end{cases} \tag{3}$$

where the $e_{it}^{j} \sim N(0, \sigma_{ij}^{2})$, $j = x, y$, are white noise.

To have a fully general dependence set-up we assume f^x to be non-stationary, thereby inducing cointegration across units in the X's as well as in the Y's in case of cointegration between X_i and Y_i ($|\phi_i| < 1$)[3]. In this case the common factor f^y is stationary and induces short-run correlation between the Y's across units, while when cointegration does not hold ($\phi_i = 1$) it is non-stationary so that cross-units cointegration in the Y's still holds.

GH proposed to test for no cointegration allowing for breaks at unknown periods by taking the minimum of a sequence of tests computed for all possible breakpoints. A panel test along the lines of Pedroni's (1999) group mean test may be simply defined as the mean (or some robust statistic, such as the median or an $\alpha-$trimmed mean) of the individual statistics[4] computed for the individual units.

Similarly to the case of panel cointegration tests, the bootstrap is a natural candidate for solving the problem of inference under the general set-up of dependent units. To this end, we need to design a resampling scheme delivering pseudodata:

(*i*) reproducing both the autocorrelation and cross-correlation properties of the data;

(*ii*) accounting for the break;

(*iii*) obeying the null hypothesis of no cointegration.

As mentioned above, as in Fachin (2006), the algorithm we propose is based upon Paparoditis and Politis (2001) "Continuous-Path Block Bootstrap" (CBB), a resampling scheme designed to construct non-stationary pseudo-series from data of the same type. Requirement (*i*) is readily satisfied by resampling together all the units (columns), while (*iii*) by resampling separately the X's and the Y's, thus constructing pseudodata matching a given time observation for the former with a different one for the latter. We thus need to discuss only (*ii*). The key point here is that our implementation of the CBB, which assumes that the series are non-stationary with a constant drift (which can then be consistently estimated). Now, if X is $I(1)$ with drift, a slope change in (1b) is easily seen to imply a change in the mean of the $\Delta Y's$. Therefore, the CBB must be applied separately to the time samples before and after the breakpoint t_i^b as estimated by some $\widehat{t}_i^b$. However, since to satisfy (*i*) the resampling scheme is to applied to all

[3]This set-up is likely to be representive of many empirical applications: for instance, in the case of regional consumption and income f^x will be the national stochastic GDP trend causing income, and hence consumption, to be cointegrated across regions.

[4]Although GH examined the most common residual-based no cointegration statistics, Phillips' Z_α and Z_t and the ADF, we will concentrate, without much loss of generality but some computational advantage, only on the latter.

columns (units) for a given block of rows (time observations), we need to impose the constrain $\widehat{t}_i^b = \widehat{t}^b \; \forall i$. A natural choice is $\widehat{t}^b = median(\widehat{t}_1^b, \ldots, \widehat{t}_N^b)$. To account for estimation error in $\widehat{t}^b$, and, to some extent, heterogeneous timing of the breaks, we will take the pseudodata of the block centred on $\widehat{t}^b$ to be equal to the actual data.

Summing up, the bootstrap testing procedure we propose to implement is the following:

1. compute for each unit $i = 1, \ldots, N$, of the data set under study, $\{X_1 \ldots X_N, Y_1 \ldots Y_N\}_{t=1}^T$ the no cointegration statistic $ADF_i^* = \inf ADF_i(t^b)$ for $t^b[\tau_1 T, \tau_2 T]$; the trimming coefficients τ_1, τ_2 must chosen to ensure stability of the statistic at the endpoints;

2. compute the Group statistic as, *e.g.*, the mean or median of the N individual statistics: $\widehat{G}_m = \sum_{i=1}^N \widehat{ADF}_i / N$, or $\widehat{G}_{me} = median(\widehat{ADF}_1, \ldots, \widehat{ADF}_N)$;

3. estimate the breakpoints for each unit; a natural choice is $\widehat{t}_i^b = \arg\min(ADF_i(t^b))$;

4. estimate a common breakpoint $\widehat{t}^b$, e.g., $\widehat{t}^b = median(\widehat{t}_1^b, \ldots, \widehat{t}_N^b)$;

5. apply the CBB with block length s (assumed to be even; the choice of s is discussed in more detail below) separately to the X's and the Y's over the intervals $[1, \ldots, \widehat{t}^b - \frac{s}{2} - 1]$ and $[\widehat{t}^b + \frac{s}{2} + 1, \ldots, T]$, obtaining four matrices of pseudodata: $\{X_1^* \ldots X_N^*\}_{t=1}^{\widehat{t}^b - \frac{s}{2} - 1}$, $\{X_1^* \ldots X_N^*\}_{t=\widehat{t}^b + \frac{s}{2} + 1}^{T^*}$, $\{Y_1^* \ldots Y_N^*\}_{t=1}^{\widehat{t}^b - \frac{s}{2} - 1}$, $\{Y_1^* \ldots Y_N^*\}_{t=\widehat{t}^b + \frac{s}{2} + 1}^{T^*}$; note that $T^* < T$, as some observations are lost in the chaining;

6. construct the pseudo-dataset for the entire sample, joining the pseudodata constructed before and after the break with a central block of actual data: $\mathbf{X}_i^* = \left[x_{i1}^* \ldots x_{i\widehat{t}^b - \frac{s}{2} - 1}^* x_{i\widehat{t}^b - \frac{s}{2}} \ldots x_{i\widehat{t}^b + \frac{s}{2}} x_{i\widehat{t}^b + \frac{s}{2} + 1}^* \ldots x_{iT^*}^* \right]'$, $i = 1, \ldots, N$, and analogously for the Y's;

7. compute the Group statistics G^* (steps 1-2) for the pseudo-data set, $\{X_1^* \ldots X_N^*, Y_1^* \ldots Y_N^*\}_{t=1}^{T^*}$;

8. repeat steps (5) to (7) a large number (say, B) of times;

9. compute the boostrap significance level; assuming that the rejection region is the left tail of the distribution, $p^* = prop(G^* < \widehat{G})$.

Although exploratory simulations showed the results to be quite robust to the choice of block length, in principle this is a critical point of the algorithm. Here for computational convenience we applied a simple rule-of-thumb, fixing it at $T/10$. In future work we plan to implement Politis and White's (2003) algorithm.

3 Monte Carlo Experiment

3.1 Design

The simulation experiment, based upon the DGP (1a)-(3) is obviously quite complex, and the tests to be evaluated computationally demanding (this issue is discussed in more detail below). Rather than aiming at the unfeasible task of a complete design we will define as a base case an empirically relevant set-up and then explore a few interesting variations. Let us first discuss the design parameters common to all experiments. Similarly to GH we take the model as correctly specified; with no loss of generality we set both constant and slope to 2 before the break, with the slope halved after it. The factor loadings are chosen so to ensure substantial cross-correlation in the Y's and cross-cointegration of the the X's: $\gamma_i^j \sim Uniform(-1, 6)\ \forall i, j$. In the power simulations we consider on the average rather slow adjustment to equilibrium, but allowing for some heterogeneity: $\phi_i \sim Uniform(0.6, 0.8)$. Finally, further heterogeneity across units is given by the noise variances: $\sigma_{ij}^2 \sim Uniform(0.5, 1.5),\ \ j = x, y$. Given the rather short time series analysed in most experiments, in order to ensure computational stability we fixed the trimming coefficient at 25%. Let us now examine the various experiments (four altogether) in some detail.

1. *Base case*: $T = 40$, $N = 5, 10, 20, 40$; break date Uniform over units in $[0.5T \pm 3] = [17, 23]$. The time span is medium in terms of annual data, but pretty small with a quarterly frequency. It may thus be considered relevant for actual empirical applications (note that it is smaller than those considered in the simulation studies on the other cointegration tests with break available in the literature). The breaks are distributed over six periods centred in the middle of the time sample, with the testing procedure searching over the interval $[10, 30]$.

2. *Late break*: as Base case, except break date Uniform over units in $[0.75T \pm 3] = [27, 33]$. Note that this a demanding set-up, as half of the interval in which the breaks may fall is outside the searching interval.

A critical point of the bootstrap algorithm described above is the partition of the samples before and after the estimated breakpoint, constrained to be homogenous across units at the median of the individual estimates. This is intuitively acceptable if we assume all units to be affected by breaks stemming from a common cause. However, even assuming each unit to be affected by at most one break over the period of interest, these may be widely disperse over units, for instance because they stem from different causes. The following case is designed to investigate this scenario:

3. *Twin breaks:* as Base case, but break date Uniform in $[0.3T \pm 3] = [9, 15]$ in the odd-numbered units, and in $[0.6T \pm 3] = [25, 31]$ in the even-numbered ones. Note that in both cases the break may take place marginally outside the search interval, $[10, 30]$.

To evaluate the improvements (in terms of both power gains and reduction in size bias) which could be expected by moving from a standard time series to a panel set-up we also computed the average rejection rates of the asymptotic tests based on GH asymptotic critical values computed for all individual units involved in each experiment. Note that the comparison between the average performance of the asymptotic test on individual series and that of the panel tests with a smaller number of units should be taken as merely suggestive of a pattern, as the units involved are not the same.

The last issue to be discussed is the number the number of Monte Carlo replications. In all simulation exercises this is chosen trying to strike a balance between the contrasting requirements of precision in the results and control of the cost and time scale of the experiment. Here this balance is particularly difficult to achieve, as the panel structure of the data and the recursive nature of the statistics evaluated make the simulations computationally very demanding: the number of loops executed is the product of bootstrap redrawings, units, periods included in the searching interval, and number of Monte Carlo replications. With 1000 redrawings, 40 units and search over 20 periods the product of first three terms is equal to 800.000; fixing the Monte Carlo replications to 500 will thus require the execution of 400 million loops for each experiment, with an approximate confidence interval $p \pm 2\sqrt{p(1-p)/500}$ for, *e.g.*, 5% equal to $[3.1\%, 6.9\%]$. Unfortunately, costs and precision increase at different rates, with the former dominating the latter: doubling the number of replications to 1000 will increase the number of loops to the considerable figure of 800 million while producing only a marginal precision gain, as the interval becomes $[3.6\%, 6.4\%]$. We thus decided that 500 replications was a reasonable choice.

3.2 Results

The results are reported in Tables 1-4 below. In the Base case ($T = 40$, $N = 5, 10, 20, 40$), consistently with GH we find that the asymptotic test on individual series is considerably biased against the null. On the contrary, both the mean and median panel cointegration bootstrap tests are somehow slightly undersized but, taking into account Monte Carlo estimation error, essentially converge to nominal levels. Power is always high both for the mean and median tests provided the significance level and cross-section sample size are not too small (in practice, in our simulations $\alpha \geq 5\%$ or $N \geq 10$). Hence, exploiting the panel dimension using the proposed bootstrap procedure seems to provide a reliable solution for cointegrating testing with small time samples.

When either the breaks take place at the end of the sample (possibily outside the searching interval) or are clustered in two different intervals we find that the rejection rates of the bootstrap tests fall under both H_0 and H_1, so that they are strongly undersized; taking into account the size bias, power appears nevertheless acceptable.

Table 1

Base Case: $T = 40$, N from 5 to 40

Rejection Rates$\times 100$

	N								
	1	5	10	20	40	5	10	20	40
α	*Asy*		*Boot-Mean*				*Boot-Median*		
				A. Size: $\phi_i = 1\ \forall i$					
1.0	22.0	0.0	0.0	0.3	0.8	0.0	1.0	1.0	1.0
5.0	34.5	0.5	3.8	2.0	3.0	3.0	9.8	4.0	3.3
10.0	42.5	1.5	14.0	5.0	7.8	5.0	20.5	8.0	7.8
				B. Power: $\phi_i \sim Uniform(0.6, 0.8)$					
1.0	70.5	45.0	75.0	87.3	94.5	41.3	74.0	90.3	96.5
5.0	83.0	79.7	96.5	98.5	99.8	73.3	96.5	99.0	100.0
10.0	87.5	90.0	99.5	100.0	100.0	87.0	98.8	100.0	100.0

DGP: eqs. (1a)-(3), $t_i^b \sim Uniform(17, 23)$;
search interval: $[10, 30]$;
Asy: average rejection rates of invidual no cointegration tests over all 40 units, Gregory and Hansen (1996) asymptotic critical values;
Boot-mean/median: bootstrap test on the mean/median across units of the no cointegration statistics;
Bootstrap: 1000 redrawings, block size $T/10$;
Montecarlo: 500 replications.

Table 2

Late break: $T = 40$, N from 5 to 40

Rejection Rates$\times 100$

	N								
	1	5	10	20	40	5	10	20	40
α	*Asy*		*Boot-Mean*				*Boot-Median*		
				A. Size: $\phi_i = 1\ \forall i$					
1.0	11.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5.0	21.5	0.0	0.0	0.0	0.3	0.5	0.5	0.5	0.5
10.0	28.8	0.5	1.5	0.5	1.3	1.3	1.5	1.0	0.8
				B. Power: $\phi_i \sim Uniform(0.6, 0.8)$					
1.0	59.3	28.3	38.2	72.2	87.8	21.0	32.5	67.0	86.0
5.0	74.0	61.3	81.3	94.3	98.5	52.0	70.0	92.0	98.5
10.0	80.8	78.8	94.0	98.3	100.0	69.2	89.3	97.5	99.5

$t_i^b \sim Uniform(27, 33)$; search interval: $[10, 30]$.

All other abbreviations and definitions: see table 1.

Table 3

Twin breaks: $T = 40$, N from 5 to 40

Rejection Rates$\times 100$

	N								
	1	5	10	20	40	5	10	20	40
α	*Asy*		*Boot-Mean*				*Boot-Median*		
				A. Size: $\phi_i = 1\ \forall i$					
1.0	8.8	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5.0	20.5	1.3	0.0	0.0	0.3	0.5	0.0	0.0	0.3
10.0	29.3	1.5	0.0	1.3	0.3	1.8	0.0	1.5	1.3
				B. Power: $\phi_i \sim Uniform(0.6, 0.8)$					
1.0	62.8	19.0	30.5	67.8	70.3	9.3	29.3	70.5	76.0
5.0	76.8	50.0	64.3	87.0	90.3	28.3	57.0	87.5	90.8
10.0	83.8	69.0	79.0	93.0	94.0	40.5	68.8	93.7	94.8

$$t_i^b \sim \begin{cases} Uniform(9, 15) & i = 1, 3, \ldots, N-1 \\ Uniform(25, 31) & i = 2, 4, \ldots, N \end{cases}$$

search interval: $[10, 30]$;

All other abbreviations and definitions: see table 1.

4 Empirical illustration: the Feldstein-Horioka Puzzle

One the major empirical puzzles of contemporary macroeconomics (six altogether according to Obstfeld and Rogoff, 2000) is with no doubt the evidence supporting the existence of a long-run link between the investment (I) and savings (S) to GDP (Y) ratios in advanced economies, where high capital mobility may allow the current account to be unbalanced for long periods. Banerjee and Carrion-i-Silvestre (2004) investigated the issue on a data set including 14 European economies (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Portugal, Spain, Sweden, UK) over the period 1960-2002 using panel cointegration tests allowing for a single break in the cointegrating coefficients. Although Banerjee and Carrion-i-Silvestre were not able to reach a clear conclusion, their findings appear on the whole rather favourable to the cointegration-with-break hypothesis, indeed plausible from the plots reported in Figs. 1-2. However, there are reasons to have some suspects on the reliability of these results, as the bootstrap procedure used (abandoned in the revised version of the paper, Banerjee and Carrion-i-Silvestre, 2006) implied fitting an AR model to a MA process with a unit root under no cointegration. Confirming the doubts on their empirical limits, neither of the cointegration tests with breaks available in the literature may be applied. Since savings (the right-hand side variable) are generally correlated in the short-run, and in some cases cointegrated, across economies (results not reported here available on request) the assumptions underlying Westerlund and Edgerton's (2006b) test are not satisfied, and the small sample size as well as the likely presence of heterogenous breaks advises against use of Banerjee and Carrion-i-Silvestre's (2006) procedure. It is thus of some interest to find out the results of applying the procedure proposed in this paper.

As a first step of our analysis we computed ADF tests to check the properties of the series, choosing the order of the autoregression on the basis of the significance of the last lag (maximum four). The results, reported in Table 4, suggest that the Savings/GDP ratio may be stationary in Finland and Portugal. We thus excluded these two countries and proceed to compute the individual and panel cointegration tests. Examining the individual statistics (Table 5; since essentially similar results have been obtained with 25% and 12.5% trimming we report only the latter) we find that, consistently with theoretical expectations and somehow contrary to those formed on the basis of visual inspection of the plots, only in five countries (Netherlands, Denmark, France, Spain and UK) the $Min(ADF)$ tests reject the null hypothesis of no cointegration according to the asymptotic critical values. However, the failure to reject is likely to be a mere consequence of low power: the bootstrap panel cointegration tests, reported in table 6, provided rather strong evidence in support of the hypothesis that in the majority of cases the relationship between investements and savings is a long-run equilibrium which might have undergone a structural break during the period of interest.

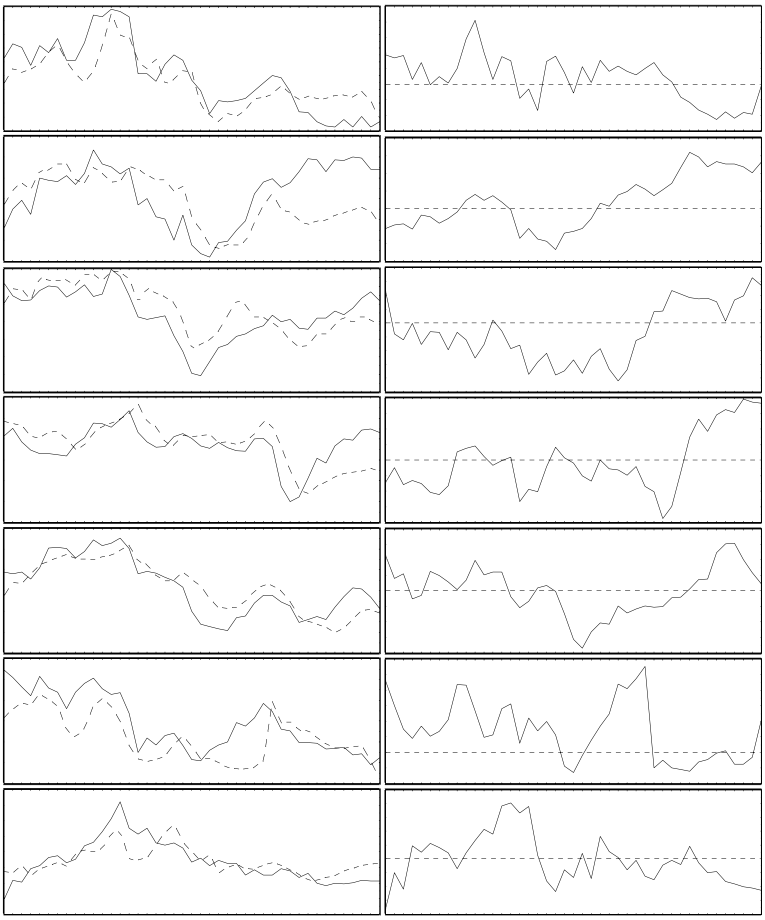

Fig. 1. Savings (S) and Investments (I) to GDP (Y) ratios dynamics, 1960-2002. Top to bottom: Austria, Belgium, Denmark, Finland, France, Germany, Greece. Left Column: S/Y (solid line) and I/Y (dotted line). Right Column: Current Account/GDP = $(S-I)/Y$ (solid line) and zero (dotted line).

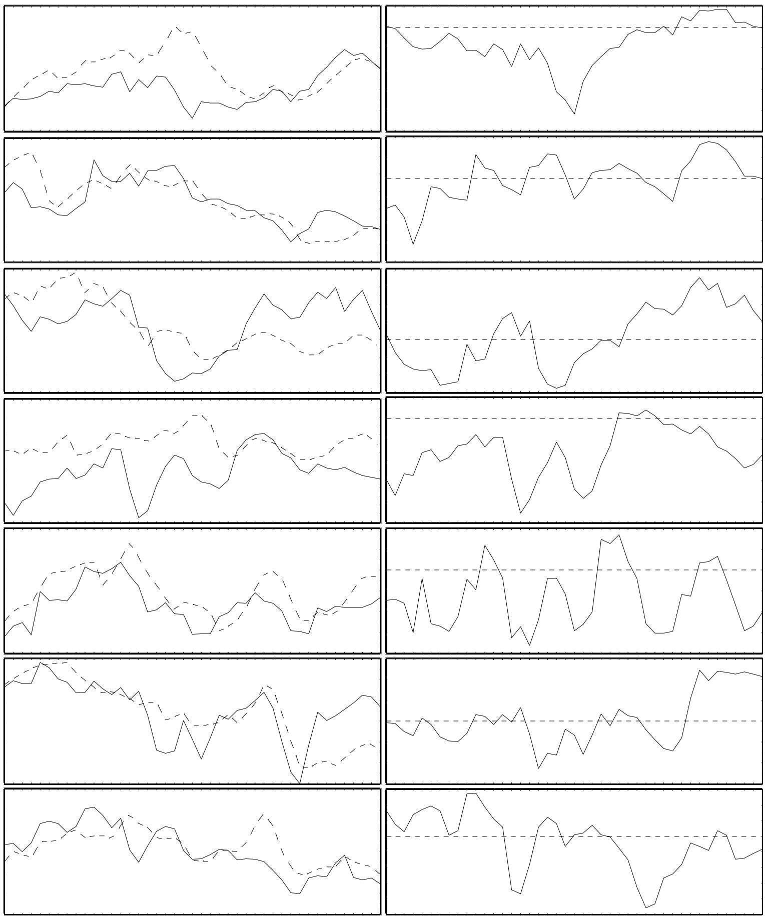

Fig. 2. Savings (S) and Investments (I) to GDP (Y) ratios dynamics, 1960-2002. Top to bottom: Ireland, Italy, Netherlands, Portugal, Spain, Sweden, UK.Left Column: S/Y (solid line) and I/Y (dotted line). Right Column: Current Account/GDP = $(S-I)/Y$ (solid line) and zero (dotted line).

Table 4

Investment and Savings to GDP ratios: ADF Unit Root Tests

	Austria	Belgium	Denmark	Finland	France	Germany	Greece
I	−1.31	−1.57	−1.42	−2.16	−1.26	−2.00	−2.13
S	−0.62	−1.46	−2.03	−3.34*	−1.49	−1.66	−1.11
	Ireland	Italy	Netherlands	Portugal	Spain	Sweden	UK
I	−2.34	−1.32	−1.11	−3.42	−2.93	−1.54	−2.16
S	−1.85	−1.29	−1.94	−4.87**	−2.42	−2.06	−0.79

*: significant at 5%; **: 1%.

Table 5

Investment and Savings, 1960-2002

Min(ADF) Cointegration Tests with Break

and estimated breakpoints

	Austria	Belgium	Denmark	France	Germany	Greece
Min(ADF)	-4.28	-4.44	-5.96***	-4.95***	-4.22	-4.45
Break	1991	1982	1990	1994	1994	1999
	Ireland	Italy	Netherlands	Spain	Sweden	UK
Min(ADF)	-3.36	-4.25	-5.50***	-6.01***	-3.97	-4.92*
Break	1987	1981	1976	1979	1970	1996

trimming: 12.5% (searching interval: 1965-1998); *break*: argmin(ADF);
critical values: $1\% : -5.47; 5\% : -4.95; 10\% : -4.68$.
*: significant at 10%; **: 5%;***: 1%.%: -5.47 ***

Table 6

Investment and Savings, 1960-2002

Bootstrap Panel Cointegration Tests

p-values $\times 100$

Trimming	$Mean\ (p - value)$	$Median\ (p - value)$
12.5%	-4.69 (0.01)	-4.44 (0.10)
25%	-4.62 (0.03)	-4.41 (0.14)

panel: Austria, Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Netherlands, Spain, Sweden, UK;
Mean/Median: mean/median of the individual Min(ADF) statistics;
bootstrap: 1000 redrawings.

5 Conclusions

Testing panel cointegration in dependent panels allowing for breaks at unknown periods is a challenging task, as two forms of dependence (between the tests computed with the break fixed at different periods and for different units) must be accounted for. Building upon Fachin (2006), in this paper we propose to solve this problem using the bootstrap. Simulation results suggest that the proposed panel testing procedures improve considerably on the performances of pure time series Gregory and Hansen (1996) tests. This findings are confirmed by an empirical application to the Feldstein-Horioka Investment-Savings Puzzle for a panel of 12 european countries. While, consistently to theoretical expectations and somehow contrary to those formed on the basis of visual inspection of the plots, the majority of individual Gregory and Hansen tests fail to reject the null of no cointegration, the evidence of the panel tests is rather strongly against it. From Table 5 we can see that the estimated break periods appear widely dispersed between the late '70's (one in 1970) and the late 90's. Since from our simulations we know that in this circumstances the panel cointegration tests may have a strong negative size bias we should adopt rather high nominal significance levels; hence, three of the four p-values in reported in Table 6 can be regarded as highly significant, with the last one close to the rejection region.

6 References

Banerjee, A., Carrion-i-Silvestre J.L. (2004) "Breaking Panel Cointegration", *Mimeo*, European University Institute.

Banerjee, A., Carrion-i-Silvestre J.L. (2006) "Cointegration in Panel Data With Breaks and Cross-Section Dependence", *Working Paper Series* n. 591, European Central Bank.

Breitung, J., and Pesaran M. H. (2006) "Unit Roots and Cointegration in Panels" Forthcoming in L. Matyas, and P. Sevestre (Eds.), *The Econometrics of Panel Data*, Kluwer Academic Publishers, Dordrecht (NL).

Engle, R.F and C.W.J. Granger (1987) "Co-integration and Error Correction: Representation, Estimation and Testing", *Econometrica*, 55, 251-176.

Fachin, S. (2004) "Bootstrap inference on Fully Modified Estimates of Cointegrating Coefficients: A Comment", *Economics Bulletin*, Vol. 3, No. 13 pp. 1-8.

Fachin, S. (2006) "Long-Run Trends in Internal Migrations in Italy: a Study in Panel Cointegration with Dependent Units", *Journal of Applied Econometrics*, forthcoming.

Gutierrez, L. (2005) "Tests for Cointegration in Panels with Regime Shifts", *Mimeo,* Department of Agricultural Economics, University of Sassari.

Gregory, A.W. and Hansen, B.E. (1996) "Residual-based tests for cointegration in model with regime shifts", *Journal of Econometrics*, 70, 99-126.

Kao, C. (1999) "Spurious regression and residual-based tests for cointegration in panel data" *Journal of Econometrics,* 90,1-44.

Obstfeld, M. and Rogoff, K. (2000) "The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?" *NBER Working Paper Series* n. 7777.

Paparoditis E, Politis DN. (2001) "The Continuous-Path Block Bootstrap" In *Asymptotics in Statistics and Probability. Papers in honor of George Roussas.* Madan Puri (ed.). VSP Publications: Zeist (NL).

Pedroni, P. (1999) "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors", *Oxford Bulletin of Economics and Statistics*, 61, 653-670.

Politis, D. N., White, H. (2003) "Automatic Block-Length Selection for the Dependent Bootstrap", *Econometric Reviews*, 23, 53-70.

Westerlund, J. (2006a) "Testing for panel cointegration with a level break", *Economics Letters*, 91, 27-33.

Westerlund, J. (2006b) "Testing for Panel Cointegration with Multiple Structural Breaks", *Oxford Bulletin of Economics and Statistics*, 68, 101-132.

Westerlund, J., Edgerton, D. (2006a) "A Panel Bootstrap Cointegration Test" *Mimeo*, Lund University.

Westerlund, J., Edgerton, D. (2006b) "Simple Tests for Cointegration in Dependent Panels with Structural Breaks" *Mimeo*, Lund University.