Testing for Cointegration with Breaks in Dependent Panels, with an application to the Feldstein-Horioka puzzle

Francesca Di Iorio
University of Naples Federico II

Stefano Fachin
University of Rome "La Sapienza"

Abstract

We propose a bootstrap panel cointegration test with breaks robust to cross-section dependence and use it to examine the long-run relationship between investment and savings (the Feldstein-Horioka puzzle) in a panel of 18 OECD countries over the 1970-2007 period. The test, shown by simulation to have good size and power properties, suggests that for the panel as a whole the evidence in favour of an investment-savings long-run relationship is rather weak and due mainly to a subset of the economies examined.

Keywords: Panel cointegration, stationary bootstrap, breaks, Feldstein-Horioka Puzzle.
JEL codes: C23, C15

Revised February 2010
Preliminary and uncomplete - Please do not quote without permission

1 Introduction[1]

The evidence suggesting the existence of a long-run link between the investment and savings to GDP ratios in many advanced economies, christened Feldstein-Horioka (FH) puzzle after Feldstein and Horioka (1980), is at the same time one of the six major empirical puzzles of contemporary macroeconomics (Obstfeld and Rogoff, 2000) and a very good example of the high demands posed on testing procedure by applied work. As the issue is if there is *widespread* evidence of such a relationship, from the empirical point of view the question is naturally tackled in a panel cointegration framework. Since the puzzle comes from the fact that high capital mobility allows the current account to be unbalanced (and thus investment to be larger than savings) for long periods, we need not to overlook the fact that capital movements are a strongly regulated matter likely to be subject to structural breaks. In other terms, the trend towards a greater worldwide financial integration is likely to have caused the link between savings and investments to become weaker over time (Frankel, 1992). For a sample of European economies possible breakpoints in the savings-investments relationship include the late '70's and early 1980's, when barriers were lifted in many countries (see e.g., Özmen and Parmaks, 2003, Juselius, 2006), and the late 1990's, when exchange rate risks were greatly lessened by the introduction of the Euro. Thus, careful testing of the FH puzzle requires a panel cointegration test allowing for structural breaks. Further, to minimize the risk of multiple breaks, which would complicate the analysis considerably, the test should deliver acceptable performances with small or moderate time samples. Finally, short- and long-run dependence across units is to be expected for both the right-and left-hand side variables (respectively, savings and investment).

Now, it is immediately seen that none of the currently available tests, including those applied by Banerjee and Carrion-i-Silvestre (2004) and Gutierrez (2009) in their efforts to investigate the FH puzzle, fully satisfies all these requirements. More specifically, the tests by Gutierrez (2009) and Westerlund (2006a,b) assume full cross-section independence[2], and are thus out of question. Both Banerjee and Carrion-i-Silvestre (2006) and Westerlund and Edgerton (2008) allow only for special types common factors. The former assume the common factors of the right-hand side variables to be orthogonal

[1] This is completely revised version of a paper previously circulated with the same title. 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 helpful discussions.
Correspondence to: s.fachin@caspur.it, fdiiorio@unina.it.

[2] Westerlund (2006b) does propose also a robust bootstrap procedure. This, however, does not appear to be advisable, as it entails simple resampling of the FMOLS or DOLS cointegrating residuals, weakly dependent (if cointegration holds) or even non-stationary (if it does not).

to those of the dependent variable, while the latter allow for common factors in the cointegrating residuals only (not in the variables). Even assuming either of these conditions are satisfied, to obtain good performances really large sample sizes seem to be needed. For instance, the power of Westerlund and Edgerton's LM test with 20 cross-section units is acceptable for $T = 200$, but disappointing (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 hypothesis 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% even with rather large sample sizes such as $T = 100$ and 40 cross-section units

We must thus conclude that a panel cointegration test with breaks and full cross-section dependence delivering an acceptable small sample performance, hence suitable for our aim of investigating the FH puzzle, is not available yet. In this paper we tackle the problem extending to the case of unknown breaks the procedure put forth in Di Iorio and Fachin (2008), hereafter DIF, who report promising results for a panel cointegration test with a break at a known date.

As we will see, the proposed procedure, based upon Paparoditis Politis and Parker's (2006) Residual-based Stationary Bootstrap (RSB) can account for fully general forms of dependence and delivers satisfactory small sample size and power properties. Since the small sample power of the standard Gregory and Hansen (1996) cointegration tests with breaks tend to be rather poor, our test may be of considerable help in applied work. It must be stressed since the interest here is on small sample testing and more general dependence structures the proposed procedure is seen as complementary, rather than alternative, to those based on factor methods; no comparison will be therefore be carried out.

We shall now first (section 2) introduce the set-up and outline the testing procedure, then (section 3) present the design and results of a Monte Carlo experiment, and examine (section 4) the FH puzzle in a panel of european countries. Some conclusions are finally drawn 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 N units and T time periods, as usual indexed respectively by i and t. In each unit X and Y are linked by a linear, not necessarily cointegrating, relationship with a break in period t_i^b:

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

A panel cointegration test allowing for breaks may be defined very simply, following Pedroni's (1999) group mean test approach, as a summary statistic of the cointegration statistics with break computed for the individual units[3]. This problem has been first tackled by Gregory and Hansen (1996), who proposed to compute a no-cointegration statistic $\theta(t_i^b)$ on the residuals of model (1) for all possible break points t_i^b and, assuming the rejection region is the left tail (as in the case of the popular ADF and Z tests), take the minimum. Hence, the cointegration with break statistic is

$$\Theta = \underset{t_i^b \in [\delta T, (1-\delta)T]}{Min} (\theta(t_i^b))$$

where the trimming factor δ is chosen to ensure computational stability, with 0.15 or 0.20 popular choices. The break point is thus implicitly estimated as

$$\widehat{t_i^b} = \underset{t_i^b \in [\delta T, (1-\delta)T]}{Arg \min} (\theta(t_i^b)). \tag{2}$$

Another option, favoured by Westerlund and Edgerton (2008) and which we will follow, is to estimate the breakpoint on the basis of a least square criterion (which for stationary variables is consistent even under multiple breaks):

$$\widehat{t_i^b} = \underset{t_i^b \in [\delta T, (1-\delta)T]}{Arg \min} (\sum_{t=1}^{T} \epsilon_{it}^2(t_i^b)). \tag{3}$$

where the notation $\epsilon_{it}(t_i^b)$ for the residuals emphasises their dependence on the breakpoint used in the estimation of model (1). For each unit the set of residuals $\left\{\epsilon_{it}(\widehat{t_i^b})\right\}_{t=1}^{T}$ is then by definition optimal in a least square sense, and can be used to compute a no-cointegration test with break. In fact, let us consider a first order autoregressive equation for the cointegrating residuals:

$$\epsilon_{it} = \rho_i \epsilon_{it-1} + \nu_{it}. \tag{4}$$

When H_0 : "no cointegration" holds $\rho_i = 1$,while when it does not $|\rho_i| < 1$. The hypothesis of no cointegration is then equivalent to H_0: $\rho_i = 1$, and that of no panel cointegration as the same hypothesis for mean or median of this individual statistics, with the latter reflecting more closely the usual

[3]For simplicity we will refer to a summary statistic of the individual cointegration tests as a "panel test", although in Pedroni's terminology this term is reserved for tests obtained imposing an homogeneity assumption.

definition of the panel null hypothesis (no cointegration in the majority of the units; for a detailed discussion see DIF).

Two important remarks are in order here.

First, (4) is *not* a model of the cointegrating residuals; its purpose is only to define a parameter expressing the null hypothesis of interest. For computational convenience we will not examine Z tests, and to evaluate the impact of studentising we shall use both the popular ADF statistic (ADF_i) and the coefficient statistic $\tau_i = T(\rho_i - 1)$, used *e.g.*, by Palm, Smeekes and Urbain (2008). The ADF lag length is chosen following Ng and Perron (1995).

Second, the $\nu_{it}'s$ are always stationary, either H_0 holds or not: they can thus be resampled via the stationary bootstrap. A bootstrap testing algorithm along the lines put forth in PPP and already exploited in DIF, may then proceed as follows:

1. Estimate the breaking cointegrating model (1), with $\widehat{t}_i^b$ given by (3), on the dataset $\{y_{it}, x_{it}\}$, $i = 1, \ldots, N$, obtaining for each unit estimates of the coefficients $(\widehat{\mu}_{ji}, \widehat{\beta}_{ji}, j = 0, 1)$ and of the optimal cointegrating residuals $\{\widehat{\epsilon}_{it}\}$;

2. Compute the N individual no cointegration statistics $\widehat{\theta}_i$, $\widehat{\theta} = \widehat{ADF}$ and $\widehat{\tau}$, on the basis of the the optimal cointegrating residuals $\left\{\widehat{\epsilon}_{it}(\widehat{t}_i^b)\right\}$;

3. Compute the Group statistics $\widehat{\theta}_{mean} = N^{-1}\sum_{i=1}^{N}\widehat{\theta}_i$, $\widehat{\theta}_{median} = median(\widehat{\theta}_1 \ldots \widehat{\theta}_N)$;

4. Compute $\widehat{\nu}_{it} = \widehat{\epsilon}_{it} - \widehat{\rho}_i\widehat{\epsilon}_{it-1}$, where $\{\widehat{\epsilon}_{it}\}$ are the estimated cointegrating residuals and $\widehat{\rho}_i$ is a consisten estimate (e.g., OLS) estimate of ρ_i;

5. Resample the series $\{\widehat{\nu}_{it}\}$ via the stationary bootstrap, obtaining $\{\nu_{it}^*\}$;

6. Cumulate $\{\nu_{it}^*\}$ obtaining pseudoresiduals $\{\epsilon_{it}^*\}$ obeying the null hypothesis of no cointegration;

7. For each unit i compute the pseudodata under the null hypothesis of no cointegration and break in $\widehat{t}_i^b$:

$$y_{it}^* = \begin{cases} \widehat{\mu}_{0i} + \widehat{\beta}_0 x_{it} + \epsilon_{it}^*, & t \le \widehat{t}_i^b \\ \widehat{\mu}_{1i} + \widehat{\beta}_1 x_{it} + \epsilon_{it}^*, & t > \widehat{t}_i^b \end{cases}$$

8. Using the datasets $\{y_{it}^*, x_{it}\}$ estimate the breaking cointegrating model (1) with breakpoints $\widehat{t}_i^{b*}$ given by (3), obtaining sets of estimates of the optimal cointegrating residuals $\{\widehat{\epsilon}_{it}^*\}$;

9. Compute the individual no cointegration statistics θ_i^*;

10. Compute the Group statistics θ_h^* ($h = mean, median$);

11. Repeat 5-11 B times;

12. Compute the boostrap significance level of the statistics: $p(\theta)^* = prop(\theta_h^* < \theta_h)$, $h = mean,\ median$.

Two remarks are in order here as well. The first, possibly obvious but nevertheless crucial, is that in step 8 the breakpoints are re-estimated for each bootstrap sample. Second, 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. While in future work we plan to investigate the issue in detail, here for computational convenience we fixed the block length at $T/10$, a simple choice which nevertheless delivered good results both in Paparoditis and Polits's (2003) and in our own exploratory simulations, and at $1.75\sqrt[3]{T}$ as in Palm, Urbain and Smeekes (2008). This rule yields block sizes respectively slightly larger than $T/10$ for rather small sample sizes (*e.g.*, 6 for $T = 40$), smaller for large sample sizes (*e.g.,* 10 for $T = 160$). For moderate sample sizes the two rules suggest approximately the same block sizes (in fact exactly the same for $T = 80$). Note that since we will not use optimal block sizes we will somehow underrate the properties of the proposed test.

3 Monte Carlo Experiment

3.1 Design

We will base our simulations on a Data Generation Process (DGP) which is essentially a generalisation to the case of dependent panels of the classical bivariate DGP adopted by, *e.g.*, Engle and Granger (1987) and Gonzalo (1994). It is very similar to that considered by Kao (1999), and it has been recently adopted by Gengenbach *et al.* (2006) and DIF. As in the previous discussion, we assume two variables linked by a linear, not necessarily cointegrating, relationship with a break in period t_i^b :

$$y_{it} = \begin{cases} \mu_{0i} + \beta_0 x_{it} + \epsilon_{it}^y, & t \le t_i^b \\ \mu_{1i} + \beta_1 x_{it} + \epsilon_{it}^y, & t > t_i^b \end{cases} \qquad (5)$$

$$\epsilon_{it}^y = \rho_i \epsilon_{it-1}^y + e_{it}^y, \qquad e_{it}^y \sim N(0, \sigma_{iy}^2) \qquad (6)$$

where $i = 1, \ldots, N$, $t = 1, \ldots, T$; When X_i and Y_i are not cointegrated $\rho_i = 1$, while $|\rho_i| < 1$ when instead they are; in the power simulations ρ_i will be generated as $Uniform(0.6, 0.8)$ across units to mimick a generally rather slow adjustment to equilibrium. To ensure some heterogeneity across units $\sigma_{iy}^2 \sim Uniform(0.5, 1.5)$, while with no loss of generality $\mu_{0i} = \beta_{0i} = 1$ $\forall i$; as in Westerlund and Edgerton (2008) we set both constant and slope to 5 after the break in all units. This is admittedly an extremely large break, of virtually no empirical relevance. However, setting the break to such a large value will permit us to evaluate the properties of the procedure independently on modelling difficulties.

Long-run growth of X is assumed to be driven by a non-stationary factor common across units (F_1), with short-run deviations caused by a second stationary common factor (F_2) and by an idiosyncratic stationary noise (ϵ^x_{it}):

$$x_{it} = \gamma_1 F_{1t} + \gamma_2 F_{2t} + \epsilon^x_{it} \tag{7}$$

Following Pesaran (2007) the factor loadings are chosen so to ensure substantial cross-correlation in the $X's$: $\gamma_i \sim Uniform(-1,3)\ \forall i$. The common factors are generated as follows:

$$\begin{bmatrix} F_{1t} \\ F_{2t} \end{bmatrix} = \begin{bmatrix} F_{1t-1} \\ 0.4F_{2t-1} \end{bmatrix} + \begin{bmatrix} f_{1t} \\ f_{2t} \end{bmatrix} \tag{8}$$

where, as in Gengenbach *et al.* (2006), both the common and idiosyncratic shocks are assumed to have a MA(1) structure:

$$\begin{bmatrix} f_{1t} \\ f_{2t} \end{bmatrix} = \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix} + \begin{bmatrix} \vartheta_1 & 0 \\ 0 & \vartheta_2 \end{bmatrix} \begin{bmatrix} \eta_{1t-1} \\ \eta_{2t-1} \end{bmatrix} \tag{9}$$

$$\epsilon^x_{it} = e^x_{it} + \varphi e^x_{it-1}, \tag{10}$$

where $\eta_{it} \sim N(0,1), i = 1,2,$ and $e^x_{it} \sim N(0, \sigma^2_{ix})$, with $\sigma^2_{ix} \sim Uniform(1, 1.4)$. Both φ and the $\vartheta's$ are generated as Uniform deviates in the range [0.5,0.7].

The simulation framework outlined above is very complex, and the tests to be evaluated computationally demanding (this issue is discussed in more detail below). Hence, 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. As standard in the literature we take the model as correctly specified.

Given that in some of our simulations the time sample is quite small to ensure computational stability we choose a 20% trimming coefficient.

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, medium in terms of annual data, but definitely small at a quarterly frequency, it is smaller than those generally considered in the simulation studies on the other cointegration tests with breaks 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 $[8, 32]$.

2. *Medium time sample:* $T = 80$, $N = 5, 10, 20$; break date Uniform over units in $[0.5T \pm 3] = [37, 43]$. The time span is now large in terms of annual data, but pretty common for quarterly data, so to make it relevant for actual empirical applications. Since we need the results

from this experiment to be closely comparable to those from the Base case we mimick a situation in which more observations (more precisely, 20) become available at both ends of the sample; hence, the breaks are distributed over the same sets of periods centred in the middle of the time sample, with the procedure searching over the interval $[16, 64]$. Finally, since the interest here is on the behaviour of the test when the time dimension grows for computational convenience we limited the experiments to a most 20 units.

3. *Large time sample*: $T = 160$, $N = 5, 10$; break date Uniform over units in $[0.5T \pm 3] = [77, 83]$. The time span is long in terms of annual data, but medium with a quarterly frequency, so that it is still relevant for actual empirical applications. The sample is extended at both ends as in the previous case, and the search is over the interval $[32, 128]$. In this case also for computational convenience we consider only small cross-section sample sizes.

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 because of the combined effects of the the panel structure of the data and the recursive nature of the statistics evaluated: 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 500 bootstrap redrawings, 40 units and search over 28 periods, as in the Base Case, the product of first three terms is equal to 560.000. Fixing the Monte Carlo replications to 1000 will thus require the execution of over half a billion loops for each experiment, with, *e.g.*, for a rejection rate $p = 5\%$ an approximate confidence interval $p \pm 2\sqrt{p(1-p)/1000}$ equal to $[3.6\%, 6.4\%]$. Reducing the length of the interval even marginally to $[4.0\%, 6.0\%]$ requires a disproportionate effort, as the number of replications and hence that of loops would double. We thus decided that 1000 replications is a reasonable choice.

3.2 Results

The results are reported in Tables 1-5 below, and rapidly summarised. First of all, there seem to be very little differences between the performances of the different tests (ADF and coefficient, group mean and median) and, quite remarkably, block sizes considered. Hence, our comments can be expressed in general terms. Second, with a small time sample ($T = 40$) the Type I errors are somehow smaller than nominal sizes, but converge rapidly to the latter for $T = 80$. Finally, power can be disappointing for T and N both small, but increases rapidly with T and, most importantly, N. With the DGP used in our simulations the rejection rate of the false null of no cointegration is 100% already in almost all cases for $T = 80$ and always for $T = 160$ (hence, a separate table for this case is redundant). Clearly, these findings are conditional to speed of adjustment and signal-noise ratio. In systems with slower adjustment and more noise the power performances will not be as satisfactory, but this is not an issue: the key message is that with a small time sample a panel procedure may grant considerable power gains, while ensuring good size control.

Table 1
$T = 40$, N *from* 5 *to* 40
Size

		N							
		5	10	20	40	5	10	20	40
Block	α	*Boot-Mean*				*Boot-Median*			
A. *ADF*									
4	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
	0.05	0.03	0.02	0.01	0.01	0.03	0.02	0.01	0.01
	0.10	0.07	0.05	0.03	0.03	0.06	0.04	0.04	0.03
6	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.05	0.02	0.01	0.01	0.00	0.02	0.02	0.01	0.01
	0.10	0.06	0.04	0.03	0.01	0.05	0.04	0.03	0.02
B. $T(\rho - 1)$									
4	1.0	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
	5.0	0.03	0.03	0.02	0.01	0.04	0.03	0.02	0.01
	10.0	0.07	0.07	0.05	0.03	0.08	0.06	0.06	0.05
6	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
	0.05	0.03	0.02	0.01	0.00	0.03	0.02	0.01	0.01
	0.10	0.07	0.05	0.04	0.03	0.07	0.04	0.04	0.03

DGP: eqs. (5)-(10),
$t_i^b \sim Uniform(0.5T \pm 3)$, search interval: $[0.2T, 0.8T]$;
Boot-mean/median: bootstrap test on the mean/median across units of the no cointegration statistics;
Bootstrap: 500 redrawings; block size: $4 = 0.10T$, $6 = 1.75\sqrt[3]{T}$.
Montecarlo: 1000 replications. Approximate confidence intervals:
0.01: $[0.004, 0.016]$,0.05: $[0.036, 0.064]$,0.10: $[0.081, 0.119]$.

Table 2

$T = 80$, *N from 5 to 20*

Size

	N					
	5	10	20	5	10	20
α	*Boot-Mean*			*Boot-Median*		
	A. *ADF*					
0.01	0.00	0.00	0.00	0.00	0.00	0.00
0.05	0.03	0.03	0.03	0.03	0.04	0.04
0.10	0.08	0.08	0.07	0.09	0.08	0.08
	B. $T(\rho - 1)$					
1.0	0.00	0.00	0.00	0.00	0.00	0.00
5.0	0.05	0.04	0.04	0.05	0.04	0.04
10.0	0.10	0.09	0.09	0.10	0.09	0.09

DGP: eqs. (5)-(10),

Block size: $8 \simeq 0.10T \simeq 1.75\sqrt[3]{T}$

All definitions and abbreviations: see Table 1

Table 3

$T = 160$, N *=5,10*

Size

		N			
		5	10	5	10
Block	α	*Boot-Mean*		*Boot-Median*	
		A. *ADF*			
16	0.01	0.00	0.00	0.00	0.00
	0.05	0.01	0.02	0.04	0.02
	0.10	0.06	0.05	0.08	0.06
10	0.01	0.00	0.00	0.01	0.00
	0.05	0.02	0.02	0.03	0.03
	0.10	0.05	0.05	0.07	0.06
		B. $T(\rho - 1)$			
16	0.01	0.01	0.00	0.01	0.01
	0.05	0.04	0.02	0.05	0.05
	0.10	0.07	0.07	0.11	0.10
10	0.01	0.01	0.00	0.01	0.01
	0.05	0.04	0.02	0.06	0.05
	0.10	0.07	0.06	0.10	0.10

DGP: eqs. (5)-(10),

Block size: $16 = 0.10T, 10 = 1.75\sqrt[3]{T}$.

All definitions and abbreviations: see Table 1

Table 4
$T = 40$, *N from* 5 *to* 40
Power

		N							
		5	10	20	40	5	10	20	40
Block	α	*Boot-Mean*				*Boot-Median*			
		A. *ADF*							
4	0.01	0.28	0.52	0.82	0.98	0.15	0.35	0.63	0.89
	0.05	0.59	0.83	0.97	1.00	0.43	0.71	0.92	0.98
	0.10	0.77	0.93	0.98	1.00	0.62	0.84	0.97	1.00
6	0.01	0.24	0.46	0.74	0.94	0.13	0.32	0.56	0.83
	0.05	0.56	0.79	0.95	1.00	0.40	0.66	0.88	0.98
	0.10	0.73	0.91	0.98	1.00	0.57	0.82	0.96	0.99
		B. $T(\rho - 1)$							
4	1.0	0.30	0.58	0.88	0.99	0.19	0.47	0.79	0.95
	5.0	0.65	0.89	0.98	1.00	0.54	0.82	0.97	1.00
	10.0	0.81	0.95	0.99	1.00	0.72	0.93	0.99	1.00
6	0.01	0.28	0.55	0.83	0.98	0.16	0.43	0.73	0.92
	0.05	0.62	0.86	0.98	1.00	0.50	0.78	0.95	1.00
	0.10	0.79	0.95	0.99	1.00	0.69	0.91	0.99	1.00

DGP: eqs. (5)-(10), $\rho_i \sim Uniform(0.6, 0.8)$
All definitions and abbreviations: see Table 1

Table 5
$T = 80$, *N from 5 to 20*
Power

	N					
	5	10	20	5	10	20
α	*Boot-Mean*			*Boot-Median*		
	A. *ADF*					
0.01	0.99	1.00	1.00	1.00	1.00	1.00
0.05	1.00	1.00	1.00	1.00	1.00	1.00
0.10	1.00	1.00	1.00	1.00	1.00	1.00
	B. $T(\rho - 1)$					
1.0	1.00	1.00	1.00	0.99	1.00	1.00
5.0	1.00	1.00	1.00	1.00	1.00	1.00
10.0	1.00	1.00	1.00	1.00	1.00	1.00

DGP: eqs. (5)-(10), $\rho_i \sim Uniform(0.6, 0.8)$;
Block size: $8 \simeq 0.10T \simeq 1.75\sqrt[3]{T}$;
All definitions and abbreviations: see Table 1

4 Testing the Feldstein-Horioka Puzzle

As discussed in the Introduction, the evidence supporting the existence of a long-run link between the investment (I) and savings (S) to GDP (Y) ratios in advanced economies is still a largely unsolved puzzle. 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. However, these results may not be entirely reliable, as the bootstrap procedure used (abandoned in the revised version of the paper, Banerjee and Carrion-i-Silvestre, 2006) under no cointegration implied fitting an AR model to a MA process with a unit root, which does not have a finite AR representation. Further, from the economic point of view an obvious remark is that in order to assess the current validity of the puzzle we should look at a more updated sample. To this end we shall examine a slighlty larger panel including 18 OECD economies (Austria, Australia, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Portugal, Spain, Sweden, UK, USA) for the slightly later period 1970-2007 (38 annual observations covering eight years of the Euro era[4]). Details on the dataset are provided in the Appendix, and plots in Fig. 1 and 2.

From the methodological point of view we have to take into account two main points: (*i*) both investment and savings are generally correlated in the short-run, and in some cases cointegrated, across economies (results not reported here available on request); (*ii*) the available time sample is rather small. As already briefly discussed, this rules out all currently available cointegration tests with breaks available in the literature, which require either restrictive assumptions on the cross-section dependence structure or large sample sizes (or both). The procedure proposed and examined in the previous sections, on the other hand, does appear potentially suitable for this task.

Examining the individual statistics (Table 6) we find that, consistently with theoretical expectations though somehow contrary to those formed on the basis of visual inspection of the plots, only in five countries (Australia, denmark, Japan, Netherlands, Sweden) out of 18 the Gregory-Hansen $Min(ADF)$ tests reject the null hypothesis of no cointegration according to the asymptotic critical values. This evidence (or, better, lack of) should however be evaluated keeping in mind that with such a small sample size

[4] All the european countries of our sample except Denmark, Sweden and UK (a total of 11) adopt the european common currency.

power is likely to be very low (Gregory and Hansen, 1996, report rejection rates around 50% for $T = 50$). Hence, the failure to reject cannot be taken as a conclusive piece of evidence.

We thus turn to the bootstrap panel cointegration tests, reported in Table 7. We first test for panel cointegration without breaks using the procedure proposed in Di Iorio and Fachin (2008); the results (panel A) are mostly not favourable to the hypothesis of cointegration. The p-values for the mean and median of individual ADF tests are respectively equal to 13.3% and 21.9%, and those of of the $T(\rho - 1)$ statistics respectively 8.9% and 20.4%. Hence, the only hint of a relationship is given by the mean of the $T(\rho - 1)$ statistics, while the medians, definitely more interesting summary statistics, are definitely consistent with the hypothesis of no long-run $I - S$ relationship.

We then turn to the tests allowing for breaks. Unsurprisingly, all p-values are now smaller. For both statistics they are marginally higher than 5% in mean (about 6%), but definitely higher than this traditional threshold in median: 17.4% for the ADF and 10.2% for $T(\rho - 1)$. There is then some evidence in favour of the existence of a long-run relationship with breaks in the coefficients, but it is rather weak, as it comes mostly from a subset of the economies included in the panel.

Finally, from Table 9 we can see that the breaks fall in the 1980's in 6 cases, while in the remaining in the 1990's or in 2000, the last observation of the search period (1977-2000).

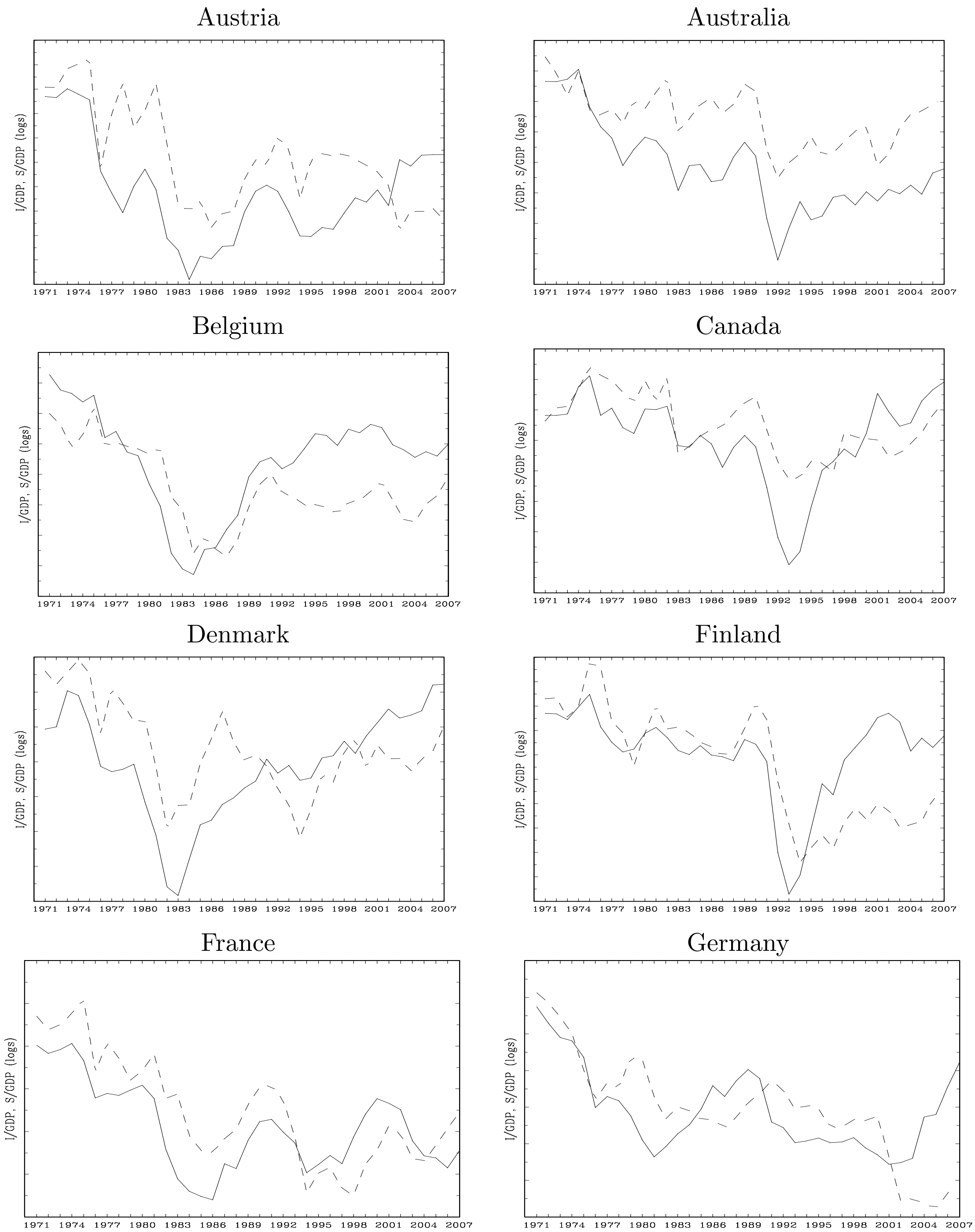

Fig. 2 Savings/GDP (solid line) and Investment/GDP (dashed line), 1970-2007 (logs).

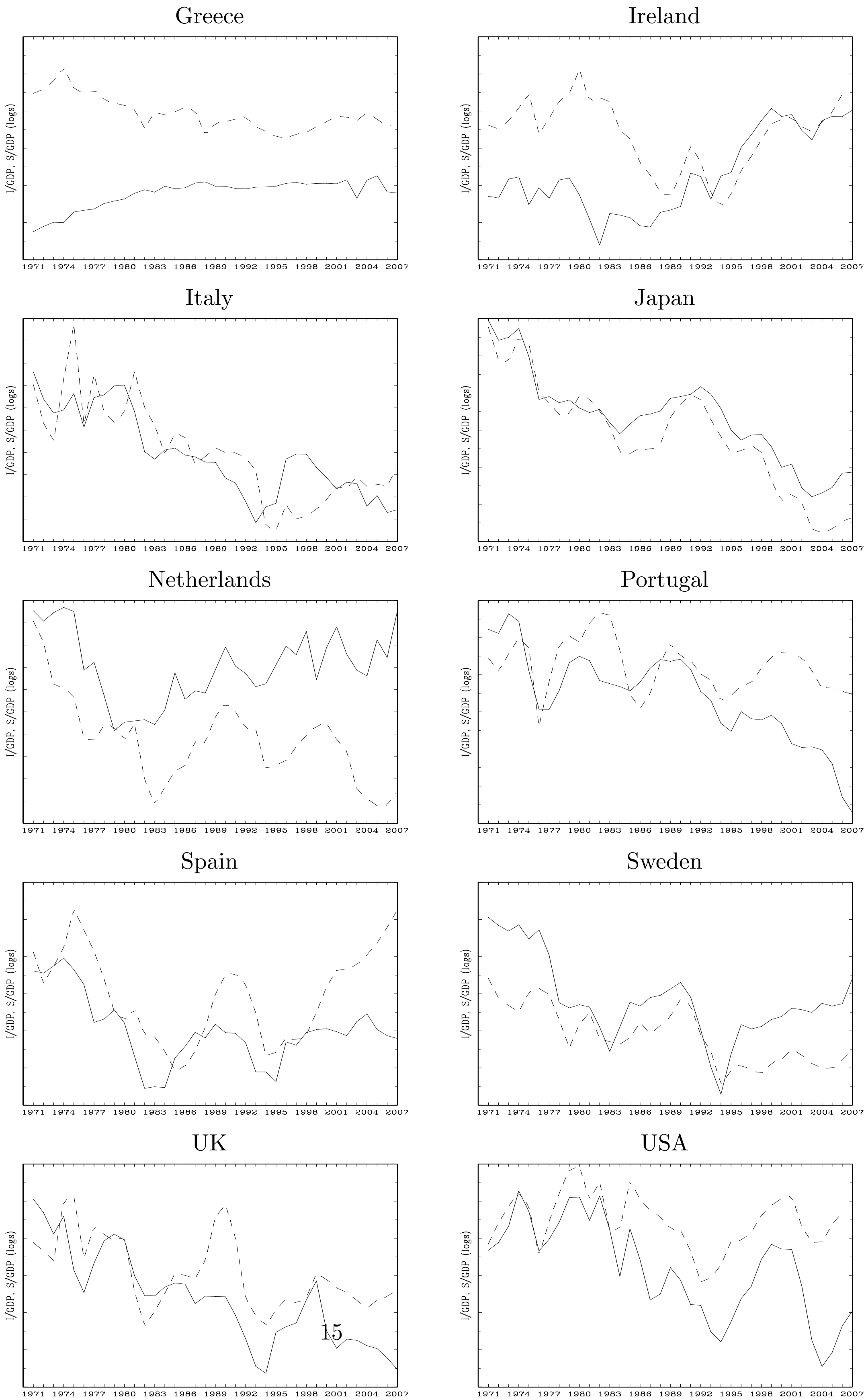

Fig. 3 Savings/GDP (solid line) and Investment/GDP (dashed line), 1970-2007 (logs).

Table 6
Investment and Savings, 1970-2007
Min(ADF) Cointegration Tests with Unknown Break

Austria	Australia	Belgium	Canada	Denmark	Finland
-1.85	-5.28^{**}	-4.45	-3.55	-5.11^{**}	-3.61
France	Germany	Greece	Ireland	Italy	Japan
-3.21	-3.83	-3.29	-3.12	-4.42	-6.03^{**}
Netherlands	Portugal	Spain	Sweden	UK	USA
-4.54^{**}	-3.92	-4.23	-4.84^{*}	-3.12	-3.58

trimming: 20% (searching interval: 1977-2000)
critical values (Gregory and Hansen, 1996):
$1\% : -5.47; 5\% : -4.95; 10\% : -4.68$.
***: significant at 1%; **: 5%;*: 1%.

Table 7
Investment and Savings Feldstein-Horioka equation
Bootstrap Panel Cointegration Tests, 1970-2007

	A. Without breaks	
	$Mean$ ($p^* \times 100$)	$Median$ ($\times 100 p^*$)
ADF	-2.30 [13.3]	-2.23 [21.9]
$T(\rho-1)$	27.64 [8.6]	29.25 [20.4]
	B. With Unknown Breaks	
ADF	-4.00 (5.4)	-3.88 (15.8)
$T(\rho-1)$	11.35 (5.9)	13.04 (10.0)

Mean/Median: mean/median of the individual statistics;
p^*: bootstrap p-values $\times 100$, 1000 redrawings.
break: $Argmin(RSS)$, 20% trimming at each end
block size: $1.75\sqrt[3]{T} \simeq 6$.

Table 8
Investment and Savings, 1970-2007
Argmin(RSS) Estimated Breakpoints

Austria	Australia	Belgium	Canada	Denmark	Finland
1995	1981	1984	1996	1991	1994
France	Germany	Greece	Ireland	Italy	Japan
1993	2000	1988	1986	1994	1984
Netherlands	Portugal	Spain	Sweden	UK	USA
1994	1998	2000	1994	1981	1999

trimming: 20% (searching interval: 1977-2000)

5 Conclusions

Investigating the FH puzzle requires tackling the challenging task of testing panel cointegration in dependent panels allowing for breaks at unknown periods. To carry out this task we must accounted for two forms of dependence, between the tests computed with the break fixed at different periods and for different units. Further, since the probability of multiple breaks, which complicate considerably the tests, increases with the time sample size acceptable small sample performances are particularly important. Building upon Di Iorio and Fachin (2008), in this paper we propose to solve this problem using the bootstrap; simulation results suggest that the proposed panel testing procedures may improve considerably on the small sample power performances of pure time series Gregory and Hansen (1996) tests.

The good small sample performances obtained suggest it as a potentially useful addition to the toolbox for non-stationary panel analysis, complementary to asymptotic factor methods procedures (Banerjee and Carrion-i-Silvestre, 2006, Westerlund and Edgerton, 2008). These impose some restrictive assumptions on the form of the cross-section dependence and require large time sample sizes, but may shed some light on the common factor structure (of interest in itself), and, in the case of Westerlund and Edgerton (2008), may allow for multiple breaks.

These expectations are confirmed by the analysis of an Investment and dataset for a panel of 18 european countries over the period 1970-2007. Assuming free capital movements there are no reasons to expect any long-run relationship between these two variables to hold, but the visual inspection of the plots suggests that this may instead possibly have been the case, provided breaks (plausible, in view of changes in capital movements regulations) are allowed. It is thus somehow surprising to discover that the large majority of individual Gregory and Hansen tests for individual countries fail to reject the null of no cointegration, consistently with theoretical expectations but somehow at odds with the graphical empirical evidence. The bootstrap panel tests, however, confirm that in the majority of the economies the evidence in support of a relationship between savings and investments is weak even if we use a powerful statistical procedure.

Clearly, much methodological and empirical work is still needed. With respect to the latter, the natural next step forward is to test if the breaks actually took place; this is the subject of Di Iorio and Fachin (2007). On the former, just to mention a few: data-driven choice of the block size, multivariate DGP's and models, partially cointegrated panels.

6 Appendix

6.1 Data source and definitions

All data, in national currency at current prices, have been downloaded from the OECD.stat database on 26 June 2009. Definitions are as follows:

Investment: Gross capital formation (transaction code: P5S1).

Savings: Net savings (transaction code B8NS1) plus Consumption of fixed capital (transaction code K1S1).

Gross Domestic Product: transaction code B1_GS1.

7 References

Bai, J., and P. Perron (1998) "Estimating and Testing Linear Models with Multiple Structural Changes", *Econometrica*, 66, 47-78.

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. (Revised December, 2006).

Di Iorio, F. and Fachin, S. (2007) ”Testing for Breaks in Cointegrated Panels” *Economics* 1-14.

Di Iorio, F. and Fachin, S. (2008) "Testing for cointegration in dependent panels via residual-based bootstrap methods" *MRPA Working Paper* n. 3139 .

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

Feldstein, M. and C. Horioka (1980) ”Domestic Saving and International Capital Flows”, *Economic Journal*, 90, 314-329.

Frankel, J.A. (1992) ”Measuring International Capital Mobility: A Review”, *American Economic Review*, 82, 197-202.

Gengenbach, C., F. C. Palm, and J.-P. Urbain (2008) "Panel unit root tests in the presence of cross-sectional dependencies: Comparison and implications for modelling" *Econometric Reviews*, forthcoming.

Gonzalo, J. (1994) ”Five Alternative Methods of Estimating Long-Run Equilibrium Relationships”, *Journal of Econometrics,* 60, 203-233.

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

Gutierrez, L. (2009) "Simple Tests for Cointegration in Panels with Structural Breaks", *Applied Economics Letters,* forthcoming.

Juselius, K. (2006) *The Cointegrated VAR Model: Econometric Methodology and Macroeconomic Applications*, Oxford University Press, Oxford

Kao, C. (1999) "Spurious Regression and Residual-based Tests for Cointegration in Panel Data" *Journal of Econometrics,* 90,1-44.

Ng, S., Perron, P. (1995) "Unit Root Tests in ARMA Models With Data Dependent Methods for Selection of the Truncation Lag" *Journal of the American Statistical Association,* 90, 268-281.

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

Özmen, E. and K. Parmaks (2003) "Policy regime change and the FeldsteinHorioka puzzle: the UK evidence" *Journal of Policy Modeling*, 25, 137-149.

Palm, F. C., S. Smeekes, and J.-P. Urbain (2008) "Cross-sectional dependence robust block bootstrap panel unit root tests". *METEOR Research Memorandum* 08/048, Maastricht University.

Parker, C., Paparoditis E., Politis D.N. (2006) "Unit root testing via the stationary bootstrap" *Journal of Econometrics* 133, 601-638.

Paparoditis, E., Politis, D.N. (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.

Pesaran, M.H. (2007) "A Simple Panel Unit Root Test in the Presence of Cross Section Dependence" J*ournal of Applied Econometrics*, 22, 265-312.

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. (2008) "A Simple Test for Cointegration in Dependent Panels with Structural Breaks", *Oxford Bulletin of Economics and Statistics*, 70, 665-703.