MPRA
Munich Personal RePEc Archive

Savings and Investments in the OECD, 1970-2007: a Panel Cointegration test with breaks

Di Iorio, Francesca and Fachin, Stefano

University of Rome "La Sapienza"

25 November 2010

Online at https://mpra.ub.uni-muenchen.de/26781/
MPRA Paper No. 26781, posted 26 Nov 2010 20:09 UTC

Savings and investments in the OECD, 1970-2007: a panel cointegration test with breaks

Francesca Di Iorio
University of Naples Federico II

Stefano Fachin
University of Rome "La Sapienza"

Abstract

In this paper we test for the existence of a long-run relationship between investment and savings (the Feldstein-Horioka puzzle) in a panel of 18 OECD countries, 1970-2007, allowing for heterogenous breaks in the coefficients. For this purpose we develop a bootstrap panel cointegration test with breaks robust to cross-section dependence. The test suggests that, even allowing for breaks in the countries where capital control regulations changed in the sample period, there is no evidence of an investment-savings long-run relationship for the panel as a whole.

Keywords: Feldstein-Horioka Puzzle, Investment, Savings, Panel cointegration, stationary bootstrap, breaks, OECD.

JEL codes: C23, C15

Number of words: about 7000

Revised November 2010

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 (Feldstein-Horioka, FH, puzzle after Feldstein and Horioka, 1980) is considered one of the six major empirical puzzles of contemporary macroeconomics (Obstfeld and Rogoff, 2000). In a world with no barriers, capitals should seek the most profitable investment opportunities regardless of national borders. Thus, in any individual country, capital formation (I) should not be constrained by domestic savings (S), as the difference between the two aggregates is the current account balance (B), which in such a world does not need to be zero. Hence, the evidence of a savings-investment relationship was indeed puzzling. After thirty years, during which the issue has been investigated in a huge literature (for a recent review, see Apergis and Tsoumas, 2009, who list nearly 200 references), the puzzle is still essentially open, as existing studies leave much to be desired. Tests of the long-run validity of the puzzle based on cointegration methods have as a null hypothesis the absence of a relationship, *i.e.*, no puzzle. Hence, the prevailing conclusion that no puzzle exists (*e.g.*, Kim, 2001) may be simply due to the notiorous lack of power of cointegration tests with the small or moderate samples typically used. Panel cointegration studies, such as Pelgrin and Schich (2008), do have higher power, but are based on tests valid only for independent units. If this assumption does not hold these tests are strongly biased against the null hypothesis of no relationship (Banerjee, Marcellino and Osbat, 2004), so that the reported rejections of the hypothesis of no relationship may be spurios. In fact, in Di Iorio and Fachin (2010) we showed that, using a bootstrap panel cointegration test valid for dependent units, evidence of a saving-investment long-run relationship cannot be found for a panel of 18 economies including the core of OECD[2], but only some of them[3]. Hence, a saving-investment relationship can be present in some circumstances, but it does not seem to be a law of general validity for the advanced economies.

However, as Frankel (1992) pointed out, we need not to overlook the fact that capital movements regulations underwent many changes over the last few decades. Strict controls were widely applied in the first decades after the second world war, but gradually lifted in most countries since the late '70's (see OECD, 2002). For instance, in the UK capital movements were completely liberalised in 1979 by the first Tatcher government, with the Netherlands following shortly afterwards (Bakker and Chapple, 2002). Within the European Union (EU) all capital controls were removed in 1990[4], and in 1992 the Maastricht treaty went as far as prohibiting to member countries the application of capital controls with third countries also. Finally, in the late 1990's exchange rate risks, which descourage capital movements, were greatly reduced by the introduction of the Euro. Hence, Di Iorio and Fachin (2010)'s failure to reject the hypothesis of no savings-investment relationship may be due to the assumption of constant parameters. To reach a reliable conclusion we need to apply a more general test, allowing for changes in the cointegrating coefficients. Unfortunately, (no) cointegration tests with breaks have many shortcomings. The power of the standard Gregory and Hansen (1996) procedure tends to be rather poor, while the available panel cointegration tests (*e.g.*, Banerjee and Carrion-i-Silvestre, 2010, Gutierrez, 2009, Westerlund, 2006a,b) require either independence across the units of the panel or rather large time samples. Neither of these conditions are satisfied by the OECD dataset studied by Di Iorio and Fachin (2010). In this paper we therefore develop a new test, which is

[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, to the participants at the Amsterdam 2010 Panel data conference and seminars in Maastricht and the University of Rome "Tor Vergata" for helpful comments and suggestions. The usual disclaimers apply.
Correspondence to: s.fachin@caspur.it, fdiiorio@unina.it.

[2]Austria, Australia, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Portugal, Spain, Sweden, UK, USA. The period is 1970-2007.

[3]More precisely, ten: Portugal, Japan, United Kingdom, Australia, United States, Italy, Denmark, Netherlands, Canada and Sweden.

[4]Council Directive 88/361/EEC of June 24, 1988.

the extention to the case of unknown breaks of a bootstrap panel cointegration test shown by Di Iorio and Fachin (2010) to be asymptotically valid. As we will see, this new testing procedure can account for fully general forms of dependence and delivers satisfactory small sample size and power properties, provided some care is applied. It must be stressed that, since the interest here is on small sample testing and more general dependence structures, the proposed procedure should be 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) examine in more detail the task of testing the Feldstein-Horioka puzzle and the dataset of interest, then (section 3) outline the proposed testing procedure and present the results of a Monte Carlo experiment, finally (section 4) discuss the results delivered by the new test on our panel of OECD countries. Some conclusions are finally drawn in section 5.

2 Testing the Feldstein-Horioka Puzzle

The long-run FH equation is

$$i_t = \mu + \beta s_t + \epsilon_t \tag{1}$$

where $i = log(I/GDP)$, $s = log(S/GDP)$, and the coefficient β is known in the literature as saving retention ratio. In closed economies investments (fixed capital formation plus changes in inventories) are, by definition, equal to savings, so that β is constrained to be equal to 1, the constant μ to zero, and the residuals $\{\epsilon\}_{t=1}^{T}$, which can only reflect errors of measurement, are stationary. On the other hand, in open economies with no barriers to capital movements there are no constraints on the coefficients β and μ, and the residuals are stationary only if savings and investments are linked by a long-run equilibrium relationship. Even in absence of legal barriers such a relationship may exist as a consequence of market imperfections inducing home bias.

As we argued above, this basic model may need to be generalised to allow for changes in the coefficients. Since with small to moderate time samples the number of such changes which is practically manageable is limited to one, we write this generalised FH equation as

$$i_{jt} = \begin{cases} \mu_{0j} + \beta_{0j} s_{jt} + \epsilon_{jt}, & t \leq t_j^b \\ \mu_{1j} + \beta_{1j} s_{jt} + \epsilon_{jt}, & t > t_j^b \end{cases} \tag{2}$$

where we introduced a country index $j = 1, \ldots, N$, and t_j^b is the break point for country j. We shall test the validity of (2) as a long-run relationship in the panel already studied by Di Iorio and Fachin (2010), which includes Austria, Australia, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Portugal, Spain, Sweden, UK and USA (hence $N = 18$), for the period 1970-2007 (hence, $T = 38$, with the sample including eight years of the Euro era[5]).

Details on the dataset are provided in the Appendix, and plots in Fig. 1 and 2. As we can see, the two variables (which can be considered I(1): see Di Iorio and Fachin, 2010) often do seem to follow closely related paths (this is certainly the case, *e.g.*, in Italy and Japan). Greece is a peculiar case, with investments consistently much higher than savings; in fact, the two variables are *inversely* related (see Di Iorio and Fachin, 2010). Of course, in the light of the recent debt crisis (European Commission, 2010) this is not surprising.

However, in several cases the association seems to break, generally in the direction of a weakening of the saving-investment link. Without going into a detailed discussion of the individual cases, this for instance clearly appears to be the case of Belgium in the early 1990's. Since breaks are likely to be the consequence of changes in the degree of openness in capital markets it is

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

instructive to examine the well-known Chinn-Ito index of capital openness (Chinn and Ito, 2008). The first differences of the index are plotted for each country in Figs. 3-4. As to be expected, the index is generally constant (hence, the difference is zero) in most of the period, with some large jumps following regulations changes. This is particularly evident for the cases of UK, where the differences are zero or negligible in all years except 1979, and the Netherlands, where the same holds for 1982. In the case of Belgium the jump is in 1990, when the EU liberalisation came into effect, and it may thus explain the particularly evident fall in the association between the two variables mentioned above. The only exceptions to this picture are Canada, USA and Germany, countries of well-known liberal tradition (OECD, 2002, p. 63) where, as a consequence, the index has been fixed for all (or almost all) the period at the maximum level, and the differences at zero.

Summing up, with few exceptions, allowing for breaks seems necessary when testing for the existence a long-run relationship. Although the changes in the capital movements regulations help locating plausible breakpoints, these cannot be assumed as precisely known *a priori.* Rather, they should be considered as unknown parameters to be estimated endogenously. The classical reference for cointegration testing with breaks at unknown dates is Gregory and Hansen (1996), who proposed to compute a no-cointegration statistic (say, $\theta(t^b)$) for all possible break points t^b and, assuming the rejection region is the left tail (as in the case of the popular ADF and Z tests), take the minimum:

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

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

$$\widehat{t^b} = Arg \min_{t^b \in [\delta T, (1-\delta)T]}(\theta(t^b)) \tag{3}$$

which, though intuitevely appealing (it is the break maximising the probability of rejection) it is not necessarily the best choice. As we will see below, a natural alternative is a least square criterion.

The natural first step of our study is thus to compute the standard Gregory-Hansen $Min(ADF)$ cointegration tests separetely for each country of our panel. For complete generality, in this first step we allow for a break in all countries of our panel. Consistently with the theoretical expectations, though somehow contrary to those formed on the basis of visual inspection of the plots, only in seven countries (Australia, Denmark, Finland, Italy, Japan, Netherlands, Sweden) out of 18 the $Min(ADF)$ tests reject the null hypothesis of no cointegration according to the asymptotic critical values (see Table 1). Since all these countries except Finland are included in the group for which Di Iorio and Fachin (2010) found the Feldstein-Horioka relationship to hold, the Gregory-Hansen tests essentially do not add any evidence to what we already knew from the panel cointegration tests with no breaks. However, this evidence (or, better, lack of) should be evaluated keeping in mind that with our sample size power is likely to be very low: the rejection rates reported by Gregory and Hansen (1996) for $T = 50$ are around 50% (Gregory and Hansen, 1996, table 2). Hence, the failure to reject (which would imply no puzzle) cannot be taken as a conclusive piece of evidence[6].

[6]It is interesting to recall that for the UK Özmen and Parmaksiz (2003) with the sample 1948-1998, which includes almost three decades of strict capital controls, obtain a $Min(ADF)$ test only marginally significant at 10%.

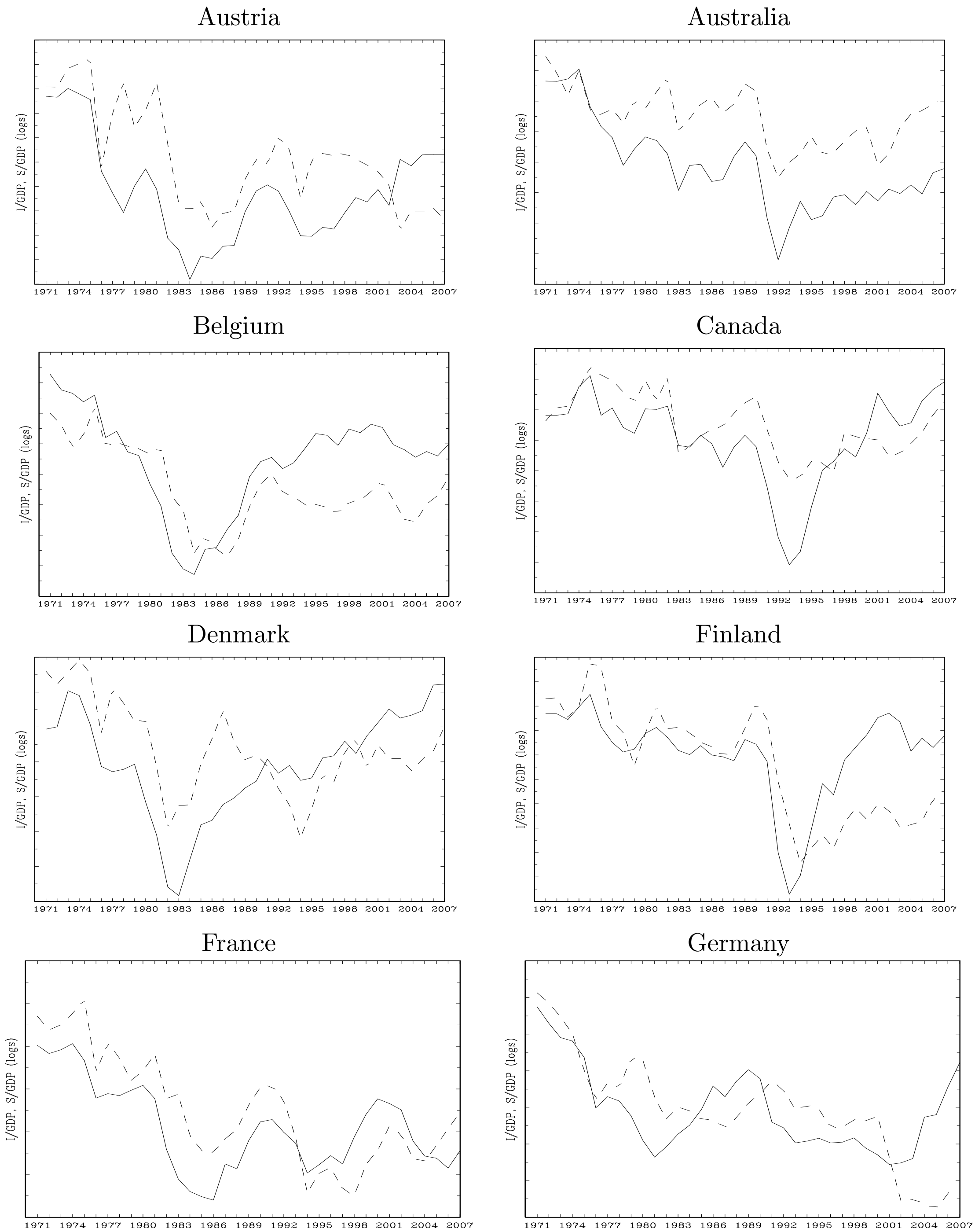

Fig. 1 Savings/GDP (solid line) and Investment/GDP (dashed line), 1970-2007 (logs).

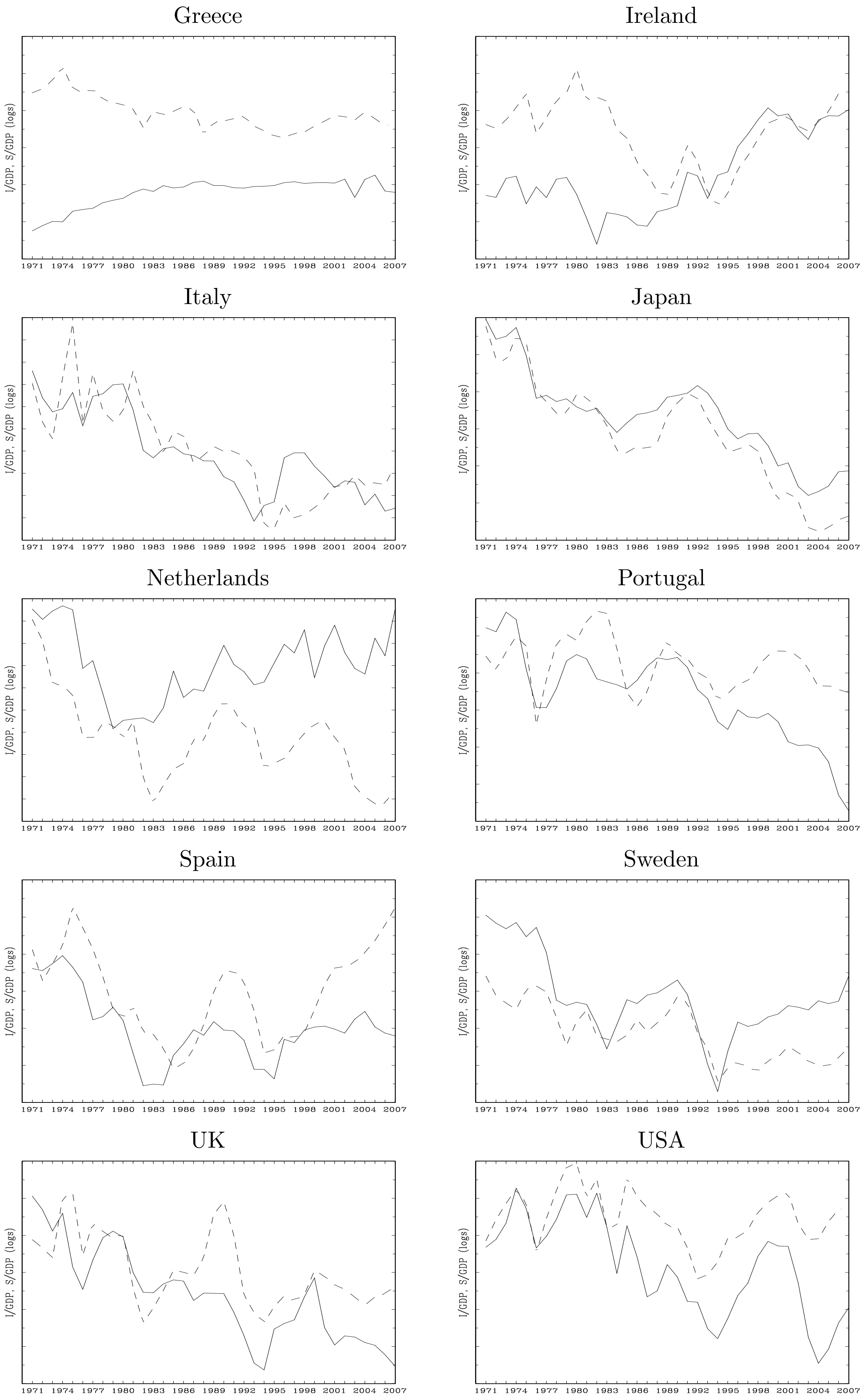

Fig. 2 Savings/GDP (solid line) and Investment/GDP (dashed line), 1970-2007 (logs).

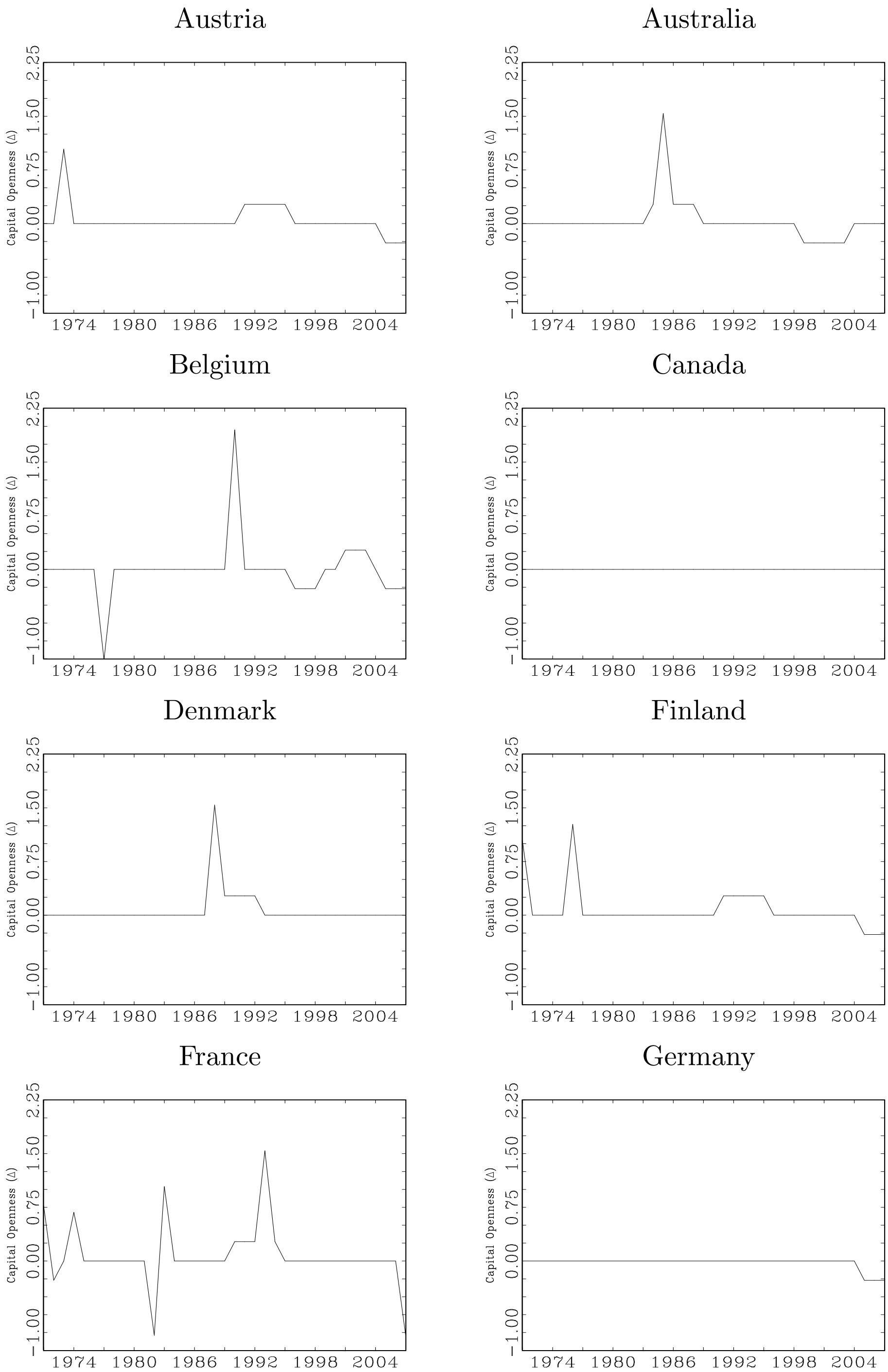

Fig. 1, First differences in Chinn-Ito Capital Openness index, 1971-2007.

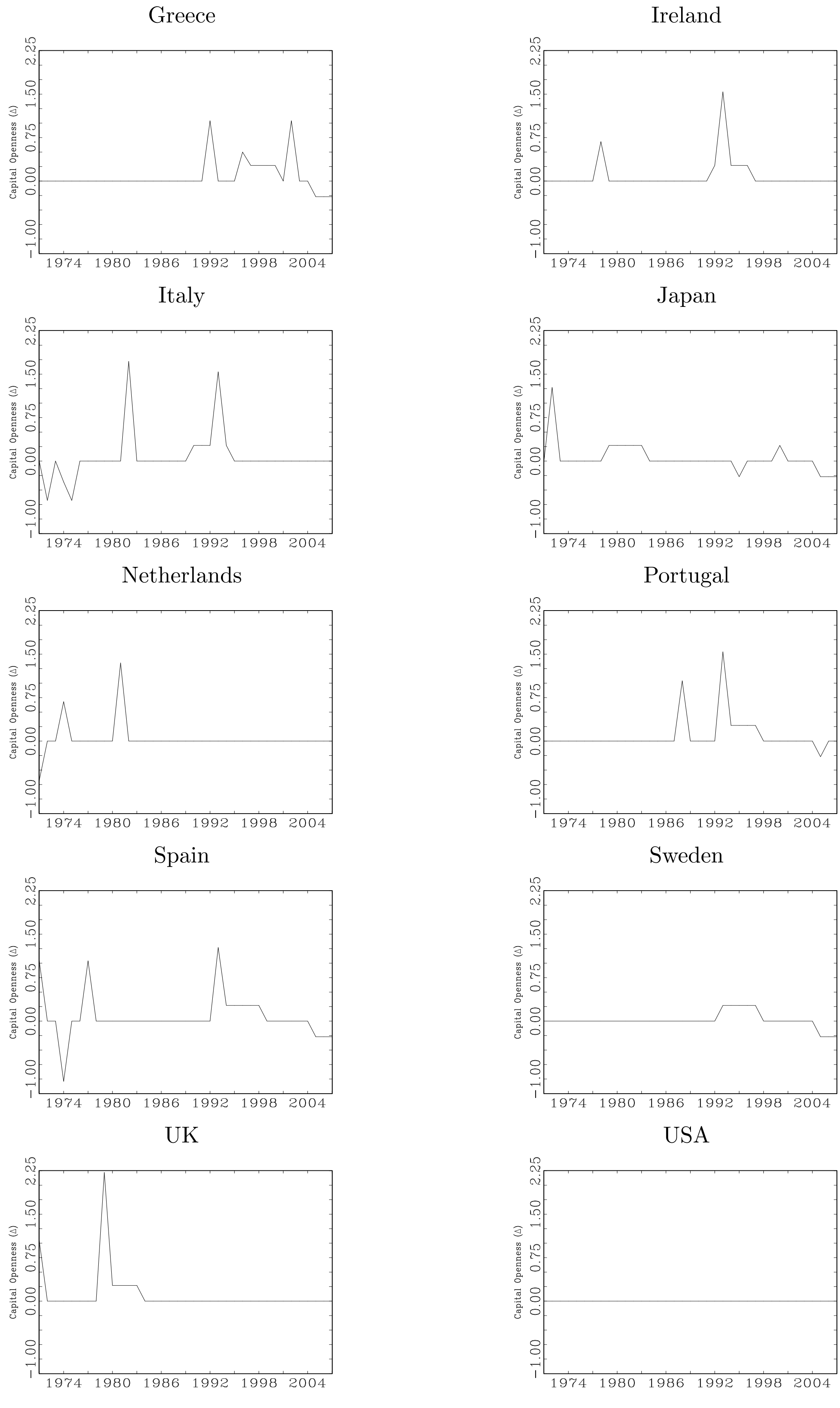

Fig. 2 First differences of Chinn-Ito Capital Openness index, 1971-2007. Note: data for 1975-1981 for the Netherlands, not available, set to zero (no change).

Table 1
Investment and Savings, 1970-2007
Min(ADF) Cointegration Tests with Unknown Break

Austria	Australia	Belgium	Canada	Denmark	Finland
-3.56 [1988]	-5.54^{**} [1979]	-4.81 [1984]	-3.73 [1997]	-5.11^{**} [1990]	-4.81^{*} [1997]
France	**Germany**	**Greece**	**Ireland**	**Italy**	**Japan**
-3.33 [1992]	-3.83 [2000]	-4.50 [1976]	-3.88 [1987]	-5.13^{**} [1996]	-6.03^{**} [1984]
Netherlands	**Portugal**	**Spain**	**Sweden**	**UK**	**USA**
-5.29^{**} [1998]	-4.11 [1986]	-4.23 [2000]	-5.37^{**} [1993]	-3.56 [1978]	-3.88 [1997]

trimming: $\delta = 0.15$ (searching interval: 1976-2001)
critical values (Gregory and Hansen, 1996): $1\% : -5.47; 5\% : -4.95; 10\% : -4.68$,
***: significant at 1%; **: at 5%;*: at 10%;
Estimated breaks in brackets underneath the statistics.

To enhance the power of our testing procedure it is natural to try exploit the panel dimension, ignored so far. In designing a panel cointegration test suitable for our empirical task we have to take into account three 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; (*iii*) the breaks are heteregenous across countries. It is easy to check that none of the currently available tests (including those already applied for the same purpose of testing the FH puzzle by Banerjee and Carrion-i-Silvestre, 2004, and Gutierrez, 2009), fully satisfies these requirements. More specifically, the tests by Gutierrez (2009) and Westerlund (2006a,b) assume full cross-section independence[7], so that they are out of question. Both Banerjee and Carrion-i-Silvestre (2010) 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 independent from to those of the dependent variable[8], while the latter allow for common factors in the cointegrating residuals, but 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 for $T = 100$ (at best slightly higher than 50%; see Westerlund and Edgerton, 2008, tables 2 and 3). 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 tests for non-stationarity of the estimated idiosyncratic residuals and the estimated common factors *jointly* reject. The critical point is that, as Banerjee and Carrion-i-Silvestre themselves point out, the latter relies only on the time dimension. Hence, not surprisingly "we require T to be large for the statistic [for non-stationarity of the estimated common factors] to show good properties in terms of empirical size and power" (Banerjee and Carrion-i-Silvestre, 2010, p. 14), so that the power of the procedure to detect cointegration in the traditional sense will be determined by the size of the time sample.

We must conclude that a panel cointegration test with breaks and full cross-section dependence delivering an acceptable performance with small time samples is not available yet. In the next section we will tackle the task of developing it.

[7] 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).

[8] Although Banerjee and Carrion-i-Silvestre (2010) remark that this assumption may be relaxed, their available results depend on it.

3 Testing for Panel Cointegration with Breaks

3.1 Set-up

Let us consider for simplicity a standard bivariate panel cointegration set-up, with the right- and left-handside variables, denoted for generality by X and Y, observed over N units and T time periods, as usual indexed respectively by j and t. In the base case, dating back to Engle and Granger (1987), each unit X and Y are believed to be linked by a linear, but not necessarily cointegrating, relationship:

$$y_{jt} = \mu_{0j} + \beta_{0j} x_{jt} + \epsilon_{jt} \tag{4}$$

Equation (4) may be generalised to allow for time-varying coefficients with a break in period t_j^b:

$$y_{jt} = \begin{cases} \mu_{0j} + \beta_{0j} x_{jt} + \epsilon_{jt}, & t \leq t_j^b \\ \mu_{1j} + \beta_{1j} x_{jt} + \epsilon_{jt}, & t > t_j^b \end{cases} \tag{5}$$

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[9]. This can either be taken as the standard Gregory and Hansen $Min(ADF)$ tests, or the statistics corresponding to breakpoints estimated according to some other criterion. Westerlund and Edgerton (2008) suggest to estimate the breakpoint on the basis of the least squares criterion[10], which for stationary variables is consistent even under multiple breaks:

$$\widehat{t}_j^b = \underset{t_j^b \in [\delta T, (1-\delta)T]}{Arg \min} \left(\sum_{t=1}^{T} \widehat{\epsilon}_{jt}^2 (t_j^b) \right). \tag{6}$$

where the notation $\widehat{\epsilon}_{jt}(t_j^b)$ for the residuals emphasises their dependence on the breakpoint used in the estimation of model (5). For each unit the set of residuals $\left\{ \widehat{\epsilon}_{jt}(\widehat{t}_j^b) \right\}_{t=1}^{T}$ is by definition optimal in a least square sense, and can be used to compute a no-cointegration test with break alternative to $Min(ADF_j)$. Now, consider a first order autoregressive equation for the optimal cointegrating residuals:

$$\widehat{\epsilon}_{jt}(\widehat{t}_j^b) = \rho_{jt} \widehat{\epsilon}_{jt-1}(\widehat{t}_j^b) + \nu_{jt}. \tag{7}$$

When H_0 : "no cointegration" holds $\rho_j = 1$, while under cointegration $\left| \rho_j \right| < 1$. The hypothesis of no cointegration is then equivalent to $H_0 : \rho_j = 1$, and that of no panel cointegration as the same hypothesis for mean or median of this individual statistics. The latter arguably reflects more closely the usual definition of the panel null hypothesis as "no cointegration in the majority of the units" (for a detailed discussion of the relationship between summary statistic and alternative hypothesis, see Di Iorio and Fachin, 2010).

Two important remarks are in order here.

First, (7) is *not* a model of the cointegrating residuals; its purpose is only to define a parameter expressing the null hypothesis of interest. Second, and most important, the $\nu'_{jt}s$ are always stationary, either H_0 holds or not. They can thus be resampled via any resampling scheme valid for weakly dependent units, such as the stationary bootstrap (Politis and Romano, 1994).

A bootstrap testing algorithm along the lines put forth in Parker, Paparoditis and Politis (2006) and already exploited in Di Iorio and Fachin (2010), may then proceed as follows:

[9] 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.

[10] Using the the standard abbreviation for residual sum of squares, below we will refer to this criterion as *Argmin(RSS)*.

1. Estimate the breaking cointegrating model (5), with $\widehat{t}_j^b$ given by (6), on the dataset $\{y_{jt}, x_{jt}\}$, $j = 1, \ldots, N$, obtaining for each unit j estimates of the coefficients $(\widehat{\mu}_{rj}, \widehat{\beta}_{rj}, r = 0, 1)$ and of the optimal cointegrating residuals $\left\{\widehat{\epsilon}_{jt}(\widehat{t}_j^b)\right\}$;

2. Compute the N individual no cointegration statistics $\widehat{\theta}_j$ on the basis of the the optimal cointegrating residuals $\left\{\widehat{\epsilon}_{jt}(\widehat{t}_j^b)\right\}$;

3. Compute the summary statistics of interest, e.g. $\widehat{\theta}_{mean} = N^{-1}\sum_{i=1}^{N}\widehat{\theta}_i$, $\widehat{\theta}_{median} = median(\widehat{\theta}_1 \ldots \widehat{\theta}_N)$, according the alternative hypothesis of interest;

4. Compute $\widehat{\nu}_{jt} = \widehat{\epsilon}_{jt}(\widehat{t}_j^b) - \widehat{\rho}_j\widehat{\epsilon}_{jt-1}(\widehat{t}_j^b)$, where $\{\widehat{\epsilon}_{jt}\}$ are the optimal cointegrating residuals and $\widehat{\rho}_j$ is a consistent estimate (*e.g.*, OLS) of ρ_j;

5. Resample the series $\{\widehat{\nu}_{jt}\}$ via the stationary bootstrap:
   - generate $L_1, \ldots, L_T$ i.i.d. from a geometric distribution with parameter $\xi = 1/(1+B)$, $B =$ mean block size;
   - for each $t \in [1, T-1]$ let $K_t = \inf\{k : L_1 + \ldots + L_T \geq t\}$ and $M_t = L_1 + \ldots + L_{K_t}$;
   - generate $m_1, \ldots, m_K$ i.i.d. from a uniform distribution on $\{2, \ldots, T\}$;
   - for all $t \in [1, K]$ set $\nu_t^* = \widehat{\nu}_{\left[(m_{K_t}+(t-M_t)) \bmod (T-1)\right]+2}$.

6. Cumulate $\left\{\nu_{jt}^*\right\}$ obtaining pseudoresiduals $\left\{\epsilon_{jt}^*\right\}$ obeying the null hypothesis of no cointegration:
$$\epsilon_{jt}^* = \sum_{i=1}^{t}\nu_{jt}^* \quad t = 1, \ldots, T \tag{8}$$

7. For each unit j construct the pseudodata under the null hypothesis of no cointegration and break in $\widehat{t}_j^b$:
$$y_{jt}^* = \begin{cases} \widehat{\mu}_{0j} + \widehat{\beta}_{0j}x_{jt} + \epsilon_{jt}^* & t \leq \widehat{t}_j^b \\ \widehat{\mu}_{1j} + \widehat{\beta}_{1j}x_{jt} + \epsilon_{jt}^* & t > \widehat{t}_j^b \end{cases} \tag{9}$$

8. Using the datasets $\left\{y_{jt}^*, x_{jt}\right\}$ estimate for each unit j the breaking cointegrating model (5) for all possible breakpoints t_j^{b*},obtaining the corresponding sets of residuals $\left\{\widehat{\epsilon}_{jt}^*(t_j^{b*})\right\}$; estimate the optimal breakpoints $\widehat{t}_j^{b*} = Arg\min(\sum_{t=1}^{T}\widehat{\epsilon}_{jt}^{*2}(t_j^{b*}))$ and the optimal cointegrating residuals $\left\{\widehat{\epsilon}_{jt}^*(\widehat{t}_j^{b*})\right\}$;

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

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

11. Repeat 5-11 B times;

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

Considering that this is the generalisation of the test shown by Di Iorio and Fachin (2010) to be asymptotically valid, and that our main interest is the empirical application in a small sampe, we will evaluate the properties of this algorithm by simulation. Here we will report only the results obtained using the popular Augmented Dickey-Fuller test, which has the actractive feature of being closely related to the standard Gregory-Hansen $Min(ADF)$ statistics; the coefficient statistic $\tau_j = T(\rho_j - 1)$, used *e.g.*, by Palm, Smeekes and Urbain (2010), gave in all cases comparable results (available on request).

A final remark is that, although exploratory simulations showed the results to be quite robust to the choice of mean 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 mean block length at $T/10$, a simple choice which nevertheless delivered good results both in Paparoditis and Politis'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 mean block sizes respectively slightly larger and smaller than $T/10$ for small and large sample sizes (*e.g.*, 6 for $T = 40$ and 10 for $T = 160$). For moderate sample sizes the two rules suggest approximately the same lengths (exactly the same for $T = 80$). Note that since we will not use optimal mean block sizes we will somehow underrate the properties of the proposed test.

3.2 Monte Carlo Experiment

3.2.1 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 Di Iorio and Fachin (2010). Since we do not know if breaks took place or not we shall evaluate the performances of the test in both circumstances. We then first of all consider two variables linked by a constant parameter linear, not necessarily cointegrating, relationship:

$$
\begin{aligned}
y_{jt} &= \mu_{0j} + \beta_{0j} x_{jt} + \epsilon^y_{jt}, && (10)\\
\epsilon^y_{jt} &= \rho_j \epsilon^y_{jt-1} + e^y_{jt}, \qquad e^y_{jt} \sim N(0, \sigma^2_{jy}) && (11)
\end{aligned}
$$

where $j = 1, \ldots, N$, $t = 1, \ldots, T$. When X_j and Y_j are not cointegrated $\rho_j = 1$, while $|\rho_j| < 1$ when instead they are; in the power simulations ρ_j 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 $\sigma^2_{jy} \sim Uniform(0.5, 1.5)$, while with no loss of generality $\mu_{0j} = \beta_{0j} = 1\ \forall j$. In the more general case of heterogenous breaks in the coefficients equation (10) is replaced by

$$
y_{jt} = \begin{cases} \mu_{0j} + \beta_{0j} x_{jt} + \epsilon^y_{jt}, & t \leq t^b_j \\ \mu_{1j} + \beta_{1j} x_{jt} + \epsilon^y_{jt}, & t > t^b_j \end{cases} \tag{12}
$$

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_{jt}):

$$
x_{jt} = \gamma_{1j} F_{1t} + \gamma_{2j} F_{2t} + \epsilon^x_{jt} \tag{13}
$$

Following Pesaran (2007) the factor loadings are chosen so to ensure substantial cross-correlation in the $X's$: $\gamma_{rj} \sim Uniform(-1, 3)\ \forall j$ and $r = 1, 2$. The common factors are generated as follows:

$$
\begin{bmatrix} F_{1t} \\ F_{2t} \end{bmatrix} = \begin{bmatrix} F_{1t-1} \\ 0.4F_{2t-1} \end{bmatrix} + \begin{bmatrix} f_{1t} \\ f_{2t} \end{bmatrix} \tag{14}
$$

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

$$\begin{bmatrix} f_{1t} \\ f_{2t} \end{bmatrix} = \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix} + \begin{bmatrix} \vartheta_1 & 0 \\ 0 & \vartheta_2 \end{bmatrix} \begin{bmatrix} \eta_{1t-1} \\ \eta_{2t-1} \end{bmatrix} \tag{15}$$

$$\epsilon_{jt}^x = e_{jt}^x + \varphi e_{jt-1}^x, \tag{16}$$

where $\eta_{rt} \sim N(0,1), r = 1,2$, and $e_{jt}^x \sim N(0, \sigma_{jx}^2)$, with $\sigma_{jx}^2 \sim Uniform(1, 1.4)$. 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; all these cases will be simulated with and without breaks in the DGP for the $Y's$.

Given that in some of our simulations the time sample is quite small to ensure computational stability we choose a trimming coefficient $\delta = 0.20$.

1. *Base case:* $T = 40$, $N = 5, 10, 20, 40$; break date Uniform over units in $[0.5T \pm 3] = [17, 23]$. This is the time sample of our empirical dataset. 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 still 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.

3.2.2 Results

The results of the simulations are reported in Tables 2-7 below, and rapidly summarised. Within each set of simulations (with or without breaks in the coefficients) there seem to be very little differences between the performances of the mean and median tests; second, the results are quite similar across the different block sizes as well. Hence, our comments can be expressed in general terms.

First of all, when breaks are present in the DGP (tables 2-4), with a small time sample ($T = 40$) Type I errors are somehow smaller than nominal sizes, to which they converge rapidly for $T = 80$. Power can be disappointing when T and N are both small, but it also increases rapidly with T and, most importantly, with N: in our simulations the rejection rate of the false null of no cointegration is 100% in almost all cases for $T = 80$, and always for $T = 160$. 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 when breaks are present in the DGP our panel cointegration test with breaks may grant power incrasing with the cross-section dimension, while ensuring good size control.

Of course, in practice we do not know the DGP, so that we may wrongly apply this test allowing for breaks to a dataset generated by a DGP with no breaks. In these circumstances the performances of the test turn out to be in some cases quite different (Tables 5-7). With small or moderate time samples (in our set-up T up to 80) the size bias is massive, and even worse, it increases with N. The bias, however, completely disappears for $T = 160$, with power always high. The explanation of these findings is quite simple. When the DGP is equation (10), which has no breaks in the systematic part, model (5), which allows for time-varying coefficients, is misspecified, with explanatory power spuriously shifted from the residuals to the systematic part. In fact, with small time samples a single break in the coefficients may improve the fit of the model enough to produce residuals observationally equivalent to realisations of stationary processes. The no cointegration ADF statistics will then be spuriously large in absolute value. On the other hand, the bootstrap DGP (9) does have time-varying coefficients. Hence, the breaking model (5) is correctly specified for the bootstrap data, and its residuals will approximate well the true non stationary bootstrap noise (8). The bootstrap ADF statistics will then be generally small. Asymptotically a single break is instead no more sufficient to produce spurious mean reversion in the residuals of a model estimated with non cointegrating data, so that the size bias disappears.

An important remark here is that Gregory and Hansen's (1996) *Min(ADF)* test also ovverejects when the are no breaks in the DGP. In fact, our results are quite consistent with Gregory and Hansen's. For instance, for a *Min(ADF)* test with $\alpha = 0.05$ they report a Type I error of 0.17 and power of 0.49 ($T = 50$; Gregory and Hansen, 1996, table 2). With a similar time sample size ($T = 40$) and N from 5 to 40 we have Type I error and power ranging respectively from 0.10 to 0.39 and from to 0.54 to 1 (mean block 6, values very similar with mean block 4; see table 5). Summing up, with small to moderate time samples our test will have correct size if breaks are actually present in the DGP, but, similarly to Gregory and Hansen's, will be oversized if there are not. This implies that with small samples no rejections can definitely be considered as reliable, while rejections may be spurious. Hence, the application of the test requires some care.

Table 2

Size and Power of bootstrap panel cointegration tests with breaks

$T = 40$, *N from* 5 *to* 40, *breaks in DGP*

		N							
		5	10	20	40	5	10	20	40
	α	*Mean*				*Median*			
Mean block		*Size*							
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
		Power							
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

DGP: with breaks; eqs. (11)-(16);
size: $\rho_i = 1\ \forall i$;
power: $\rho_i \sim Uniform(0.6, 0.8)$;
$t_i^b \sim Uniform(0.5T \pm 3)$, search interval: $[0.2T, 0.8T]$;
Mean/median: bootstrap test on the mean/median across units of the no cointegration statistics;
Bootstrap: 500 redrawings; mean block: $4 = 0.10T$, $6 = 1.75\sqrt[3]{T}$.
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 3

Size and Power of bootstrap
panel cointegration tests with breaks

$T = 80$, *N from* 5 *to* 20, *breaks in DGP*

	N					
	5	10	20	5	10	20
α	*Mean*			*Median*		
	Size					
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
	Power					
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

DGP: eqs. (11)-(16),
Mean block: $8 = 0.10T \simeq 1.75\sqrt[3]{T}$
All definitions and abbreviations: see Table 2.

Table 4

Size and Power of bootstrap
panel cointegration tests with breaks
$T = 160$, N =5 *and* 10, *breaks in DGP*

		N			
		5	10	5	10
Block	α	*Mean*		*Median*	
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

DGP: eqs.(11)-(16);
Mean block: $16 = 0.10T, 10 = 1.75\sqrt[3]{T}$;
Power: 1.00 in all cases;
All definitions and abbreviations: see Table 2.

Table 5

Size and Power of bootstrap panel cointegration tests with breaks
$T = 40$, N *from* 5 *to* 40, *no breaks in DGP*

		N							
		5	10	20	40	5	10	20	40
		Mean				*Median*			
Mean block	α	*Size*							
4	0.01	0.03	0.03	0.06	0.16	0.02	0.03	0.04	0.09
	0.05	0.09	0.15	0.25	0.42	0.10	0.11	0.17	0.32
	0.10	0.20	0.26	0.39	0.59	0.17	0.22	0.32	0.50
6	0.01	0.03	0.03	0.06	0.12	0.02	0.03	0.05	0.10
	0.05	0.10	0.14	0.25	0.39	0.10	0.12	0.19	0.30
	0.10	0.21	0.25	0.38	0.58	0.19	0.22	0.33	0.48
		Power							
4	0.01	0.24	0.48	0.84	0.99	0.16	0.35	0.69	0.92
	0.05	0.56	0.78	0.97	1.00	0.44	0.67	0.90	0.99
	0.10	0.72	0.89	0.99	1.00	0.61	0.79	0.96	1.00
6	0.01	0.23	0.46	0.82	0.98	0.15	0.34	0.67	0.91
	0.05	0.54	0.76	0.96	1.00	0.42	0.65	0.89	0.99
	0.10	0.70	0.88	0.98	1.00	0.59	0.78	0.96	1.00

DGP: eqs. (10)-(11),(13)-(16);
All definitions and abbreviations: see Table 2.

Table 6

Size and Power of bootstrap

panel cointegration tests with breaks

$T = 80$, *N from* 5 *to* 20, *no breaks in DGP*

	N					
	5	10	20	5	10	20
α	*Mean*			*Median*		
	Size					
0.01	0.06	0.09	0.12	0.05	0.07	0.13
0.05	0.17	0.22	0.30	0.15	0.22	0.29
0.10	0.26	0.35	0.45	0.26	0.32	0.42
	Power					
0.01	0.96	1.00	1.00	0.89	0.99	1.00
0.05	1.00	1.00	1.00	0.99	1.00	1.00
0.10	1.00	1.00	1.00	0.99	1.00	1.00

DGP: eqs. (10)-(11),(13)-(16);
Mean block: $8 = 0.10T \simeq 1.75\sqrt[3]{T}$;
All definitions and abbreviations: see Table 2.

Table 7

Size and Power of bootstrap

panel cointegration tests with breaks

$T = 160$, N =5 *and* 10, *no breaks in DGP*

		N			
		5	10	5	10
Mean block	α	*Mean*		*Median*	
10	0.01	0.00	0.01	0.01	0.01
	0.05	0.03	0.05	0.04	0.04
	0.10	0.09	0.10	0.08	0.09
16	0.01	0.00	0.01	0.01	0.01
	0.05	0.03	0.05	0.04	0.04
	0.10	0.09	0.10	0.08	0.10

DGP: eqs. (10)-(11),(13)-(16);
Mean block: $10 \simeq 1.75\sqrt[3]{T}; 16 = 0.10T$;
Power: 1.00 in all cases;
All definitions and abbreviations: see Table 2.

4 A breaking panel cointegration analysis of the Feldstein-Horioka puzzle

We can now apply the testing procedure developed in the previous section to our data. Before computing the tests it is important to check if the estimated breakpoints are compatible with the changes in capital openness, the most likely source of breaks in the savings-investment relationship. Comparing the estimated breakpoints (Table 8; note that as discussed before, the search was not performed for Canada, Germany and USA, where capital movements had been already completely liberalised by 1970) with the changes in the Chinn-Ito index (Figs. 3-4) we can see that this seems to be generally the case. The only exception is Belgium, where the estimated break, 1984, falls a few years before the large increase in the openness index associated with the 1990 EU liberalisation. In most european countries (Austria, Finland, France, Italy, Netherlands, Sweden) the breaks fall between 1993 and 1995, with a plausible lag of a few years after the 1990-92 EU liberalisation directives. An analogous lag of two years is also found in the UK, where the break is placed in 1981 and the policy change took place in 1979.

We then proceed to compute the summary panel cointegration statistics, combining ADF statistics computed on the residuals of the long-run FH equation without breaks, equation (1), for Canada, Germany and USA, and with heterogenous breaks, equation (2), for all the other countries. We will then check for the robustness of the results with respect to Belgium. Since, consistently with the results reported in Di Iorio and Fachin (2010), we find the relationship to be spurious (inverse) in Greece this country is excluded from the panel (estimates not reported here, available on request). As it can be appreciate from Table 9, even allowing for a break in Belgium the bootstrap p-values are greater than 0.10 (marginally in mean, 0.11, clearly so in median, 0.17). As to be expected, with no break in this country both p-values increase, respectively to 0.14 (mean) and 0.18 (median).

From the discussion in the previous section we know that these tests have size close to nominal when there are breaks in the DGP, and larger than it when there are not. In either case, no rejection, the conclusion reached for our dataset, is a reliable outcome. We can then conclude that, even allowing for a break in all countries where the degree of capital openness changed within our study period, there is no evidence of a FH puzzle for our panel as a whole.

This does not completes our study, though. The time-series tests discusses in Section 2 suggest, consistently with Di Iorio and Fachin (2010), that a savings-investment relationship may hold in some countries of the panel. To investigate the issue further we estimated by FM-OLS equation (2) for these countries, with breaks chosen to minimise the residual sum of squares. The estimates of the saving-retention ratios are reported in Table 10. A first remark is that except in one case (Japan, after the break) all values are smaller than 1. Further, in all cases except Australia and Japan, the coefficients after the breaks are, as expected, lower than in the first part of the sample. In fact, in Finland and the Netherlands the coefficients after the break are not significant (implying no relationship), but given the very small size of the subsamples it is safer not to put too much weight on this finding. In Italy, where the break is placed in 1994, the relationship becames even inverse. From Fig. 2 we can see that indeed in this country savings increased rapidly in the first 1990's, then declined, while investments first fell and then rose slowly; both paths are to a large extent the consequence of the fiscal policies implemented during the period. More precisely, the increase of the early '90's is associated with the 1992-1995 currency crises, which forced the Italian autorities to implement very restrictive policies[11]. Expenditure cuts and tax increases led in 1992, for the first time after many years of large deficits, to a primary surplus. This reached its maximum level (6.7%) in 1997, when the local maximum of the saving rate also falls. The marked improvements in the state of public finances allowed Italy to participate in 1998 to the formation of the European Economic and Monetary Union. Thanks to this, interest rates, and therefore debt service, declined significantly (debt service as a GDP ratio fell by -4.3% between 1997 and 2006: see Marino, Momigliano and Rizza, 2008), causing the need of large primary surpluses to become less stringent. As a result, these, although in principle still a target of fiscal policy, in practice also declined, reaching a minimum of only 0.5% of GDP in 2007. Obviously, the shift towards less restrictive fiscal policies and the fall in interest rates also explain the growth in the investment rate after the mid-90's.

In both Australia and Japan we instead find saving retention ratios unexpectedly larger in the second part of the samples (breaks are estimated to fall respectively in 1981 and 1984). However, the two cases are quite different. In Japan the coefficient before the break is very small and negative, which is puzzling in view of both the *a priori* expectations and the *ex-post* graphical evidence. However, the variance of the estimate is very high, so that it is safer not to draw any conclusions in this case either.

In Australia the explanation is the different reaction of the two variables to the 1974 and early

[11] Bewtween 1991 and 1995 the *Lira* devaluated by nearly 30%. Since at the same time capital controls were being dismantled the risk of a public debt crisis was very high (see *e.g.*, Rossi, 2007).

1990's recessions. The saving rate fell rapidly in both cases, while investments only in the latter, when monetary policy was considerably tighter (see *e.g.*, Nelson, 2004). It is then not surprising that the elasticity turns out to be higher after 1984 then beforehand. An important remark is that, how it can be appreciated from the plots in Fig. 1, Australia is the only country where investment rate generally exceed the saving rate[12], after the mid-'70's largely so. Hence, although the a long-run relationship linking savings and investment did seem to exist, the latter have not been constrained by the former, as they would under no capital mobility (strict FH hypothesis[13]).

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

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

trimming: $\delta = 0.15$ (searching interval: 1976-2001)

Table 9
Investment and Savings, 1970-2007
Bootstrap panel cointegration tests with breaks

$Mean(ADF)$	$Median(ADF)$
-3.68 (0.11)	-3.72 (0.17)

Mean/Median: mean/median of the individual statistics;
In brackets: bootstrap p-values, 1000 redrawings;
Mean block size: $6 \simeq 1.75\sqrt[3]{T}$;
Panel: Austria, Australia, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, Portugal, Spain, Sweden, UK, USA;
Break: $Argmin(RSS)$ with trimming: $\delta = 0.15$, in all countries except Canada, Germany and USA.
Lag selection: Ng and Perron (1995)

[12] In Denmark and Finland this happened only before the estimated breakpoints (see Table 12).

[13] Note that this extreme view was never held by Feldstein and Horioka themselves: "The findings of Feldstein and Horioka [...] do not imply that there is no capital mobility [...] it is reasonable to intepret the Feldstein and Horioka findings as evidence that there are substantial imperfections in the international capital market and that a very large share of domestic savings tend to remain in the home country" (Feldstein, 1982, p. 3)

Table 11

Saving and Investments in the long-run:

FM-OLS estimates of the long-run saving retention ratio

	Australia	Denmark	Finland	Italy
β_0	0.38 [0.13]	0.76 [0.17]	0.54 [0.90]	0.69 [0.10]
β_1	0.98 [0.10]	0.37 [0.42]	−0.15 [0.25]	−0.63 [0.26]
	Japan	Netherlands	Sweden	
β_0	−0.02 [0.35]	0.84 [0.33]	0.38 [0.05]	
β_1	1.34 [0.11]	0.05 [0.76]	0.27 [0.16]	

β_0, β_1 : long-run saving retention ratio before and after the breaks, see model (2)
Breaks: see Table 8;
Standard errors: in brackets.

Table 12

Average Saving/GDP and Investments/GDP ratios

	Australia		Denmark		Finland	
	1970-1981	*1982-2007*	*1970-1991*	*1991-2007*	*1970-1994*	*1995-2007*
I/GDP	0.28	0.25	0.22	0.20	0.27	0.19
S/GDP	0.26	0.20	0.19	0.22	0.25	0.26
	Italy		Japan		Netherlands	
	1970-1994	*1995-2007*	*1970-1984*	*1985-2007*	*1970-1994*	*1995-2007*
I/GDP	0.23	0.20	0.33	0.25	0.23	0.21
S/GDP	0.23	0.21	0.34	0.27	0.26	0.27
			Sweden			
			1970-1994	*1995-2007*		
		I/GDP	0.21	0.17		
		S/GDP	0.25	0.23		

5 Conclusions

In a world with free capital movements there are no reasons to expect any long-run relationship between savings and investments. Available empirical tests of this apparently very simple statement are however not conclusive, either because based on procedures likely to have low power, hence possibly unable to reject false null hypotheses of no relationship (the time series no cointegration test applied by, *e.g.*, by Kim, 2001) or, on the opposite, strongly oversized, hence possibly rejecting true null hypotheses of no relationship (the first generation panel cointegration tests, assuming independence across units, applied, *e.g.*, by Pelgrin and Schich, 2008). In Di Iorio and Fachin (2010) we showed that a carefully designed bootstrap panel cointegration test with good power properties does support the expectation of no relationship for a panel of 18 OECD economies (Austria, Australia, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Portugal, Spain, Sweden, UK, USA) over the period 1970-2007. However, this does not exclude the existence of a long-run relationship with breaks in the coefficients, plausible in view of the changes in capital movements regulations which took place in those decades. The standard Gregory-Hansen (1996) time series test of no cointegration with breaks rejects the hypothesis of no relationship only in a minority of this panel (Australia, Denmark, Finland, Italy, Japan, Netherlands, Sweden), but, since it is known to have rather low power, cannot be taken as a conclusive piece of evidence. A more powerful procedure is needed.

Building upon Di Iorio and Fachin (2010), in this paper we propose to solve this challenging task through a novel bootstrap panel cointegration test with breaks. Simulation results suggest that when breaks are present in the DGP this new test has good size and power properties. When there are no breaks it is (like Gregory and Hansen's) severely oversized with small to moderate time samples, but asymptotically it has correct size. Power is always satisfactory. Provided it is used with some care, taking into account that allowing for redundant breaks will bias the test towards rejection, it can thus be a potentially useful addition to the toolbox for non-stationary panel analysis[14].

The conclusions brought by the application of this new test to our panel of 18 OECD economies for the period 1970-2007 are quite clear-cut. Even allowing for a break in all countries where the degree of capital openness changed within our study period (all but Canada, Germany and USA) the null hypothesis of no cointegration is not rejected. Since this is always a reliable conclusion, both when there actually are breaks in the DGP and not, we may safely conclude that there is no evidence of a FH puzzle in the core of OECD as a whole over the last four decades.

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

Apergis, N. and C. Tsoumas (2009) "A survey of the Feldstein Horioka puzzle: What has been done and where we stand" *Research in Economics,* 63, 64-76.

Bakker, A. and B. Chapple (2002) "Advanced Country Experiences with Capital Account Liberalization" *Occasional Paper* n. 214, International Monetary Fund, Washington D.C. (USA).

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

Banerjee, A. and J.L. Carrion-i-Silvestre (2010) "Cointegration in Panel Data With Breaks and Cross-Section Dependence", *Mimeo,* University of Birmingham.

Banerjee A. M. Marcellino, and C. Osbat (2004) "Some Cautions on the Use of Panel Methods for Integrated Series of Macro-Economic Data" *The Econometrics Journal,* 7: 322-340.

Chinn, M. and H. Ito (2008) "A New Measure of Financial Openness", *Journal of Comparative Policy Analysis*, 10, 309-322.

[14] More precisely, it should be considered as complementary to asymptotic factor methods procedures (Banerjee and Carrion-i-Silvestre, 2010, 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.

Cooray, A.V. and B. Felmingham (2007) "The cyclical and trend behaviour of Australian investment and savings" University of Wollongong, http://ro.uow.edu.au/commpapers/583.

Di Iorio, F. and S. Fachin (2010) "A Panel Cointegration study of the long-run relationship between Savings and Investments in the OECD economies, 1970-2007" *MRPA Working Paper* n. 25873.

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

European Commission (2010) *Report on Greek government deficit and debt statistics* Brussells (BE).

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

Feldstein, M. (1982) "Domestic Savings and International Capital Movements in the Long Run and in the Short Run" *NBER Working Paper Series* n. 947.

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 (2010) "Panel unit root tests in the presence of cross-sectional dependencies: Comparison and implications for modelling" *Econometric Reviews*, 29, 111-145.

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

Gregory, A.W. and B.E. Hansen (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,* 17, 197-200.

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

Kim, S.H. (2001) "The saving–investment correlation puzzle is still a puzzle" *Journal of International Money and Finance,* 20, 1017–1034.

Marino, M.R., S. Momigliano and P. Rizza (2008) "I conti pubblici nel decennio 1998-2007: fattori temporanei, tendenze di medio periodo, misure discrezionali" Banca d'Italia, *Questioni di Economia e Finanza*, n. 15.

Nelson, E. (2004) "Monetary Policy Neglect and the Great Inflation in Canada, Australia, and New Zealand" Federal Reserve Bank of St. Louis, *Working Paper* 2004-008A.

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

OECD (2002) *Forty Years' Experience with the OECD Code of Liberalisation of Capital Movements* OECD, Paris.

Özmen, E. and K. Parmaksiz (2003) "Policy regime change and the Feldstein-Horioka puzzle: the UK evidence" *Journal of Policy Modelling,* 25, 139-149.

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

Paparoditis, E., and D.N. Politis (2003) "Residual-based block bootstrap for unit root testing" *Econometrica* 71, 813-855.

Pelgrin, F. and S. Schich (2008) "International capital mobility: What do national saving-investment dynamics tell us?" *Journal of International Money and Finance,* 27, 331-344.

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.

Politis, D.N., and J.P. Romano (1994) "The Stationary Bootstrap" *Journal of American Statististical Association,* 89, 1303-1313.

Rossi, S. (2007) *La politica economica italiana 1968-2007* Laterza, Roma-Bari.

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