Testing for cointegration in dependent panels via residual-based bootstrap methods

Di Iorio, Francesca and Fachin, Stefano

University of Rome "La Sapienza"

9 May 2007

Testing for cointegration in dependent panels via residual-based bootstrap methods

Francesca Di Iorio
University of Naples Federico II

Stefano Fachin
University of Rome "La Sapienza"

Abstract

We address the issue of panel cointegration testing in dependent panels, showing by simulations that tests based on the stationary bootstrap deliver good size and power performances even with small time and cross-section sample sizes and allowing for a break at a known date. They can thus be an empirically important alternative to asymptotic methods based on the estimation of common factors. Potential extensions include test for cointegration allowing for a break in the cointegrating coefficients at an unknown date.

Keywords: Panel Cointegration, Stationary Bootstrap, Wild Bootstrap.
JEL codes: C23, C15

Revised November 2008

1 Introduction[1]

The rate of expansion of the literature on the analysis of non-stationary panels, as revealed *e.g.,* by a comparison of the list of references in the surveys by Banerjee (1999) and Breitung and Pesaran (2006) is impressive. This growing interest is due to good reasons: first of all, many important economic questions are naturally framed in a panel perspective (for instance, the Purchasing Power Parity issue, Pedroni, 2004, and migrations, Fachin, 2007). Further, when only small time samples are available, adding the cross-section dimension grants considerable improvements of the small samples properties of testing procedures, provided the possible linkages across units are properly accounted for. This issue is currently actively investigated in the literature, with two[2] main solutions being suggested: (i) modelling the linkages as due to unobserved common factors; these can be estimated by principal components methods (Bai and Ng, 2004) and then removed from the data so to apply simple procedures for independent panels (Banerjee and Carrion-i-Silvestre 2006, Gegenbach, Urbain and Palm, 2006, Westerlund 2008); (ii) apply bootstrap algorithms designed to deliver estimates of the distribution of the statistics of interest conditional on the cross-section linkages as present in the dataset at hand. Concentrating on (no-) cointegration tests, to the best of our knowledge two bootstrap approaches have been put forth so far. Fachin (2007) applies the Continuous-Path Block bootstrap (Paparoditis and Politis, 2001, 2003) separately to the right- and the left-hand side variables, hence generating unrelated pseudoseries obeying the null hypothesis of no cointegration, while Westerlund and Edgerton (2007) develop a sieve bootstrap procedure for testing the null of cointegration.

Unfortunately, neither the common factor nor the existing bootstrap approaches are fully satisfactory. Let us discuss them in turn. A first problem with the common factor approach is that, as Gegenbach, Urbain and Palm (2006) explicitly admit, it requires large samples. Thus, although investigating the possible common factor structure of the data could be very important for its own sake, in many empirical applications the available information set may simply be not rich enough. A second problem is that it hinges upon a series of assumptions which may be very restrictive: Banerjee and Carrion-i-Silvestre (2006) and Westerlund (2008) allow for common factors

[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. Comments and suggestions from participants to the conferences on "Bootstrap and Time Series" (Kaiserlautern, June 2008) and " Factor Structures for Panel and Multivariate Time Series Data" (Maastricht, September 2008) are gratefully acknowledged. Special thanks to Joakim Westerlund for help with his dataset. Correspondence to: s.fachin@caspur.it, fdiiorio@unina.it.

[2] We are excluding the full information maximum likelihood approach (Groen and Kleibergen, 2003) which, requiring the time dimension to be much larger than the cross-section dimension, is of very limited empirical interest.

in the cointegrating residuals but not in the variables themselves; this more general model is adopted by Gegenbach, Urbain and Palm (2006), who however assume a full rank, block-diagonal matrix of loadings, hence ruling out the empirically relevant case of a single source of non-stationarity common across units and variables. For instance, in the case of regional consumption and income this may be a stochastic trend in national GDP.

Block bootstrap, model-free methods were showed by Fachin (2007) to be empirically useful tools in tackling the problems at hand. However his algorithm destroys *any* relationship between the modelled variables, not only long-run ones. On the other hand, the sieve bootstrap (shown to be valid for inference on cointegrating regressions by Chang, Park and Song, 2006) hinges upon the assumption of a linear structure of the cointegrating residuals[3].

In this paper we shall try to improve on the existing bootstrap methods. Our main conjecture is that Parker, Paparoditis and Politis' (2006) Residual-based Stationary Bootstrap test for unit roots may be applied to the estimated cointegrating residuals. In fact, the potential of block bootstrap methods in this field is stressed by Chang *et al.* (2006). A further viable route which we will explore is to extend Herwartza and Neumann's (2005) Wild Boostrap (WB) procedures for inference on cointegrating coefficients to panel cointegration testing. In this paper we will thus first outline both approaches (section 2), evaluate their small sample performances by simulation (section 3) and finally draw some conclusions (section 4).

2 Single-equation panel cointegration testing via residual-based bootstrap: set-up

Parker, Paparoditis and Politis (2006), henceforth PPP[4], developed a bootstrap unit root test based on the stationary bootstrap (Politis and Romano, 1994), a resampling method suitable for weakly dependent series. In this method the resampling is carried out chaining blocks of observations of the originary series of random length starting at random locations, and thus reproduce the weak dependence properties of the latter. The extension of PPP Residual-based Stationary Bootstrap (RSB) unit root tests to single-equation cointegration testing is straightforward.

Consider for simplicity two $I(1)$ variables, X and Y, linked by a linear

[3]Further, the sieve bootstrap cointegration test proposed by Westerlung and Edgerton (2007) may deliver poor power in small samples. Their procedure involves estimating the sieve through the Yule-Walker equations, so to obtain stationary bootstrap residuals obeying the null of cointegration. However, under no cointegration the bootstrap residuals, though stationary, will have a root arbitrarily close to 1, very difficult to distinguish from a unit root in small samples.

[4]Not to be confused with the acronym for Purchasing Power Parity.

relationship

$$y_t = \mu + \beta x_t + \epsilon_t, t = 1, \ldots, T \tag{1}$$

Consider then the equation

$$\epsilon_t = \rho\epsilon_{t-1} + \nu_t \tag{2}$$

It is immediately seen that when H_0 : no cointegration holds $\rho = 1$,while when it does not $|\rho| < 1$. The hypothesis of no cointegration is then equivalent to H_0: $\rho = 1$. Two important remarks are in order here. First, (2) is *not* a model of the cointegrating residuals; its purpose is only to define a parameter expressing the null hypothesis of interest. Second, the $\nu_t' s$ are always stationary, either H_0 holds or not: they can thus be resampled via the stationary bootstrap.

An algorithm along the lines put forth in PPP, mean zero case, may then proceed as follows:

1. Compute $\widehat{\nu}_t = \widehat{\epsilon}_t - \widehat{\rho}\widehat{\epsilon}_{t-1}$, where $\{\widehat{\epsilon}_t\}$ are the estimated residuals and $\widehat{\rho}$ is the OLS estimate of ρ;
2. Resample the series $\{\widehat{\nu}_t\}$ via the stationary bootstrap, obtaining $\{\nu_t^*\}$;
3. Cumulate $\{\nu_t^*\}$ obtaining pseudoresiduals $\{\epsilon_t^*\}$ obeying the null hypothesis of no cointegration;
4. Compute $y_t^* = \widehat{\mu} + \widehat{\beta}x_t + \epsilon_t^*$;
5. Estimate the cointegrating regression on the dataset $\{y_t^*, x_t\}$: $y_t^* = \widehat{\mu}^* + \widehat{\beta}^* x_t + \widehat{\epsilon}_t^*$;
6. Estimate ρ^* applying (2) to the residuals $\widehat{\epsilon}_t^*$;
7. Repeat 2-6 B times;
8. Test the hypothesis $\mathrm{H}_0 : \rho = 1$ on the basis of the distribution of the $\rho^{*\prime}s$, which obey it. Note that the consistency results reported in PPP are in fact general enough to allow the use of more general statistics function of ρ, such as the ADF.

The model-based sieve bootstrap, applied by Westerlund and Edgerton (2007), replaces (2) with a linear autoregressive model (in fact, the Augmented Dickey-Fuller equation):

$$\epsilon_t = \rho\epsilon_{t-1} + \sum_{j=1}^{p} \phi_j \Delta\epsilon_{t-j} + \widetilde{\nu}_t \tag{3}$$

so to obtain empirically white noise $\widetilde{\nu}_t's$ on which simple resampling may be applied. Note that the autoregressive polynomial plays here the same role of the block length in the RSB algorithm, *i.e.* capturing the memory of the process. The bootstrap residuals $\{\epsilon_t^*\}$ are then constructed recursively on the basis of (3).

A test following a closely related approach can be constructed building upon Herwartza and Neumann's (2005) WB procedures for inference on cointegrating coefficients. More precisely, the algorithm would replace steps 1-3 above with the following:

1. Compute $\widehat{\widetilde{\nu}}_t = \widehat{\epsilon}_t - \widehat{\rho}\widehat{\epsilon}_{t-1} - \sum_{j=1}^{p} \widehat{\phi}_j \Delta\widehat{\epsilon}_{t-j}$;
2. Resample the series $\left\{\widehat{\widetilde{\nu}}_t\right\}$ via WB, obtaining $\{\nu_t^*\}$;
3. Construct recursively $\{\nu_t^*\}$ from the AR equation in [1] under $H_0 : \rho = 1$, obtaining pseudoresiduals $\{\epsilon_t^*\}$ obeying the null hypothesis of no cointegration.

Let us now introduce the panel dimension, ignored so far. The basic idea of unit root and panel cointegration tests is that of achieving power gains by pooling or averaging (the latter approach being more general, as no homogeneity constrains are imposed) the information from N individual units, indexed by i in the following discussion. However, the null and alternative hypothesis of these tests are a delicate issue worth a careful discussion (see also Pedroni, 2004). Given the null hypothesis of no cointegration in all units, *i.e.* $H_0 : \rho_i = 1$ for $i = 1, \ldots, N$, we can define four different alternative hypothesis:

(i) $\mathrm{H}_1 : \rho_i < 1$ in all units;

(ii) $\mathrm{H}_1 : \rho_i < 1$ in at least one unit;

(iii) $\mathrm{H}_1 : \rho_i < 1$ in most of the units;

(iv) $\mathrm{H}_1 : \rho_i < 1$ in most of units *or* $\rho_i << 1$ in a smaller number of units;

The choice of the alternative hypothesis dictates the statistic which should be used to summarise the $\rho_i's$. More precisely, the are easily seen to be the following:

(i) $G = Max(\rho_i)$;

(ii) $G = Min(\rho_i)$;

(iii) $G = Median(\rho_i)$;

(iv) $G = Mean(\rho_i)$;

Clearly, the same holds for transformations of ρ such as the ADF.

The alternative hypothesis (i) and (ii) are obviously of little interest, while the main merit of (iv) can be argued to be that of justifying the use of the mean of the individual statistics as panel cointegration statistic. In fact, the alternative hypothesis best reflecting the idea that the testing procedure should find which hypothesis best describes the panel *as a whole* is clearly (iii). The median of the individual statistics should then be considered the basic panel statistic. However, this statistic is of notoriously difficult treatment by asymptotic methods. As a consequence, with the only exception of Fachin (2007) who examine the performance of a median-based test, case (iv) is the only one considered in the literature (for instance as "Group Mean test" in Pedroni, 1999). This stresses again the strong potential of bootstrap methods in panel cointegration testing, and leads us to the next question: how to extend the algorithms outlined above to panel data sets? In fact, the task turns out to be easily accomplished. As we have seen in the Introduction an essential feature to be taken into account is dependency across units. In order to reproduce it in the pseudoseries, in both cases we simply need to apply the resampling algorithm to the entire cross-sections. In this way the (short- and long-run) cross-correlation structure of the data is exactly reproduced in the bootstrap data. More precisely, letting $\widehat{\nu}_{it} = \widehat{\epsilon}_{it} - \widehat{\rho}_i \widehat{\epsilon}_{it-1}$, in step 2 of the RSB algorithm we apply the stationary boostrap to the entire $T \times N$ matrix of the residuals $\mathbf{V} = [\widehat{\boldsymbol{\nu}}_1 \ldots \widehat{\boldsymbol{\nu}}_N]$, where $\widehat{\boldsymbol{\nu}}_j = [\widehat{\nu}_{1j} \ldots \widehat{\nu}_{Tj}]'$, while in the same step of the WB algorithm the pseudoresiduals are generated as $\boldsymbol{\nu}_t^* = \zeta_t \widetilde{\boldsymbol{\nu}}_t$, where $\zeta_t \tilde{\ } IID(0,1)$, $\widetilde{\boldsymbol{\nu}}_t = [\widetilde{\nu}_{t1} \widetilde{\nu}_{t2} \ldots \widetilde{\nu}_{tN}]$. In the final step the statistic of interest becames either the median or the mean of the cointegration ADF statistics computed for each of the N units, so that the bootstrap estimate of the significance level of the test is $p^* = prop(S^* < \widehat{S})$, $S = Median(\mathbf{ADF})$ or Mean($\mathbf{ADF}$) where $\mathbf{ADF} = [ADF_1 \ldots ADF_N]$.

3 Monte Carlo

3.1 Design

We will base our simulations on a DGP which is essentially a generalisation of the classical Engle and Granger (1987) DGP to the case of dependent panels, with the design of the panel structure related to those used by Kao (1999), Fachin (2007), and Gegenbach *et al.* (2006)[5]. Before discussing the details of the design a remark in order is that since panel DGPs are inevitably very complex, simulation experiments are computationally very demanding. Hence, rather than aiming at the unfeasible task of a complete design our aim will be that of defining an empirically relevant set-up.

[5] Several parameters are in fact fixed at the values used by the latter.

In our base case in the spirit of conditional modelling we assume a variable of interest, Y, known to be linked by a linear, possibly cointegrating, relationship to a right-hand side variable[6] X:

$$\begin{cases} y_{it} = \mu_{0i} + \beta_i x_{it} + \epsilon^y_{it} \\ \epsilon^y_{it} = \rho_i \epsilon^y_{it-1} + e^y_{it}, \qquad e^y_{it} \sim N(0, \sigma^2_{iy}) \end{cases} \tag{4}$$

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 $\sigma^2_{iy} \sim Uniform(0.5, 1.5)$, , while with no loss of generality $\mu_{0i} = \beta_i = 1\ \forall i$.

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}):

$$x_{it} = \gamma_1 F_{1t} + \gamma_2 F_{2t} + \epsilon^x_{it} \tag{5}$$

Following Pesaran (2006) 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:

$$\begin{bmatrix} F_{1t} \\ F_{2t} \end{bmatrix} = \begin{bmatrix} F_{1t-1} \\ 0.4F_{2t-1} \end{bmatrix} + \begin{bmatrix} f_{1t} \\ f_{2t} \end{bmatrix} \tag{6}$$

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

$$\begin{bmatrix} f_{1t} \\ f_{2t} \end{bmatrix} = \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix} + \begin{bmatrix} \vartheta_1 & 0 \\ 0 & \vartheta_2 \end{bmatrix} \begin{bmatrix} \eta_{1t-1} \\ \eta_{2t-1} \end{bmatrix} \tag{7}$$

$$\epsilon^x_{it} = e^x_{it} + \varphi e^x_{it-1}, \tag{8}$$

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].

Testing for cointegration with regime shifts is an important issue which, as we will see, arises in the empirical illustration on the Fisher effect also. To shed some light on the potential of our procedure for this purpose, along the base case we will thus also consider that of a relationship with a break in both constant and slope:

$$y_{it} = \begin{cases} \mu_{0i} + \beta_0 x_{it} + \epsilon^y_{it}, & t \leq t_b \\ \mu_{1i} + \beta_1 x_{it} + \epsilon^y_{it}, & t > t_b \end{cases} \tag{9}$$

[6]Exploratory simulations showed the performances of the test to be independent on the number of independent variables.

where $\beta_0 = 1$ and $\beta_1 = 1.5$ Since the delicate problem of the estimation of the break point is outside the scope of this paper in the tests we will assume it known, and fix it with no loss of generality at $T/2$ in all units.

Some remarks are in order. First of all, this DGP violates the assumptions of Gegenbach et al. (2006), who exclude the possibility of a single source of non stationarity common to both the left- and the right-hand side variables, and Westerlund (2008), who assumes common factors to be present in the residuals of the cointegrating equation only. However, it is likely to be representative of many empirical applications: an obvious example is the case of regional consumption and income, with the common factors given by the trend and cycle in national GDP. To shed some light on the performances that can be expected from common factors methods in this type of set-up we shall examine the performance of Westerlund's Durbin-Hausmann group mean DH_g test.

Second, cointegration across units in the X's always hold, while in the Y's it does only in case of cointegration within units between X_i and Y_i ($|\rho_i| < 1$). For simplicity we are ruling out the possibility of cointegration holding in some units only, but the design could be easily generalised further to include this case also.

The sample sizes considered in the experiment are also chosen trying to reproduce empirically relevant conditions. Hence, we assume the data set to cover up to $N = 40$ cross-section units and $T = 160$ time observations, but with the full time sample available only at aggregate level (average over all units), and the fully disaggregated sample only for $T = 20, 40$. We shall thus first compute the test on the aggregate data for $T = 20, 40, 160$ and then evaluate the gains possibly delivered by adding the panel dimension. An intermediate case is represented by the first five units on which $T = 80$ time observations are assumed to be available.

In principle an important, and still largely unsettled, aspect of block bootstrap methods is the choice of block size (mean size in the case of the stationary bootstrap, when the length is random). In practice according to the simulation results reported by PPP the RSB unit root tests appear to be quite robust to this parameter. We will thus fix it at either $0.10T$ (as in Paparoditis and Politis, 2001) or $0.15T$ with a minimum of 4, leaving implementation of data-based methods for future research. In the WB we use $\zeta_t \sim NID(0, 1)$ and fix the length of the autoregressive polynomial at the true value (one); hence, these results are of interest mostly as a benchmark for those of the RSB tests.

Finally, to strike a balance between experimental precision and computing costs the number of both Monte Carlo simulations and bootstrap redrawings has been set to 1000.

3.2 Results

The results are reported in tables 1-8 below. First of all, from the aggregate tests we can see that the cointegration RSB tests with the two different block sizes deliver essentially the same results. This is consistent with the performance of the unit root RSB tests in PPP and good news from the practitioner's point of view. Hence, to save space, for the panel tests we will report results only for block size $0.10T$, with those for block size $0.15T$ available on request. Second, the results for the RSB tests are rather close to those of the WB tests. Since the latter are based on the true order of the autoregressive polynomial of the errors this is a rather remarkable performance. With a few exceptions the following comments are thus generally valid for both type of tests. The results of the tests on the aggregate data (Tables 1-2), show that all bootstrap tests deliver performances essentially comparable to those of traditional tests which compare the ADF statistics with MacKinnon (1991) critical values. Unfortunately, while Type I errors are always close to nominal values, power is rather disappointing in empirically relevant sample sizes: for instance, with $T = 40$ only slightly higher than 50% for a 5% test.

Can adding the panel dimension help? The results in Tables 3-4 suggest it can indeed. First of all, Type I errors are generally close to nominal, except for some overrejection of the RSB test with $T = 20$ (recall that with 1000 Monte Carlo simulations approximate 95% confidence intervals around 5% and 10% are respectively 4%-6% and 8%-12%). Second, the power performance is very good: more than the high values of the rejection rates (which are conditional on the specific DGP and signal/noise ratio at hand), the important evidence here is their rapid growth with the cross-section dimension.

The good behaviour of the panel tests is confirmed by the results with $T = 80$ and the first five units (Table 5): Type I errors are essentially equal to nominal size and power reaches 100%.

Allowing for a break fixed at the DGP break date (Table 6-7) the tests show a tendency to underreject. Overall both the size bias and the power loss are however limited, suggesting that extending the procedure to the empirically more relevant case of unknown break points along the lines of Gregory and Hansen (1996) may be a worthwhile research topic.

Finally, in Table 8 we report the Type I errors of the Durbin-Hausman group mean test DH_g by Westerlund (2008). We stress again that the application of the test is obviously wrong here; a careful common factor analysis of the data would conclude that the residuals have no common factor, while the right-hand side variable does. Though largely expected, the results are nevertheless instructive of the possible consequences of a automatic application of the method: since the common factor procedure fails to remove the dependence across units, the test heavily overrejects. In fact, when the X is generated according to the full specification (5)-(8) with two common factors the true null of no cointegration is *always* rejected by DH_g test. Letting $\gamma_2 = 0$ so that there is only one, non stationary common factor, the size bias falls but it is still very large, and, though shrinking with the time dimension, it worsens with the cross-section one for a fixed time sample. The problem is that, since the bias is exactly in the direction most welcome by practitioners (against H_0: no cointegration, hence in favour of the existence of a cointegrating relationship), they will probably be too happy of the results delivered by a routine application of the test to check carefully the validity of its assumptions.

Table 1
Asymptotic and Bootstrap
Aggregate Cointegration Tests
Size

				RSB	
				Block size	
T	α	MK	WB	$0.10T$	$0.15T$[1]
20	0.01	0.00	0.01	–	0.01
	0.05	0.02	0.06	–	0.06
	0.10	0.03	0.13	–	0.14
40	0.01	0.00	0.01	0.01	0.01
	0.05	0.05	0.08	0.04	0.04
	0.10	0.06	0.13	0.12	0.12
160	0.01	0.03	0.02	0.02	0.02
	0.05	0.05	0.06	0.06	0.07
	0.10	0.09	0.11	0.12	0.12

DGP: $\overline{X} = N^{-1} \sum_{i=1}^{N} X_i, \overline{Y} = N^{-1} \sum_{i=1}^{N} Y_i$
X_i : cf. (5)-(8)
Y_i : cf. (4), $\rho_i = 1 \forall i$
H_0 : No cointegration
$test$: ADF on cointegrating residuals
MK: MacKinnon (1991) critical values
WB: Wild Bootstrap
RSB: Residual-Based Stationary Bootstrap
[1] :block size set to 4 when $T = 20$.

Table 2
Asymptotic and Bootstrap
Aggregate Cointegration Tests
Power

				RSB	
				Block size	
T	α	MK	WB	$0.10T$	$0.15T^1$
20	0.01	0.02	0.09	–	0.26
	0.05	0.16	0.30	–	0.45
	0.10	0.28	0.46	–	0.66
40	0.01	0.08	0.21	0.27	0.19
	0.05	0.42	0.52	0.61	0.49
	0.10	0.62	0.69	0.77	0.66
160	0.01	0.93	0.99	1.00	0.99
	0.05	1.00	1.00	1.00	1.00
	0.10	1.00	1.00	1.00	1.00

DGP: see Table 1; $\rho_i \sim Uniform(0.6, 0.8)$
H_0 :No cointegration; H_1 : cointegration;
test: ADF on cointegrating residuals.
^{1} :block size set to 4 when $T = 20$.

Table 3
Bootstrap Panel Cointegration Tests
Size

			Units			
			5	10	20	40
T		α	$Median(ADF)$			
20	RSB	0.05	0.08	0.08	0.10	0.10
		0.10	0.14	0.16	0.20	0.20
	WB	0.05	0.05	0.05	0.05	0.03
		0.10	0.11	0.12	0.12	0.09
40	RSB	0.05	0.06	0.07	0.07	0.06
		0.10	0.10	0.14	0.14	0.12
	WB	0.05	0.04	0.05	0.04	0.03
		0.10	0.08	0.10	0.09	0.09
T		α	$Mean(ADF)$			
20	RSB	0.05	0.08	0.09	0.12	0.10
		0.10	0.14	0.17	0.22	0.20
	WB	0.05	0.04	0.03	0.04	0.01
		0.10	0.10	0.09	0.08	0.05
40	RSB	0.05	0.05	0.07	0.06	0.05
		0.10	0.12	0.13	0.13	0.12
	WB	0.05	0.05	0.05	0.04	0.02
		0.10	0.09	0.11	0.08	0.06

DGP:
X_i : cf. (5)-(8)
Y_i : cf. (4), $\rho_i = 1 \forall i$
H_0 :No cointegration
Median(ADF): H_1 : cointegration in most units
Mean(ADF): H_1 : cointegration in a large number of units *or* strong cointegration in a smaller number of units.

Table 4
Bootstrap Panel Cointegration Tests
Power

			Units			
			5	10	20	40
T		α	$Median(ADF)$			
20	RSB	0.05	0.74	0.94	1.00	1.00
		0.10	0.85	0.97	1.00	1.00
	WB	0.05	0.97	1.00	1.00	1.00
		0.10	1.00	1.00	1.00	1.00
40	RSB	0.05	0.99	1.00	1.00	1.00
		0.10	1.00	1.00	1.00	1.00
	WB	0.05	0.99	1.00	1.00	1.00
		0.10	1.00	1.00	1.00	1.00
T		α	$Mean(ADF)$			
20	RSB	0.05	0.57	0.86	0.99	1.00
		0.10	0.83	0.97	1.00	1.00
	WB	0.05	0.99	1.00	1.00	1.00
		0.10	1.00	1.00	1.00	1.00
40	RSB	0.05	0.98	1.00	1.00	1.00
		0.10	1.00	1.00	1.00	1.00
	WB	0.05	0.98	1.00	1.00	1.00
		0.10	1.00	1.00	1.00	1.00

DGP:
X_i : cf. (5)-(8)
Y_i : cf. (4), $\rho_i \sim Uniform(0.6, 0.8)$
H_0, H_1 : see Table 3.

Table 5
Bootstrap Panel Cointegration Tests
$T = 80$, $N = 5$

		α	0.01	0.05	0.10
RSB	$Median(ADF)$	Size	0.01	0.07	0.14
		Power	1.00	1.00	1.00
	$Mean(ADF)$	Size	0.01	0.07	0.14
		Power	1.00	1.00	1.00
WB	$Median(ADF)$	Size	0.02	0.06	0.11
		Power	0.97	1.00	1.00
	$Mean(ADF)$	Size	0.01	0.06	0.10
		Power	0.99	1.00	1.00

DGP: see Table 3;
Size: $\rho_i = 1\ \forall i$; Power: $\rho_i \sim Uniform(0.6, 0.8)$.
H_0, H_1 : see Table 3.

Table 6
Bootstrap Panel Cointegration Tests
with a Known Break
$T = 40$, break at $t = 20$
Size

		Units			
		5	10	20	40
	α	$Median(ADF)$			
RSB	0.05	0.04	0.03	0.03	0.02
	0.10	0.09	0.08	0.07	0.04
WB	0.05	0.03	0.03	0.03	0.02
	0.10	0.08	0.08	0.07	0.04
	α	$Mean(ADF)$			
RSB	0.05	0.04	0.03	0.03	0.01
	0.10	0.08	0.07	0.07	0.03
WB	0.05	0.03	0.03	0.03	0.01
	0.10	0.07	0.08	0.06	0.03

DGP:
X_i : cf. (5)-(8)
Y_i : cf. (9), $\rho_i = 1 \forall i$
H_0, H_1 : see Table 3.

Table 7
Bootstrap Panel Cointegration Tests
with a Known Break
$T = 40$, break at $t = 20$
Power

		Units			
		5	10	20	40
	α	$Median(ADF)$			
RSB	0.05	0.57	0.84	0.95	0.99
	0.10	0.72	0.93	0.98	1.00
WB	0.05	0.03	0.03	0.03	0.02
	0.10	0.08	0.08	0.07	0.04
	α	$Mean(ADF)$			
RSB	0.05	0.69	0.91	0.99	1.00
	0.10	0.2	0.96	1.00	1.00
WB	0.05	0.03	0.03	0.02	0.01
	0.10	0.07	0.08	0.06	0.03

DGP:
X_i : cf. (5)-(8)
Y_i : cf. (9),
$\rho_i \sim Uniform(0.6, 0.8)$
H_0, H_1 : see Table 3.

Table 8
Durbin-Hausman Common Factors
Group Mean DH_g
Panel Cointegration Test
Size

		Units			
T	α	5	10	20	40
20	0.05	5.6	42.4	69.8	88.5
	0.10	6.5	46.8	74.6	91.0
40	0.05	2.6	24.8	35.0	48.8
	0.10	4.1	29.8	42.8	58.1
80	0.05	4.8	8.1	8.6	11.0
	0.10	6.1	13.4	13.3	16.5

DGP:
X_i : cf. (5)-(8),$\gamma_2 = 0$.
Y_i : cf. (4), $\rho_i = 1 \forall i$
H_0 : No cointegration.

4 Empirical illustration: the Fisher effect

The so-called "Fisher effect" dates back to Fisher (1930), who put forth the hypothesis that the nominal interest rate (i) adjusts to the sum of expected real interest rate (r^*) and expected inflation rate (p^*):

$$i_t = r_t^* + p_t^* \tag{10}$$

Of course, (10), which involves unobserved variables, cannot be directly tested; however, it suggests an observable direct relationship with unit coefficient between the nominal interest rate and the actual inflation rate ("full Fisher effect"). In practice, this reasonable hypothesis never found consistent support from the data (recent evidence in this direction is provided, *inter alia*, by Rose, 1988, MacDonald and Murphy, 1989, Bonham, 1991, King and Watson, 1997), although more general specifications with coefficients different from one ("partial Fisher effect") or breaks were shown to be compatible with the data (*e.g.*, Garcia and Perron, 1996). However, as Westerlund (2008) points out, the available empirical studies are weak under two important aspects. First, most studies examined US data only. Second, in the case of long-run studies the economic hypothesis is rejected when the statistical null hypothesis of no cointegration is not. Hence, low power of the statistical procedure used may lead to erroneously reject the economic hypothesis of interest, exactly as it happens with the Purchasing Power Parity theory[7].

Ghazali and Ramlee (2003) and Koustas and Serletis (1999) did examine panels of countries, but applied cointegration tests to each of them separately. Hence, they meet the first objection but not the second. To tackle both, Westerlund (2008) applied his Durbin-Hausmann panel cointegration tests to a panel of 20 OECD countries[8] for the period 1980:1-2004:4. With a p-value equal to 0.000, the group mean DH_g test provides extremely strong evidence in favour of panel cointegration between interest rates and inflation (which appear to be non-stationary on the basis of both univariate and panel unit root tests). Since the estimated coefficients are different from one the conclusion is that the hypothesis of a partial Fisher effect holding in the examined panel as a whole cannot be rejected. Although this is a certainly reasonable conclusion, in view of the uncertainty prevailing in the literature its strength is somehow suspect. The simulation reported in Table 8 suggests that in the case of common factors in the right-hand side variable, rather than in the residuals as assumed by the test, the DH_g panel cointegration test can be severely oversized. In fact, applying our bootstrap procedure[9]

[7] In fact, this empirical issue was an important motivation of the early developments of panel cointegration methods: see e.g. O'Connell (1998).

[8] Australia, Belgium, Canada, Switzerland, Germany, Denmark, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Luxembourg, Netherlands, Norway, New Zealand, Portugal, Sweden, United States.

[9] Mean block size $T/10$, 1000 redrawings. The results are robust to mean block size.

we estimate the p-value of the mean ADF cointegration statistic as 0.03, and that of the median ADF as 0.13. Applying the conventional 5% significance level the hypothesis of no panel cointegration is rejected in mean (hence, in favour of the alternative hypothesis of cointegration in most of the units *or* strong cointegration in a smaller number of units) but not in median (when the alternative hypothesis is cointegration in most of the units). At the 1% level mean and median tests agree to suggest no rejection of the null hypothesis of no Fisher effect in the panel as a whole. Our conclusion is therefore not entirely at odds with Westerlund's, but considerably more cautious and thus in line with the previous literature: there is some evidence in favour of a partial Fisher effect, but (*i*) it is weaker than suggested by the Durbin-Hausman DH_g test, and, (*ii*) it seems to come from some subset of the examined panel of OECD economies. Clearly, as suggested by Garcia and Perron (1996), allowing for breaks may strengthen the evidence in favour of a Fisher effect.

5 Conclusions

The key contribution of this paper is to put forth a test for panel cointegration in dependent panels based upon a residual based unit root test recently proposed by Parker, Paparoditis and Politis (2006). The test procedure is shown by simulation to deliver good size and power performances in panels with long- and short-run dependence due to common factors in the variables examined. The power gains with respect to aggregate tests appear particularly valuable. Applying the procedure to test the Fisher hypothesis on the Westerlund (2008) data we find some weak evidence in favour of a partial Fisher effect; our conclusions are therefore more cautious than Westerlund's. Future research will try to address the issue of data-based choice of block size, the asymptotic properties of the test, as well as generalising the procedure to allow for breaks at an unknown date.

6 References

Bai J., Ng S. (2004) "A PANIC Attack on Unit Roots and Cointegration" *Econometrica* 72, 1127-1177.

Banerjee, A. (1999) "Unit Roots and Cointegration: an Overview" *Oxford Bulletin of Economics and Statistics*, 61, 697-629.

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.

Bonham, C.S. (1991) "Correct Cointegration Tests of the Long-Run Relationship between Nominal Interest and Inflation" *Applied Economics*,23, 1487-1592.

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

Chang Y., Park J.Y., Song K. (2006) "Bootstrapping cointegrating regressions" *Journal of Econometrics* 133, 703-739.

Engle, R. F. and Granger, C. W. J. (1987) "Co-integration and Error Correction: Representation, Estimation, and Testing" *Econometrica*, 55, 251-276.

Evans, P., Wang X. (2008) "A Tale of Two Effects" *Review of Economics & Statistics*, 90 147-157.

Fachin, S. (2007) "Long-Run Trends in Internal Migrations in Italy: a Study in Panel Cointegration with Dependent Units" *Journal of Applied Econometrics* 22, 401-428.

Fisher, I. (1930) *The Theory of Interest* Macmillan, New York.

Garcia, R., Perron P. (1996) "An Analysis of the Real Rate of Interest Under Regime Shifts." *Review of Economics and Statistics*, 78, 111–125.

Gengenbach, C., Palm F., Urbain J.P. (2006) "Panel Cointegration Testing in the Presence of Common Factors" *Oxford Bulletin of Economics and Statistics* 68:S1, 683-719.

Ghazali, N.A., Ramlee S. (2003) "A long memory test of the long-run Fisher effect in the G7 countries" *Applied Financial Economics,* 13, 763–769.

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

Groen, J.J.J., F. Kleibergen (2003) "Likelihood-Based Cointegration Analysis in Panels of Vector Error-Correction Models" *Journal of Business and Economic Statistics*, 21, 295-318.

Herwartza, H., Neumann M.H. (2005)"Bootstrap inference in systems of single equation error correction models" *Journal of Econometrics*, 128, 165-193.

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

King, R.G., and Watson M.W. (1997) "Testing Long-Run Neutrality" *Federal Reserve Bank of Richmond Economic Quarterly*, 83, 69-101

Koustas, Z., Serletis A. (1999) "On the Fisher effect" *Journal of Monetary Economics,* 44, 105–130.

MacDonald, R. and Murphy P.D. (1989) "Testing for the Long Run Relationship between Nominal Interest Rates and Inflation Using Cointegration Techniques" *Applied Economics*, 21, 439-447.

MacKinnon, J. (1991), "Critical Values for Cointegration Tests" in *Long-Run Economic Relationships*, R.F. Engle and C.W.J. Granger (Eds.), Oxford University Press, Oxford (UK).

Mishkin, F.S. (1992) "Is the Fisher Effect for Real? A reexamination of the relationship between inflation and interest rates " *Journal of Monetary Economics* 30, 195–215.

O'Connell, P.G.J (1998) "The Overvaluation of Purchasing Power Parity" J*ournal of International Economics*, 44, 1-19.

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).

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

Pedroni, P. 1999. "Critical Values For Cointegration Tests In Heterogeneous Panels With Multiple Regressors". *Oxford Bulletin of Economics and Statistics* 61: 653-670.

Pedroni, P. (2004) "Panel Cointegration, Asymptotic and Finite Sample Properties of Pooled Time Series tests with an Application to the PPP hypothesis" *Econometric Theory* 20, 597-625.

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

Rose, A.K. (1988) "Is the Real Interest Rate Stable" *Journal of Finance*, 43, 1095-1112.

Westerlund, J. (2008) "Panel Cointegration Tests of the Fisher Effect" *Journal of Applied Econometrics* 23, 193-233.

Westerlund, J., Edgerton D. (2007) "A Panel Bootstrap Cointegration" Test *Economics Letters* vol. 97 pp.185-190.