Breaking panel data cointegration

Anindya Banerjee
Department of Economics
European University Institute

Josep Lluís Carrion-i-Silvestre*
Department of Econometrics, Statistics and Spanish Economy
University of Barcelona

February 1, 2005

Abstract

Misspecification errors due to the presence of unattended structural breaks can affect the power of standard panel cointegration statistics. We propose modifications to allow for one structural break when testing the null hypothesis of no cointegration that retain good properties in terms of empirical size and power. Response surfaces to approximate the finite sample moments that are required to implement the statistics are provided. Since panel cointegration statistics rely on the assumption of cross-section independence, a generalisation to the common factor framework is carried out. Moreover, for those situations where the common factor model is not suitable we suggest the applicatication of a sieve bootstrap method to compute the empirical distribution of the statistics.

Keywords: Panel cointegration, structural break, common factors, sieve bootstrap, cross-section dependence
JEL Codes: C12, C22

1 Introduction

The theory of cointegration establishes that there exist linear combinations of non-stationary variables that cancel out common stochastic trends. This phenomenon gives rise to equilibrium relationships amongst non-stationary variables, which means that in the long-run these variables follow each other. The concept of cointegration does not prevent that neither the vector of

*Corresponding author: Av. Diagonal, 690. 08034 Barcelona. E-mail: carrion@ub.edu. Part of the reseach was written while the author was visiting the European University Institute in Florence in the spring 2004. I would like to thank the Department of Economics for its hospitality during my stay. The author also acknowledges financial support from the Generalitat de Catalunya.

cointegration nor the deterministic component of the long-run relationship might change along the analysed time period. In fact, Hansen (1992), and Quintos and Phillips (1993) propose test statistics to assess the stability of the cointegration relationship. More interestingly, it is well known that the inference about the presence of cointegration can be affected by misspecification errors that do not account for changes in the parameters of the model, which can bias conclusions towards the non-stationarity –see Campos, Ericsson and Hendry (1996), and Gregory and Hansen (1996). All these considerations have driven to design procedures to test for cointegration allowing for structural breaks. Thus, Gregory and Hansen (1996) generalised the standard cointegration approach in Engle and Granger (1987) to allow for the presence of structural breaks that might affect either the deterministic component or the cointegration vector of the long-run relationship. Hao (1996), Bartley, Lee and Strazicich (2001), and Carrion-i-Silvestre and Sansó (2004) use the multivariate version of the KPSS statistic in Harris and Inder (1994), and Shin (1994) to test for the null of cointegration with one structural break. Finally, Hansen and Johansen (1993), and Busetti (2002) propose methods to estimate the cointegration rank in a multivariate framework. These proposals obey to requirements that arise in empirical modelling since there is some empirical applications in the literature that test for cointegration allowing for structural breaks. For instance, Gregory and Hansen (1996) and Gabriel, Da Silva and Nunes (2002) investigate the long-run money demand for the U.S. and Portugal, respectively. Busetti (2002) conducts two illustrations using road casualties in Great Britain, and some macroeconomic data for the UK. Finally, Clemente, Marcuello, Montañés and Pueyo (2004) focus on health care expenditure demand functions. The main conclusion that arises from these applications is that inference on cointegration analysis can be affected by the presence of structural breaks.

Non-stationary panel data econometrics literature has experienced a rapid development since 1990s. The main reason that has popularised the use of the panel data techniques is the idea that power of unit root and cointegration testing might increase due to the combination of the information that comes from both the cross-section and the time dimensions. As a result, new statistics to assess the stochastic properties of panel data sets have appeared in the literature –see Banerjee (1999), Baltagi and Kao (2000), and Baltagi (2001) for an overview of the field. Surprisingly, instability has not received too much attention in panel data cointegration framework. In this regard, Kao and Chiang (2000) analyse instability in cointegration relationships assuming that cointegration is present, with an homogeneous cointegrating vector for all individuals –although it is possible to split the panel in two sub-panels using bootstrap– and a common break point. Besides, Breitung (2002) proposes a VAR-based panel data cointegration procedure that allows introducing dummy variables outside the long-run relationship. Finally, Westerlund (2004) extends the LM statistic in McCoskey and Kao (1998) allowing for one structural break.

As can be seen, there are not many contributions in the literature that addresses the panel data cointegration hypothesis testing allowing for structural breaks. In this paper we address this concern and generalise the approach in Pedroni (1999, 2004) to account for one structural break that affects the long-run relationship in different ways. Pedroni proposes seven statistics depending on the way that the individual information is combined to define the panel tests. Moreover, the statistics can also be grouped in either parametric or non-parametric statistics, depending on the way that autocorrelation and endogeneity bias is accounted for. In this paper we only focus on the parametric statistics. One important feature of all these proposals is cross-section dependence matter. Thus, all these panel data statistics assume cross-section independence. In this paper we address this concern in two different ways. First, we generalise the proposal in Pedroni (2004) dealing with an approximate common factor model as in Bai and Ng (2004). The limiting distribution of the statistics is derived and new sets of critical values are computed when required. Second, we propose to carry out a sieve bootstrap to obtain the empirical distribution of the statistics for those cases in which the factor model should not be appropriate.

The paper proceeds as follows. In section 2 the interest of our proposal is motivated through Monte Carlo simulations. Section 3 presents the models and statistics for the null hypothesis of no cointegration with power against the alternative of broken cointegration. The moments that are required for the computation of the panel data statistics are computed in this section. In this regard, we estimate response surfaces to approximate these moments for whichever sample size. Section 4 extends the approach to the common factor framework. Section 5 focuses on the finite sample properties of the statistics. In section 6 we illustrate the proposal analysing the Feldstein-Horioka puzzle. Finally, section 7 concludes with some remarks. All proofs are collected in the Appendix.

2 Motivation

Pedroni (1999, 2004) proposes seven statistics to test the null hypothesis of no cointegration using single equation methods based on the estimation of static regressions. Since the statistics are based on single equation methods the cointegrating rank for each unit is either 0 or 1, with a heterogeneous cointegrating vector for each individual. After conducting the estimation of the individual static regressions, the cointegrating residuals are used to compute one of the statistics. The seven statistics can be classified in two different groups depending on whether they are within-dimension-based statistics –homogeneity is assumed when computing the cointegration test statistic– and between-dimension-based statistics –heterogeneous behaviour is allowed for the statistic. In order to correct for the endogeneity bias, Pedroni (1999, 2004) suggests applying the FM-OLS estimation method for the non-parametric statistics, although DOLS estimation method can be applied as well –see Kao and Chiang (2000), and Mark and Sul (2003). Notwithstanding, the statistics that use the parametric way to correct for the presence of autocorrelation does not correct for the endogeneity bias.

As mentioned in the introduction, we are only concerned with the parametric version of the statistics, *i.e.* the normalised bias and the pseudo t-ratio statistics. To motivate our proposal we analyse the effects of structural breaks on the parametric group Pedroni statistics through Monte Carlo simulations. First, we focus on the case where there is cointegration but the deterministic component changes at a point in time. In a second stage we consider the case of unstable cointegrating vector. The DGP is given by:

$$y_{i,t} = f_i(t) + \alpha_{i,t} x_{i,t} + z_{i,t}$$
$$\Delta x_{i,t} = \varepsilon_{i,t}$$
$$z_{i,t} = \rho_i z_{i,t-1} + v_{i,t}$$
$$\zeta_{i,t} = (\varepsilon_{i,t}, v_{i,t})' \sim iid\ N(0, I_2),$$

where $f_i(t)$ denotes the deterministic component.

For the first case we have $f_i(t) = \mu_i + \theta_i DU_{i,t}$ with $DU_{i,t} = 1$ for $t > T_{bi}$ and 0 otherwise, where $T_{bi} = \lambda_i T$, $\lambda_i \in (0,1)$, denotes the date of the break. The parameter set is given by $\mu_i = 1$, $\theta_i = \{0, 1, 3, 5, 10\}$, $\alpha_{i,t} = \alpha_i = 1$, and $\lambda_i = \{0.25, 0.5, 0.75\}$. The autoregressive parameter is set equal to $\rho_i = \{0, 0.95\}$. The sample size is $T = \{100, 200\}$, the number of individuals is $N = \{20, 40\}$ and 1,000 replications are carried out. For ease of simplicity but without loss of generality, in all simulations we have specified a common break point for all individuals. The model that has been estimated to compute the pseudo t-ratio Pedroni panel data cointegration test statistics includes a constant term (individual effects) as deterministic component. Results reported in Table 1 show that the effect of level shift only matters in those situations where the magnitude of the shift is large and the break point is located at the end of the time period. Therefore, we can conclude that for small and moderate level shifts the misspecification error of the deterministic component does not damage the power of Pedroni statistic.

In the second stage we have analysed the case where the structural break changes both the level and the slope of the time trend. The deterministic function is given by $f_i(t) = \mu_i + \theta_i DU_{i,t} + \xi_i t + \gamma_i DT^*_{i,t}$, where $\mu_i = 1$, $\theta_i = 3$, $\xi_i = 0.3$ and $DT^*_{i,t}$ is the dummy variable defined above. Note that in this case the pseudo t-ratio statistic has been computed using a time trend as the deterministic component. Table 1 shows that in this situation consequences of the misspecification error are more serious, since the empirical power approaches zero as the magnitude of the slope shift (γ_i) increases when the break point is placed either in the middle $(\lambda_i = 0.5)$ or at the end $(\lambda_i = 0.75)$ of the period.

The third situation analyses the effects of both the change in the level and in the cointegrating vector. As before, the deterministic component is $f_i(t) = \mu_i + \theta_i DU_{i,t}$, with $\mu_i = 1$ and $\theta_i = \{0, 3\}$. Now we focus on the change in the cointegrating vector specifying $\alpha_{i,t} = \alpha_{i,1} = 1$ for $t \leq T_{bi}$ and $\alpha_{i,t} = \alpha_{i,2} = \{0,\ 2,\ 3,\ 4,\ 5,\ 10\}$ for $t > T_{bi}$. The model that has been estimated to compute the (pseudo t-ratio) Pedroni panel data cointegration statistic includes a constant term as deterministic component. Table 2 indicates that for the empirical power to diminish the change in the cointegrating vector has to be either moderate or large, and be located in the middle $(\lambda_i = 0.5)$ or at the end $(\lambda_i = 0.75)$ of the period. Notice that this conclusion is reached irrespective of the level shift that affects the constant term.

Finally, the fourth case of study considers the change in the time trend that defines the deterministic component and the change in the cointegrating vector. In this case $f_i(t) = \mu_i + \theta_i DU_{i,t} + \xi_i t + \gamma_i DT^*_{i,t}$, with $\mu_i = 1$, $\theta_i = 3$, $\xi_i = 0.3$, $\gamma_i = 0.5$, and $\alpha_{i,t} = \alpha_{i,1} = 1$ for $t \leq T_{bi}$ and $\alpha_{i,t} = \alpha_{i,2} = \{0,\ 2, 3, 4, 5,\ 10\}$ for $t > T_{bi}$. The model that has been estimated to compute the pseudo t-ratio Pedroni panel data cointegration statistic includes individual and time effects. Table 3 reports that the change in the slope implies further reductions on the empirical power of the statistic when the break point is located in the middle and at the end of the period.

In all, we can conclude that misspecification errors due to the lack of accounting for a structural break can reduce the power of the panel data cointegration test in Pedroni (2004) in those cases where the break point is placed in the middle or at the end of the time period. Therefore, we have observed a bias towards the spurious non-rejection of the null hypothesis of no cointegration. A relevant feature is that the power distortions appear when break changes either the slope of the time trend or the cointegrating vector, but no effects are to be expected when the break only affects the constant term.

3 Models and test statistics

Let $\{Y_{i,t}\}$ be a $(m \times 1)$-vector of non-stationary stochastic process whose elements are individually I(1). Moreover, let us assume that the DGP that describes $Y_{i,t}$ is given by the following triangular representation

$$\begin{aligned} \Delta x_{i,t} &= \varepsilon_{i,t} \\ y_{i,t} &= f_i(t) + x'_{i,t}\delta_{i,t} + e_{i,t} \end{aligned}$$

where $Y_{i,t} = \left(y'_{i,t}, x'_{i,t}\right)'$ is conveniently partitioned into two vectors of dimension $y_{i,t}$ $((m-r) \times 1)$ and $x_{i,t}$ $(r \times 1)$ respectively, $i = 1, \ldots, N$, $t = 1, \ldots, T$. The disturbance terms $\xi_{i,t} = \left(\varepsilon'_{i,t}, e'_{i,t}\right)'$ are assumed to satisfy the strong-mixing conditions in Phillips (1987) and Phillips and Perron (1988). The $(m \times r)$ matrix of r cointegrating vectors is $\delta_{i,t} = (-\alpha_{i,t}, I_r)'$ where $\alpha_{i,t}$ is the $((m-r) \times 1)$ submatrix of parameters to be estimated and I_r is the identity matrix. At this stage and in order to set the analysis in a simplified framework, let us assume that $\{\varepsilon_{i,t}\}$ and $\{e_{i,t}\}$ are independent –if we weaken this assumption, then DOLS estimation method should be applied in order to account for the endogeneity bias.

The general functional form for the deterministic term $f(t)$ is given by

$$f_i(t) = \mu_i + \beta_i t + \theta_i DU_{i,t} + \gamma_i DT^*_{i,t},$$

where

$$DU_{i,t} = \begin{cases} 0 & t \leq T_{bi} \\ 1 & t > T_{bi} \end{cases} ; DT^*_{i,t} = \begin{cases} 0 & t \leq T_{bi} \\ (t - T_{bi}) & t > T_{bi} \end{cases},$$

with $T_{bi} = \lambda_i T$, $\lambda_i \in (0, 1)$, denoting the time of the break for the i-th individual, $i = 1, \ldots, N$.

Note also that the cointegrating vector is specified as a function of time so that

$$\alpha_{i,t} = \begin{cases} \alpha_{i,1} & t \leq T_{bi} \\ \alpha_{i,2} & t > T_{bi} \end{cases}.$$

Using these elements, we propose up to six different model specifications:

- Model 1. Constant term with a level shift but stable cointegrating vector:

$$y_{i,t} = \mu_i + \theta_i DU_{i,t} + x'_{i,t}\delta_i + e_{i,t} \tag{1}$$

- Model 2. Time trend with a level shift but stable cointegrating vector:

$$y_{i,t} = \mu_i + \beta_i t + \theta_i DU_{i,t} + x'_{i,t}\delta_i + e_{i,t} \tag{2}$$

- Model 3. Time trend with both level and slope shifts but stable cointegrating vector:

$$y_{i,t} = \mu_i + \beta_i t + \theta_i DU_{i,t} + \gamma_i DT^*_{i,t} + x'_{i,t}\delta_i + e_{i,t} \tag{3}$$

- Model 4. Constant term with both level and cointegrating vector shift:

$$y_{i,t} = \mu_i + \theta_i DU_{i,t} + x'_{i,t}\delta_{i,t} + e_{i,t} \tag{4}$$

- Model 5. Time trend with both level and cointegrating vector shift (the slope does not shifts):

$$y_{i,t} = \mu_i + \beta_i t + \theta_i DU_{i,t} + x'_{i,t}\delta_{i,t} + e_{i,t} \tag{5}$$

- Model 6. The time trend and the cointegrating vector shifts:

$$y_{i,t} = \mu_i + \beta_i t + \theta_i DU_{i,t} + \gamma_i DT^*_{i,t} + x'_{i,t}\delta_{i,t} + e_{i,t} \tag{6}$$

Using one of these specifications we propose to test the null hypothesis of no cointegration against the alternative hypothesis of cointegration using the ADF test statistic applied to the residuals of the cointegration regression as in Engle and Granger (1987) and Gregory and Hansen (1996) but in the panel data framework developed in Pedroni (1999, 2004).

Our proposal can be described in the following steps. First and following Gregory and Hansen (1996), we proceed to the OLS estimation of one of the models given in (1) to (6) and, then, we run the following ADF type-regression equation on the estimated residuals ($\hat{e}_{i,t}(\lambda_i)$):

$$\Delta\hat{e}_{i,t}(\lambda_i) = \rho_i \hat{e}_{i,t-1}(\lambda_i) + \sum_{j=1}^{k} \phi_{i,j} \Delta\hat{e}_{i,t-j}(\lambda_i) + \varepsilon_{i,t}. \tag{7}$$

Note that the notation that is used refers to the break fraction (λ_i) parameter, which in most of cases is unknown. In order to get rid of the break fraction parameter, Gregory and Hansen (1996) suggest to estimate the models given in (1) to (6) for all possible break dates, obtain the estimated OLS residuals and compute the corresponding ADF statistic. With the sequence of ADF statistics at hands, we can estimate the break point for each individual as the date that minimises the sequence of individual ADF test statistics –either the t-ratio, $t_{\hat{\rho}_i}(\lambda_i)$, or the normalised bias, $T\hat{\rho}_i(\lambda_i)$. Gregory and Hansen (1996) derive the limiting distribution of $t_{\hat{\rho}_i}\left(\hat{\lambda}_i\right) = \inf_{\lambda_i \in (0,1)} t_{\rho_i}(\lambda_i)$ and $T\hat{\rho}_i\left(\hat{\lambda}_i\right) = \inf_{\lambda_i \in (0,1)} T\hat{\rho}_i(\lambda_i)$, which are shown not to depend on the break fraction parameter. Note that the estimation of the break point $\hat{T}_{bi}$ is conducted as

$$\begin{aligned} \hat{T}_{bi} &= \underset{\lambda_i \in (0,1)}{\arg\min} t_{\hat{\rho}_i}(\lambda_i) \\ \hat{T}_{bi} &= \underset{\lambda_i \in (0,1)}{\arg\min} T\hat{\rho}_i(\lambda_i), \end{aligned}$$

$\forall i = 1, \ldots, N$. At this point we could either follow Gregory and Hansen (1996) and test the null hypothesis for each individual or decide to combine the individual information in a panel data statistic.

The panel statistics in which we are going to focus the null hypothesis testing are given by the parametric $Z_{\hat{\rho}_{NT}}$ and $Z_{\hat{t}_{NT}}$ tests in Pedroni (1999, 2004), which can be thought as the panel data version of the rho-statistic and t-statistic tests in Phillips and Ouliaris (1990). These test statistics are defined by pooling the individual ADF tests, so that they belong to the class of between dimension test statistics. Specifically, they are computed as:

$$\begin{aligned} TN^{-1/2} Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right) &= TN^{-1/2} \sum_{i=1}^{N} \frac{\sum_{t=1}^{T} \hat{e}_{i,t-1}\left(\hat{\lambda}_i\right) \Delta\hat{e}_{i,t}\left(\hat{\lambda}_i\right)}{\sum_{t=1}^{T} \hat{e}_{i,t-1}^2\left(\hat{\lambda}_i\right)} = TN^{-1/2} \sum_{i=1}^{N} T\hat{\rho}_i\left(\hat{\lambda}_i\right) \\ N^{-1/2} Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right) &= N^{-1/2} \sum_{i=1}^{N} \frac{\sum_{t=1}^{T} \hat{e}_{i,t-1}\left(\hat{\lambda}_i\right) \Delta\hat{e}_{i,t}\left(\hat{\lambda}_i\right)}{\left(\sum_{t=1}^{T} \hat{s}_i^2\left(\hat{\lambda}_i\right) \hat{e}_{i,t-1}^2\left(\hat{\lambda}_i\right)\right)^{1/2}} = N^{-1/2} \sum_{i=1}^{N} t_{\hat{\rho}_i}\left(\hat{\lambda}_i\right). \end{aligned}$$

Note that in this framework we allow for a high degree of heterogeneity since the cointegrating vector, the short run dynamics and the break point estimate might be differing amongst individuals. The use of the panel data cointegration test aims to increase the power of the statistical inference when testing the null hypothesis of no cointegration, but some heterogeneity is preserved conducting the estimation of the parameters individually.

Following Pedroni (1999, 2004), the panel test statistics are shown to converge to standard Normal distributions once they have been properly standardizes, *i.e.*

$$\begin{aligned} TN^{-1/2} Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right) - \Theta_1 \sqrt{N} &\Rightarrow N(0, \Psi_1) \\ N^{-1/2} Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right) - \Theta_2 \sqrt{N} &\Rightarrow N(0, \Psi_2), \end{aligned}$$

where $\Rightarrow$ denotes weak convergence. The moments of the limiting distributions, $\Theta_1, \Psi_1, \Theta_2$ and Ψ_2, are approximated by Monte Carlo simulation for the different specifications and allowing up to seven stochastic regressors in the cointegrating relationship –*i.e.* the dimension of the $Y_{i,t}$ $(m \times 1)$-vector goes from (2×1) to (8×1). Since the limit distribution of the tests can provide a poor approximation in finite samples, we have approximated the moments of the test statistics for different values of the sample size, specifically $T = \{30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 400, 500, 1,000\}$. In addition, the finite sample distributions depend on the procedure that is applied when selecting the order (k) of the parametric correction in (7), so that the finite sample distributions are obtained in two different ways: (i) assuming the value of k to be fixed, for which we have specified three values $k = 0$, $k = 2$ and $k = 5$, and (ii) selecting the lag length using the t-sig criterion in Ng and Perron (1995) with a $k_{\max} = 5$ as the maximum order of lags. In all simulations $r = 10,000$ replications were done. Table 4 presents the moments of the limit distributions based on $T = 1,000$. As can be seen, the moments of the distribution depends both on the specification and the number of stochastic regressors.

Reporting the moments of the finite sample distribution for the different values of T and different procedures for the selection of k will take a lot of space. Instead, in order to summarise all these results we have estimated response surfaces to model the moments of each test statistic as a function of T and the number of stochastic regressors $p = (m-1)$, *i.e.* $M_j = g(T_j, p_j)$, $j = 1, \ldots, (15 * 7)$. The general functional form that has been essayed is

$$g(T_j, p_j) = \sum_{l=0}^{2} \left(\beta_{0,l} + \beta_{1,l} \frac{1}{T_j} + \beta_{2,l} \frac{1}{T_j^2} + \beta_{3,l} \frac{1}{T_j^3} \right) p_j^l.$$

These functions have been estimated by OLS using the Newey-West robust covariance disturbance matrix to assess the individual significance of the regressors –the level of significance is the 10%. Tables 5 to 8 report the estimated coefficients of the response surfaces. A GAUSS code is available from the authors to compute the statistics and corresponding moments.

4 Common factors in panel cointegration

Previous derivations are valid under the assumption that individuals are cross-section independent. However, this requirement is hardly satisfied in empirical economic applications where countries or regions depend each other. In order to generalise the framework of the paper we have extended our approach to account for the presence of common factors as in Bai and Ng

(2004). In this situation the model is given in structural form as:

$$y_{i,t} = f_i(t) + x'_{i,t}\beta_{i,t} + u_{i,t} \tag{8}$$

$$u_{i,t} = F_t\pi_i + e_{i,t}, \tag{9}$$

$$(I-L)F_t = C(L)w_t \tag{10}$$

$$(1-\rho_i L)e_{i,t} = H_i(L)\varepsilon_{i,t}, \tag{11}$$

$t = 1,\dots,T$, $i = 1,\dots,N$, where $C(L) = \sum_{j=0}^{\infty} C_j L^j$, and $f_i(t)$ denotes the deterministic component, F_t denotes a $(1\times r)$-vector containing the common factors, with π_i the vector of loadings. Despite the operator $(1-L)$ in equation (10), F_t does not have to be I(1). In fact, F_t can be I(0), I(1), or a combination of both, depending on the rank of $C(1)$. If $C(1) = 0$, then F_t is I(0). If $C(1)$ is of full rank, then each component of F_t is I(1). If $C(1) \neq 0$, but not full rank, then some components of F_t are I(1) and some are I(0). Our analysis is based on the same set of assumptions in Bai and Ng (2004). Let $M < \infty$ be a generic positive number, not depending on T and N:

Assumption A: (i) for non-random π_i, $\|\pi_i\| \le M$; for random π_i, $E\|\pi_i\|^4 \le M$, (ii) $\frac{1}{N}\sum_{i=1}^{N}\pi_i\pi'_i \xrightarrow{p} \Sigma_\Pi$, a $(l\times l)$ positive matrix.

Assumption B: (i) $w_t \sim iid(0,\Sigma_w)$, $E\|w_t\|^4 \le M$, and (ii) $Var(\Delta F'_t) = \sum_{j=0}^{\infty} C_j\Sigma_w C'_j > 0$, (iii) $\sum_{j=0}^{\infty} j\|C_j\| < M$; and (iv) $C(1)$ has rank l_1, $0 \le l_1 \le l$.

Assumption C: (i) for each i, $\varepsilon_{i,t} \sim iid(0,\Sigma_\varepsilon)$, $E|\varepsilon_{i,t}|^8 \le M$, $\sum_{j=0}^{\infty} j|H_{i,j}| < M$, $\omega_i^2 = H_i(1)^2\sigma_i^2 > 0$; (ii) $E(\varepsilon_{i,t}\varepsilon_{j,t}) = \tau_{i,j}$ with $\sum_{i=1}^{N}|\tau_{i,j}| \le M$ for all j;
(iii) $E\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[\varepsilon_{i,s}\varepsilon_{i,t} - E(\varepsilon_{i,s}\varepsilon_{i,t})]\right|^4 \le M$, for every (t,s).

Assumption D: The errors $\varepsilon_{i,t}$, w_t, and the loadings π_i are three mutually independent groups.

Assumption E: $E\|F_0\| \le M$, and for every $i = 1,\dots,N$, $E|e_{i,0}| \le M$.

Assumption A ensures that the factor loadings are identifiable. Assumption B establishes the conditions on the short and long-run variance of ΔF_t –*i.e.* positive definite short-run variance and long-run variance that can be of reduced rank in order to accommodate linear combinations of $I(1)$ factors to be stationary. Assumption C(i) allows for some weak serial correlation in $(1-\rho_i L)e_{i,t}$, whereas C(ii) and C(iii) allow for weak cross-section correlation. Finally, Assumption E defines the initial conditions.

The estimation of the common factors are done as in Bai and Ng (2004). If we compute the first difference:

$$\Delta y_{i,t} = \Delta f_i(t) + \Delta x'_{i,t}\beta_{i,t} + \Delta F_t\pi_i + \Delta e_{i,t},$$

and take the orthogonal projections:

$$\begin{aligned} M_i\Delta y_{i,t} &= M_i\Delta F_t\pi_i + M_i\Delta e_{i,t} \\ &= f_t\pi_i + z_{i,t}, \end{aligned} \tag{12}$$

with $M_i = I - \Delta x_i^d \left(\Delta x_i^{d\prime} \Delta x_i^d\right)^{-1} \Delta x_i^{d\prime}$ the idempotent matrix, and $f_t = M_i \Delta F_t$ and $z_{i,t} = M_i \Delta e_{i,t}$. The superscript d in Δx_i^d indicates that there are deterministic elements. The estimation of the common factors and factor loadings can be done as in Bai and Ng (2004) using principal components. Specifically, the estimated principal component of $f = (f_2, f_3, \ldots, f_T)$, denoted as $\tilde{f}$, is $\sqrt{T-1}$ times the r eigenvectors corresponding to the first r largest eigenvalues of the $(T-1) \times (T-1)$ matrix $y_i^* y_i^{*\prime}$, where $y_{i,t}^* = M_i \Delta y_{i,t}$. Under the normalization $\tilde{f}\tilde{f}' / (T-1) = I_r$, the estimated loading matrix is $\tilde{\Pi} = \tilde{f}' y_i^* / (T-1)$. Therefore, the estimated residuals are defined as

$$\tilde{z}_{i,t} = y_{i,t}^* - \tilde{f}_t \tilde{\pi}_i. \tag{13}$$

We can recover the idiosyncratic disturbance terms through cumulation, *i.e.* $\tilde{e}_{i,t} = \sum_{j=2}^{t} \tilde{z}_{i,j}$, and test the unit root hypothesis ($\alpha_{i,0} = 0$) using the ADF regression equation

$$\Delta \tilde{e}_{i,t}\left(\hat{\lambda}_i\right) = \alpha_{i,0} \tilde{e}_{i,t-1}\left(\hat{\lambda}_i\right) + \sum_{j=1}^{k} \alpha_{i,j} \Delta \tilde{e}_{i,t-j}\left(\hat{\lambda}_i\right) + \varepsilon_{i,t}.$$

When $r = 1$ we can use the ADF type equation to analyse the order of integration of F_t as well. However, we should proceed in two steps. In the first step we regress $\tilde{F}_t$ on the deterministic specification and the stochastic regressors. In the second step we estimate the ADF regression equation using the detrended common factor $\left(\tilde{F}_t^d\right)$, *i.e.* the residuals of the first step:

$$\Delta \tilde{F}_t^d = \delta_0 \tilde{F}_{t-1}^d + \sum_{j=1}^{k} \delta_j \Delta \tilde{F}_{t-j}^d + u_t,$$

and test if $\delta_0 = 0$.

Finally, if $r > 1$ we should use one of the two statistics proposed in Bai and Ng (2004) to fix the number of common stochastic trends (q). As before, let $\tilde{F}_t^d$ denote the detrended common factors. Start with $q = r$ and proceed in three stages –we reproduce these steps here for completeness:

1. Let $\tilde{\beta}_\perp$ be the q eigenvectors associated with the q largest eigenvalues of $T^{-2} \sum_{t=2}^{T} \tilde{F}_t^d \tilde{F}_t^{d\prime}$.
2. Let $\tilde{Y}_t^d = \tilde{\beta}_\perp \tilde{F}_t^d$, from which we can define two statistics:
   (a) Let $K(j) = 1 - j/(J+1)$, $j = 0, 1, 2, \ldots, J$:
      i. Let $\tilde{\xi}_t^d$ be the residuals from estimating a first-order VAR in $\tilde{Y}_t^d$, and let

$$\tilde{\Sigma}_1^d = \sum_{j=1}^{J} K(j) \left(T^{-1} \sum_{t=2}^{T} \tilde{\xi}_t^d \tilde{\xi}_t^{d\prime}\right).$$

      ii. Let $\tilde{v}_c^d(q) = \frac{1}{2}\left[\sum_{t=2}^{T}\left(\tilde{Y}_t^d \tilde{Y}_{t-1}^{d\prime} + \tilde{Y}_{t-1}^d \tilde{Y}_t^{d\prime}\right) - T\left(\tilde{\Sigma}_1^d + \tilde{\Sigma}_1^{d\prime}\right)\right]\left(T^{-1}\sum_{t=2}^{T} \tilde{Y}_{t-1}^d \tilde{Y}_{t-1}^{d\prime}\right)^{-1}$.

iii. Define $MQ_c^d(q) = T\left[\tilde{v}_c^d(q) - 1\right]$.

(b) For p fixed that does not depend on N and T:

i. Estimate a VAR of order p in $\Delta\tilde{Y}_t^d$ to obtain $\tilde{\Pi}(L) = I_q - \tilde{\Pi}_1 L - \ldots - \tilde{\Pi}_p L^p$. Filter $\tilde{Y}_t^d$ by $\tilde{\Pi}(L)$ to get $\tilde{y}_t^d = \tilde{\Pi}(L)\tilde{Y}_t^d$.

ii. Let $\tilde{v}_f^d(q)$ be the smallest eigenvalue of

$$\Phi_f^d = \frac{1}{2}\left[\sum_{t=2}^{T}\left(\tilde{Y}_t^d\tilde{Y}_{t-1}^{d\prime} + \tilde{Y}_{t-1}^d\tilde{Y}_t^{d\prime}\right)\right]\left(T^{-1}\sum_{t=2}^{T}\tilde{Y}_{t-1}^d\tilde{Y}_{t-1}^{d\prime}\right)^{-1}.$$

iii. Define the statistic $MQ_f^d(q) = T\left[\tilde{v}_f^d(q) - 1\right]$.

3. If $H_0 : r_1 = q$ is rejected, set $q = q - 1$ and return to the first step. Otherwise, $\tilde{r}_1 = q$ and stop.

The following Theorem offers the main results concerning these statistics.

Theorem 1 *Let $\{y_{i,t}\}$ the stochastic process with DGP given by (8) to (11). The following results hold as $N, T \to \infty$. Let k be the order of autoregression chosen such that $k \to \infty$ and $k^3/\min[N,T] \to 0$.*
(1) Under the null hypothesis that $\rho_i = 1$ in (11),
(1.a) for the specification that does not include time trend with or without level shift(s):

$$ADF_{\hat{e}}^c(i) \Rightarrow \frac{\frac{1}{2}\left(W(1)^2 - 1\right)}{\left(\int_0^1 W(r)^2\, dr\right)^{1/2}},$$

(1.b) for those specifications including a time trend with or without level shift(s):

$$ADF_{\hat{e}}^\tau(i) \Rightarrow -\frac{1}{2}\left(\int_0^1 V(r)^2\, dr\right)^{-1/2},$$

where $V(r) = W(r) - rW(1)$.
(1.c) for those specifications including a time trend with slope shift(s):

$$ADF_{\hat{e}}^\gamma(i) \Rightarrow -\frac{1}{2}\left(\sum_{j=1}^{l+1}(\lambda_j - \lambda_{j-1})^2\int_0^1 V(b_j)^2\, dr\right)^{-1/2},$$

for $j = 1, \ldots, l$ structural breaks, where $V(b_j) = W(b_j) - b_j W(1)$, with $b_j = (r - \lambda_{j-1})/(\lambda_j - \lambda_{j-1})$ so that $0 < b_j < 1$, and $T_{j-1}^b < t \leq T_j^b$ with $\lambda_j = T_j^b/T$, $\lambda_0 = 0$ and $\lambda_{l+1} = 1$.

(2) When $r = 1$, under the null hypothesis that F_t has a unit root and no slope shift(s)

$$ADF_{\tilde{F}}^{d} \Rightarrow \frac{\int_0^1 W_w^d(r)\, dW_w^d(r)}{\left(\int_0^1 W_w^d(r)^2\, dr\right)^{1/2}},$$

where $W_w^d(r)$ denotes the detrended Brownian motion, while when we allow for slope shift(s)

$$ADF_{\tilde{F}}^{d}(\lambda) \Rightarrow \frac{\int_0^1 W_w^d(r,\lambda)\, dW_w^d(r,\lambda)}{\left(\int_0^1 W_w^d(r,\lambda)^2\, dr\right)^{1/2}},$$

where $W_w^d(r,\lambda)$ is the detrended Brownian motion and λ denotes either the break fraction parameter –if there is only one slope shift– or the break fraction vector –if there are more than one slope shift.
(3) When $r > 1$, let W_q be a q-vector of standard Brownian motion and W_q^d the detrended counterpart. Let $v_*^d(q)$ be the smallest eigenvalues of the statistic computed for a model that does not include slope shift(s)

$$\Phi_*^d = \frac{1}{2}\left[W_q^d(1)\, W_q^d(1)' - I_p\right]\left[\int_0^1 W_q^d(r)\, W_q^d(r)'\, dr\right]^{-1},$$

and let $v_*^d(q,\lambda)$ be the smallest eigenvalues of the statistic computed for the model that includes slope shift(s)

$$\Phi_*^d(\lambda) = \frac{1}{2}\left[W_q^d(1,\lambda)\, W_q^d(1,\lambda)' - I_p\right]\left[\int_0^1 W_q^d(r,\lambda)\, W_q^d(r,\lambda)'\, dr\right]^{-1},$$

(3.1) Let J be the truncation lag of the Bartlett kernel, chosen such that $J \to \infty$ and $J/\min\left[\sqrt{N}, \sqrt{T}\right] \to 0$. Then, under the null hypothesis that F_t has q stochastic trends, $T\left[\tilde{v}_c^d(q) - 1\right] \xrightarrow{d} v_*^d(q)$ and $T\left[\tilde{v}_c^d(q,\lambda) - 1\right] \xrightarrow{d} v_*^d(q,\lambda)$.
(3.2) Under the null hypothesis that F_t has q stochastic trends with a finite VAR($\bar{p}$) representation and a VAR(p) is estimated with $p \geq \bar{p}$, $T\left[\tilde{v}_f^d(q) - 1\right] \xrightarrow{d} v_*^d(q)$ and $T\left[\tilde{v}_f^d(q,\lambda) - 1\right] \xrightarrow{d} v_*^d(q,\lambda)$.

The proof of the Theorem is outlined in the Appendix. Some remarks are in order. First, note that the definition of the common factors framework implies that the matrix of projections M_i that is used above cannot depend on i, which means that all elements that are defined in Δx_i^d should be the same across i. There are two different kind of elements in Δx_i^d: (i) the deterministic regressors and (ii) the stochastic regressors. Regarding the latter, we have shown in the Appendix that the limiting distribution of the statistics do not depend on the presence of stochastic regressors, so that we can ignore the effect of these elements when defining M_i. Unfortunately, this is not true for the deterministic regressors. Thus, to warrant that M_i does not (asymptotically) depend on i we have to assume common break dates, *i.e.* we assume that the break points are the same for all individuals. This restriction can be seen as a limitation of our analysis, but in fact it is due to the definition of the common factors framework. Thus, (12) specifies a common factor structure for all individuals, so that f_t cannot depend on i. If we look at the definition of $f_t = M_i \Delta F_t$ we can see that the specification of heterogeneous structural breaks implies that the idempotent matrix M_i depends on i. The only way to overcome this situation is to impose $M_i = M \ \forall i$ so that the structural breaks are the same for all individuals. That is the reason why in Theorem 1 we do not have included any subscript on λ for the individuals.

Second, the limiting distribution of the ADF statistic for the idiosyncratic disturbance term does not depend on the presence of stochastic regressors. Moreover, the presence of level shifts do not affect the limiting distribution of the ADF statistic that is computed using the idiosyncratic disturbance term.

Finally, the distribution of the statistics that focus on the common factors depend on some elements that define the deterministic component although, surprisingly, they do not depend on the number of stochastic regressors. Specifically, the presence of level shifts do not affect the limiting distribution of the ADF and Φ_*^d statistics, although this is not true when there are slope shifts. For the latter, the test statistics depends on the number and location of the structural breaks. Moreover, in this case we have to assume that these structural breaks are common to all individuals. The limiting distribution for the ADF statistic when there is one structural break can be found in Perron (1989) for the specification denoted as Model C. For the $\Phi_*^d(\lambda)$ we have simulated asymptotic critical values that depend both on the number of stochastic common trends and on the break fraction. Note that the critical values reported in Table 9 correspond to the case of only one structural breaks, though our approach can be easily extended to multiple slope shifts.

The individual ADF statistic for the idiosyncratic disturbance terms can be pooled to define a panel data cointegration test. Thus, following the steps given in previous section we can define

$$N^{-1/2} Z^e_{\hat{t}_{NT}} \left(\hat{\lambda}\right) - \Theta_2^e \sqrt{N} \Rightarrow N\left(0, \Psi_2^e\right),$$

where the superscript e denotes the idiosyncratic disturbance term. As for the previous statistics, we have approximated the moments Θ_2^e and Ψ_2^e by simulation. These moments depend on the deterministic specification that is used and, except for the case of slope changes, they are the same as the ones for the statistics in Bai and Ng (2004) –note that these authors prefer to combine individual p-values instead of using these moments.

5 Monte Carlo simulation

We have analysed the finite sample performance of the statistics that have been proposed in the paper conducting a simulation experiment. The empirical size of the tests is studied regressing two independent random walks, which have been generated as the cumulated sum of *iid* $N(0,1)$ processes. The sample size has been set equal to $T = \{50, 100, 250\}$ and the number of individuals at $N = \{20, 40\}$. The results reported in Table 10 are obtained from $r = 5,000$ replications, assuming that the break point is unknown and using the estimated response surfaces of the previous section. As can be seen, the empirical size of both the normalised bias and the pseudo t-ratio statistics is close to the nominal size irrespective of T and N.

The empirical power of the statistics is assessed using the DGP given by:

$$\begin{aligned} y_{i,t} &= \mu_i + \theta_i DU_{i,t} + \xi_i t + \gamma_i DT^*_{i,t} + x'_{i,t}\alpha_{i,t} + z_{i,t} \\ z_{i,t} &= \rho_i z_{i,t-1} + v_{i,t}, \end{aligned}$$

where $v_{i,t} \sim iid\ N(0,1)\ \forall i\,,\ i = 1, \ldots, N$. The specification of the values of the parameters depends on the model under consideration. In general, the constant and, when required, the slope of the trend are set equal to $\mu_i = 1$ and $\xi_i = 0.3$, respectively. When there is a change in the level the magnitude is set equal to $\theta_i = 3$, while for the slope shift we consider $\gamma_i = 0.5$. The change in the cointegrating vector is given by $\alpha_{i,t} = \alpha_{i,1} = 1$ for $t \leq T_{bi}$ and $\alpha_{i,t} = \alpha_{i,1} = 3$ for $t > T_{bi}$, for a break point located at $\lambda_i = 0.5$, $\forall i\,,\ i = 1, \ldots, N$ –the same results are obtained when $\lambda_i = 0.25$ and $\lambda_i = 0.75$. The autoregressive coefficient is set at $\rho_i = 0.5$. The computation of the statistics controls the autocorrelation in the disturbance term including up to $k_{\max} = 5$ lags using the t-sig criterion to select the order of the autoregressive correction. Results reported in Table 11 indicates that the empirical power of both statistics equals one in all situations. The results contrast with the ones in Table 3 where it has been shown that structural breaks, when not accounted for, reduces the power of the statistics.

Let us now deal with the situation with common factors. The DGP is given by a bivariate system:

$$\begin{gathered} y_{i,t} = f_i(t) + x'_{i,t}\delta_{i,t} + u_{i,t} \\ u_{i,t} = F_t\pi_i + e_{i,t} \\ F_t = \phi F_{t-1} + \sigma_F w_t \\ e_{i,t} = \rho_i e_{i,t-1} + \varepsilon_{i,t} \\ \Delta x_{i,t} = v_{i,t}, \end{gathered}$$

where $(w_t, \varepsilon_{i,t}, v_{i,t})'$ follow a mutually *iid* standard multivariate Normal distribution for $\forall i, j$ $i \neq j$ and $\forall t, s$ $t \neq s$. In this paper we consider two different situations depending on the number of common factors, *i.e.* $r = \{1, 3\}$, and specify three values for the autoregressive parameters $\phi = \{0.8, 0.9, 1\}$ and $\rho_i = \{0.8, 0.9, 1\}$ $\forall i$. Note that these values allows to analyse both the empirical size and power of the statistics. The importance of the common factors is controlled through the specification of $\sigma^2_F = \{0.5, 1, 10\}$. The number of common factors is estimated using the panel BIC information criterion in Bai and Ng (2002) with $r_{\max} = 6$ as the maximum number of factors. We consider $N = 40$ individuals and $T = \{50, 100, 250\}$ time observations.

Table 12 reports the results for the constant and time trend cases without structural break. As can be seen, the empirical size of either the ADF pooled idiosyncratic t-ratio statistic $\left(Z^e_{\hat{t}_{NT}}\right)$ and the ADF statistic of the common factor is close to the nominal size, which is set at the 5% level of significance. As expected the power of the tests increases as the autoregressive parameter moves away from unity. Moreover, the power of the $Z^e_{\hat{t}_{NT}}$ test is higher or equal to the power shown by the $ADF^d_{\hat{F}}$ test. Note that these conclusions are obtained irrespective of the deterministic specification.

Tables 13 and 14 show that these results do not change when specifying three common factors for the constant case. Thus, the $Z^e_{\hat{t}_{NT}}$ test shows a correct empirical size and good power. Regarding the $MQ^d_c(q)$ test, it shows correct empirical size, while as expected the test has low power for large values of the autoregressive parameter –the bandwidth for the Bartlett spectral window is set as $J = 4ceil\left[\min\left[N, T\right]/100\right]^{1/4}$. Simulations available upon request indicate that similar conclusions are reached for the time trend case, and when using the parametric approach for the MQ test.

Similar results are obtained when we introduce one structural in the model. At this stage of the analysis we assume that the break point is known and located at $\lambda_i = \{0.25, 0.5, 0.75\}\ \forall i$. Table 15 reports results for the empirical size and power for the model that allows for one level shift with $\lambda_i = 0.5$ and one common factor. It should be mentioned that there are no variations for neither the model that includes a slope shift nor for the other values of λ_i –these results are available upon request. On the one hand, the panel data unit root test on the idiosyncratic disturbance terms show good properties in terms of empirical size and power. On the other hand, the ADF statistic for the common factor shows right size although, as expected, it has low power when the autoregressive parameter is close to unity and the sample size is small.

6 Empirical illustration

The correlation between investment and savings as a ratio of the GDP has devoted huge amount of literature aim reconcile the observation of significant correlation with the idea of capital mobility. The fact that the domestic investment has to be financed by domestic saving goes against the conventional wisdom that in a world of perfect capital mobility, where capital flows among countries should act to equalise the yields to investors, such correlations should not be observed. Thus, (high) capital mobility implies that investment does not need to be correlated with saving. Therefore, the idea of capital mobility and the correlation between investment and saving rates is known as Feldstein-Horioka Puzzle.

There have been different attempts in the literature to assess if such correlation is significant. Some analyses have followed a cross-sectional approach using a sample of countries for which average values of investment and saving ratios in a given time period are analysed. However, most of the analyses have applied time series techniques to assess the extent of the correlation.

In this regard, since both investment and saving ratios are found to be non-stationary processes the presence of correlation requires that cointegration has to be met. Cointegration has been tested from either country-by-country framework or panel data framework.

We contribute to this literature through the analysis of the Feldstein-Horioka Puzzle allowing for one structural break. The selection of this topic for our analysis is not only due to the great attention that has received in recent times, but because there is empirical and theoretical evidence that this correlation might change along time. For instance, Coakley, Kulasi and Smith (1998) note that the coefficient on saving has shown some tendency to decline over recent years for developed countries. Jansen (2000) finds that long-run correlation decreases smoothly over time, which is consistent with the notion of increased international capital mobility. Banerjee and Zanghieri (2003) analyse fourteen European countries and reporting that long run association drops quickly starting from the mid-80, when most European countries fully liberalised their external accounts. Finally, Westerlund (2004) illustrates the LM cointegration test statistic with one structural break using the Feldstein-Horioka puzzle concluding that the null hypothesis of cointegration cannot be rejected once the presence of structural breaks is taken into account.

In this section we are going to investigate the Feldstein-Horioka puzzle through the application of Pedroni cointegration statistics and the modifications that have been suggested in the paper. To the best of our knowledge, this is the first time that the null of no cointegration is tested including one structural break in the panel data set. The data set is the one in Banerjee and Zanghieri (2003) and is taken from the European Commision's Annual Macroeocomic Database of the Directorate General for Economic and Financial Affairs (AMECO), that combines data obtained from national sources as well as from the IMF and OECD. The data measured at an annual frequency covers from 1960 to 2002 for fourteen countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Portugal, Spain, Sweden and United Kingdom.

Figure 1 presents pictures of the investment and saving shares for four countries. As can be seen, there is evidence that the relationship between investment and saving might has changed the pattern along the time period. Table 18 reports the results on the Pedroni (2004) statistic specifying a constant as the deterministic component. All computations have been carried out using GAUSS. The order of the autoregressive correction that is required in (7) is selected with the t-sig criterion in Ng and Perron (1995) with $k_{\max} = 7$ lags for the maximum order. Different conclusions are reached depending on the statistic. On the one hand, when using the $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ statistic the null hypothesis of no cointegration cannot be rejected neither at the 5% not at the 10% level of significance. On the other hand, the $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ statistic rejects the null at the 5% level. However, these conclusions might not be valid if there is some dependence amongst individuals. The computation of Pedroni statistics assumes cross-section independence across i, an assumption that is difficult to be hold in empirical applications. Banerjee, Marcellino and Osbat (2004, 2005) show that one of the crucial assumption underlying all the tests of panel cointegration, namely the absence of cointegration across the units of the sample is likely to be violated in many macroeconomic time series. In fact, Banerjee and Zanghieri (2003) report that there is cross-section cointegration between the individuals of the panel sets that we are considering here. In order to take into account cross-section dependence when carrying out the cointegration analysis, we have decided to compute the bootstrap distribution of the statistics. Some cautions about the method that is used to bootstrap cointegration relationships are required, since not all available procedures lead to consistent estimates. In this regard, we have followed Phillips (2001), Park (2002), and Chang, Park and Song (2002), and we have decided to use sieve bootstrap. Our proposal is a modified version of the sieve bootstrap described in the papers mentioned above. Specifically, it consist of the following steps:

- **Step 1:** Fit one of the regressions in (1) to (6) by OLS to obtain $\hat{e}_{i,t}\left(\hat{\lambda}_i\right)$, and define $w_{i,t} = \left(\hat{u}_{i,t}\left(\hat{\lambda}_i\right), v'_{i,t}\right)'$ where $\hat{u}_{i,t}\left(\hat{\lambda}_i\right) = \Delta\hat{e}_{i,t}\left(\hat{\lambda}_i\right)$ and $v_{i,t} = \Delta y^1_{i,t}$.

- **Step 2:** Apply the sieve estimation method to the following VAR(q):

$$w_{i,t} = \Phi_1 w_{i,t-1} + \cdots + \Phi_q w_{i,t-q} + \varepsilon_{i,t},$$

  where the order of the VAR(q) is approximated using the BIC criterion with the maximum order given by $q_{\max} = T^{1/2}$. Obtain $\varepsilon^*_{i,t}$ by resampling the centered fitted residuals $\tilde{\varepsilon}_{i,t} - \frac{1}{T}\sum_{t=1}^{T}\tilde{\varepsilon}_{i,t}$, and construct the bootstrap samples $w^*_{i,t}$ recursively using

$$w^*_{i,t} = \Phi_1 w^*_{i,t-1} + \cdots + \Phi_q w^*_{i,t-q} + \varepsilon^*_{i,t},$$

  given the initial values $w^*_{i,t} = w_{i,t}$ for $t = 0, \ldots, 1-q$.

- **Step 3:** Define $w^*_{i,t} = \left(u^*_{i,t}\left(\hat{\lambda}_i\right), v^{*\prime}_{i,t}\right)'$ analogously as $w_{i,t} = \left(u_{i,t}\left(\hat{\lambda}_i\right), v'_{i,t}\right)'$. Obtain the bootstrap samples $e^*_{i,t}\left(\hat{\lambda}_i\right)$ and $y^{1*}_{i,t}$ by integrating $u^*_{i,t}\left(\hat{\lambda}_i\right)$ and $v^*_{i,t}$ respectively, *i.e.* $e^*_{i,t}\left(\hat{\lambda}_i\right) = e^*_{i,0}\left(\hat{\lambda}_i\right) + \sum_{j=1}^{t} u^*_{i,j}\left(\hat{\lambda}_i\right)$ and $y^{1*}_{i,t} = y^{1*}_{i,0} + \sum_{j=1}^{t} v^*_{i,t}$, with $e^*_{i,0}\left(\hat{\lambda}_i\right) = \hat{e}^*_{i,0}\left(\hat{\lambda}_i\right)$ and $y^{1*}_{i,0} = y^1_{i,0}$. Then, generate the bootstrap samples for $y^{2*}_{i,t}$ from

$$y^{2*}_{i,t} = f_i(t) + y^{1*}_{i,t}\delta_{i,t} + e^*_{i,t}\left(\hat{\lambda}_i\right), \quad (14)$$

  where the definition of $f_i(t)$ and $\delta_{i,t}$ depends on the model under consideration.

- **Step 4:** Estimate (14) by OLS for each individual assuming unknown break point position and compute the panel cointegration statistics. In this paper we have considered 2,000 bootstrap replications.

Now, using the bootstrap critical values the null hypothesis of no cointegration cannot be rejected by any of the statistics. Therefore, we should conclude that there is no correlation between investment and saving shares, which has been interpreted in the literature as evidence of capital mobility. However, pictures given above indicate that this relationship might has experienced the effect of structural changes. If this is the case, we have shown that the power of Pedroni panel cointegration statistics can be reduced if the structural breaks do not occur at the beginning of the time period. In order to investigate the sensitivity of the cointegration analysis to the presence of structural breaks we have estimated the model that includes a level shift (Model 1), and the model that includes both level and cointegrating vector shift (Model 4). Table 18 presents the values of the statistics for these models. When Model 1 is estimated the null hypothesis of no cointegration is rejected by both statistics using the Normal distribution. This conclusion is robust to the presence of cross-section dependence, since the bootstrap critical values lead to reject the null hypothesis of no cointegration at least at the 10% level of significance. The situation is not so clear when estimating Model 4. Now the null hypothesis is still rejected by both statistics when assuming cross-section independence, but this conclusion does not hold when comparing the $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ statistic with its bootstrap distribution.

7 Conclusions

This paper has shown that inference based on parametric Pedroni panel cointegration test statistics can be affected by the presence of structural breaks. Monte Carlo evidence indicates that in some situations the power of the tests drops as the magnitude of the structural break increases. Specifically, when the structural break affects either the slope of the time trend or the cointegrating vector the power approaches zero as T, N and the magnitude of the break increases. Notwithstanding, the power of the standard parametric Pedroni panel cointegration statistics is not so much affected when the structural break only shifts the level –we require a large magnitude of structural breaks located at the end of the time period to reduce the power of the statistics.

These features have motivated our proposal, and have led us to design statistical procedures to account for the presence of structural breaks when testing for cointegration. Six different specifications have been introduced depending on the effect of structural breaks on the long-run relationship. Finite sample and asymptotic moments have been computed that allow defining panel cointegration statistics for the specifications considered.

The cross-section dependence is addressed in the paper in two different ways. First, we assume an approximate common factor structure to model the cross-section dependence. We derive the limiting distributions of statistics in two situations of interest, *i.e.* (i) for the case of no structural break, and (ii) when there are level and slope shifts. The performance of the approach is investigated through Monte Carlo simulations, from which we conclude that the statistics show good performance once structural breaks are accounted for. The paper illustrates the application of the statistics analysing the Feldstein-Horioka puzzle. Since the assumption of cross-section independence is hardly satisfied in practice, we have approximated the empirical distribution of the statistics using sieve bootstrap. This defines the second approach to cross-section dependence matter. The main conclusion is that after structural breaks are considered we find evidence that point to cointegration between investment and saving shares.

References

[1] Bai, J. and Ng, S. (2002): "Determining the Number of Factors in Approximate Factor Models", *Econometrica*, 70, 191-221.

[2] Bai, J. and S. Ng (2004): "A PANIC Attack on Unit Roots and Cointegration", *Econometrica*, 72, 4, 1127-1177.

[3] Baltagi, B. H. (2001): *Econometric Analysis of Panel Data*. John Wiley & Sons.

[4] Baltagi, B. H., and C. Kao (2000): Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey. In Baltagi (Ed.) *Nonstationary Panels, Cointegration in Panels and Dynamic Panels*, Advanves in Econometrics, 15. North-Holland, Elsevier.

[5] Banerjee, A. (1999): "Panel Data Unit Roots and Cointegration: An Overview", *Oxford Bulletin of Economics and Statistics*, Special issue, 607–629.

[6] Benerjee, A., Marcellino, M. and C. Osbat (2004): "Some Cautions on the Use of Panel Methods for Integrated Series of Macro-Economic Data", *Econometrics Journal*, 7, 2, 322-340.

[7] Benerjee, A., Marcellino, M. and C. Osbat (2005): "Testing for PPP: Should We Use Panel Methods?". *Empirical Economics*, Forthcoming.

[8] Benerjee, A. and P. Zanghieri (2003): "A New Look at the Feldstein-Horioka Puzzle using an Integrated Panel". *Working Paper* 2003-22. Centre d'Etudes Prospectives et d'Information Internacionales.

[9] Bartley, W. A., J. Lee and M. C. Strazicich (2001): "Testing the Null of Cointegration in the Presence of a Structural Break", *Economics Letters*, 73 3, 315–323.

[10] Breitung, J. (2002): "A Parametric Approach to the Estimation of Cointegration Vectors in Panel Data", *Mimeo*. Institute of Statistics and Econometrics. Humboldt University Berlin.

[11] Busetti, F. (2002): "Testing for (Common) Stochastic Trends in the Presence of Structural Breaks", *Journal of Forecasting*, 21, 81–105.

[12] Campos, J., N. Ericsson and D. F. Hendry (1996): "Cointegration Tests in the Presence of Structural Breaks", *Journal of Econometrics*, 70, 187-220.

[13] Carrion-i-Silvestre, J. Ll. and A. Sansó (2004): "Testing the Null of Cointegration with Structural Breaks", *Mimeo*. Department of Econometrics, Statistics and Spanish Economy. University of Barcelona.

[14] Chang, Y., J. Y. Park and K. Song (2002): "Bootstrapping Cointegrating Regressions". *Mimeo*. Rice University.

[15] Coakley, J., F. Kulasi and R. Smith (1998): "The Feldstein-Horioka Puzzle and Capital Mobility: A Review", *International Journal of Finance and Economics*, 3, 169-188.

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

[17] Gregory, A. W. and B. E. Hansen (1996): "Residual-based Tests for Cointegration in Models with Regime Shifts", *Journal of Econometrics*, 70, 99-126.

[18] Hansen, B. E. (1992): "Tests for Parameter Instability in Regressions with I(1) Processes", *Journal of Business and Economic Statistics*, 10, 3, 321-335.

[19] Hansen, H. and S. Johansen (1993): "recursive Estimation in Cointegrated VAR-Models", *Mimeo*. Institute of Mathematical Statistics. University of Copenhagen.

[20] Hao, K. (1996): "Testing for Structural Change in Cointegrated Regression Models: some Comparisons and Generalisations", *Econometric Reviews*, 15, 4, 401–429.

[21] Harris, D. and B. Inder (1994): "A Test of the Null Hypothesis of Cointegration". In Colin P. Hargreaves (Ed.) *Nonstationary Time Series Analysis and Cointegration*. Oxford University Press.

[22] Jansen, W. J. (2000): "International Capital Mobility", *Journal of International Money and Finance*, 19, 507-511.

[23] Kao, C. and M. H. Chiang (2000): "On the Estimation and Inference of a Cointegrated Regression in Panel Data". In *Nonstationary Panels, Panels Cointegration and Dynamic Panels. Advances in Econometrics*, vol. 15. Ed. B. H. Baltagi. Elsevier.

[24] Mark, N. C. and D. Sul (2003): "Cointegration Vector Estimation by Panel DOLS and Long-run Money Demand", *Oxford Bulletin of Economics and Statistics*, 65, 5, 655-680.

[25] McCoskey, S. and C. Kao (1998): "A Residual-Based Test of the Null of Cointegration in Panel Data", *Econometric Reviews*, 17, 57-84.

[26] Ng, S. and P. Perron (1995): "Unit Root Test in ARMA models with Data-Dependent Methods for the Selection of the Truncation Lag", *Journal of the American Statistical Association*, 90, 268-281.

[27] Park, J. Y. (2002): "An Invariance Principle for Sieve Bootstrap in Time Series", *Econometric Theory*, 18, 469-490.

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

[29] 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, 3, 597-625.

[30] Perron, P. (1989): "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis", *Econometrica*, 57, 1361-1401.

[31] Phillips, P. C. B. (1987): "Time Series Regression with a Unit Root", *Econometrica*, 55, 277-301.

[32] Phillips, P. C. B. (2001): "Bootstrapping Spurious Regressions". *Cowles Foundation Discussion Paper* num. 1330. Yale.

[33] Phillips, P. C. B. and S. Ouliaris (1990): "Asymptotic Properties of Residual Based Tests for Cointegration", *Econometrica*, 58, 165-193.

[34] Phillips, P. C. B. and P. Perron (1988): "Testing for a Unit Root in Time Series Regression", *Biometrika*, 335-346.

[35] Quintos, C. and P. C. B. Phillips (1993): "Parameter Constancy in Cointegrating Regressions", In *New Developments in Time Series Econometrics*, J. M. Dufour and B. Raj (Eds.). Studies in Empirical Economics. Physica-Verlag. Springer Verlag.

[36] Shin, Y. (1994): "A Residual-based Test of the Null of Cointegration against the Alternative of no Cointegration", *Econometric Theory*, 10, 91–115.

[37] Westerlund, J. (2004): "Testing the Null of Panel Cointegration in the Presence of a Structural Break", *Mimeo*. Department of Economics. University of Lund.

A Mathematical Appendix

A.1 Pedroni Test Statistic with constant

For sake of simplicity let us first assume that there is no structural breaks affecting the model and there is no deterministic elements in the model –note that the presence of a constant term does not change the results since it disappears when taking first differences. Let us assume the model given by (8) and (9). Alternatively, the model can be expressed as:

$$y_{i,t} = x'_{i,t}\beta_i + F_t\pi_i + e_{i,t}.$$

As can be seen, the model assumes that residuals from the static regression follows a factor structure as defined in Bai and Ng (2004). Note that if we introduce (12) in (13) we obtain

$$\begin{aligned} \tilde{z}_{i,t} &= z_{i,t} + f_t\pi_i - \tilde{f}_t\tilde{\pi}_i \\ &= z_{i,t} - v_t H^{-1}\pi_i - \tilde{f}_t d_i, \end{aligned} \tag{15}$$

where $v_t = \tilde{f}_t - f_t H$ and $d_i = \tilde{\pi}_i - H^{-1\prime}\pi_i$. The computation of the partial sum processes of (15) gives:

$$T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} = T^{-1/2}\sum_{j=2}^{t} z_{i,j} - T^{-1/2}\sum_{j=2}^{t} v_j H^{-1}\pi_i - T^{-1/2}\sum_{j=2}^{t}\tilde{f}_j d_i. \tag{16}$$

Let us analyse each element of (16) separately. The left-hand side of (16) is equal to

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} &= T^{-1/2}\sum_{j=2}^{t} M_i\Delta\tilde{e}_{i,j} \\ &= T^{-1/2}\sum_{j=2}^{t}\Delta\tilde{e}_{i,j} - T^{-1/2}\sum_{j=2}^{t}\left[P_i\Delta\tilde{e}_i\right]_j, \end{aligned} \tag{17}$$

where $\left[P_i\Delta\tilde{e}_i\right]_j$ denotes the j-th element of the matrix $P_i\Delta\tilde{e}_i$, and $P_i = I_{T-1} - M_i$. The first element on the right of (17) is equal to

$$T^{-1/2}\sum_{j=2}^{t}\Delta\tilde{e}_{i,j} = T^{-1/2}\tilde{e}_{i,t} - T^{-1/2}\tilde{e}_{i,1} = T^{-1/2}\tilde{e}_{i,t} + O_p(1),$$

so that by the invariance principle

$$T^{-1/2}\sum_{j=2}^{t}\Delta\tilde{e}_{i,j} \Rightarrow \sigma W(r).$$

The second element on the right hand of (17) is

$$T^{-1/2}\sum_{j=2}^{t}\left[P_i\Delta\tilde{e}_i\right]_j = T^{-1/2}\left(x_{i,t}-x_{i,1}\right)'\left(\Delta x_i'\Delta x_i\right)^{-1}\Delta x_i'\Delta\tilde{e}_i.$$

Note that $\left(\Delta x_i'\Delta x_i\right)^{-1}\Delta x_i'\Delta\tilde{e}_i = \left(T^{-1}\Delta x_i'\Delta x_i\right)^{-1}\left(T^{-1}\Delta x_i'\Delta\tilde{e}_i\right) = o_p(1)$, since $\left(T^{-1}\Delta x_i'\Delta x_i\right)\to^{p} Q_{\Delta x_i\Delta x_i}$, the variance and covariance matrix of $\Delta x_i'\Delta x_i$, and $T^{-1}\Delta x_i'\Delta\tilde{e}_i \to 0$ since these elements are orthogonal by definition. On the other hand, $T^{-1/2}x_{i,t} \Rightarrow \Omega_{22,i}^{1/2}W_k(r)$ and $T^{-1/2}x_{i,1}\to 0$ by assumption. These derivations lead us to

$$T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} = T^{-1/2}\tilde{e}_{i,t} + o_p(1),$$

since $T^{-1/2}x_{i,t}\left(\Delta x_i'\Delta x_i\right)^{-1}\Delta x_i'\Delta\tilde{e}_i = o_p(1)$. The same result can be achieved for $T^{-1/2}\sum_{j=2}^{t} z_{i,j}$, *i.e.*

$$T^{-1/2}\sum_{j=2}^{t} z_{i,j} = T^{-1/2}e_{i,t} + o_p(1).$$

This indicates that the presence of stochastic regressors does not have any effect on the partial sum processes. Regarding the term involving $\{v_t\}$ we see from Eq. (A.3) in Bai and Ng (2004) that

$$T^{-1/2}\sum_{j=2}^{t} v_j = O_p\left(C_{NT}^{-1}\right),$$

where $C_{NT} = \min\left\{N^{-1/2}, T^{-1/2}\right\}$. Moreover and as shown in Bai and Ng (2004), the term $d_i = O_p\left(C_{NT}^{-1}\right)$ and $T^{-1/2}\sum_{j=2}^{t}\tilde{f}_j = O_p(1)$, so that

$$T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} = T^{-1/2}\sum_{j=2}^{t} z_{i,j} + O_p\left(C_{NT}^{-1}\right).$$

From all these results it follows that

$$DF_{\tilde{e}}^{c}(i) \Rightarrow \frac{\frac{1}{2}\left(W(r)^2-1\right)}{\left(\int_0^1 W(r)^2\,dr\right)^{1/2}},$$

that is, the limiting distribution is the same derived in Bai and Ng (2004) for the constant case –see Bai and Ng (2004) for the proof. The same result is found for the ADF test. This implies that the presence of stochastic regressors does not affect the limiting distribution of the statistic.

Let us now deal with the unit root hypothesis testing when there is $r = 1$ common factor. The first difference of the model defines an idempotent matrix M_i that depends on the individual. At first sight this goes against the definition of common factor since we assume that this element is common to all individuals and, hence, it cannot depend on i. Notwithstanding, it is shown below that the elements that depend on i vanish asymptotically. Thus, note that

$$\begin{aligned}\sum_{j=2}^{t} \tilde{f}_j &= \sum_{j=2}^{t} M_i \Delta \tilde{F}_t \\ &= \tilde{F}_t - (x_{i,t} - x_{i,1})' \left(\Delta x_i' \Delta x_i\right)^{-1} \Delta x_i' \Delta \tilde{F}, \end{aligned} \tag{18}$$

since we define $\tilde{F}_1 = 0$. Note that the first element of (18) is

$$\tilde{F}_t = H\left(F_t - F_1\right) + V_t,$$

since $\Delta \tilde{F}_t = H\ \Delta F_t + v_t$ and $V_t = \sum_{j=2}^{t} v_j$. The detrended estimated factor will remove F_1:

$$\tilde{F}_t^d = H\ F_t^d + V_t^d,$$

which it can be shown that

$$T^{-1/2} \tilde{F}_t^d = H\ T^{-1/2} F_t^d + O_p\left(C_{NT}^{-1}\right),$$

since $T^{-1/2} V_t^d = O_p\left(C_{NT}^{-1}\right)$ –see Bai and Ng (2004), Lemma B.2. The second term in (18) is $T^{-1/2}\left(x_{i,t} - x_{i,1}\right)' \left(\Delta x_i' \Delta x_i\right)^{-1} \Delta x_i' \Delta \tilde{F} = o_p(1)$, since $T^{-1} \Delta x_i' \Delta x_i$ converges to the matrix of covariance of Δx_i and $T^{-1} \Delta x_i' \Delta \tilde{F} = o_p(1)$ by assumption. Since

$$\begin{aligned} T^{-1/2} \tilde{F}_t^d &\Rightarrow H\ W_w^d(r) \\ T^{-2} \sum_{t=2}^{T} \tilde{F}_{t-1}^d \tilde{F}_{t-1}^{d\prime} &\Rightarrow H^2\ \sigma_w^2 \int_0^1 W_w^d(r)^2\, dr \\ T^{-1} \sum_{t=2}^{T} \tilde{F}_{t-1}^d \Delta \tilde{F}_t &\Rightarrow H^2\ \sigma_w^2 \int_0^1 W_w^d(r)\, dW(r), \end{aligned}$$

the DF statistic converges to

$$\begin{aligned} DF_{\tilde{F}}^d &= \frac{T^{-1} \sum_{t=2}^{T} \tilde{F}_{t-1}^d \Delta \tilde{F}_t}{\left(\tilde{\sigma}_u^2 T^{-2} \sum_{t=2}^{T} \left(\tilde{F}_{t-1}^d\right)^2\right)^{1/2}} \\ &\Rightarrow \frac{\int_0^1 W_w^d(r)\, dW(r)}{\left(\int_0^1 W_w^d(r)^2\, dr\right)^{1/2}}, \end{aligned} \tag{19}$$

where $W_w^d(r)$ denotes the detrended Brownian motion and $\tilde{\sigma}_w^2 \xrightarrow{p} H^2\ \sigma_w^2$. The ADF statistic has the same limiting distribution provided that the order of the autoregressive correction is selected such that $k \to \infty$ and $k^3 / \min[N, T] \to 0$.

The limiting distribution of the test statistic that is used when there are more than one common factor ($r > 1$) is the same as the one derived in Bai and Ng (2004) for the constant case. We address the reader to their paper for the proof of this part of the Theorem.

A.2 Pedroni Test Statistic with constant term and level shifts

The specification that includes level shifts does not affect the limiting distribution of the statistic, that is, we obtain the same limiting distribution derived above for the constant case. Let us consider the simplest situation in which there is only one level shift, although the derivations can be extended to multiple level shifts. The deterministic function is given by

$$f_i(t) = \mu_i + \theta_i DU_{i,t},$$

which implies that $\Delta f_i(t) = \theta_i D\left(T_b^i\right)_t$ and $\Delta x_{i,t}^d = \left(D\left(T_b^i\right)_t, \Delta x_{i,t}'\right)$. Note that

$$T^{-1}\Delta x_i^{d\prime}\Delta x_i^d = \begin{bmatrix} T^{-1} & T^{-1}\Delta x_{i,T_b^i+1}' \\ T^{-1}\Delta x_{i,T_b^i+1} & T^{-1}\Delta x_i'\Delta x_i \end{bmatrix},$$

where in the limit all elements but $T^{-1}\Delta x_i'\Delta x_i$ converges to zero. On the other hand, we can distinguish two elements of $T^{-1}\Delta x_i^{d\prime}\Delta\tilde{e}_i$. The first element is given by $T^{-1}D\left(T_b^i\right)'\Delta\tilde{e}_i = T^{-1/2}\left(T^{-1/2}\Delta\tilde{e}_{i,T_b^i+1}\right)$, where $T^{-1/2}\Delta\tilde{e}_{i,T_b^i+1} \Rightarrow \sigma dW(\lambda_i)$. The second set of elements is given by $T^{-1}\Delta x_i'\Delta\tilde{e}_i$ that converges to zero since we have assumed independency. Regarding the partial sum process of $\Delta x_{i,t}^d$

$$T^{-1/2}\sum_{j=2}^{t}\Delta x_{i,j}^d = \left[T^{-1/2}DU_{i,t} \quad T^{-1/2}\left(x_{i,t} - x_{i,1}\right)'\right],$$

The extra rescaling factor $T^{-1/2}$ that is not used when obtaining the limit of $T^{-1}D\left(T_b^i\right)'\Delta\tilde{e}_i$ implies that $T^{-1/2}\sum_{j=2}^{t}\Delta x_{i,j}^d\left(\Delta x_i^{d\prime}\Delta x_i^d\right)^{-1}\Delta x_i^{d\prime}\Delta\tilde{e}_i = o_p(1)$, from which is evident that the limit distribution of the statistic $ADF_{\tilde{e}}^c(i)$ is not affected by the presence of level shifts, so that $ADF_{\tilde{e}}^c(i)$ converges to the same limiting distribution as in the constant case without level shifts provided that the order of the autoregressive correction is selected such that $k \to \infty$ and $k^3/\min[N,T] \to 0$.

Regarding the situation in which there is only one common factor, $r = 1$, and we proceed to test the unit root hypothesis, we only have to analyse the order of magnitude of $T^{-1/2}\sum_{j=2}^{t}\left[P_i\Delta\tilde{F}\right]_j$, where $P_i = I_{T-1} - M_i = \Delta x_i^d\left(\Delta x_i^{d\prime}\Delta x_i^d\right)^{-1}\Delta x_i^{d\prime}\Delta\tilde{F}$. As for the idiosyncratic disturbance term analysis, $T^{-1}\Delta x_i^{d\prime}\Delta\tilde{F}$ involves two different elements. First, $T^{-1}D\left(T_b^i\right)'\Delta\tilde{F} = T^{-1}\Delta\tilde{F}_{T_b+1} = O_p\left(T^{-1/2}\right)$. Second, $T^{-1}\Delta x_i'\Delta\tilde{F} = o_p(1)$ by assumption. Therefore, using these elements and the results derived above we can see that $T^{-1/2}\sum_{j=2}^{t}\left[P_i\Delta\tilde{F}\right]_j = o_p(1)$, so that both the presence of level shifts and stochastic regressors does not affect the limiting distribution of the $ADF^d_{\tilde{F}}$ statistic, which is the same as the one derived for the constant case without level shifts.

This feature is also found for the statistic that is used when $r > 1$, which has the same limiting distribution as for the constant case in Bai and Ng (2004).

A.3 Pedroni Test Statistic with time trend

The generalisation that includes a time trend can be carried out as well. In this case the model (8) is replaced by

$$y_{i,t} = \mu_i + \beta_i t + x'_{i,t}\beta_i + u_{i,t}.$$

Note that as before we are not dealing with the structural break case since we are defining the benchmark limiting distributions. Contrary to previous specification, taking first differences does not remove the deterministic elements, since now the trend becomes a constant. This is a relevant feature since the limiting distribution of the ADF-type statistic varies. However, the asymptotic distribution of the statistic is the same as the one derived in Bai and Ng (2004) for the trend case. The proof follows similar steps above. Now the first difference of regressors defines the following idempotent matrix

$$M_i = I_{T-1} - \Delta x_i^d \left(\Delta x_i^{d\prime} \Delta x_i^d\right)^{-1} \Delta x_i^{d\prime},$$

where the Δx_i^d matrix is defined by the row vectors $\left(1, \Delta x'_{i,t}\right)'$. Note that as before the first element of (17) converges to

$$T^{-1/2} \sum_{j=2}^{t} \Delta \tilde{e}_{i,j} \Rightarrow \sigma W(r).$$

The limiting expression of the second element in (17) has to be derived in several steps. First, note that $T^{-1}\Delta x_i^{d\prime}\Delta x_i^d$ converges to variance and covariance matrix of Δx_i^d, so that all these elements are $O_p(1)$. The first element of the vector $T^{-1}\Delta x_i^{d\prime}\Delta\tilde{e}_i$ is given by $T^{-1/2}\left(T^{-1/2}\sum_{t=1}^{T}\Delta\tilde{e}_{i,t}\right) = T^{-1/2}\left(T^{-1/2}\left(\tilde{e}_{i,T} - \tilde{e}_{i,1}\right)\right)$, where $T^{-1/2}\left(\tilde{e}_{i,T} - \tilde{e}_{i,1}\right) \Rightarrow \sigma W(1)$ since $T^{-1/2}\tilde{e}_{i,1} \to 0$. Note that the extra rescaling term $T^{-1/2}$ would be used below. The rest of the elements in $T^{-1}\Delta x_i^{d\prime}\Delta\tilde{e}_i$ involve cross-products among the first difference of the stochastic regressors and $\Delta\tilde{e}_i$ that converges to zero since we have assumed independency. Therefore,

$$\left(\Delta x_i^{d\prime}\Delta x_i^d\right)^{-1}\Delta x_i^{d\prime}\Delta\tilde{e}_i = \left[\begin{array}{c} E\, T^{-1/2}\left(T^{-1/2}\left(\tilde{e}_{i,T} - \tilde{e}_{i,1}\right)\right) + o_p(1) \\ \left(-D^{-1}CE\right)\, T^{-1/2}\left(T^{-1/2}\left(\tilde{e}_{i,T} - \tilde{e}_{i,1}\right)\right) + o_p(1) \end{array}\right]$$

where $E = \left(A - BD^{-1}C\right)^{-1}$ and $A = 1, B = T^{-1}\iota'\Delta x_i, C = B'$ and $D = T^{-1}\Delta x_i'\Delta x_i$ denote the elements of the partitioned matrix $T^{-1}\Delta x_i^{d\prime}\Delta x_i^d$, with $\iota = (1, \ldots, 1)'$. The partial sum process of $\Delta x_{i,t}^{d}$ is

$$T^{-1/2}\sum_{j=2}^{t}\Delta x_{i,j}^{d}=\left[T^{-1/2}t \quad T^{-1/2}\left(x_{i,t}-x_{i,1}\right)'\right],$$

so that

$$T^{-1/2}\sum_{j=2}^{t}\Delta x_{i,j}^{d}\left(\Delta x_{i}^{d\prime}\Delta x_{i}^{d}\right)^{-1}\Delta x_{i}^{d\prime}\Delta\tilde{e}_{i}=\frac{t}{T}E\ \left(T^{-1/2}\left(\tilde{e}_{i,T}-\tilde{e}_{i,1}\right)\right)+o_{p}\left(1\right),$$

since $T^{-1}\left(x_{i,t}-x_{i,1}\right)'=o_{p}\left(1\right)$. Moreover, the matrix E can be expressed as

$$\begin{aligned}\left(A-BD^{-1}C\right)^{-1} &= A^{-1}+A^{-1}B\left(D-CA^{-1}B\right)^{-1}CA^{-1}\\ &= 1+B\left(D-B'B\right)^{-1}B'.\end{aligned}$$

Note that $B=T^{-1}\iota'\Delta x_{i}\to_{p}0$ so that $\left(A-BD^{-1}C\right)^{-1}=1+o_{p}\left(1\right)$. Therefore,

$$\begin{aligned}T^{-1/2}\sum_{j=2}^{t}\Delta x_{i,j}^{d}\left(\Delta x_{i}^{d\prime}\Delta x_{i}^{d}\right)^{-1}\Delta x_{i}^{d\prime}\Delta\tilde{e}_{i} &= \frac{t}{T}\left(T^{-1/2}\left(\tilde{e}_{i,T}-\tilde{e}_{i,1}\right)\right)+o_{p}\left(1\right)\\ &\Rightarrow r\ \sigma W\left(1\right).\end{aligned}$$

From Bai and Ng (2004), the terms $T^{-1/2}\left\|\sum_{j=2}^{t}v_{j}\right\|=O_{p}\left(C_{NT}^{-1}\right)$, $\|d_{i}\|=O_{p}\left(C_{NT}^{-1}\right)$ and $T^{-1/2}\left\|\sum_{j=2}^{t}\tilde{f}_{j}\right\|=O_{p}\left(1\right)$. These derivations lead us to

$$\begin{aligned}T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} &= T^{-1/2}\tilde{e}_{i,t}-\frac{t}{T}T^{-1/2}\tilde{e}_{i,T}+O_{p}\left(C_{NT}^{-1}\right)\\ &\Rightarrow \sigma\left(W\left(r\right)-r\ W\left(1\right)\right)\equiv\sigma V\left(r\right).\end{aligned}$$

The DF statistic is

$$DF_{\tilde{e}}^{\tau}\left(i\right)=\frac{T^{-1}\sum_{t=2}^{T}\tilde{e}_{i,t-1}\Delta\tilde{e}_{i,t}}{\left(\tilde{\sigma}^{2}T^{-2}\sum_{t=2}^{T}\tilde{e}_{i,t-1}^{2}\right)^{1/2}}.$$

Note that the following identity holds

$$T^{-1}\sum_{t=2}^{T}\tilde{e}_{i,t-1}\Delta\tilde{e}_{i,t}=\frac{\tilde{e}_{i,T}^{2}}{2T}-\frac{\tilde{e}_{i,1}^{2}}{2T}-\frac{1}{2T}\sum_{t=2}^{T}\left(\Delta\tilde{e}_{i,t}\right)^{2},$$

which shows that $T^{-1}\tilde{e}_{i,T}^{2}\Rightarrow\sigma^{2}V\left(1\right)^{2}=0$, $T^{-1}\tilde{e}_{i,1}^{2}=0$ and $T^{-1}\sum_{t=2}^{T}\left(\Delta\tilde{e}_{i,t}\right)^{2}\to_{p}\sigma^{2}$, from which it follows that $T^{-1}\sum_{t=2}^{T}\tilde{e}_{i,t-1}\Delta\tilde{e}_{i,t}\to_{p}-\sigma^{2}/2$ and $T^{-2}\sum_{t=2}^{T}\tilde{e}_{i,t-1}^{2}\Rightarrow\sigma^{2}\int_{0}^{1}V\left(r\right)^{2}dr$

–see Bai and Ng (2004), Lemma G.4. Using these elements it is straightforward to see that

$$DF_{\tilde{e}}^{\tau}(i) \Rightarrow -\frac{1}{2}\left(\int_0^1 V(r)^2\, dr\right)^{-1/2},$$

where $V(r) = W(r) - r\ W(1)$, *i.e.* the limiting distribution is the same derived in Bai and Ng (2004) for the trend case. Although the proof is more involved, the same result is achieved for the ADF test. As before, this implies that the presence of stochastic regressors does not affect the limiting distribution of the statistic. Note that this result is also achieved when there are level shifts in the model, since the impulse dummies do not affect the limiting distribution of the $ADF_{\tilde{e}}^{\tau}(i)$ statistic.

Let us now deal with the unit root hypothesis testing when there is $r = 1$ common factor. As before,

$$\sum_{j=2}^{t} \tilde{f}_j = \tilde{F}_t - \sum_{j=2}^{t}\left[P_i \Delta \tilde{F}\right]_j.$$

We can distinguish two different elements in $T^{-1}\Delta x_i^{d\prime}\Delta\tilde{F}$. The first one is $T^{-1}\sum_{t=2}^{T}\Delta\tilde{F}_t = H\ T^{-1}(F_T - F_1) = O_p\left(T^{-1/2}\right)$. The second set of elements is $T^{-1}\Delta x_i^{\prime}\Delta\tilde{F} = o_p(1)$ by assumption. Following similar steps above, it is cumbersome but straightforward to see that

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{t}\tilde{f}_j &= H\ T^{-1/2}\left(F_t - F_1 - (F_T - F_1)\frac{t}{T}\right) + O_p\left(C_{NT}^{-1}\right) \\ &= H\ T^{-1/2}F_t^d + O_p\left(C_{NT}^{-1}\right), \end{aligned}$$

where F_t^d denotes the detrended common factor, which is obtained as the residual of a regression on a constant and a time trend. Therefore, DF statistic given by (19) converges to

$$DF_{\tilde{F}}^{d} \Rightarrow \frac{\int_0^1 W_w^d(r)\, dW(r)}{\left(\int_0^1 W_w^d(r)^2\, dr\right)^{1/2}},$$

where, as before, $W_w^d(r)$ denotes the detrended Brownian motion and $\tilde{\sigma}_w^2 \xrightarrow{p} H^2\ \sigma_w^2$. The ADF statistic has the same limiting distribution provided that the order of the autoregressive correction is selected such that $k \to \infty$ and $k^3/\min[N,T] \to 0$.

A.4 Pedroni Test Statistic with time trend and slope shift(s)

The model is given by the following deterministic specification

$$f_i(t) = \mu_i + \beta_i t + \theta_i DU_{i,t} + \gamma_i DT_{i,t}^{*},$$

which implies that $\Delta f_i(t) = \beta_i + \theta_i D\left(T_b^i\right)_t + \gamma_i DU_{i,t}$ and $\Delta x_{i,t}^d = \left(1, D\left(T_b^i\right)_t, DU_{i,t}, \Delta x_{i,t}'\right)$. In order to simplify the steps of the proof, we deal with the equivalent specification that does not include the impulse dummy, *i.e.* $\Delta x_{i,t}^d = \left(1, DU_{i,t}, \Delta x_{i,t}'\right)$. This simplifies derivations, although it does not imply loss of generality. Moreover, note that the subspace spanned by $\left(1, DU_{i,t}, \Delta x_{i,t}'\right)$ is equivalent to the one spanned by $\left(DU_{i,t}^1, DU_{i,t}^2, \Delta x_{i,t}'\right)$ where $DU_{i,t}^1 = 1$ for $t \leq T_b$ and 0 otherwise, and $DU_{i,t}^2 = 1$ for $t > T_b$ and 0 otherwise. This redefinition makes $DU_{i,t}^1$ and $DU_{i,t}^2$ to be orthogonal. Note that as before the first element of (17) converges to

$$T^{-1/2}\sum_{j=2}^{t}\Delta\tilde{e}_{i,j} \Rightarrow \sigma W(r).$$

The limiting expression of the second element in (17) has to be derived in several steps. First, note that $T^{-1}\Delta x_i^{d\prime}\Delta x_i^d$ converges to variance and covariance matrix of Δx_i^d, so that all these elements are $O_p(1)$. The first element of the vector $T^{-1}\Delta x_i^{d\prime}\Delta\tilde{e}_i$ is given by $T^{-1/2}\left(T^{-1/2}\sum_{t=1}^{T_b}\Delta\tilde{e}_{i,t}\right) = T^{-1/2}\left(T^{-1/2}\left(\tilde{e}_{i,T_b} - \tilde{e}_{i,1}\right)\right)$, where $T^{-1/2}\left(\tilde{e}_{i,T_b} - \tilde{e}_{i,1}\right) \Rightarrow \sigma W(\lambda)$ since $T^{-1/2}\tilde{e}_{i,1} \to 0$. The second element is $T^{-1/2}\left(T^{-1/2}\sum_{t=T_b+1}^{T}\Delta\tilde{e}_{i,t}\right) = T^{-1/2}\left(T^{-1/2}\left(\tilde{e}_{i,T} - \tilde{e}_{i,T_b}\right)\right)$, where $T^{-1/2}\left(\tilde{e}_{i,T} - \tilde{e}_{i,T_b}\right) \Rightarrow \sigma W(1) - \sigma W(\lambda)$. Note that as before the extra rescaling term $T^{-1/2}$ would be used below. Finally, the third set of elements in the product is $T^{-1}\Delta x_i'\Delta\tilde{e}_i$ that converges to zero since we have assumed independency. Therefore,

$$\left(\Delta x_i^{d\prime}\Delta x_i^d\right)^{-1}\Delta x_i^{d\prime}\Delta\tilde{e}_i = \begin{bmatrix} E\,T^{-1/2}\left(T^{-1/2}\left(\tilde{e}_{i,T_b} - \tilde{e}_{i,1}\right), T^{-1/2}\left(\tilde{e}_{i,T} - \tilde{e}_{i,T_b}\right)\right)' + o_p(1) \\ \left(-D^{-1}CE\right)\,T^{-1/2}\left(T^{-1/2}\left(\tilde{e}_{i,T_b} - \tilde{e}_{i,1}\right), T^{-1/2}\left(\tilde{e}_{i,T} - \tilde{e}_{i,T_b}\right)\right)' + o_p(1) \end{bmatrix}$$

where $E = \left(A - BD^{-1}C\right)^{-1}$ and $A = diag(\lambda, 1-\lambda), B = T^{-1}\left[DU_i^1, DU_i^2\right]'\Delta x_i, C = B'$ and $D = T^{-1}\Delta x_i'\Delta x_i$ denote the elements of the partitioned matrix $T^{-1}\Delta x_i^{d\prime}\Delta x_i^d$. Moreover, following the steps given above $\left(A - BD^{-1}C\right)^{-1} = A^{-1} + o_p(1)$, since $B \to_p 0$. The partial sum process of $\Delta x_{i,t}^d$ for $t \leq T_b$ is

$$T^{-1/2}\sum_{j=2}^{t}\Delta x_{i,j}^d = \left[T^{-1/2}t \quad 0 \quad T^{-1/2}\left(x_{i,t} - x_{i,1}\right)'\right],$$

while for $t > T_b$ is

$$T^{-1/2}\sum_{j=2}^{t}\Delta x_{i,j}^d = \left[T^{-1/2}T_b \quad T^{-1/2}\left(t - T_b\right) \quad T^{-1/2}\left(x_{i,t} - x_{i,1}\right)'\right],$$

so that for $t \leq T_b$

$$T^{-1/2}\sum_{j=2}^{t}\Delta x_{i,j}^{d}\left(\Delta x_{i}^{d\prime}\Delta x_{i}^{d}\right)^{-1}\Delta x_{i}^{d\prime}\Delta\tilde{e}_{i} = \frac{t}{T}\frac{1}{\lambda}\left(T^{-1/2}\left(\tilde{e}_{i,T_b}-\tilde{e}_{i,1}\right)\right)+o_p\left(1\right) \Rightarrow \frac{r}{\lambda}\sigma W\left(\lambda\right),$$

since $T^{-1}\left(x_{i,t}-x_{i,1}\right)' = o_p\left(1\right)$. Therefore, for $t \leq T_b$

$$T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} = T^{-1/2}\tilde{e}_{i,t}-\frac{t}{T}T^{-1/2}\tilde{e}_{i,T}+O_p\left(C_{NT}^{-1}\right) \Rightarrow \sigma\left(W\left(r\right)-\frac{r}{\lambda}W\left(\lambda\right)\right),$$

since from Bai and Ng (2004), the terms $T^{-1/2}\left\|\sum_{j=2}^{t}v_j\right\| = O_p\left(C_{NT}^{-1}\right)$, $\|d_i\| = O_p\left(C_{NT}^{-1}\right)$ and $T^{-1/2}\left\|\sum_{j=2}^{t}\tilde{f}_j\right\| = O_p\left(1\right)$. Note that we can define $b_1 = r/\lambda$ so that $0 < b_1 < 1$, which in turn implies that

$$T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} \Rightarrow \sigma\sqrt{\lambda}W\left(b_1\right)-\sigma b_1\sqrt{\lambda}W\left(1\right) = \sigma\sqrt{\lambda}\left(W\left(b_1\right)-b_1W\left(1\right)\right)\equiv\sigma\sqrt{\lambda}V\left(b_1\right),$$

given the properties of Brownian motions. On the other hand, for $t > T_b$

$$T^{-1/2}\sum_{j=2}^{t}\Delta x_{i,j}^{d}\left(\Delta x_{i}^{d\prime}\Delta x_{i}^{d}\right)^{-1}\Delta x_{i}^{d\prime}\Delta\tilde{e}_{i} = \frac{T_b}{T}\frac{1}{\lambda}\left(T^{-1/2}\left(\tilde{e}_{i,T_b}-\tilde{e}_{i,1}\right)\right) +\frac{t-T_b}{T}\frac{1}{1-\lambda}\left(T^{-1/2}\left(\tilde{e}_{i,T}-\tilde{e}_{i,T_b}\right)\right)+o_p\left(1\right) \Rightarrow \sigma\left(W\left(\lambda\right)+\frac{r-\lambda}{1-\lambda}\left(W\left(1\right)-W\left(\lambda\right)\right)\right),$$

so that

$$T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} = T^{-1/2}\tilde{e}_{i,t}-\frac{t}{T}T^{-1/2}\tilde{e}_{i,T}+O_p\left(C_{NT}^{-1}\right) \Rightarrow \sigma\left(W\left(r\right)-W\left(\lambda\right)-\frac{r-\lambda}{1-\lambda}\left(W\left(1\right)-W\left(\lambda\right)\right)\right).$$

As before, we can define $b_2 = \left(r-\lambda\right)/\left(1-\lambda\right)$ so that $0 < b_2 < 1$, which in turn implies that

$$T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j}\Rightarrow\sigma\sqrt{1-\lambda}\left(W\left(b_2\right)-b_2W\left(1\right)\right)\equiv\sigma\sqrt{1-\lambda}V\left(b_2\right).$$

Using similar developments as in the previous proof, the numerator of the DF statistic converges to $T^{-1}\sum_{t=2}^{T}\tilde{e}_{i,t-1}\Delta\tilde{e}_{i,t}\rightarrow_p -\sigma^2/2$, while the denominator is

$$\begin{aligned} T^{-2}\sum_{t=2}^{T}\tilde{e}_{i,t-1}^2 &= T^{-2}\sum_{t=2}^{T_b+1}\tilde{e}_{i,t-1}^2 + T^{-2}\sum_{t=T_b+2}^{T}\tilde{e}_{i,t-1}^2 \\ &\Rightarrow \sigma^2\left(\lambda^2\int_0^1 V(b_1)^2\,db_1 + (1-\lambda)^2\int_0^1 V(b_2)^2\,db_2\right), \end{aligned}$$

with $V(b_1)$ and $V(b_2)$ two independent Brownian bridges. Therefore, the limiting distribution of the DF statistic is

$$DF_{\tilde{e}}^{\tau}(i)\Rightarrow -\frac{1}{2}\left(\lambda^2\int_0^1 V(b_1)^2\,db_1 + (1-\lambda)^2\int_0^1 V(b_2)^2\,db_2\right)^{-1/2}.$$

It can be shown that this limiting distribution is symmetric around $\lambda = 0.5$ since in this case we can interchange λ^2 and $(1-\lambda)^2$ and obtain the same distribution. Furthermore, this result can be extended to multiple slope shifts, since it is straightforward to see that $T^{-2}\sum_{t=2}^{T}\tilde{e}_{i,t-1}^2$ can be split in the different subperiods that define the l multiple structural changes, so that

$$DF_{\tilde{e}}^{\tau}(i)\Rightarrow -\frac{1}{2}\left(\sum_{j=1}^{l+1}(\lambda_j-\lambda_{j-1})^2\int_0^1 V(b_j)^2\,dr\right)^{-1/2},$$

where l denotes the number of structural breaks, $V(b_j) = W(b_j) - b_j W(1)$, with $b_j = (r-\lambda_{j-1})/(\lambda_j-\lambda_{j-1})$ so that $0 < b_j < 1$, and $T_{j-1}^b < t \leq T_j^b$ with $\lambda_j = T_j^b/T$, $\lambda_0 = 0$ and $\lambda_{l+1} = 1$. As before, the same limiting distribution is found for the ADF statistic.

The limiting distribution of the ADF statistic when there is one common factor is affected by the presence of slope shifts. We can distinguish three different elements in $T^{-1}\Delta x_i^{d\prime}\Delta\tilde{F}$. As in the case of the time trend, the first element is $T^{-1}\sum_{t=2}^{T}\Delta\tilde{F}_t = H\,T^{-1}(F_T - F_1) = O_p\left(T^{-1/2}\right)$. The second element is given by $T^{-1}\sum_{t=T_b+1}^{T}\Delta\tilde{F}_t = H\,T^{-1}(F_T - F_{T_b}) = O_p\left(T^{-1/2}\right)$. Finally, the third set of elements is $T^{-1}\Delta x_i'\Delta\tilde{F} = o_p(1)$ by assumption. Following similar steps as in the case of the time trend we can see that

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{t}\tilde{f}_j &= H\,T^{-1/2}\left(F_t - F_1 - (F_T - F_1)\frac{t}{T} - (F_T - F_{T_b})\frac{t-T_b}{T}1(t>T_b)\right) + O_p\left(C_{NT}^{-1}\right) \\ &= H\,T^{-1/2}F_t^d + O_p\left(C_{NT}^{-1}\right), \end{aligned}$$

where $1(t>T_b)$ is an indicator function. Now F_t^d is obtained as the residual of a regression on a constant, a time trend and the dummy variable $DT_t^* = (t-T_b)\;1(t>T_b)$. Using these elements it is straightforward to see that the DF statistic given by (19) converges to

$$DF_{\tilde{F}}^{d}(\lambda) \Rightarrow \frac{\int_0^1 W_w^d(r,\lambda)\, dW(r,\lambda)}{\left(\int_0^1 W_w^d(r,\lambda)^2\, dr\right)^{1/2}},$$

where, as before, $W_w^d(r,\lambda)$ denotes the detrended Brownian motion, λ is the break fraction parameter and $\tilde{\sigma}_w^2 \xrightarrow{p} H^2\, \sigma_w^2$. Note that this limiting distribution has been considered in Perron (1989) for the specification denoted as Model C.

Table 1: Empirical power of Pedroni cointegration statistic. The structural change affects the deterministic component

		$\rho_i = 0$				$\rho_i = 0.5$			
		$T = 100$		$T = 250$		$T = 100$		$T = 250$	
λ_i	(θ_i, γ_i)	$N = 20$	$N = 40$	$N = 20$	$N = 40$	$N = 20;$	$N = 40$	$N = 20;$	$N = 40$
0.25	$(0, 0)$	1	1	1	1	1	1	1	1
	$(1, 0)$	1	1	1	1	1	1	1	1
	$(3, 0)$	1	1	1	1	1	1	1	1
	$(5, 0)$	1	1	1	1	1	1	1	1
	$(10, 0)$	1	1	1	1	0.49	0.88	1	1
0.5	$(1, 0)$	1	1	1	1	1	1	1	1
	$(3, 0)$	1	1	1	1	1	1	1	1
	$(5, 0)$	1	1	1	1	0.99	1	1	1
	$(10, 0)$	0.94	1	1	1	0.08	0.09	0.90	0.99
0.75	$(1, 0)$	1	1	1	1	1	1	1	1
	$(3, 0)$	1	1	1	1	1	1	1	1
	$(5, 0)$	1	1	1	1	0.99	1	1	1
	$(10, 0)$	0.83	0.98	1	1	0.01	0.00	0.72	0.94
0.25	$(0, 0)$	1	1	1	1	1	1	1	1
	$(3, 0.5)$	1	1	1	1	1	1	1	1
	$(3, 0.7)$	1	1	1	1	1	1	1	1
	$(3, 1)$	1	1	0.99	1	1	1	0.99	1
0.5	$(3, 0.5)$	0.65	0.89	0.01	0	0.02	0	0	0
	$(3, 0.7)$	0.02	0.01	0	0	0	0	0	0
	$(3, 1)$	0	0	0	0	0	0	0	0
0.75	$(3, 0.5)$	0.34	0.54	0	0	0	0	0	0
	$(3, 0.7)$	0	0	0	0	0	0	0	0
	$(3, 1)$	0	0	0	0	0	0	0	0

DGP: $y_t = \mu_i + \theta_i DU_{i,t} + \xi_i t + \gamma_i DT^*_{i,t} + \alpha_i x_{i,t} + z_{i,t}$; $\Delta x_{i,t} = \varepsilon_{i,t}$ and $z_{i,t} = \rho_i z_{i,t-1} + v_{i,t}$ with $\zeta_{i,t} = (\varepsilon_{i,t}, v_{i,t})' \sim iid\ N(0, I_2)$, $\mu_i = 1$, $\xi_i = 0.3$ and $\alpha_i = 1$. The nominal size is set at the 5% level and 1,000 replications are carried out.

Table 2: Empirical power of Pedroni cointegration statistic. The structural change affects both the deterministic component and the cointegrating vector

			$\rho_i = 0$				$\rho_i = 0.5$			
			N ($T = 100$)		N ($T = 250$)		N ($T = 100$)		N ($T = 250$)	
λ_i	(θ_i, γ_i)	$(\alpha_{i,1}, \alpha_{i,2})$	20	40	20	40	20	40	20	40
0.25	$(0,0)$	$(1,0)$	1	1	1	1	1	1	1	1
	$(0,0)$	$(1,2)$	1	1	1	1	1	1	1	1
	$(0,0)$	$(1,3)$	1	1	1	1	1	1	1	1
	$(0,0)$	$(1,4)$	1	1	1	1	1	1	1	1
	$(0,0)$	$(1,5)$	1	1	1	1	0.99	1	1	1
	$(0,0)$	$(1,10)$	0.99	1	1	1	0.97	1	1	1
0.5	$(0,0)$	$(1,2)$	1	1	1	1	0.98	1	1	1
	$(0,0)$	$(1,3)$	0.98	1	0.99	1	0.50	0.77	0.76	0.94
	$(0,0)$	$(1,4)$	0.71	0.92	0.86	0.99	0.27	0.42	0.42	0.67
	$(0,0)$	$(1,5)$	0.45	0.68	0.62	0.853	0.17	0.31	0.32	0.50
	$(0,0)$	$(1,10)$	0.17	0.30	0.26	0.406	0.13	0.18	0.19	0.31
0.75	$(0,0)$	$(1,2)$	1	1	1	1	0.83	0.97	0.96	1
	$(0,0)$	$(1,3)$	0.76	0.92	0.86	0.98	0.11	0.11	0.20	0.28
	$(0,0)$	$(1,4)$	0.26	0.32	0.33	0.48	0.02	0.01	0.04	0.03
	$(0,0)$	$(1,5)$	0.09	0.10	0.12	0.13	0.01	0.01	0.02	0.01
	$(0,0)$	$(1,10)$	0.01	0	0.01	0	0.01	0	0.01	0
0.25	$(3,0)$	$(1,2)$	1	1	1	1	1	1	1	1
	$(3,0)$	$(1,3)$	1	1	1	1	0.99	1	1	1
	$(3,0)$	$(1,4)$	1	1	1	1	0.99	1	1	1
	$(3,0)$	$(1,5)$	1	1	1	1	0.98	1	1	1
	$(3,0)$	$(1,10)$	0.98	1	1	1	0.97	1	0.99	1
0.5	$(3,0)$	$(1,2)$	1	1	1	1	0.97	1	1	1
	$(3,0)$	$(1,3)$	0.97	1	1	1	0.51	0.74	0.72	0.92
	$(3,0)$	$(1,4)$	0.71	0.92	0.84	0.98	0.23	0.44	0.43	0.69
	$(3,0)$	$(1,5)$	0.44	0.66	0.63	0.88	0.18	0.29	0.29	0.50
	$(3,0)$	$(1,10)$	0.18	0.28	0.26	0.42	0.12	0.18	0.19	0.32
0.75	$(3,0)$	$(1,2)$	1	1	1	1	0.77	0.95	0.96	1
	$(3,0)$	$(1,3)$	0.74	0.91	0.86	0.98	0.11	0.10	0.18	0.26
	$(3,0)$	$(1,4)$	0.22	0.35	0.32	0.47	0.03	0.01	0.04	0.03
	$(3,0)$	$(1,5)$	0.09	0.09	0.10	0.14	0.01	0.00	0.02	0.01
	$(3,0)$	$(1,10)$	0.01	0	0.01	0.01	0	0	0.01	0

DGP: $y_t = \mu_i + \theta_i DU_{i,t} + \xi_i t + \gamma_i DT^*_{i,t} + \alpha_{i,t} x_{i,t} + z_{i,t}$; $\Delta x_{i,t} = \varepsilon_{i,t}$ and $z_{i,t} = \rho_i z_{i,t-1} + v_{i,t}$ with $\zeta_{i,t} = (\varepsilon_{i,t}, v_{i,t})' \sim iid\ N(0, I_2)$, $\mu_i = 1$, $\xi_i = 0.3$ and $\alpha_{i,t} = \alpha_{i,1}$ for $t \leq T_{b,i}$ and $\alpha_{i,t} = \alpha_{i,2}$ for $t > T_{b,i}$. The nominal size is set at the 5% level and 1,000 replications are carried out.

Table 3: Empirical power of Pedroni cointegration statistic. The structural change affects both the deterministic component and the cointegrating vector

			$\rho_i = 0$				$\rho_i = 0.5$			
			N $(T = 100)$		N $(T = 250)$		N $(T = 100)$		N $(T = 250)$	
λ_i	(θ_i, γ_i)	$(\alpha_{i,1}, \alpha_{i,2})$	20	40	20	40	20	40	20	40
0.25	$(3, 0.5)$	$(1, 2)$	1	1	1	1	0.99	1	1	1
	$(3, 0.5)$	$(1, 3)$	1	1	1	1	0.99	1	0.98	1
	$(3, 0.5)$	$(1, 4)$	1	1	1	1	0.96	1	0.95	1
	$(3, 0.5)$	$(1, 5)$	0.98	1	0.98	1	0.92	1	0.95	1
	$(3, 0.5)$	$(1, 10)$	0.85	0.98	0.95	1	0.88	0.98	0.93	1
0.5	$(3, 0.5)$	$(1, 2)$	0.43	0.72	0	0	0.01	0.01	0	0
	$(3, 0.5)$	$(1, 3)$	0.36	0.53	0.01	0	0.05	0.04	0	0
	$(3, 0.5)$	$(1, 4)$	0.28	0.41	0.03	0.01	0.08	0.09	0.01	0
	$(3, 0.5)$	$(1, 5)$	0.23	0.30	0.05	0.04	0.08	0.10	0.01	0.01
	$(3, 0.5)$	$(1, 10)$	0.14	0.21	0.08	0.13	0.12	0.19	0.09	0.10
0.75	$(3, 0.5)$	$(1, 2)$	0.71	0.89	0.04	0.02	0.04	0.02	0	0
	$(3, 0.5)$	$(1, 3)$	0.52	0.68	0.11	0.08	0.08	0.08	0.01	0
	$(3, 0.5)$	$(1, 4)$	0.28	0.34	0.09	0.08	0.08	0.05	0.01	0
	$(3, 0.5)$	$(1, 5)$	0.15	0.16	0.06	0.04	0.05	0.05	0.01	0.01
	$(3, 0.5)$	$(1, 10)$	0.04	0.03	0.03	0.01	0.05	0.03	0.03	0.01

DGP: $y_t = \mu_i + \theta_i DU_{i,t} + \xi_i t + \gamma_i DT^*_{i,t} + \alpha_{i,t} x_{i,t} + z_{i,t}$; $\Delta x_{i,t} = \varepsilon_{i,t}$ and $z_{i,t} = \rho_i z_{i,t-1} + v_{i,t}$ with $\zeta_{i,t} = (\varepsilon_{i,t}, v_{i,t})' \sim iid\ N(0, I_2)$, $\mu_i = 1$, $\xi_i = 0.3$ and $\alpha_{i,t} = \alpha_{i,1}$ for $t \leq T_{b,i}$ and $\alpha_{i,t} = \alpha_{i,2}$ for $t > T_{b,i}$. The nominal size is set at the 5% level and 1,000 replications are carried out.

Table 4: Asymptotic moments for the test statistics

	Model 1				Model 2				Model 3			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
$m-1$	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
1	-25.124	73.605	-3.558	0.388	-31.702	80.102	-4.003	0.341	-36.102	98.290	-4.276	0.366
2	-30.807	89.178	-3.943	0.392	-37.262	97.782	-4.343	0.355	-41.353	113.560	-4.581	0.374
3	-36.241	99.942	-4.285	0.373	-42.352	112.792	-4.637	0.369	-46.254	124.446	-4.853	0.364
4	-41.323	113.847	-4.580	0.373	-47.420	127.582	-4.912	0.368	-51.393	136.173	-5.124	0.364
5	-46.457	121.902	-4.865	0.365	-51.847	136.375	-5.145	0.362	-56.221	148.416	-5.366	0.366
6	-51.609	142.541	-5.131	0.384	-56.491	152.524	-5.375	0.378	-60.893	159.531	-5.593	0.365
7	-56.732	151.879	-5.389	0.372	-61.259	163.744	-5.606	0.375	-65.777	172.601	-5.820	0.369

	Model 4				Model 5				Model 6			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
$m-1$	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
1	-28.682	91.014	-3.798	0.431	-36.915	106.592	-4.324	0.393	-45.094	139.700	-4.783	0.418
2	-38.757	123.284	-4.427	0.436	-45.797	136.480	-4.821	0.408	-58.158	175.030	-5.453	0.415
3	-48.118	149.200	-4.944	0.431	-54.411	161.488	-5.271	0.415	-70.768	217.036	-6.037	0.432
4	-56.713	173.081	-5.380	0.430	-63.063	184.648	-5.687	0.410	-83.254	256.429	-6.573	0.441
5	-65.513	206.886	-5.798	0.447	-71.671	210.886	-6.081	0.417	-95.459	284.133	-7.065	0.435
6	-73.589	221.307	-6.163	0.427	-79.723	240.506	-6.425	0.434	-106.892	318.951	-7.498	0.443
7	-81.754	240.575	-6.513	0.423	-88.079	251.068	-6.771	0.417	-118.597	357.847	-7.923	0.455

Table 5: Response surfaces for $(k = 0)$

	Model 1				Model 2			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
$\hat{\beta}_{0,0}$	0.39	60.648	-3.127	-19.196	0.339	67.8	-3.684	-26.679
$\hat{\beta}_{0,1}$	5.064	-1226.67	-8.833	121.763	6.104	-1885.589	-9.439	144.172
$\hat{\beta}_{0,2}$			179.334	-2571.386		16698.79	28.308	3575.522
$\hat{\beta}_{0,3}$		196196.7	-1990.403	58983.27	1029.447			-72734.42
$\hat{\beta}_{1,0}$	-0.005	16.530	-0.429	-6.238	0.003	17.645	-0.341	-5.625
$\hat{\beta}_{1,1}$	1	-1325.654		124.468	0.902	-1543.665		180.54
$\hat{\beta}_{1,2}$	34.590	42679.5	-60.807	-1312.53	39.629	53149.58	-51.393	-4444.318
$\hat{\beta}_{1,3}$		-532567.3				-663605.3		48906.87
$\hat{\beta}_{2,0}$		-0.362	0.016	0.112		-0.39	0.01	0.067
$\hat{\beta}_{2,1}$			-0.236	3.084		6.859	-0.228	1.208
$\hat{\beta}_{2,2}$		225.078	5.935	-51.736			4.325	
$\hat{\beta}_{2,3}$								

	Model 3				Model 4			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
$\hat{\beta}_{0,0}$	0.359	91.108	-3.971	-31.767	0.43	60.884	-3.221	-19.845
$\hat{\beta}_{0,1}$	7.472	-3645.426	-8.979	442.209	3.046			318.553
$\hat{\beta}_{0,2}$	59.681	75512.06	-49.326	-10829.74	102.433	-87968.11	-110.06	-16385.62
$\hat{\beta}_{0,3}$		-777252.4		164392.7		1874059		307418.8
$\hat{\beta}_{1,0}$		14.514	-0.314	-5.334		35.776	-0.628	-10.047
$\hat{\beta}_{1,1}$	0.852	-1361.209	-2.06	124.516	3.307	-3225.963	-2.236	219.694
$\hat{\beta}_{1,2}$	42.03	47092.270		-1139.025		121345.6		-1980.416
$\hat{\beta}_{1,3}$		-562391.3				-1725484		
$\hat{\beta}_{2,0}$		-0.216	0.008	0.039	0.001	-1.033	0.023	0.136
$\hat{\beta}_{2,1}$			0.038	5.521	-0.165		-0.188	12.356
$\hat{\beta}_{2,2}$			-3.393	-128.867	11.955	797.530	-3.325	-290.951
$\hat{\beta}_{2,3}$								

	Model 5				Model 6			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
$\hat{\beta}_{0,0}$	0.364	74.286	-3.78	-27.851	0.366	87.342	-3.968	-30.483
$\hat{\beta}_{0,1}$	6.564	-2146.293	-6.974	242.942	10.855	-2699.1	-8.23	191.322
$\hat{\beta}_{0,2}$		20055.64			-266.021	28358.7	-83.457	4931.966
$\hat{\beta}_{0,3}$				-42063.93	7384.621			-123591.5
$\hat{\beta}_{1,0}$	0.008	34.679	-0.544	-9.615	0.007	33.827	-0.505	-9.373
$\hat{\beta}_{1,1}$	2.617	-3212.648	-0.868	322.608	3.982	-3213.574	-1.767	357.392
$\hat{\beta}_{1,2}$	41.638	115262.4	-43.717	-8330.38		118816.1		-9875.101
$\hat{\beta}_{1,3}$		-1488387		113929.5		-1614095		131909.7
$\hat{\beta}_{2,0}$		-1.053	0.018	0.097		-0.888	0.014	0.072
$\hat{\beta}_{2,1}$	-0.161	24.408	-0.306	11.106	-0.325		-0.189	9.476
$\hat{\beta}_{2,2}$	10.166			-273.637	15.916	730.025	-5.392	-222.194
$\hat{\beta}_{2,3}$								

Table 6: Response surfaces for $(k=2)$

	Model 1				Model 2			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
$\hat{\beta}_{0,0}$	0.415	62.309	-3.213	-19.672	0.336	69.482	-3.735	-26.724
$\hat{\beta}_{0,1}$	0.967	-104.685	4.102	91.646	2.873		1.953	-42.725
$\hat{\beta}_{0,2}$	85.478		-428.601	-6704.974	58.52	-18212.26	-286.032	421.283
$\hat{\beta}_{0,3}$			6757.605	103243.4		275990.6	5567.945	
$\hat{\beta}_{1,0}$	-0.018	15.196	-0.414	-5.961	0.005	15.915	-0.334	-5.411
$\hat{\beta}_{1,1}$	1.579	-172.849	4.4	17.452	1.368	-236.01	6.006	56.212
$\hat{\beta}_{1,2}$			-26.560	438.162	-30.879		-59.458	-245.555
$\hat{\beta}_{1,3}$					612.861			
$\hat{\beta}_{2,0}$	0.002	-0.147	0.015	0.08	-0.001	-0.195	0.010	0.05
$\hat{\beta}_{2,1}$	-0.085	-9.382		4.173			-0.138	0.614
$\hat{\beta}_{2,2}$			-2.239	-71.787				
$\hat{\beta}_{2,3}$								
	Model 3				Model 4			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
$\hat{\beta}_{0,0}$	0.353	89.831	-4.011	-31.141	0.429	66.591	-3.235	-19.246
$\hat{\beta}_{0,1}$	6.456	-173.345	5.695	25.550	1.626	-1025.367		66.531
$\hat{\beta}_{0,2}$		-6455.393	-543.11	-4224.155	100.548	30787.53	-76.879	-6527.479
$\hat{\beta}_{0,3}$			8627.961	84886.75				120101.1
$\hat{\beta}_{1,0}$	0.006	14.775	-0.317	-5.476	-0.002	29.482	-0.624	-9.880
$\hat{\beta}_{1,1}$	-1.009	-274.989	5.92	63.485	2.983	438.987	13.891	135.547
$\hat{\beta}_{1,2}$	81.566		-53.692	-245.299	-45.199	-24349.6	-374.707	-2851.772
$\hat{\beta}_{1,3}$	-631.881						4741.349	32282.1
$\hat{\beta}_{2,0}$	-0.001	-0.155	0.009	0.054			0.024	0.104
$\hat{\beta}_{2,1}$	0.181		-0.147		-0.199	-138.416	-0.304	2.734
$\hat{\beta}_{2,2}$	-5.953				6.356	3434.739		
$\hat{\beta}_{2,3}$								
	Model 5				Model 6			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
$\hat{\beta}_{0,0}$	0.380	78.16	-3.825	-27.922	0.383	91.354	-4.016	-31.322
$\hat{\beta}_{0,1}$		-1049.361	4.26	97.299	2.626		6.241	156.004
$\hat{\beta}_{0,2}$	94.123	11495.16	-98.231	-5411.362	90.144	-40668.1	-493.482	-11876.56
$\hat{\beta}_{0,3}$				88658.21		521643.9	7199.83	218847.8
$\hat{\beta}_{1,0}$	0.004	29.349	-0.524	-9.171	0.011	29.639	-0.488	-8.640
$\hat{\beta}_{1,1}$	2.825		14.206	143.512	2.271	-281.767	13.469	106.736
$\hat{\beta}_{1,2}$		-7443.609	-434.678	-3425.206		4060.874	-326.613	-744.496
$\hat{\beta}_{1,3}$			5633.642	49471.7			3497.908	
$\hat{\beta}_{2,0}$			0.017	0.047	-0.001		0.014	
$\hat{\beta}_{2,1}$	-0.136	-86.63	-0.236	5.337	-0.121	-58.02	-0.272	6.34
$\hat{\beta}_{2,2}$		1363.246		-74.929	1.486			-89.481
$\hat{\beta}_{2,3}$								

Table 7: Response surfaces for $(k = 5)$

	Model 1				Model 2			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
$\hat{\beta}_{0,0}$	0.411	61.076	-3.196	-19.09	0.327	70.537	-3.758	-26.36
$\hat{\beta}_{0,1}$		2333.282	-2.138	-251.033	1.926	3688.82	9.947	-577.95
$\hat{\beta}_{0,2}$	89.8	14804.35		-2084.949		-111525	-102.888	8408.465
$\hat{\beta}_{0,3}$	2785.821		-2085.401		5474.382	2935591		-70550.62
$\hat{\beta}_{1,0}$	-0.018	14.491	-0.419	-5.96	0.008	14.356	-0.324	-5.371
$\hat{\beta}_{1,1}$	1.282	1468.171	15.196	-102.32	0.596	1834.07	14.124	-70.904
$\hat{\beta}_{1,2}$		-14669.95	-348.385	2139.019		-25876.53	-368.115	1714.215
$\hat{\beta}_{1,3}$			4192.348				4228.759	
$\hat{\beta}_{2,0}$	0.001		0.016	0.068	-0.001		0.010	0.029
$\hat{\beta}_{2,1}$	-0.069		-0.289	1.362			-0.172	
$\hat{\beta}_{2,2}$								
$\hat{\beta}_{2,3}$								
	Model 3				Model 4			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
$\hat{\beta}_{0,0}$	0.367	89.609	-4.013	-30.645	0.435	58.969	-3.269	-19.333
$\hat{\beta}_{0,1}$		6021.446	10.749	-722.651		3904.387	8.328	-180.318
$\hat{\beta}_{0,2}$	139.06	-119796.1	-322.281	10502.56		-154668.9	-527.974	-3147.907
$\hat{\beta}_{0,3}$	4467.837	2779171		-173970.8	4215.836	3727902	6371.796	
$\hat{\beta}_{1,0}$	-0.004	13.944	-0.307	-5.325	-0.003	27.297	-0.59	-9.213
$\hat{\beta}_{1,1}$	1.052	1272.166	15.51	-75.522	1.582	3537.474	24.265	-136.207
$\hat{\beta}_{1,2}$	-11.117		-394.685	1780.818	9.137	-47364.92	-666.318	2432.973
$\hat{\beta}_{1,3}$			4719.719				8318.551	
$\hat{\beta}_{2,0}$			0.008	0.014		0.826	0.023	
$\hat{\beta}_{2,1}$		70.693	-0.27		-0.092		-0.479	
$\hat{\beta}_{2,2}$		-2726.643						81.938
$\hat{\beta}_{2,3}$								
	Model 5				Model 6			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
$\hat{\beta}_{0,0}$	0.343	60.513	-3.828	-27.858	0.378	71.175	-4.026	-31.256
$\hat{\beta}_{0,1}$	1.636	6899.273	12.998	-253.539	0.867	11057.17	13.951	-408.183
$\hat{\beta}_{0,2}$	-88.332	-322583.2	-182.882	-13556.49		-499199.5	-403.515	-10959.25
$\hat{\beta}_{0,3}$	7587.446	5934887		359446.7	6664.312	10140177		228140.6
$\hat{\beta}_{1,0}$	0.016	32.949	-0.496	-8.767	0.011	29.816	-0.451	-8.12
$\hat{\beta}_{1,1}$	1.581	2404.486	25.834	-154.416	1.509	3570.033	24.749	-194.826
$\hat{\beta}_{1,2}$			-842.643	6727.948		-40673.95	-789.654	8114.972
$\hat{\beta}_{1,3}$			9999.87	-124097.4			10535.88	-121465.5
$\hat{\beta}_{2,0}$	-0.001		0.017	-5.136	-0.001	0.754	0.012	-0.078
$\hat{\beta}_{2,1}$	-0.083	187.542	-0.574	247.073	-0.074		-0.331	
$\hat{\beta}_{2,2}$		-7067.742	9.204					120.421
$\hat{\beta}_{2,3}$								

Table 8: Response surfaces for the automatic lag length selection method ($k_{max} = 5$)

	Model 1				Model 2			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
$\hat{\beta}_{0,0}$	0.41	56.823	-3.218	-19.638	0.372	71.034	-3.778	-26.654
$\hat{\beta}_{0,1}$	10.777	2079.863	-34.87	-97.193	1.676	1730.194	-42.359	-392.725
$\hat{\beta}_{0,2}$	-284.429		737.622	-3103.602	97.645	40207.55	1018.228	4124.832
$\hat{\beta}_{0,3}$	4332.145		-11377.84				-13147.4	
$\hat{\beta}_{1,0}$	-0.004	18.14	-0.442	-6.027	0.005	13.145	-0.351	-5.628
$\hat{\beta}_{1,1}$	-2.036		1.628	-68.79		1293.969	3.225	
$\hat{\beta}_{1,2}$	55.887	28710.63		1511.876		-18644.32	-36.265	
$\hat{\beta}_{1,3}$								
$\hat{\beta}_{2,0}$		-0.748	0.017	0.081	-0.001		0.01	0.064
$\hat{\beta}_{2,1}$	0.165	205.976	-0.114	0.967		48.061	-0.161	-5.218
$\hat{\beta}_{2,2}$		-5962.404			5.98			140.98
$\hat{\beta}_{2,3}$								

	Model 3				Model 4			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
$\hat{\beta}_{0,0}$	0.389	72.251	-4.061	-31.033	0.517	70.453	-3.286	-19.519
$\hat{\beta}_{0,1}$	5.779	7427.681	-43.941	-465.591	1.919		-26.176	-101.832
$\hat{\beta}_{0,2}$	-225.895	-177465.4	921.364	1330.639		44801.21	166.728	-2334.409
$\hat{\beta}_{0,3}$	5584.734	2808044	-13082.02					
$\hat{\beta}_{1,0}$		19.721	-0.335	-5.637	-0.02	26.003	-0.649	-9.78
$\hat{\beta}_{1,1}$	-0.798		3.616			2162.096	3.806	-26.883
$\hat{\beta}_{1,2}$	56.865	37740.3	-30.174		72.559		-45.143	
$\hat{\beta}_{1,3}$								
$\hat{\beta}_{2,0}$	0.001	-0.737	0.009	0.059	0.001		0.025	0.086
$\hat{\beta}_{2,1}$	0.093	190.91	-0.194	-6.067	0.176	275.749	-0.227	-8.473
$\hat{\beta}_{2,2}$		-5491.499		146.45		-8513.873		292.56
$\hat{\beta}_{2,3}$								

	Model 5				Model 6			
	$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	
	Θ_1	Ψ_1	Θ_2	Ψ_2	Θ_1	Ψ_1	Θ_2	Ψ_2
$\hat{\beta}_{0,0}$	0.399	109.977	-3.875	-27.694	0.424	87.103	-4.071	-31.407
$\hat{\beta}_{0,1}$		-8193.521	-35.047	-296.345		5713.206	-41.846	-243.518
$\hat{\beta}_{0,2}$	119.632	607421	681.665	1996.116	147.021	-286832.2	938.021	-12550.07
$\hat{\beta}_{0,3}$		-9915995	-8374.721			7973939	-14997.98	240521.8
$\hat{\beta}_{1,0}$	0.011	7.772	-0.549	-9.262	0.011	19.269	-0.509	-8.675
$\hat{\beta}_{1,1}$	0.937	6841.871	4.83	-9.079	0.858	4385.407	5.806	-66.601
$\hat{\beta}_{1,2}$		-256154.3	-74.577	-540.416		-58727.81	-111.692	3590.481
$\hat{\beta}_{1,3}$		3915350					1563.76	-69517.66
$\hat{\beta}_{2,0}$	-0.001	1.661	0.018	0.048	-0.001	1.239	0.014	
$\hat{\beta}_{2,1}$			-0.235	-8.318			-0.178	-7.49
$\hat{\beta}_{2,2}$	13.846			283.209	15.245		-3.999	271.886
$\hat{\beta}_{2,3}$								

Table 9: Asymptotic critical values for the MQ tests

	$\lambda = 0.1$			$\lambda = 0.2$			$\lambda = 0.3$		
r	1%	5%	10%	1%	5%	10%	1%	5%	10%
1	-32.163	-23.629	-19.865	-34.858	-26.091	-22.144	-36.123	-27.562	-23.619
2	-43.372	-34.321	-30.056	-46.436	-37.139	-32.688	-46.773	-37.778	-33.492
3	-53.648	-44.378	-39.748	-55.828	-46.232	-41.766	-57.136	-47.511	-42.775
4	-63.359	-53.470	-48.595	-65.206	-55.582	-50.645	-65.570	-55.883	-51.370
5	-73.691	-62.796	-57.434	-74.601	-64.165	-59.199	-75.573	-64.731	-59.919
6	-81.346	-71.238	-65.663	-83.575	-72.562	-67.309	-83.921	-73.247	-67.908
	$\lambda = 0.4$			$\lambda = 0.5$			$\lambda = 0.6$		
r	1%	5%	10%	1%	5%	10%	1%	5%	10%
1	-36.635	-28.147	-24.140	-36.775	-28.226	-24.419	-36.805	-28.178	-24.176
2	-47.134	-38.391	-34.282	-48.148	-38.907	-34.553	-47.611	-38.587	-34.246
3	-57.176	-47.642	-43.088	-56.753	-47.715	-43.333	-57.230	-47.865	-43.200
4	-67.481	-56.958	-52.039	-65.752	-56.418	-51.708	-67.094	-56.599	-51.785
5	-75.603	-65.386	-60.204	-75.378	-65.302	-60.251	-75.182	-64.986	-60.057
6	-84.718	-73.703	-68.372	-83.902	-73.746	-68.222	-84.059	-73.136	-67.973
	$\lambda = 0.7$			$\lambda = 0.8$			$\lambda = 0.9$		
r	1%	5%	10%	1%	5%	10%	1%	5%	10%
1	-36.302	-27.751	-23.890	-35.249	-26.722	-22.713	-32.918	-24.712	-20.896
2	-47.383	-38.223	-34.045	-46.572	-37.227	-33.085	-43.959	-35.248	-31.190
3	-56.908	-47.282	-42.693	-55.960	-46.442	-41.998	-54.568	-45.183	-40.623
4	-66.869	-56.270	-51.337	-65.833	-55.750	-50.890	-63.920	-53.985	-49.399
5	-75.074	-64.828	-59.867	-74.046	-64.430	-59.290	-74.177	-63.063	-57.839
6	-85.434	-73.646	-68.332	-83.244	-72.857	-67.721	-82.664	-71.518	-66.449

Table 10: Empirical size of the tests (nominal size = 5%)

		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ statistic					
N	T	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
20	50	0.039	0.046	0.043	0.033	0.054	0.045
	100	0.055	0.049	0.053	0.059	0.048	0.050
	250	0.050	0.053	0.046	0.052	0.056	0.059
40	50	0.040	0.049	0.046	0.030	0.044	0.056
	100	0.047	0.047	0.057	0.066	0.051	0.047
	250	0.056	0.061	0.047	0.044	0.046	0.055
		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ statistic					
N	T	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
20	50	0.044	0.045	0.049	0.047	0.050	0.045
	100	0.050	0.050	0.045	0.046	0.043	0.053
	250	0.043	0.047	0.043	0.040	0.049	0.053
40	50	0.045	0.051	0.055	0.048	0.041	0.052
	100	0.041	0.047	0.047	0.044	0.046	0.043
	250	0.048	0.053	0.046	0.032	0.045	0.048

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 11: Empirical power of the normalised bias and pseudo t-ratio statistics for $\lambda_i = 0.5$ (nominal size = 5%)

		$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ statistic					
N	T	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
20	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	250	1	1	1	1	1	1
40	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	250	1	1	1	1	1	1
		$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ statistic					
N	T	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
20	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	250	1	1	1	1	1	1
40	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	250	1	1	1	1	1	1

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 12: Empirical size and power when there is one common factor ($N = 40$)

			Constant case: $f_i(t) = \mu_i$						Trend case: $f_i(t) = \mu_i + \beta_i t$					
			$\sigma_F^2 = 0.5$		$\sigma_F^2 = 1$		$\sigma_F^2 = 10$		$\sigma_F^2 = 0.5$		$\sigma_F^2 = 1$		$\sigma_F^2 = 10$	
T	ρ_i	ϕ	$Z^e_{\hat{t}_{NT}}$	$ADF^d_{\hat{F}}$	$Z^e_{\hat{t}_{NT}}$	$ADF^d_{\hat{F}}$	$Z^e_{\hat{t}_{NT}}$	$ADF^d_{\hat{F}}$	$Z^e_{\hat{t}_{NT}}$	$ADF^d_{\hat{F}}$	$Z^e_{\hat{t}_{NT}}$	$ADF^d_{\hat{F}}$	$Z^e_{\hat{t}_{NT}}$	$ADF^d_{\hat{F}}$
50	1	1	0.049	0.067	0.039	0.045	0.052	0.065	0.054	0.077	0.054	0.058	0.058	0.062
100	1	1	0.056	0.047	0.054	0.058	0.058	0.048	0.069	0.056	0.062	0.042	0.063	0.052
250	1	1	0.053	0.047	0.050	0.052	0.055	0.055	0.059	0.048	0.048	0.049	0.060	0.052
50	1	0.9	0.039	0.141	0.041	0.110	0.032	0.121	0.053	0.108	0.033	0.116	0.052	0.104
100	1	0.9	0.049	0.288	0.042	0.305	0.053	0.345	0.054	0.227	0.038	0.200	0.053	0.191
250	1	0.9	0.039	0.853	0.047	0.928	0.052	0.959	0.052	0.773	0.050	0.829	0.042	0.838
50	1	0.8	0.043	0.294	0.038	0.366	0.048	0.337	0.042	0.226	0.061	0.212	0.045	0.220
100	1	0.8	0.057	0.763	0.042	0.817	0.041	0.854	0.070	0.611	0.059	0.642	0.052	0.661
250	1	0.8	0.054	0.995	0.046	1	0.054	1	0.053	1	0.043	1	0.044	1
50	0.9	1	1	0.047	1	0.049	1	0.060	0.965	0.068	0.971	0.084	0.875	0.071
100	0.9	1	1	0.054	1	0.061	1	0.048	1	0.071	1	0.053	1	0.061
250	0.9	1	1	0.056	1	0.060	1	0.038	1	0.056	1	0.054	1	0.074
50	0.9	0.9	1	0.145	1	0.139	1	0.145	0.96	0.123	0.945	0.099	0.861	0.093
100	0.9	0.9	1	0.338	1	0.300	1	0.310	1	0.208	1	0.211	1	0.203
250	0.9	0.9	1	0.968	1	0.971	1	0.978	1	0.840	1	0.829	1	0.833
50	0.9	0.8	1	0.367	1	0.342	1	0.338	0.961	0.238	0.948	0.224	0.834	0.231
100	0.9	0.8	1	0.852	1	0.861	1	0.883	1	0.67	1	0.675	1	0.671
250	0.9	0.8	1	1	1	1	1	1	1	1	1	1	1	1
50	0.8	1	1	0.061	1	0.063	1	0.049	1	0.063	1	0.058	1	0.057
100	0.8	1	1	0.058	1	0.067	1	0.050	1	0.053	1	0.061	1	0.058
250	0.8	1	1	0.048	1	0.042	1	0.047	1	0.065	1	0.053	1	0.057
50	0.8	0.9	1	0.126	1	0.128	1	0.127	1	0.100	1	0.107	1	0.115
100	0.8	0.9	1	0.366	1	0.324	1	0.327	1	0.193	1	0.221	1	0.217
250	0.8	0.9	1	0.974	1	0.959	1	0.963	1	0.858	1	0.847	1	0.83
50	0.8	0.8	1	0.386	1	0.337	1	0.352	1	0.227	1	0.246	1	0.232
100	0.8	0.8	1	0.861	1	0.882	1	0.867	1	0.68	1	0.668	1	0.669
250	0.8	0.8	1	1	1	1	1	1	1	1	1	1	1	1

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 13: Empirical size and power. Constant case with three common factors ($N = 40$)

T	ρ_i	α	σ_F^2	$Z^e_{\hat{t}_{NT}}$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	1	1	0.5	0.082	0.006	0.167	0.340	0.484
100	1	1	0.5	0.057	0.003	0.024	0.186	0.784
250	1	1	0.5	0.050	0.001	0.02	0.128	0.848
50	1	0.9	0.5	0.117	0.021	0.139	0.312	0.525
100	1	0.9	0.5	0.061	0.086	0.053	0.206	0.652
250	1	0.9	0.5	0.051	0.771	0.017	0.075	0.134
50	1	0.8	0.5	0.121	0.066	0.090	0.302	0.539
100	1	0.8	0.5	0.051	0.509	0.041	0.122	0.325
250	1	0.8	0.5	0.061	0.986	0.007	0.003	0.001
50	1	1	1	0.061	0	0.003	0.030	0.967
100	1	1	1	0.052	0	0.013	0.063	0.921
250	1	1	1	0.050	0	0.010	0.078	0.909
50	1	0.9	1	0.030	0.001	0.006	0.045	0.945
100	1	0.9	1	0.036	0.093	0.033	0.134	0.737
250	1	0.9	1	0.034	0.844	0.008	0.041	0.104
50	1	0.8	1	0.033	0.039	0.010	0.062	0.886
100	1	0.8	1	0.048	0.56	0.025	0.095	0.317
250	1	0.8	1	0.052	0.994	0.001	0.001	0.001
50	1	1	10	0.060	0	0.002	0.015	0.979
100	1	1	10	0.049	0.001	0.006	0.059	0.931
250	1	1	10	0.060	0.004	0.009	0.084	0.900
50	1	0.9	10	0.044	0.008	0.001	0.027	0.957
100	1	0.9	10	0.053	0.116	0.030	0.133	0.717
250	1	0.9	10	0.042	0.904	0.006	0.022	0.065
50	1	0.8	10	0.030	0.034	0.012	0.059	0.886
100	1	0.8	10	0.049	0.651	0.014	0.076	0.256
250	1	0.8	10	0.043	0.994	0.001	0.001	0.001

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 14: Empirical size and power. Constant case with three common factors ($N = 40$)

T	ρ_i	α	σ^2_F	$Z^e_{\hat{t}_{NT}}$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	ρ_i	$Z^e_{\hat{t}_{NT}}$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	0.9	1	0.5	0.918	0.014	0.224	0.378	0.384	0.8	0.929	0.018	0.299	0.372	0.311
100	0.9	1	0.5	0.988	0.004	0.049	0.249	0.698	0.8	0.989	0.007	0.071	0.282	0.640
250	0.9	1	0.5	0.998	0.002	0.020	0.191	0.787	0.8	0.998	0.003	0.029	0.223	0.745
50	0.9	0.9	0.5	1	0.053	0.168	0.357	0.422	0.8	1	0.058	0.224	0.351	0.367
100	0.9	0.9	0.5	1	0.111	0.082	0.276	0.531	0.8	1	0.134	0.095	0.292	0.479
250	0.9	0.9	0.5	1	0.946	0.008	0.028	0.018	0.8	1	0.954	0.006	0.022	0.018
50	0.9	0.8	0.5	1	0.134	0.121	0.299	0.446	0.8	1	0.155	0.145	0.305	0.395
100	0.9	0.8	0.5	1	0.690	0.031	0.120	0.159	0.8	1	0.742	0.031	0.110	0.117
250	0.9	0.8	0.5	1	1	0	0	0	0.8	1	1	0	0	0
50	0.9	1	1	1	0	0.004	0.056	0.94	0.8	0.996	0	0.004	0.064	0.932
100	0.9	1	1	1	0	0.007	0.098	0.895	0.8	1	0	0.007	0.105	0.888
250	0.9	1	1	1	0.001	0.010	0.127	0.862	0.8	1	0	0.01	0.127	0.863
50	0.9	0.9	1	1	0.004	0.008	0.071	0.917	0.8	1	0.005	0.013	0.083	0.899
100	0.9	0.9	1	1	0.087	0.057	0.226	0.630	0.8	1	0.098	0.057	0.223	0.622
250	0.9	0.9	1	1	0.935	0.007	0.032	0.026	0.8	1	0.943	0.006	0.031	0.020
50	0.9	0.8	1	1	0.036	0.025	0.123	0.816	0.8	1	0.039	0.026	0.119	0.816
100	0.9	0.8	1	1	0.693	0.031	0.109	0.167	0.8	1	0.708	0.033	0.113	0.146
250	0.9	0.8	1	1	1	0	0	0	0.8	1	1	0	0	0
50	0.9	1	10	0.937	0.003	0.002	0.032	0.963	0.8	0.985	0.004	0.002	0.033	0.961
100	0.9	1	10	1	0.001	0.007	0.095	0.897	0.8	1	0.001	0.008	0.098	0.893
250	0.9	1	10	1	0.001	0.009	0.116	0.874	0.8	1	0	0.008	0.125	0.867
50	0.9	0.9	10	0.936	0.008	0.008	0.058	0.926	0.8	0.983	0.005	0.011	0.06	0.924
100	0.9	0.9	10	1	0.082	0.058	0.230	0.630	0.8	1	0.091	0.055	0.225	0.629
250	0.9	0.9	10	1	0.938	0.006	0.032	0.024	0.8	1	0.942	0.007	0.031	0.020
50	0.9	0.8	10	0.929	0.041	0.021	0.105	0.833	0.8	0.979	0.041	0.023	0.108	0.828
100	0.9	0.8	10	1	0.698	0.031	0.117	0.154	0.8	1	0.699	0.028	0.113	0.160
250	0.9	0.8	10	1	1	0	0	0	0.8	1	1	0	0	0

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 15: Empirical size and power. One level shift, known break point ($\lambda_i = 0.5$) and one common factor ($N = 40$)

T	ρ_i	α	σ_F^2	$Z^e_{\hat{t}_{NT}}$	$ADF^d_{\hat{F}}$	ρ_i	$Z^e_{\hat{t}_{NT}}$	$ADF^d_{\hat{F}}$	ρ_i	$Z^e_{\hat{t}_{NT}}$	$ADF^d_{\hat{F}}$
50	1	1	0.5	0.050	0.058	0.9	1	0.059	0.8	1	0.060
100	1	1	0.5	0.053	0.053	0.9	1	0.058	0.8	1	0.055
250	1	1	0.5	0.046	0.051	0.9	1	0.051	0.8	1	0.053
50	1	0.9	0.5	0.042	0.121	0.9	1	0.128	0.8	1	0.138
100	1	0.9	0.5	0.049	0.275	0.9	1	0.324	0.8	1	0.316
250	1	0.9	0.5	0.047	0.837	0.9	1	0.948	0.8	1	0.948
50	1	0.8	0.5	0.042	0.282	0.9	1	0.303	0.8	1	0.319
100	1	0.8	0.5	0.049	0.695	0.9	1	0.782	0.8	1	0.803
250	1	0.8	0.5	0.050	0.981	0.9	1	1	0.8	1	1
50	1	1	1	0.041	0.057	0.9	1	0.059	0.8	1	0.060
100	1	1	1	0.050	0.058	0.9	1	0.053	0.8	1	0.056
250	1	1	1	0.050	0.049	0.9	1	0.048	0.8	1	0.053
50	1	0.9	1	0.041	0.119	0.9	1	0.137	0.8	1	0.128
100	1	0.9	1	0.054	0.292	0.9	1	0.307	0.8	1	0.308
250	1	0.9	1	0.042	0.889	0.9	1	0.949	0.8	1	0.953
50	1	0.8	1	0.039	0.304	0.9	1	0.310	0.8	1	0.316
100	1	0.8	1	0.048	0.748	0.9	1	0.797	0.8	1	0.798
250	1	0.8	1	0.053	0.994	0.9	1	1	0.8	1	1
50	1	1	10	0.048	0.058	0.9	1	0.060	0.8	1	0.057
100	1	1	10	0.054	0.057	0.9	1	0.054	0.8	1	0.052
250	1	1	10	0.053	0.045	0.9	1	0.049	0.8	1	0.052
50	1	0.9	10	0.038	0.113	0.9	1	0.122	0.8	1	0.130
100	1	0.9	10	0.046	0.288	0.9	1	0.287	0.8	1	0.296
250	1	0.9	10	0.049	0.941	0.9	1	0.944	0.8	1	0.951
50	1	0.8	10	0.038	0.289	0.9	1	0.291	0.8	1	0.290
100	1	0.8	10	0.045	0.791	0.9	1	0.790	0.8	1	0.793
250	1	0.8	10	0.044	1	0.9	1	0.999	0.8	1	1

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 16: Empirical size and power with three common factors. One level shift, known break point ($\lambda = 0.5$, $N = 40$)

T	ρ_i	α	σ_F^2	$Z^e_{\hat{t}_{NT}}$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	1	1	0.5	0.082	0.011	0.179	0.349	0.461
100	1	1	0.5	0.064	0.002	0.039	0.196	0.763
250	1	1	0.5	0.063	0.001	0.013	0.130	0.856
50	1	0.9	0.5	0.117	0.032	0.137	0.332	0.499
100	1	0.9	0.5	0.070	0.047	0.061	0.206	0.686
250	1	0.9	0.5	0.052	0.653	0.079	0.117	0.151
50	1	0.8	0.5	0.126	0.077	0.104	0.274	0.545
100	1	0.8	0.5	0.055	0.361	0.104	0.184	0.351
250	1	0.8	0.5	0.054	0.930	0.066	0.004	0
50	1	1	1	0.050	0	0.001	0.034	0.965
100	1	1	1	0.056	0.001	0.004	0.066	0.929
250	1	1	1	0.051	0.001	0.009	0.092	0.898
50	1	0.9	1	0.039	0.002	0.006	0.052	0.940
100	1	0.9	1	0.042	0.039	0.042	0.157	0.762
250	1	0.9	1	0.042	0.770	0.038	0.089	0.103
50	1	0.8	1	0.034	0.014	0.015	0.071	0.900
100	1	0.8	1	0.036	0.408	0.080	0.179	0.333
250	1	0.8	1	0.047	0.989	0.011	0	0
50	1	1	10	0.054	0.001	0.001	0.020	0.976
100	1	1	10	0.054	0	0.004	0.060	0.935
250	1	1	10	0.051	0	0.009	0.093	0.898
50	1	0.9	10	0.038	0.003	0.005	0.038	0.950
100	1	0.9	10	0.046	0.047	0.046	0.166	0.74
250	1	0.9	10	0.050	0.855	0.019	0.055	0.071
50	1	0.8	10	0.032	0.013	0.013	0.071	0.896
100	1	0.8	10	0.044	0.486	0.070	0.152	0.291
250	1	0.8	10	0.048	1	0	0	0

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 17: Empirical size and power with three common factors. One level shift, known break point ($\lambda = 0.5$, $N = 40$)

T	ρ_i	α	σ_F^2	$Z^e_{\hat{t}_{NT}}$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	ρ_i	$Z^e_{\hat{t}_{NT}}$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	0.9	1	0.5	0.092	0.010	0.184	0.346	0.460	0.8	0.089	0.007	0.184	0.356	0.453
100	0.9	1	0.5	0.063	0.002	0.031	0.187	0.780	0.8	0.063	0.001	0.031	0.191	0.777
250	0.9	1	0.5	0.060	0	0.014	0.138	0.848	0.8	0.057	0	0.011	0.131	0.858
50	0.9	0.9	0.5	0.111	0.033	0.141	0.308	0.518	0.8	0.118	0.034	0.139	0.318	0.509
100	0.9	0.9	0.5	0.058	0.045	0.061	0.215	0.679	0.8	0.066	0.046	0.060	0.222	0.672
250	0.9	0.9	0.5	0.055	0.642	0.086	0.125	0.147	0.8	0.052	0.654	0.077	0.127	0.142
50	0.9	0.8	0.5	0.113	0.073	0.101	0.281	0.545	0.8	0.115	0.068	0.102	0.282	0.548
100	0.9	0.8	0.5	0.056	0.351	0.103	0.186	0.360	0.8	0.068	0.352	0.096	0.186	0.366
250	0.9	0.8	0.5	0.049	0.935	0.062	0.003	0	0.8	0.053	0.935	0.061	0.004	0
50	0.9	1	1	0.060	0	0.002	0.036	0.962	0.8	0.061	0	0.003	0.036	0.961
100	0.9	1	1	0.052	0	0.005	0.070	0.925	0.8	0.060	0.001	0.003	0.071	0.925
250	0.9	1	1	0.054	0	0.006	0.085	0.909	0.8	0.052	0	0.006	0.090	0.905
50	0.9	0.9	1	0.035	0.002	0.005	0.051	0.942	0.8	0.032	0.001	0.005	0.046	0.948
100	0.9	0.9	1	0.040	0.037	0.043	0.157	0.763	0.8	0.044	0.041	0.040	0.163	0.756
250	0.9	0.9	1	0.051	0.747	0.042	0.100	0.111	0.8	0.048	0.752	0.046	0.090	0.112
50	0.9	0.8	1	0.030	0.013	0.013	0.073	0.901	0.8	0.032	0.011	0.014	0.074	0.901
100	0.9	0.8	1	0.038	0.405	0.081	0.167	0.347	0.8	0.042	0.409	0.083	0.175	0.333
250	0.9	0.8	1	0.045	0.987	0.012	0.001	0	0.8	0.046	0.987	0.013	0	0
50	0.9	1	10	0.053	0.002	0.001	0.019	0.978	0.8	0.054	0.005	0.002	0.024	0.969
100	0.9	1	10	0.055	0	0.003	0.067	0.93	0.8	0.055	0.001	0.005	0.072	0.922
250	0.9	1	10	0.053	0	0.006	0.089	0.906	0.8	0.052	0	0.008	0.084	0.908
50	0.9	0.9	10	0.035	0.007	0.005	0.038	0.95	0.8	0.036	0.008	0.006	0.038	0.948
100	0.9	0.9	10	0.043	0.050	0.042	0.172	0.736	0.8	0.042	0.049	0.041	0.163	0.747
250	0.9	0.9	10	0.045	0.857	0.022	0.055	0.066	0.8	0.048	0.850	0.022	0.062	0.066
50	0.9	0.8	10	0.033	0.023	0.015	0.071	0.891	0.8	0.032	0.022	0.014	0.065	0.899
100	0.9	0.8	10	0.043	0.485	0.070	0.151	0.294	0.8	0.041	0.479	0.069	0.154	0.298
250	0.9	0.8	10	0.049	0.999	0.001	0	0	0.8	0.046	1	0	0	0

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Figure 1. Investment and saving for some countries

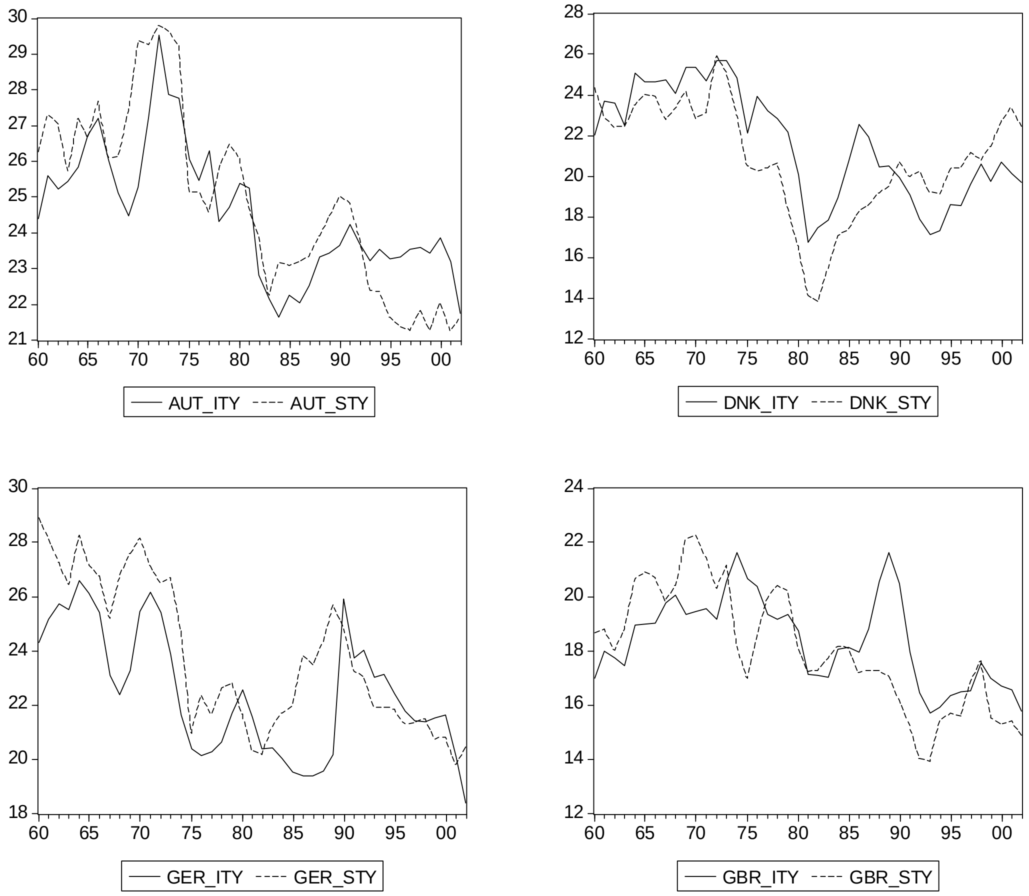

Table 18: Panel cointegration statistics

Pedroni model (individual effects)

	Test	p-val	Bootstrap distribution 1%	2.5%	5%	10%
$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	-1.167	0.122	-4.203	-3.652	-3.170	-2.678
$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$	-2.021	0.022	-4.818	-4.188	-3.539	-2.827

Model 1 (level shift)

	Test	p-val	Bootstrap distribution 1%	2.5%	5%	10%
$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	-1.973	0.024	-2.661	-2.240	-1.989	-1.644
$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$	-4.011	0.000	-4.746	-4.179	-3.719	-3.275

Model 4 (level and cointegrating vector shift)

	Test	p-val	Bootstrap distribution 1%	2.5%	5%	10%
$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$	-1.937	0.026	-3.019	-2.547	-2.161	-1.729
$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$	-2.999	0.001	-5.257	-4.693	-4.033	-3.408

The bootstrap is based on 2,000 replications.